The effects of syllable structure on consonantal timing and vowel compression in child and adult speakers of German

Sandra Michaela Peters

München 2015

The effects of syllable structure on consonantal timing and vowel compression in child and adult speakers of German

Sandra Michaela Peters

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

vorgelegt von
Sandra Michaela Peters
aus München

München, den 24.03.2015

Erstgutachter: Prof. Dr. Jonathan Harrington
Zweitgutachter: Prof. Dr. Philip Hoole
Tag der mündlichen Prüfung: 03.07.2015

Contents

List of Figures

List of Tables

Zusammenfassung auf Deutsch

Einleitung

Die vorliegende Dissertation beinhaltet experimentelle Untersuchungen, welche die Effekte des Konsonantentimings in der Sprachproduktion untersuchen. Dies bezieht sich sowohl auf komplexe Silben, welche Konsonantverbindungen (Cluster) entweder im Silbenonset (CCVC) oder in der Silbencoda (CVCC) beinhalten, als auch auf einfache Silben, welche nur einen Konsonanten (Singleton) in beiden Silbenpositionen (CVC) enthalten. Das Hauptaugenmerk liegt hierbei auf der Tatsache, dass sich Cluster von Singletons hinsichtlich ihres Timings mit dem Vokal unterscheiden. Laut Artikulatorischer Phonologie (Browman und Goldstein 1986; 1988; 1990) werden Cluster im Silbenonset global mit dem Vokal gekoppelt. Dies bedeutet, dass hinzugefügte Konsonanten im Onset sich in den Vokal hinein verschieben. Es entsteht somit ein stabiler *C-center*, d.h. der Mittelpunkt zweier Konsonanten in komplexen Onsets ist gleich dem konsonantischen Mittelpunkt eines Konsonanten in einfachen Onsets. Cluster in der Silbencoda hingegen werden sequentiell angeordnet. Dies bedeutet, dass hinzugefügte Konsonanten aneinander gereiht werden, was wiederum nicht zu einer Verschiebung zum Vokal hin führt (Marin und Pouplier, 2010). Durch dieses unterschiedliche Timing von Onset- und Codaclustern, macht die Artikulatorische Phonologie direkte Aussagen über die Vokaldauer einer Silbe mit einer ansteigenden Anzahl von Konsonanten: da sich Onsetcluster in den Vokal hinein verschieben, kommt es zu einer kürzeren akustischen Vokaldauer in CCVC Silben im Vergleich zu CVC Silben. Eine solche Vokalkürzung (im Folgenden kompensatorische Vokalkürzung) kann laut den Annahmen der artikulatorischen Phonologie jedoch nicht durch Codacluster ausgelöst werden, da es bei CVCC Silben eine sequentielle Anordnung der Konsonanten gibt, so dass die akustische Vokaldauer gleich bleibt (Nam et al., 2009).

Ein solch unterschiedliches Timing von Onset- und Codaclustern wurde für viele Sprachen gefunden, z.B. für das Rumänische (Marin, 2013) für das Englische (Marin und Pouplier, 2010) und für das Deutsche (Pouplier, 2012). Jedoch zeigen neuere Studien, dass eine C-center Organisation nicht generell für alle Onsetcluster zutrifft und dass sich Sprachen hinsichtlich des (C)CV Timings unterscheiden (siehe Hermes et al., (2008; 2012) für *'impure-s'* Cluster im Italienischen; Shaw et al., (2009) für Marokkanisch Arabisch sowie Goldstein et al., (2007a) und Hermes et al., (2011) für Tashlhiyt Berber). Des Weiteren zeigte Byrd (1995), dass auch Codacluster (zumindest bei manchen Sprechern) eine C-center Organisation aufweisen können. Eine solche C-center Organisation wurde auch für /lp/ und /lk/ Cluster in der Silbencoda gefunden (Marin und Pouplier, 2010). Hinzu kommt, dass es akustische Studien gibt, welche kompensatorische Vokalkürzung bedingt durch Codacluster aufweisen, z.B. für das Holländische (Waals, 1992), für das Englische (Katz, 2012; Munhall et al., 1992), für das Rumänische (Marin und Shigemori, 2014) und für das Schwedische (Lindblom und Rapp, 1973). Jedoch gibt es bezüglich der codabedingten Vokalkürzung einige unterschiedliche Ergebnisse, welche von der Sprache, der Artikulationsart des darauf folgenden Codakonsonanten und von dem Vokal selbst abhängen. Im Englischen fanden manche Forscher beispielsweise keine Anzeichen für kompensatorische Vokalkürzung (Crystal und House, 1982; 1988; 1990), andere allerdings schon (Katz, 2008, 2012). Außerdem gibt es Studien, welche kompensatorische Vokalkürzung nur in Vokalen gefunden haben, welche Sonoranten vorausgehen, jedoch nicht Obstruenten (Katz, 2008, 2012; Marin und Pouplier 2010), wohingegen andere Studien durchaus kompensatorische Vokalkürzung in Wörtern mit Obstruenten in der Coda gefunden haben (Munhall et al., 1992; Shaiman, 2001). Zudem bewirken Laterale in der Coda eine starke kompensatorische Vokalkürzung von phonologisch langen Vokalen im Holländischen, wohingegen Obstruenten in der Coda eine geringere Kürzung auslösen (Waals, 1992).

Wie schon erwähnt, scheint eine globale Organisation von Onsetclustern von gewissen Faktoren abzuhängen. Eine artikulatorische Studie zur Silbenstruktur im Rumänischen (Marin, 2013) hat beispielsweise gezeigt, dass die Artikulationsart entschieden dazu beträgt, ob ein Onsetcluster eine C-center Organisation aufweist oder nicht. So zeigte sich, dass das Onsetcluster /sk/ eine C-center Organisation aufweist, nicht aber das Onsetcluster /ps/. Diese clusterspezifischen Unterschiede könnten auch durch Unterschiede im interkonsonantalen Timing bedingt sein: Bombien et al. (2013) und Hoole et al. (2009) zeigten beispielsweise, dass die Gesten des Onsetclusters /kl/ mehr überlappen als die Gesten des Onsetclusters /kn/. Brunner et al. (2014) hat wiederum angemerkt (in Bezug auf Hermes et al., 2008), dass Vokalhöhe und Vokallänge einen Einfluss auf die Organisation von Onsetclustern haben. Des Weiteren zeigten Studien zum Spracherwerb, dass Kinder Codacluster eher erwerben als Onsetcluster (Levelt et al., 2000; Lleó und Prinz 1996). Der spätere Erwerb von Onsetclustern könnte artikulatorische Gründe haben aufgrund des unterschiedlichen Timings von Onset- und Codaclustern (siehe auch Nam et al., 2009).

Das Ziel dieser Arbeit war, die Effekte der Silbenstruktur auf das Konsonantentiming und auf die Vokalkomprimierung zu untersuchen. Kapitel 2 richtet sich an kompensatorische Vokalkürzung bedingt durch Codacluster sowohl in der Sprachproduktion, als auch in der Perzeption. Unter der Benutzung von EMA (Elektromagnetische Artikulographie) befasst sich Kapitel 3 mit kompensatorischer Vokalkürzung – welche sowohl durch Onsetcluster (/kn/ and /kl/), als auch durch Codacluster (/pt/ and /mt/) bedingt ist – und untersucht mögliche zugrunde liegende artikulatorische Mechanismen. Kapitel 4 untersucht die Stabilität des Gestentimings innerhalb Silben unter prosodischen Einflüssen. Diese Studie beinhaltet sowohl komplexe und einfache (/kn/ und /n/) Onsets, als auch gespannte und ungespannte Vokale (/aː/ and /a/) in akzentuiertem und deakzentuiertem Kontext. Zu guter Letzt untersucht Kapitel 5 die Silbenorganisation während des Spracherwerbs, wobei sowohl das Konsonant-zu-Konsonant Timing, als auch das Konsonant-zu-Vokal Timing im Silbenonset (/kl/, /kn/, /pl/) und in der Silbencoda (/kl/, /kn/, /ks/) miteinbezogen werden.

Kompensatorische Vokalkürzung in Produktion und Perzeption

Ziel der durchgeführten Studie in Kapitel 2 war festzustellen, ob es kompensatorische Vokalkürzung bedingt durch Codacluster auch im Deutschen gibt und wie Hörer eine solche Kürzung wahrnehmen. Hierfür wurden einsilbige Echtwörter von zehn Sprechern des (Standard-)Deutschen sowohl in akzentuiertem, als auch in deakzentuiertem Kontext aufgenommen. Die Zielwörter beinhalteten gespannte und ungespannte Vokale mit komplexen Onsets. Hierfür wurden die Onsetcluster /kl/ und /kn/ gewählt, da zuvor gezeigt wurde, dass sich eben diese Cluster hinsichtlich ihres interkonsonantalen Timings unterscheiden (Bombien et al., 2013; Hoole et al., 2009), was sich wiederum auf die Dauer des darauf folgenden Vokals auswirken könnte. Die Zielwörter hatten entweder eine einfache oder eine komplexe Coda (z.B. /knal/ - /knalt/ oder /klɪk/ - /klɪkt/).

Die Hypothesen waren, dass Codacluster generell kompensatorische Vokalkürzung auslösen (**H1**) und dass dieser Effekt stärker in Wörtern mit sonoranter Coda, in Wörtern mit gespannten Vokalen und in akzentuierten Wörtern (**H2**) ist. Hypothese **H2** ergab sich aus den Ergebnissen früherer Studien (z.B. Katz, 2012; Waals, 1992), welche vermuten lassen, dass überwiegend längerer Segmente komprimiert werden.

Die akustische Analyse des Produktionsexperiments zeigte, dass es generell nur eine sehr geringe Vokalkürzung gibt. Dennoch gab es eine Tendenz zur Vokalkürzung (1) in Wörtern mit ungespannten Vokalen, welche Sonoranten vorausgehen und (2) in Wörtern mit gespannten Vokalen, welche jeweils in akzentuiertem Kontext produziert wurden. Zudem scheint eine codebedingte Kürzung sprecherabhängig zu sein: manche Sprecher kürzten den Vokal, andere hingegen nicht. Hypothesen **H1** und **H2** konnten also nur zu einem gewissen Maß bestätigt werden. Außerdem zeigte sich, dass der dem Vokal folgende Codakonsonant (C1) gleich gehalten werden muss (bei dem Wortpaar /kni:t, kni:st/ kam es zu einer Vokallängung in komplexer Coda, welche möglicherweise auf den unterschiedlichen Codakonsonanten zurückzuführen sein könnte). Zusätzlich zeigte die akustische Analyse eine Tendenz zur C1 Kürzung. Dieses Ergebnis könnte mitunter ein Grund dafür sein, dass Vokalkürzung weniger stark ausgeprägt war.

Sowohl der Effekt von Vokalkürzung, als auch der Effekt von C1 Kürzung wurde in einem Perzeptionsexperiment getestet. Dieses beinhaltete natürliche Stimuli aus dem eben beschriebenen Produktionsexperiment und wurde mit 21 Hörern durchgeführt. Hierbei wurde den Hörern jeweils ein solches Minimalpaar (z.B. /knɪk/ vs. /knɪkt/) präsentiert, welches entweder Vokalkürzung und C1 Kürzung aufwies, oder sich in der Dauer dieser beiden Silbenkomponenten nicht unterschied. Die Aufgabe der Hörer war, zu entscheiden, in welchem der beiden Wörter der Vokal länger sei. Die Hypothese war, dass Hörer für kompensatorische Kürzung kompensieren (**H3**) und dies zu einem größeren Ausmaß in Kontexten, welche Vokalkürzung begünstigen (**H4**). Wenn Hörer für diese Kürzung kompensieren, dann nehmen sie gekürzte Vokale als länger wahr in Wörtern mit Codaclustern. Die Ergebnisse zeigten jedoch, dass die Hörer generell nicht für kompensatorische Kürzung kompensiert haben, da sie in etwa zwei Drittel aller Fälle den Vokal in Wörtern mit komplexen Codas immer als kürzer wahrgenommen haben, ganz gleich ob eine Kürzung vorlag oder nicht. Aus diesem Grund weisen die Ergebnisse vielmehr auf eine perzeptive Kürzung des Vokals hin und widersprechen Hypothese **H3**. Jedoch zeigte sich eine Tendenz zur Kompensation in akzentuierten Wörtern mit sonoranter Coda und in Wörtern mit gespannten Vokalen. Dies wiederum befürwortet zum Teil Hypothese **H4**. Daher zeigte diese Studie, dass in Kontexten, welche grundsätzlich mehr Kürzung begünstigen, auch tendenziell kompensiert wird.

Zusammenfassend lässt sich sagen, dass längere Segmente in größerem Maße akustische gekürzt werden (siehe auch Turk und Shattuck-Hufnagel, 2000; White und Turk, 2010), und dass Hörer in genau diesen Kontexten zur Kompensation tendieren.

Artikulatorische Mechanismen, welche kompensatorischer Vokalkürzung zugrunde liegen

In Kapitel 3 sollte die Frage geklärt werden, ob Vokalkürzung sowohl durch Codacluster, als auch durch Onsetcluster gefunden werden kann. Die Hypothese war daher, dass Vokale in komplexen Silben akustisch kürzer sind als in einfachen Silben (**H1**). Sollten solche Kürzungen gefunden werden, stellt sich die Frage nach den zugrunde liegenden artikulatorischen Mechanismen. Um diese Frage zu klären, wurden zwei in der Literatur vorgeschlagene zeitliche (Katz, 2012) und drei räumliche Mechanismen untersucht. Die zeitlichen Mechanismen umfassen eine erhöhte Überlappung des Vokals mit dem benachbarten Konsonanten und eine Komprimierung der Vokalgeste (Katz, 2012) in komplexen Silben verglichen mit einfachen Silben. Die räumlichen Mechanismen umfassen eine größere Steifheit der Vokalgeste (*stiffness*) sowie koartikulatorische Einflüsse sowohl auf die vertikale Zungenposition (*target undershoot*), wie auch auf die vertikale Kieferposition. Katz (2012) nimmt an, dass die Vokalgeste in komplexen Silben im Vergleich zu einfachen Silben komprimiert ist (**H2**). Eine komprimierte Geste resultiert dann in kürzerer akustische Dauer. Des Weiteren nimmt Katz (2012) an, dass es zu mehr CV Überlappung in Wörtern mit komplexen Onsets und zu mehr VC Überlappung in Wörtern mit komplexer Coda kommt (**H3**). Eine solch größere Überlappung zwischen Vokal und Konsonant in komplexen verglichen mit einfachen Onsets beziehungsweise Codas, führt dazu, dass der Vokal maskiert wird und somit akustisch kürzer ist. Im Sinne der Artikulatorischen Phonologie (Browman und Goldstein, 1990, 1992) ist eine erhöhte CV Überlappung in komplexen Onsets generell zu erwarten, nicht jedoch eine erhöhte VC Überlappung in komplexen Codas. Zusätzlich zu diesen beiden zeitlichen Mechanismen könnten allerdings auch Unterschiede in der Stiffness der Vokalgeste akustische Vokalkürzung bedingen. Stiffness spiegelt die relative Geschwindigkeit wieder, die zur Ausführung einer Geste genutzt wird, unabhängig von ihrer physikalischen Auslenkung im Vokaltrakt (Guenther, 1995; Ostry und Munhall, 1985). Bisherige Studien haben gezeigt, dass kürzere Segmente eine höhere Stiffness aufweisen als längere Segmente (Löfqvist, 2004; Ostry und Munhall, 1985). Daher wurde angenommen, dass die Vokalgeste in komplexen Silben mit einer größeren Stiffness produziert wird, was in einem akustisch kürzerem Segment resultiert (**H4**). Viele Arten akustischer Kürzung können zudem auf *vowel target undershoot* zurückgeführt werden, d.h. die Geste eines Vokals erreicht nicht ihre Zielposition. Stevens und House (1963) zeigten, dass ungespannte Vokale eher zu undershoot neigen als gespannte Vokale und dies sei auch der Grund dafür, dass ungespannte Vokale akustisch kürzer sind als gespannte Vokale. Des Weiteren wird vowel undershoot häufig in hypoartikulierenden Kontexten – wie beispielsweise schnellem Sprachtempo oder in deakzentuierten Kontexten – gefunden (Guenther, 1995;

Moon und Lindblom, 1994). In solchen Kontexten wird stärker koartikuliert (Lindblom, 1990) und vor allem die Vokaldauer ist akustisch gekürzt. Daher wurde angenommen, dass es in komplexen Silben zu vowel target undershoot (d.h. zu einer weniger weit abgesenkten Zungenposition während des Vokals) kommt, resultierend in kürzeren akustischen Dauern (**H5**). Außerdem scheint auch die Kieferposition eine wichtige Rolle in Bezug auf Segmentdauern zu spielen. Munhall et al. (1992) erklären ihre codabedingte Vokalkürzung mit einer Trunkierung der Vokalgeste durch die Kiefergeste. Eine solche Trunkierung und eine höhere Kieferposition erklären zudem kürzere Vokaldauern in deakzentuierten Kontexten (siehe Beckman et al., 1992; Harrington et al., 1995). Eine höhere Kieferposition während des Vokals in komplexen Silben wurde daher erwartet, was wiederum in einer akustisch kürzeren Vokaldauer resultiert (**H6**).

Das Sprachmaterial in Kapitel 3 beinhaltete überwiegend Nichtwörter, mit einfachen und komplexen Onsets, sowie einfachen und komplexen Codas. Für onsetbedingte Kürzung wurde die akustische Vokaldauer in den Wortpaaren /knaːp/ - /naːp/, /klaːp/ - /laːp/ und /knaːp/ - /naːp/ berechnet; für codabedingte Kürzung in den Wortpaaren /naːp/ - /naːpt/, /laːp/ - /laːpt/ oder /laːm/ - /laːmt/. Ein solches Sprachmaterial zu wählen erschien wichtig, um mögliche andere Einflüsse auf die akustische Vokaldauer auszuschließen. Die Aufnahmen erfolgten unter der Nutzung von Elektromagnetischer Artikulographie (EMA). Hierfür wurden Sensoren auf Zungenspitze, Zungenrücken und den Lippen der Sprecher angebracht, um Überlappungen sowohl zwischen dem Vokal und den umgebenden Konsonanten zu messen (für CV und VC Überlappung in Abhängigkeit von Komplexität), als auch zwischen den Konsonanten innerhalb der zwei mit einbezogenen Onsetcluster /kn/ und /kl/. Grund hierfür waren die Ergebnisse vorheriger Studien von Bombien et al. (2013) und Hoole et al. (2009), welche mehr Überlappung zwischen /k/ und /l/ als zwischen /k/ und /n/ gefunden haben. Solche interkonsonantalen Überlappungsunterschiede können sich durchaus auf die CV Überlappung und damit auf die akustische Vokaldauer auswirken (siehe beispielsweise Marin, 2013). Die oben genannten Zielwörter waren in Trägersätzen eingebettet, welche für eine optimale Messung aller Zungenbewegungen innerhalb des Zielwortes ausgewählt wurden. Aus syntaktischen Gründen mussten zwei unterschiedliche

Trägersätze gewählt werden. „Melanie's Omi [Zielwort] imitiert ein Lied." wurde für alle Onsettokens und substantivähnlichen Codatokens genutzt. Da Zielwörter mit komplexer Coda ein Silbenfinales -t beinhalteten, welches im Deutschen ein Verb kennzeichnet, wurde folgender Trägersatz für diesen Teil der Zielwörter genutzt: „Melanie's Omi [Zielwort] ihm einmal". Die /i/-Umgebung wurde gewählt, um eine aktive Zungen- und Kieferbewegung während des Zielvokals messen zu können. Analysiert wurden folgende Parameter: die akustischen Vokaldauern, der CC, CV und VC lag, die Vokalplateaudauern, die Stiffness, die vertikale Zungen- und Kieferposition. Diese Variablen wurden jeweils in Abhängigkeit von Onset- bzw. Codakomplexität gemessen.

Die akustische Analyse zeigte, dass sowohl Onset-, als auch Codacluster kompensatorische Vokalkürzung bedingen. Hypothese **H1** konnte somit bestätigt werden.

Was die zugrunde liegenden Mechanismen betrifft, werden zunächst die Ergebnisse zu den beiden zeitlichen Parametern beschrieben, gefolgt von den drei räumlichen Parametern. Die Dauer des Vokalplateaus änderte sich weder durch eine zunehmende Anzahl von Onsetkonsonanten, noch durch eine zunehmende Anzahl von Codakonsonanten. Daraus kann geschlossen werden, dass eine Komprimierung der Vokalgeste nicht der Auslöser für kompensatorische Vokalkürzung sein kann. Hypothese **H2** konnte somit nicht bestätigt werden. Bezüglich der Gestenüberlappung, wurden zunächst die Ergebnisse von Bombien et al. (2013) und Hoole et al. (2009) repliziert: es gab mehr CC Überlappung innerhalb des Onsetclusters /kl/ als innerhalb des Onsetclusters /kn/. Ausgehend von diesem Ergebnis könnte eine geringere Verschiebung zum Vokal in komplexen Onsets bei /kl/ vs. /l/ erwartet werden, da bereits eine größere Konsonant-zu-Konsonant Überlappung vorliegt. Diese Überlegung konnte bestätigt werden: es gab nur eine Verschiebung hin zum Vokal in /kn/ vs. /n/ (d.h. mehr CV Überlappung in Wörtern mit komplexen Onsets als in Wörtern mit einfachen Onsets), nicht jedoch bei /kl/ vs. /l/. Diese Verschiebung bei /kn/ deutet auf eine C-center Organisation hin und könnte somit die akustische Vokalkürzung bedingen. Die nicht vorhandene Verschiebung bei /kl/ könnte darauf zurückzuführen sein, dass der Lateral eher an den velaren Konsonant als an den darauf folgenden Vokal gekoppelt ist. Bezüglich der VC Überlappung bei Codaclustern zeigte sich, dass es bei /l/-Wörtern

(/laːp/ - /laːpt/ und /laːm/ - /laːmt/) zu einer Verschiebung des Codakonsonanten hin zum Vokal kommt. Eine solche Verschiebung konnte jedoch nicht bei /n/-Wörtern (/naːp/ - /naːpt/) gefunden werden, obwohl der Codakonsonant stets ein Labial (/p/ und /m/) war. Die gefundene erhöhte VC Überlappung in komplexen Codas mit /l/ im Onset könnte ein artikulatorischer Mechanismus sein, der zur kompensatorischen Vokalkürzung führt. Somit konnte Hypothese **H3** nur teilweise für beide Silbenpositionen bestätigt werden. Sowohl die nicht vorhandene erhöhte CV Überlappung in /kl/, als auch die beobachtete Verschiebung hin zum Vokal bei einem Teil der Codaclustern widersprechen den Annahmen der Artikulatorischen Phonologie.

Die Analyse zur Stiffness zeigte, dass ausschließlich komplexe Onsets einen Einfluss auf diesen Parameter haben. Vokale in Silben mit komplexen Onsets wurden deutlich schneller produziert, als in einfachen Silben. Hypothese **H4** traf also für Onsets zu, nicht jedoch für Codas. Dieser Unterschied in der Vokalstiffness bei Onsets könnte die gefundene kompensatorische Vokalkürzung ausgelöst haben.

Wie oben erwähnt, wurde auch die vertikale Zungen- und Kieferposition während des Vokals gemessen. Erneut zeigten sich Unterschiede zwischen Onsets und Codas. Ein zusätzlicher Konsonant im Silbenonset führte zu einer geringeren Absenkung des Zungenrückens während des Vokals /aː/, d.h. es kam zu vowel target undershoot. Es waren keine Unterschiede in der Kieferposition zu erkennen. Was komplexe Codas betrifft, so zeigten die Ergebnisse sowohl eine geringere Absenkung der Zunge, als auch des Kiefers, verglichen mit einfachen Codas. Diese Unterschiede in Zungen- und Kieferposition können codabedingte Vokalkürzung erklären. Hypothese **H5**, welche besagte, dass es zu vowel target undershoot in komplexen Silben kommt, konnte somit für Onsets und Codas bestätigt werden.[1] Hypothese **H6**, welche besagte, dass der Kiefer weniger abgesenkt ist in Wörtern mit komplexen Onsets und Codas, konnte nur für Codas bestätigt werden. Onsetcluster scheinen hingegen direkt mit der Vokalgeste verbunden zu sein, unabhängig von der Kieferposition.

[1] Es ist allerdings nicht ganz klar, inwiefern die beobachtete geringere Zungenposition in komplexen Codas nur ein Nebenprodukt der geringeren Kieferöffnung ist.

Die Ergebnisse dieser Studie haben gezeigt, dass die Annahmen der Artikulatorischen Phonologie nicht grundsätzlich zutreffen: nicht alle Onsetcluster sind global und nicht alle Codacluster sind sequentiell mit dem Vokal getimed. Des Weiteren scheinen unterschiedliche artikulatorische Mechanismen zur kompensatorischen Vokalkürzung zu führen. Onsetbedingte Vokalkürzung ist am Wahrscheinlichsten durch eine erhöhte Stiffness der Vokalgeste und durch vowel target undershoot bedingt. Codabedingte Kürzung wiederum ist am ehesten durch die geringere Absenkung des Kiefers während des Vokals zu erklären. Zudem zeigte diese Studie Unterschiede zwischen Silbenonsets und Silbencodas nicht nur hinsichtlich ihres Timingverhaltens, sondern auch hinsichtlich ihres räumlichen Verhaltens. Erstens wurde nur für das Onsetcluster /kn/ eine Verschiebung hin zum Vokal gefunden, nicht jedoch für das Onsetcluster /kl/. Für Cluster in der Silbencoda wurde eine Verschiebung hin zum Vokal gefunden, allerdings nur für Wörter mit /l/ im Onset, nicht jedoch für Wörter mit /n/ im Onset. Zweitens beeinflussten nur Onsetcluster die Stiffness der Vokalgeste, wohingegen Codacluster keinen Einfluss auf diesen Parameter hatten. Drittens beeinflussten nur Codacluster die Position des Kiefers während des Vokals, nicht jedoch Onsetcluster.

Zusammenfassend lässt sich sagen, dass zusätzliche Konsonanten im Silbenonset vor allem die Zungenbewegung des folgenden Vokals zu beeinflussen scheinen: die Vokalgeste wurde mit einer größeren Stiffness produziert und der Zungenrücken wurde weniger abgesenkt, d.h. es kam zu koartikulationsbedingtem undershoot bei Wörtern mit Onsetclustern. Zusätzliche Konsonanten in der Silbencoda hingegen scheinen vornehmlich mit der Kieferbewegung während des Vokals zusammen zu hängen. Das bedeutet, dass der Kiefer eine wichtige und aktive Rolle in codabedingter Vokalkürzung spielt (siehe hierzu auch Munhall et al., 1992). Die Ergebnisse dieses Kapitels erzielten weitere Evidenzen dafür, dass vermutlich nur Onsetcluster mit dem Vokal gekoppelt sind, wie es im Coupled Oscillator Model (Goldstein et al., 2007a; Nam und Saltzman, 2003) angenommen wird. Speziell die aktive Rolle des Kiefers bei Codaclustern muss in dem Coupled Oscillator Model und in der Artikulatorischen Phonologie berücksichtigt werden.

Stabilität des Gestentimings

In Kapitel 4 sollte der Einfluss von Kürzungsfaktoren auf das Silbentiming geprüft werden. Dies umfasste einerseits den Effekt von Deakzentuierung, andererseits den Effekt von ungespannten Vokalen. Deakzentuierte Kontexte bedingen Kürzung von Segmenten (Cho, 2004; Lehiste, 1970; Pierrehumbert und Talkin, 1992) und im Sinne von Lindblom's (1990) H&H-Theorie kommt es in deakzentuierten Kontexten zu mehr Überlappung und undershoot. In dieser Studie sollte gezeigt werden, inwiefern sich die Timingverhältnisse innerhalb einer Silbe ändern, je nachdem ob diese in akzentuiertem oder deakzentuiertem Kontext produziert wurde. Des Weiteren sollte untersucht werden, ob eine Verschiebung hin zum Vokal in komplexen Onsets verstärkt oder verringert wird bedingt durch prosodische Schwächung. Aus diesem Grund wurden nur die Wortpaare mit /kn/ im Onset in die Analyse mit einbezogen, da für diese im vorherigen Kapitel (Kapitel 3) eine solche Verschiebung in akzentuiertem Kontext gefunden wurde. Des Weiteren sollte untersucht werden, ob eine solche Verschiebung hin zum Vokal auch in Wortpaaren mit einem ungespannten Vokal zu sehen ist. In einer Studie zur zeitlichen Organisation von Onsetclustern im Deutschen zeigten Brunner et al. (2014) beispielsweise, dass der C-center stabiler bei Wörtern mit gespannten Vokalen als mit ungespannten Vokalen oder Diphthongen ist. Aus diesem Grund wurden in der vorliegenden Studie Silben mit gespanntem /aː/ und Silben mit ungespanntem /a/ miteinbezogen.

Bezüglich prosodischer Einflüsse auf die Akustik und Artikulation gab es bislang einige Untersuchungen. Akustische Studien zeigten, dass vor allem Vokale gelängt werden, wenn sie mit einem Wortakzent assoziiert sind (Cho, 2004; Lehiste, 1970; Pierrehumbert und Beckman, 1988), aber auch Konsonanten können gelängt werden (Crystal und House, 1988; Ingrisano und Weiser, 1979; Klatt, 1974; Stathopoulus und Weiser, 1983, Umeda, 1977). Es wird daher angenommen, dass die Silbenkomponenten in ihrer akustischen Dauer kürzer sind in deakzentuierten Wörtern im Vergleich zu akzentuierten Wörtern (**H1**).

Artikulatorische Studien zeigten längere und stärkere (größerer linguopalataler Kontakt in EPG Studien) Gesten (Bombien et al., 2007; Byrd und Choi 2010; Cho, 2004, 2006) und weniger Gestenüberlappung (siehe z. B. Byrd, 2000; Byrd et al., 2000; Byrd und Saltzman, 1998; Cho, 2004) unter dem Einfluss höher prosodischer Grenzen. Mücke et al. (2008) zeigten jedoch, dass es weniger Überlappung in Wörtern, die mit Fokus produziert wurden, gab als in unakzentuierten Wörtern. Sie erklärten dies durch Plateaukürzungen im Gegensatz zu indirekter Kürzung ausgelöst durch erhöhte Überlappung. Der Einfluss von (De)akzentuierung auf die Gestenüberlappung ist also nicht ganz klar und es könnte einen Zusammenhang zwischen Überlappung und Plateaudauer geben. Es wurde daher angenommen, dass es in deakzentuierten Wörtern zu mehr Überlappung kommt (**H2**); kommt es jedoch zu mehr Überlappung in diesen Wörtern, so sollten die Plateaudauern gekürzt sein (**H3**). Des Weiteren wurde die Stiffness aller Silbenkomponenten gemessen, um zu sehen ob Gesten schneller produziert werden, wenn im deakzentuierten Kontext hypoartikuliert wird (**H4**).

Die akustische Analyse zeigte, dass alle Silbenkomponenten ($C1_{on}$, $C2_{on}$, der Vokal (V) und $C1_{off}$) durch Deakzentuierung gekürzt werden. Was die Vokale betrifft, so zeigte sich, dass nur gespannte Vokale gekürzt werden, nicht jedoch ungespannte Vokale (siehe z.B. Harrington et al., 2015; Mooshammer und Fuchs, 2002; Mooshammer und Geng, 2008, für ähnliche Ergebnisse). Somit konnte Hypothese **H1** bestätigt werden.

Die artikulatorischen Analysen zeigten, dass die einzelnen Silbenkomponenten auf unterschiedliche Art und Weise von prosodischer Schwächung beeinflusst werden. Während die CC und VC Überlappung in diesem Kontext größer wurde, so wurde die CV Überlappung geringer. Diese geringere Überlappung in CV Komponenten könnte dadurch bedingt sein, dass die Plateaus von $C2_{on}$ und V gekürzt wurden. Die Ergebnisse für größere Überlappung in deakzentuierten Wörtern in CC und VC Sequenzen bestätigten Hypothese **H2**, welche besagte, dass es zu mehr Gestenüberlappung im deakzentuiertem Kontext kommt. Was jedoch die geringere Überlappung in CV Sequenzen betrifft, so konnte Hypothese **H3** bestätigt werden, welche besagte, dass die Plateaudauern gekürzt sind, wodurch es zu weniger Überlappung im deakzentuiertem Kontext kommt. Sprecher scheinen also unterschiedliche Strategien anzuwenden, um akzentuierte und deakzentuierte Wörter zu differenzieren. Dies beinhaltet einerseits weniger Überlappung, andererseits aber auch längere Gestenplateaus in akzentuierten Wörtern.

Des Weiteren zeigte sich, dass ungespannte Vokale durchaus artikulatorisch komprimiert werden können, ohne akustisch gekürzt zu sein (siehe z.B. Mooshammer und Geng, 2008, für räumliche aber nicht zeitliche Komprimierung in ungespannten Vokalen).

Was den Einfluss von Deakzentuierung auf die Stiffness betrifft, so war nur $C1_{off}$ betroffen. Dieser wurde in deakzentuierten Wörtern mit einer erhöhten Stiffness produziert. Eine genauere Betrachtung der beiden Komponenten Geschwindigkeit und Auslenkung zeigte, dass es zu einer geringeren Auslenkung in deakzentuierten Wörtern kam. Dies könnte durch eine geringere Kieferöffnung im deakzentuierten Kontext bedingt sein.

Wenden wir uns nun dem Einfluss der Deakzentuierung auf das globale Timing des Onsetclusters /kn/ zu. Eine Verschiebung des $C2_{on}$ hin zum Vokal war in deakzentuierten Wörtern nicht mehr vorhanden. Der Einfluss von Deakzentuierung muss also in Modellen, welche die Silbenstruktur beschreiben – wie die Artikulatorische Phonologie – berücksichtigt werden.

Silbentiming während des Spracherwerbs

Bisherige Studien zum Spracherwerb haben gezeigt, dass Cluster im Silbenonset später erlernt und weniger akkurat produziert werden als Cluster in der Silbencoda (Levelt et al., 2000; Lleó und Prinz, 1996; Templin, 1957). Diese Befunde sind überraschend, da Singletons früher im Onset (CV Struktur) als in der Coda (VC Struktur) erworben werden (z. B. Fikkert, 1994; Levelt et al., 2000; Salidis und Johnson, 1997; Vihman und Ferguson, 1987). Des Weiteren ist die CV Struktur der einzig universelle Silbentypus in den Sprachen der Welt (Bell und Hooper, 1978), d.h. fast jede Sprache hat CV Silben[2]. Ein Grund für den späteren Erwerb von Onsetclustern könnte artikulatorisch bedingt sein. Denn Onset- und Codacluster unterscheiden sich hinsichtlich ihres Timings mit dem Vokal: laut dem Coupled Oscillator Model (Nam und Saltzman 2003; Nam et al., 2009; Goldstein et al., 2006), werden beide Konsonanten innerhalb eines Onsetclusters *in-phase* mit dem folgenden Vokal

[2]Eine Studie von Breen und Pensalfini (1999) allerdings zeigte, dass Arrernte keine Silbenonsets hat.

koordiniert und *anti-phase* miteinander. Diese konkurrierende Kopplung führt zur globalen Organisation von Onsetclustern im Sinne der Artikulatorischen Phonologie (z.B. Browman und Goldstein, 1986; 1988; 1990). Ein solch komplexes Timing von Onsetclustern könnte für Kinder schwieriger zu erlernen sein, da die *in-phase* Kopplung dominanter ist und sich die *anti-phase* Kopplung erst stabilisieren muss Nam et al. (2009). Dies ist nicht nötig bei Codaclustern, da hier die Gesten lediglich *anti-phase* zueinander koordiniert sind. Es wird also angenommen, dass Kinder kein globales Timing bei Onsetclustern aufweisen (**H1**).

Hinzu kommt, dass Onsetcluster zu einem geringeren Umfang miteinander überlappen, als Codacluster (Byrd, 1996; Chitoran et al., 2002; Hardcastle und Roach, 1979; Pouplier, 2012). Dies wurde darauf zurückgeführt, dass Onsetcluster perzeptiv wahrnehmbar sein müssen, da – im Gegensatz zu Codaclustern – kein anderer akustischer Hinweis zur perzeptiven Enkodierung vorhanden ist (Byrd, 1996; Chitoran et al., 2002). Eine größere Überlappung innerhalb Codacluster, deren beiden konsonantischen Komponenten dennoch von Erwachsenen wahrgenommen werden, könnte ein weiterer Grund sein, weshalb für diese Silbenposition ein früherer Clustererwerb beschrieben wird. Es wird also angenommen, dass Codacluster mehr überlappen als Onsetcluster (**H2**).

Um diese Hypothesen zu überprüfen, wurden Äußerungen von vier Kindern im Alter von 5 bis 6 Jahren aufgenommen. Wie bereits in den Kapiteln 3 und 4 erfolgten die Aufnahmen unter der Nutzung von Elektromagnetischer Artikulographie (EMA). Das Sprachmaterial umfasste abbildbare Wörter mit einfachen und komplexen Onsets (z.B. „Note” vs. „Knoten”). Die Onsetcluster waren /kl/, /kn/ und /pl/. Diese Cluster wurden ausgewählt, da zuvor bei Erwachsenen Sprechern unterschiedliche Timingmuster festgestellt wurden (Bombien, 2011; Bombien et al., 2013) und sich die Frage stellte, wie Kinder diese Cluster koordinieren. Zusätzlich wurden abbildbare Wörter mit einfacher oder komplexer Coda aufgenommen (z.B. „Sack” vs. „Dackel”). Die Codacluster waren /kl/, /kn/ und /ks/.

Die artikulatorischen Ergebnisse bezüglich Onsetcluster zeigten mehr Überlappung zwischen /k/ und /l/ als zwischen /k/ und /n/. Des Weiteren gab es tendentiell mehr Überlappung innerhalb /kl/ als innerhalb /pl/. Diese Ergebnisse sind konform mit bisherigen Ergebnissen zu Erwachsenen Sprechern (Bombien, 2011; Bombien et al., 2013; Hoole et al., 2009). Dies zeigt, dass das Konsonant-zu-Konsonant Timing bereits erlernt wurde. Des Weiteren zeigte sich, dass es keine Verschiebung hin zum Vokal gab in komplexen Onsets verglichen mit einfachen Onsets: das Timing zwischen beispielsweise /n/ und dem Vokal in „Note" und „Knoten" unterschied sich nicht. Hypothese **H1** konnte somit bestätigt werden. Die Tatsache, dass Kinder keine konkurrierende Kopplung bei Onsetclustern produzieren, könnte eine Erklärung dafür sein, dass Onsetcluster später erworben werden als Codacluster. Des Weiteren wurde die akustische Vokaldauer nicht von Onsetclustern beeinflusst.

Der Grad der interkonsonantalen Überlappung unterschied sich nicht in Bezug auf die Silbenposition: /kl/ im Silbenonset weiß denselben Grad an Überlappung auf wie in der Silbencoda. Hypothese **H2** konnte somit nicht bestätigt werden.

Was Codacluster betrifft, so gab es eine sequentielle Anordnung von /kl/ und /ks/, jedoch kam es zu einer Verschiebung weg vom Vokal in /kn/. Die akustische Analyse zeigte hier, dass es eine Vokalkürzung nur bei /kn/ Clustern gab.

Zusammenfassung

Im Rahmen der Artikulatorischen Phonologie wird für Onsetcluster grundsätzlich angenommen, dass diese einen C-center Effekt aufweisen. Codacluster hingegen sind laut dieser Theorie sequentiell angeordnet. Allerdings scheinen speziell die Annahmen bezüglich Onsets nicht generell zuzutreffen (siehe z.B. Marin, 2013; Pouplier und Beňuš, 2011). Diese Dissertation befürwortet diese Befunde, dass nicht alle Onsetcluster ein solch globales Timing mit dem folgendem Vokal aufweisen. Während es mehr CC Überlappung in dem Onsetcluster /kl/ als in /kn/ gab, scheint nur das letztere Cluster eine C-center Organisation aufzuweisen. Der Onsetkonsonant einer Silbe scheint zudem auch einen Einfluss auf das VC(C) Timing zu haben: nur in Wörtern mit /l/ im Onset, gab es eine Verschiebung des Codakonsonanten hin zum Vokal in Wörtern mit komplexer Coda verglichen mit Wörtern mit einfacher Coda . Dies war nicht der Fall in Wörtern mit /n/ im Silbenonset, obwohl der dem Vokal folgende Codakonsonant stets ein labialer Konsonant war. Codacluster sind also nicht grundsätzlich sequentiell koordiniert (siehe auch Byrd 1995; Marin und Pouplier, 2010). Die Studie in Kapitel 3 zeigte auch, dass das Konsonant-zu-Konsonant Timing wiederum das Konsonant-zu-Vokal Timing beeinflusst, aber es gab keine Unterschiede zwischen den beiden Onsetclustern was ihren Einfluss auf die Stiffness der Vokalgeste betrifft. Daher scheint die beobachtete erhöhte Stiffness (welche durch eine größere Geschwindigkeit zustande kam) in Wörtern mit komplexen Onsets unabhängig von einer C-center Organisation zu sein. Aus diesem Grund kann rückgeschlossen werden, dass Onsetcluster, welche keine Verschiebung hin zum Vokal aufweisen, trotzdem mit dem folgenden Vokal verbunden sein können. Zudem zeigte sich, dass sich Onset- und Codacluster recht unterschiedlich innerhalb einer Silbe verhalten, nicht nur was ihre zeitliche Organisation betrifft, sondern auch ihre räumliche Organisation mit dem folgenden bzw. vorangehenden Vokal. Daher müssen zusätzlich zu den zeitlichen Aspekten auch räumliche Parameter, welche von der Silbenposition abhängen, in der Artikulatorischen Phonologie und dem Coupled Oscillator Model beachtet und miteinbezogen werden.

Kapitel 4 zeigte, dass Sprecher unterschiedliche Arten von artikulatorischen Strategien anzuwenden scheinen, um einen Phrasenakzent zu produzieren oder zu kennzeichnen: weniger Gestenüberlappung oder eine Längung der Geste. Auch die Koordinationsmuster innerhalb CCVC Silben unterschieden sich hinsichtlich der Silbenposition: es gab mehr CC und VC Überlappung aber weniger CV Überlappung in deakzentuierten Wörtern. Die geringere CV Überlappung in deakzentuierten Wörtern könnte eine Konsequenz der gefundenen Plateaukürzung beider entsprechenden Gesten sein. Des Weiteren zeigte die Studie in Kapitel 4, dass die globale Organisation von Onsetclustern in deakzentuierten Wörtern nicht mehr vorhanden war. Dieser Einfluss der Prosodie auf die globale Organisation eines Onsetclusters muss in der Artikulatorischen Phonologie berücksichtigt werden.

Die abschließende experimentelle Untersuchung in Kapitel 5 zeigte, dass Kinder Onsetcluster mit einer sequentiellen Koordinierung produzieren. Während des Spracherwerbs werden also Onsetcluster mit dem gleichen Timingmuster koordiniert wie Codacluster. Die Vorhersagen der Artikulatorischen Phonologie scheinen also von dem Spracherwerb beziehungsweise von der physischen Entwicklung von Sprechern abzuhängen.

Abschließend lässt sich sagen, dass das Konsonant-zu-Konsonant Timing das Konsonant-zu-Vokal Timing beeinflusst. Außerdem beeinflusst ein zusätzlicher Onsetkonsonant die Plateaudauer des $C2_{on}$ und $C1_{off}$, d.h. während die Plateaudauer des Vokals nahezu unbeeinflusst bleibt, so sind die Plateaus der vokalumgebenden Konsonanten komprimiert bedingt durch einen zusätzlichen Onsetkonsonanten. Zusätzlich bewirkt nicht nur das Timing eines Konsonanten mit dem Vokal eine akustische Kürzung der Vokaldauer (wie in der Artikulatorischen Phonologie angenommen), sondern auch der Einfluss von Clusters auf räumliche Parameter.

Diese Dissertation war ein erster Ansatz die Konsonant-zu-Vokal Überlappung direkt zu messen. Bisherige Studien nutzten immer einen Ankerpunkt: entweder den Codakonsonanten in Onsetanalysen oder den Onsetkonsonanten in Codaanalysen (z.B. Browman and Goldstein, 1988, 2000; Byrd, 1995; Marin, 2013; Marin and Pouplier, 2010; Pouplier, 2012). Diese Studien nutzten somit immer eine indirekte Messung, um die Vorhersagen der Artikulatorischen Phonologie zu prüfen. Der Vorteil der direkten Messung von Überlappung ist, dass konfundierende Faktoren (wie Kürzung von Segmenten) ausgeschlossen werden können. Der Nachteil ist jedoch, dass dadurch die messbaren Elemente limitiert sind. So können nur Wörter untersucht werden, welche unterschiedliche Artikulatoren in jedem Segment aufweisen. Zusätzlich muss ein Vokal vorhanden sein, welcher eine aktive Zungenbewegung benötigt. Die Cluster, welche für die artikulatorischen Analysen dieser Dissertation genutzt wurden, waren beschränkt auf solche, für die zuvor Unterschiede im interkonsonantalen Timing gezeigt wurden. Aus diesem Grund wurden nur /kl, kn/ in Kapitel 3 und /kl, kn, pl/ in Kapitel 5 einbezogen. Überdies, wurde nur ein Onsetcluster (/kn/) in Kapitel 4 analysiert. Deshalb müssen die Ergebnisse dieser Dissertation für andere Sprachen, für andere Cluster und für weitere Vokale (wie /i/, /o/ or /u/) bestätigt werden.

Chapter 1

Introduction

This thesis is an experimental investigation of how syllable structure influences consonantal overlap and the vowel. This approach includes acoustic, perceptual and articulatory analyses of child and adult speakers of German. Further, it investigates possible (C)CV timing variations concerning the German syllable-initial clusters /kn/ and /kl/ for which Bombien et al. (2013) and Hoole et al. (2009) found differences in interconsonantal timing. These two clusters have also been shown to differ in their occurrence in the languages of the world that allow consonant clusters. For example, the onset cluster /kl/ is frequently attested in many languages, whereas /kn/ appears to be less common (Ladefoged and Maddieson, 1996). In addition, /k/ is more likely to be deleted diachronically before /n/ than before /l/ (cf. Vennemann, 2000). The clusters /kn/ and /gn/ are also more likely to be reduced to /n/ in children's early cluster productions (Fox, 2003; Romonath, 1991), whereas /kl, gl/ are more likely to be reduced to /k/ (Lleó and Prinz, 1996; Ohala, 1999). The question that arises is whether consonant-to-consonant timing differences may in turn affect the consonant-to-vowel timing and maybe also acoustic vowel duration. Furthermore, diachronic long vowel shortening before coda clusters is a confirmed sound change in Germanic languages (e.g. 'kept' and 'soft' were produced with a long vowel in Old English). There is also some synchronic observation that vowels are shorter in syllables with complex codas compared with simplex codas (e.g. Katz, 2012; Munhall et al., 1992). This effect is known as coda-driven incremental compensatory vowel shortening. The question that arises is whether this compression effect also exists in German and what the underlying articulatory mechanisms are.

In order to get to the bottom of these questions section 1.1 will review some background information on the structure of syllables with some emphasis on historical changes. A brief introduction to the acquisition of the syllable will be given in section 1.2. The results presented later in this thesis will frequently touch concepts of the framework of Articulatory Phonology and the coupled oscillator model. Therefore an introduction to these models will be given in section 1.3. Section 1.4 will provide an overview of recent articulatory research on consonant clusters in syllable onset and coda position that is especially relevant to the theme of this work. The research aims will be summarized in section 1.5.

1.1 Basic observations in syllabes

Within syllables a sequence of consonants (C) and vowels (V) can be identified. Generally, it has been claimed that every language has syllables with a CV structure, while there are languages that may or may not permit syllables with a VC structure, i.e. there is a clear cross-linguistic preference for CV syllables (Clements, 1990; Clements and Keyser, 1983)[1]. In addition, studies focusing on speech development show that CV structures are typically acquired before VC structures (e.g. Fikkert, 1994; Levelt et al., 2000; Salidis and Johnson, 1997; Vihman and Ferguson, 1987). The frame/content theory (e.g. Davis and MacNeilage, 2004; MacNeilage, 1998) relates the cross-linguistic preference for and the earlier acquisition of the CV pattern to the opening and closing movements of the mouth during babbling. Nam et al. (2009), on the other hand, attribute these facts to a stronger coupling within the CV syllable (see section 1.3 for more details). Ohala (1990) and Ohala and Kawasaki (1984), in turn, attribute the CV preference in the languages of the world to the perceptual robustness of word-initial compared with word-final consonants. For syllables containing

[1]However, it has to be considered that Arrernte has been shown not to have syllable onsets (Breen and Pensalfini, 1999), which causes doubt on this universality.

consonant clusters (combinations of consonants without an intervening vowel or speech pause), however, there is no clear preference for one syllable position. Their occurrence is language-dependent, e.g. Finnish has no clusters in onset position and Spanish has no clusters in coda position, but they are equally distributed and frequent from a cross-linguistic point of view (Blevins, 1995).

Usually, a syllable contains an onset and a rhyme, which again consists of the nucleus and the coda (see e.g. Haugen, 1956; Hockett, 1955; Pike and Pike, 1947). The onset consists of a syllable-initial consonant or a consonant cluster; the nucleus is the most sonorous sound in the sequence and usually consists of the vowel (with the exception of syllabic consonants) and is considered as the peak of the syllable; and the coda consists of a syllable-final consonant or a consonant cluster. Support for the idea that onsets and also codas are cohesive units within a syllable can be seen in studies on speech errors. For example, when spoonerisms occur, and one consonant is substituted for another, this only occurs in same syllable position: initial consonants are swapped for initial consonants whereas final consonants are swapped for final consonants. Thus "a cherry tart" is produced "a terry chart" or "Ouch, I have a stiff neck" is produced "Ouch, I have a stick neff" (Shattuck-Hufnagel, 1992, p. 214). While speech errors including the rhyme are observed, errors including the CV part of a syllable are very uncommon (Shattuck-Hufnagel, 1983). Phonological evidence for the rhyme as time unit can be seen in the compression effect compensatory shortening: the more consonants in the rhyme, the shorter the vowel, e.g. "lieb" > "liebt" > "liebst" (> meaning vowel is longer). In addition, in German the rhyme cannot consist only of a lax vowel and there are only lax vowels before /ŋ/ as well as before many syllable final clusters (Kohler, 1977). Historically, a vowel and a consonant within the rhyme are often merged resulting in a long vowel and loss of the consonant (Morin, 1992). This effect is known as compensatory lengthening and its directionality is always right-to-left, i.e. only the rhyme of a syllable is affected, not the onset (Hayes, 1989). Thus VC (but not CV) becomes diachronically often a long V (Old Teutonic /fɪmf/ → Old English /fiːf/; Old English /nixt/ → Middle English /niːt/; Proto-Indio-European /nisdo/ → Sanskrit /niːda/) (Kavitskaya, 2001).

Each language has phonotactic restrictions. In German, for example, there is only one consonant, the post-alveolar fricative /ʃ/, which is allowed before /pr, tr/ (e.g. ‘Sprache’, ‘Straße’) and /m/ is the only nasal consonant, which is allowed to occur before /p/ in syllable final or medial position (e.g. ‘Lampe’, ‘plump’) (Kohler, 1977). In many languages of the world onset clusters do not contain the same place of articulation: /tl/ is rarer in the world’s languages than /tr/; likewise /pw/ is rarer than /pl/ (Philipp, 1974; Wiese, 1996). In English, for example, the onset clusters /pl/ and /tw/ are allowed but /tl/ and /pw/ are disallowed as onset clusters (cf. Selkirk, 1984). However, phonotactic restrictions are language-dependent, for example, some onset clusters occur in certain languages while others do not: /kn, gn/ exist as onset clusters in German (‘Knoten’, ‘Gnade’) but not in English; /sp/ exists as an onset cluster in English (‘speak’, ‘spoon’) but not in German. Further, languages differ in syllable structures they permit. For instance, German permits complex codas as well as complex onsets (Kohler, 1977), while Hawaiian only allows a single consonant in the onset and none in the coda (Elbert and Pukui, 1979) and Standard Chinese allows only nasal consonants in coda position (Duanmu, 2002).

Sequences of consonants and vowels are organized into syllables based on the sonority principle. This means that syllables tend to follow the sonority sequencing principle (SSP) (see Clements, 1990) or, reversely, scales of consonantal strength, so that phonemes at the edges of a syllable are least vocalic or sonorous. Sonority depends on the amount of spectral energy, but it is also closely linked to the constriction of the vocal tract (Lindblom, 1983). Thus low vowels have the greatest sonority, whereas voiceless stops have least sonority.

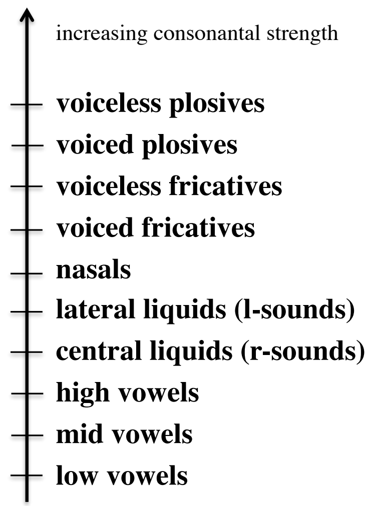


Figure 1.1: *Scale of consonantal strength following Vennemann (1988).*

Figure 1.1 illustrates the scale of increasing consonantal strength as presented by Vennemann (1988). Syllables tend to follow such scales in that consonantal strength falls from the onset to the nucleus of a syllable and it increases from the nucleus to the coda of a syllable (cf. *Head Law and Coda Law* in Vennemann, 1988). Syllables structured in this way are said to conform to the sonority profile. If they conform to the sonority profile, consonants clusters in syllable onsets increase in sonority from left to right and consonant clusters in syllable codas decrease in sonority from left to right. Therefore, a language is more likely to have a syllable like /pla/ than /lpa/, because in /pla/ the sonority increases from its lowest /p/ for /l/, and reaching a peak with /a/. Similarly, a language is more likely to have /alp/ than /apl/. However, Ohala (1992) and Ohala and Kawasaki-Fukumori (1997) criticize sonority sequencing as being only descriptive and not explanatory. For example, they argue that circularity arises as soon as a boundary position has to be derived from sonority scales (e.g. whether the medial cluster in “scoundrel” has to be divided into /n#dr/ or whether it may also be an onset cluster /#ndr/), and that onset clusters such as /sk/ or /st/ as well as coda clusters such as /ks/ and /ts/ are against this sonority hierarchy. The sonority profile is therefore a general tendency which determines many, but by no means, all phonotactic constraints. Ohala and Kawasaki-Fukumori (1997) therefore propose to replace the non-empirical parameter "sonority" by acoustic parameters (such as amplitude, periodicity, spectral shape, and f0) describing the degree of modulation. This measure should then predict which sequences exist and are more preferred in the world's languages. Mattingly (1981), on the other hand, assumes that the transmission of information is supported by the concurrent availability of information on multiple sounds caused by coarticulation or overlapping gestures. Such a parallel transmission can be maximal in CV structures as the articulatory movements for the consonant and the vowel are initiated simultaneously (see Öhman, 1967). This is possible since for the vowel the vocal tract is not critically constricted and nothing hides the emergence of acoustic correlates. For consonant clusters, however, simultaneously initiated consonants are likely to block each other in their acoustic outcome. Therefore, overlap between consonants within clusters must be less than in CV sequences. Mattingly (1981) links the extent to which sounds block each other to constriction strength (as in Figure 1.1) referring to manner of articulation. Thus for example, clusters containing two voiceless plosives are more likely to block each other and have to overlap less than clusters containing a voiceless plosive and a lateral liquid.

The release of consonants (especially plosives) leads to strong acoustic modifications in CV but not in VC structures. Listeners are sensitive to these acoustic modifications and the consonant is therefore perceptually more salient in CV than in VC within VCCV sequences (Ohala, 1990; Repp, 1978). In CVC syllables listeners identify initial consonants to a better degree than final consonants (Redford and Diehl, 1999). Further, the CV structure has a precise gestural synchronisation differing from the VC structure (Krakow, 1999). For example, in *'mam'* the temporal coordination of lip closure and velum lowering for /m/ differs with respect to syllable position: these two gestures are synchronous in onset position but in coda position the velum lowering occurs much earlier than the lip closure. Because of this greater overlap the consonantal identity is more likely to get lost in perceptual terms in syllable-final position (Krakow, 1999). Such differences in gestural coordination depending on syllable position can also be observed for laterals: the tongue dorsum gesture occurs earlier with respect to the tongue tip closure when the /l/ occurs syllable-finally (Browman and Goldstein, 1995; Sproat and Fujimura, 1993). The implications of these greater temporal overlap patterns in VC structures can be observed both synchronically and diachronically. For example, the coarticulatory influence of a nasal on a preceding vowel is greater than on a following vowel (Krakow, 1999). Also historically, the emergence of contrastive vowel nasalization can be traced back to a loss of a final nasal and its nasalization remaining on the preceding vowel (Greenberg, 1966; Kawasaki, 1986). Likewise, syllable-final /l/s are often velarized influencing the quality of the preceding vowel. For example, in many dialects of English "milk" may be realized as [mɪɤk], [mɪʊk] and even as [mʊːk] (Hardcastle and Barry, 1989). Also in central Bavarian there is /l/-vocalization influencing the vowel quality, e.g. "viel" [fʊɪ], "Wald" [vɔɪd] (König et al., 1991).

Historical processes also indicate a preference for initial consonants over final consonants. For example, a common sound change is final consonant weakening and loss (Hock, 1986). Concerning assimilation, initial consonants influence final consonants more likely than reversely (synchronically: "fixed cars" → [fɪkskɑrs] (Byrd, 1992); diachronically: L. Latin "nocte" > Italian "notte") (Ohala, 1990). Further, CV syllables are more stable than VC syllables, which can be seen in the occurrence of phase transitions with increased speech rate (Tuller and Kelso, 1991): when speakers produce for example /ip/, they shift to /pi/ as rate increases, but not vice versa. All these facts may be reasons why the CV syllable is preferred in the languages of the world: while there are various languages containing CV syllables without having VC syllables, there are barely languages containing VC syllables without having also CV syllables (Jakobson and Halle, 1956).

1.2 Syllable acquisition

At the earliest stages of language acquisition, the syllable is the most basic unit (Jakobson, 1941). Children first produce CV combinations, then VC followed by CVC combinations (e.g. Fikkert, 1994; Levelt et al., 2000; Salidis and Johnson, 1997; Vihman and Ferguson, 1987). Thus, children's early productions are basic syllabic elements. Generally, children do show a tendency to simplify the structure of a syllable: there is final consonant deletion (e.g. *Buch* [buː]), deletion of unstressed syllables (e.g. *Banane* [naːnə]), reduplication (e.g. *Mandarine* [nini]) and cluster reduction (e.g. *Spiegel* [piːgəl]) (Kauschke, 2012, p. 36). Concerning children's development of vowel length and rhyme, Fikkert (1994) proposed four stages. First, children produce core syllables (CV) with either long and short vowels. At the second stage, the rhyme is branched into nucleus and coda and children produce obstruents earlier than sonorants in syllable-final position. At stage three, children were shown to shorten long vowels before sonorants and lengthened short vowels when the sonorant coda was deleted. No such relationship was found between vowels and obstruent codas, i.e. obstruent codas were produced after both long and short vowels. At stage four, children acquire vowel length contrast before obstruents. Also Bernhardt and Stoel-Gammon (1998), Demuth and Fee (1995) and Kehoe and Lleó (2003) reported a relationship between vowel length and coda production: only lax vowels were produced in closed syllables, whereas tense vowels appeared in open syllables. As far as the acquisition of the coda is concerned, Kehoe and Stoel-Gammon (2001) showed that final voiced obstruents are acquired after final voiceless obstruents and nasals in English speaking children. Fikkert (1994) proposes that by producing obstruents first in coda position, children prefer to produce a sonority decrease from nucleus to coda (see also Ohala, 1999). When children become older (at the age of two years), they are able to produce more complex syllables containing consonant clusters (Lleó and Prinz, 1996), but at the early stages some phonological processes like cluster simplification and reduction are frequently observed in various languages (e.g. Lleó and Prinz, 1996; McLeod et al., 2001b; Ohala, 1998; Smit, 1993; Watson and Scukanec, 1997). The mastery of adult-like realizations of words containing consonant clusters is protracted and may last even up to the age of 8 (Smit et al., 1990). In addition, the acquisition of clusters depends on their composition, for example, word-initial /kl/ is acquired earlier than word-initial /kn/ (Fox, 2003; Fox-Boyer, 2014).

Children seem to be aware of syllable structure and divide syllables into onset und rhyme. Barton et al. (1980), for example, showed that 4- to 5-year-olds identified onsets as units: they treated the first sound of "swing" as /sv/ rather than /s/, i.e. they considered the onset as a cluster. Treiman (1985) showed evidence for the rhyme in children's phonological awareness: in a game, which required the replacement of the initial CV and the retainment of the final C (e.g. the stimulus contained /mon/ and the child had to produce /lʌn/), children erroneously retained the entire VC form (e.g. they produced /lon/ instead of /lʌn/). Further, children seem to have a knowledge about the phonotactic constraints of their native language. English acquiring children aged between 3-5 years, for example, produce the non-word *kiften* more accurately than *bufkit* – because there are real-words containing /ft/ (‘fifteen’, ‘safety’) but not containing /fk/ (Edwards et al., 2004).

1.3 Articulatory Phonology

In the framework of Browman and Goldstein’s Articulatory Phonology (1986; 1988; 1990; 1992; 2000), speech gestures are defined by tract variables, referring to constriction location and constriction degree in the vocal tract. Five different tract variables are defined: lips, tongue tip (TT), tongue body (TB), velum (VEL) and glottis (GLO). These articulators are assumed to be independent of each other, which allows for separate gestural representations. Tract variables are specified for constriction location (*protrusion*, *labial*, *dental*, *alveolar*, *postalveolar*, *palatal*, *velar*, *uvular*, and *pharyngeal*), i.e. variables can constrict the vocal tract in more than one location. In addition to location, there is also constriction degree (*closure*, *critical*, *narrow*, *mid*, and *wide*). These types of different degrees are necessary for distinguishing between constrictions that can be made at the same location. For example, the constriction degree for the stop /t/ is closure, whereas it is critical for the fricative /s/, and both are articulated at constriction location alveolar. The remaining three constriction degrees (*narrow*, *mid*, and *wide*) apply to approximants and vowels.

Constriction degree is the gestural equivalent of manner of articulation – although Browman and Goldstein note that it is not acoustic in any sense – whereas constriction location is similar to place if articulation. When combining constriction location and constriction degree, for example, /k/ would be produced with tongue back constriction location velar and a tongue back constriction degree closure. A wide glottal gesture in combination with stops leads to aspiration and a narrow glottal gesture is associated with unaspirated stops.

As described above, Browman and Goldstein note that the set of articulators forming a constriction and a location of that constriction are independent gestures. Thus it may be possible for more than one articulator to make a constriction at a given location. Further, two gestural variables can be combined producing fine-grained articulatory distinctions, such as bilabial or labio-dental.

When gestures are combined into a representation of an utterance, they are considered as primitives of phonological contrast (Browman and Goldstein, 1992). The relationships between gestures at any point in the utterance is represented in a two-dimensional gestural score. This is schematically illustrated for the word "pan" in Figure 1.2. In this schema, the voiceless aspirated stop /p^h/ corresponds to both the labial closure and the wide glottal gesture. The gesture defined by TB constriction degree wide and TB constriction location pharyngeal corresponds to the vowel /æ/. Finally, the nasal /n/ is represented by the wide velum in combination with TT constriction location alveolar and TT constriction degree closure.

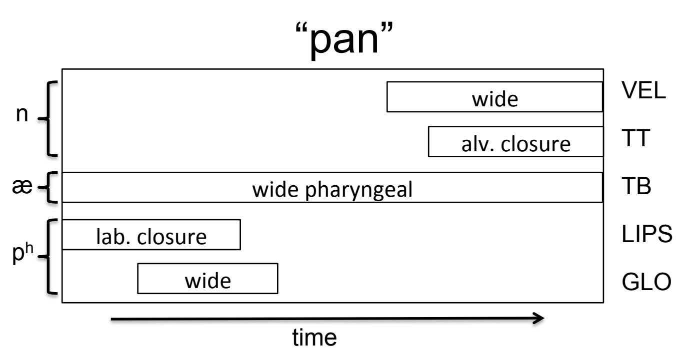


Figure 1.2: *Gestural scores for the English word "pan" [p^{h}æn].*

Browman and Goldstein (1989) claim that gestural scores are indeed lexical entries. Thus minimal contrast between lexical entries – such as between 'pan' and 'pat' – is identified by differences in the gestures that are involved. Thus while the final tongue tip closure in 'pat' is associated with a wide glottal gesture, the /n/ of 'pan' will rather have a wide velum aligned with the tongue tip closure instead. These kinds of phonological contrasts are thus directly linked to gestures and gestural coordination. Consonants and vowels are represented on separate tiers: vowels are represented on the vocalic tier, consonants on the consonantal tier. Whereas two adjacent vowels are not affected by an intervening consonantal gesture, two consonants are sensitive to an intervening vocalic gestures. This sensitivity prevents them from overlapping (Browman and Goldstein, 1990; Gafos, 1999) (see also Figure 1.2 for non-overlapping consonants in the CVC sequence, i.e. the labial closure is clearly separated from the alveolar closure). This is because consonantal gestures require a constriction at a specific domain in the vocal tract, whereas vowels are specified by a lack of constriction (Gafos, 1999). This organization on separate tiers has been used to explain, for example, coarticulatory phenomena (Keating, 1985) and why vowel harmony is a common process in the world's languages (Browman and Goldstein, 1990; Gafos, 1999). Further, vocalic and consonantal gestures are also distinguished by the tract variable of stiffness (an abstract control parameter influencing the time-space behavior of the system (Cooke, 1980)), which is greater for consonants than for vowels (Browman and Goldstein, 1992). Despite the division of consonants and vowels into separate tiers there is a relationship between vowels and consonants and between adjacent consonants. Vocalic and consonantal gestures were shown to have coordination relationships that link particular gesture positions with positions of another gesture, depending on which gesture precedes or follows the other (Browman and Goldstein, 1990).

Browman and Goldstein (1995) proposed differences between VC and CV timing relationships. The two gestures required to produce a nasal (tongue tip closure and velum lowering) or a lateral (tongue tip raising and tongue body retraction) are timed differently with the vowel depending on syllable position. In the case of a lateral, the tongue dorsum gesture occurs much earlier with respect to the tongue tip closure when in coda position than when in onset position (Browman and Goldstein, 1995; Sproat and Fujimura, 1993). A number of other articulatory studies have also shown that onset consonants have different patterns of gestural organization from coda consonants (e.g. Browman and Goldstein, 1995; Byrd, 1996; Fougeron and Keating, 1997; Kochetov, 2006; Krakow, 1989; Sproat and Fujimura, 1993).

In the coupled oscillator model (Goldstein et al., 2009; Nam and Saltzman, 2003), each gesture is associated with an abstract oscillator and any oscillating system is associated with an abstract 360° cycle. In this model, timing relations between gestures are expressed by two coupling or phasing relations: in-phase (0°) and anti-phase (180°). For example, the CV structure has been observed to be synchronously coordinated (Öhman, 1967). If two gestures start synchronously, then they are coupled in-phase. If two gestures are coupled anti-phase, they start one after the other, i.e. they are sequentially organized.

With respect to onset clusters, studies have shown that consonants exhibit a stable timing relationship among themselves, but each of the consonants is also in a timing relationship with the following vowel (Browman and Goldstein, 1988, 2000; Byrd, 1995; Honorof and Browman, 1995). Specifically, it has been observed that, as more consonants are added to the syllable onset, the vowel-adjacent consonant (e.g. /l/ in p<u>l</u>ug vs. <u>l</u>ug) shifts towards the vowel[2], while the vowel-remote consonant shifts away from the vowel, such that the singleton and the cluster line up most stably along their midpoints. It is concluded that the midpoint of the cluster – the 'c-center' – is the point in an onset cluster that has the most stable relationship with the following vowel. Coda consonants, on the other hand, have been described to be timed sequentially with the preceding vowel, so that with the addition of more coda consonants, the cluster's left edge remains stable (Browman and Goldstein, 1988, 2000; Honorof and Browman, 1995; Marin and Pouplier, 2010). The presumed c-center organization for onset clusters and the sequential organization for coda clusters are schematically illustrated in Figure 1.3 with reference to a specific predefined anchorpoint used in articulatory studies focusing on syllable timing.

[2] These studies usually use the coda consonant as an anchorpoint. Therefore, it would be more precisely to say that the vowel-adjacent consonant shifts toward the anchorpoint.

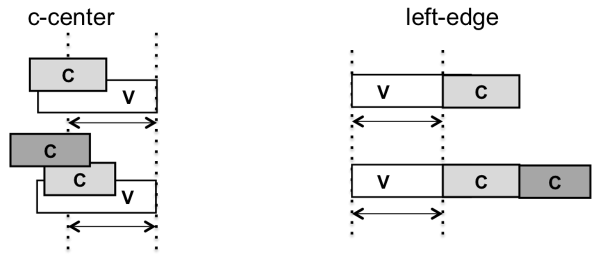


Figure 1.3: *Schematic illustration of c-center alignment for onset clusters (left) and left-edge alignment for coda clusters (right) (from Marin and Pouplier, 2010).*

In general then onset clusters are linked to the vowel in a different way than coda clusters. For onsets, the coupled oscillator model assumes that each consonant on its own is in an in-phase (0°) timing relation with the vowel. If this in-phase timing relation were the only timing relation between onset clusters and the vowel, the two consonantal gestures would be produced simultaneously (Browman and Goldstein, 2000). However, depending on the particular consonants, this might result in one of the consonants being unrecoverable (Mattingly, 1981; Silverman, 1997). Therefore, the onset consonants also have to be in an anti-phase (180°) timing relation with each other. This additional timing relation causes the consonants to be produced one after the other, and thus to be perceptible. This resulting competitive coupling of onset clusters causes the c-center effect: there is gestural overlap between the vowel-adjacent consonant and the vowel because of their in-phase timing with each other. The vowel-remote consonant, however, is not produced in synchrony with the vowel-adjacent consonant since there is additional anti-phase coupling between these consonants. This leads then to the c-center effect and to a shift of the vowel-adjacent consonant towards the vowel.

For coda clusters, on the other hand, different timing patterns are assumed: only the vowel-adjacent consonant is timed anti-phase (180°) with the vowel, and also the consonants within a coda cluster are timed anti-phase (180°) with each other. There is thus no competitive coupling and coda clusters are therefore said to be sequentially timed with the preceding vowel. This means that the cluster's left-edge (rather than the cluster's c-center) remains in a stable relationship with the preceding vowel. The two different timing relationships for onset and coda clusters are schematically illustrated in Figure 1.4.

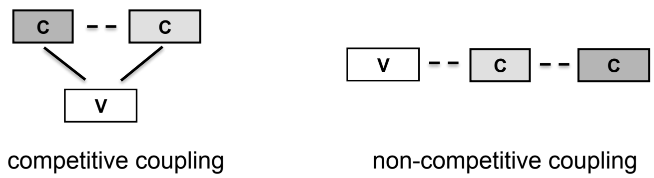


Figure 1.4: *Schematic illustration of in-phase and anti-phase coupling for onset (left) and coda clusters (right). In-phase coupling is indicated by solid lines; anti-phase coupling by dotted lines (from Marin and Pouplier, 2010).*

Since only onset clusters have competitive coupling relations, they are considered to be timed more fixedly. This allows for less timing variability and less overlap between onset than between coda clusters (Byrd, 1996; Chitoran et al., 2002; Hardcastle and Roach, 1979; Pouplier, 2012). One explanation is perceptual recoverability requirements (Byrd, 1996; Chitoran et al., 2002): greater overlap between consonants in coda clusters is possible because there is still recoverability as other acoustic cues are present in the preceding vowel and in the word-final release. In order to correctly identify onset clusters – or actually utterance initial clusters – the consonants are not allowed to overlap to a greater extent, as no other acoustic cues are available (Marin, 2013). However, Marin (2013) found no differences in intra-cluster timing as a function of syllable position. She concluded that overlap patterns seem to depend more on the exact identity of the consonants rather than on syllable position.

1.4 Recent research on onset and coda clusters

Evidence for the c-center hypothesis has been reported for English, German, Georgian and Italian. For example, c-center organization for complex onsets has been observed for /sp/, /sk/, /sm/, /pl/ and /kl/ onsets in English (Marin and Pouplier, 2010), for /sk/, /bl/, /km/ and /gm/ onsets in German (Pouplier, 2012), for /k'r/ and /t^{s}k'r/ onsets in

Georgian (Goldstein et al., 2007a) and for /pr/ and /pl/ onsets in Italian (Hermes et al., 2013). The c-center hypothesis has also been used to predict the syllable organization of word-initial clusters whose syllable affiliation is controversial, such as in Italian, Moroccan Arabic and Tashlhiyt Berber (Goldstein et al., 2007a; Hermes et al., 2008, 2012, 2013, 2011; Shaw et al., 2009, 2011). For example, Hermes et al. (2008, 2013) presented articulatory evidence for different timing patterns of Italian *impure /s/*. In Italian, clusters containing impure /s/ are treated differently from other clusters or from plain /s/ in the morphology (Davis, 1990). The findings from Hermes et al. (2008, 2013) showed that the cluster /pr/ in 'la prima' exhibits a c-center organization. However, when an impure /s/ is added to the syllable onset as in 'la sprima', then there is no shift of the following consonantal gestures towards the vowel as predicted by the c-center hypothesis. Instead, impure /s/ appears not to be part of the onset cluster, rather it seems to be added to the cluster as an additional simplex onset. Thus /s/-initial clusters in Italian exhibit a different morphological and articulatory behavior from other clusters in Italian, suggesting that they may not form complex onsets. Additionally, consonant sequences in Moroccan Arabic and Tashlhiyt Berber are suggested not to form complex onsets (however, it has been much debated whether consonant clusters in Moroccan Arabic are to be considered as complex onsets or as a sequence of simplex onsets; see Shaw et al., 2009). Shaw et al. (2009) showed that the coordination relations determined by Articulatory Phonology can be exerted to determine the syllable affiliation of Moroccan Arabic onset clusters. They assumed that complex onsets should exhibit the c-center effect, whereas a sequence of consonants should exhibit a sequential organization (i.e. right-edge alignment instead of c-center alignment). Their results showed that Moroccan Arabic onset clusters exhibit a right-edge alignment with the following vowel, supporting evidence that onsets are rather a series of simplex consonants than complex onsets in this particular language. Similarly, Goldstein et al. (2007a), comparing Georgian (which has been claimed to have complex onsets) with Tashlhiyt Berber (which has been claimed to have a series of simplex onsets), showed a c-center organization of onset clusters in Georgian but a right-edge alignment of onset consonants in Tashlhiyt Berber.

In all of these studies, an observed right-edge alignment of onsets has been taken as empirical evidence that these consonant sequences are a series of simplex consonants rather than complex onsets. It has to be considered, though, that a right-edge alignment has also been reported for Slovak (Pouplier and Beňuš, 2011) and for Romanian (Marin, 2013, at least for some clusters), two languages where word-initial consonant sequences are indisputably assumed to form complex onsets. In the study by Marin (2013), manner seems to affect the timing of onset consonants with the following. For example, she revealed evidence for a global organization of /sp/, /sk/, /sm/ onsets but not for /ps/, /ks/, /kt/ and /kn/ onsets.

As far as coda clusters are concerned, most studies report a sequential organization with the preceding vowel (Marin, 2013; Pouplier, 2012). However, Byrd (1995) reported a c-center organization for codas, at least for some speakers. In addition, Marin and Pouplier (2010) reported a c-center organization for /lp/ and /lk/ codas.

1.5 Research aims and outline of this thesis

In this work, an attempt is made to shed further light on consonant-to-vowel timing in simplex onsets/codas as well as complex onsets/codas with some emphasis on vowel compensatory shortening. This is motivated by slightly different or even contradictory previous results in this field.

The present thesis includes four experimental chapters. In chapter 2, the results of a pilot study focusing on compensatory vowel shortening caused by coda clusters in both production and perception will be reported. Using EMA (electromagnetic articulography), chapter 3 focuses on compensatory vowel shortening driven by both onset (/kn/ and /kl/) and coda (/pt/ and /mt/) clusters and the possible underlying articulatory mechanisms will be analyzed. Chapter 4 focuses on the stability of gestural timing patterns under prosodic influences. Here, the influence of prosodic weakening on the global organization of the onset cluster /kn/ is investigated. Finally, chapter 5 deals with syllable organization during language acquisition focusing on consonant-to-consonant as well as consonant-to-vowel timing in pre-school children's productions. A brief summary and discussion of the results with respect to the research aims will be given in Chapter 6.

Chapter 2

Compensatory shortening before complex coda clusters in the production and perception of German monosyllables

2.1 Introduction

The main aim of the present study is to investigate incremental coda compensatory shortening in the production and perception of German monosyllables and the influence of accentuation (accented vs. deaccented), coda manner (obstruent vs. sonorant) and vowel tensity (lax vs. tense) on the degree of compensatory shortening. This was motivated by previous studies showing greater shortening of phonologically long vowels (Waals, 1992) and of vowels preceding sonorants (Katz, 2012). In a production experiment, ten speakers produced German words like /klɪk/ and /klɪkt/, and the duration of the vowel (V) and the first coda consonant (C1) was measured depending on coda complexity. In a subsequent experiment, it was tested whether listeners perceive or compensate for vowel and/or C1 shortening effects found for some speakers. Natural utterances from the production experiment were presented to 21 subjects whose task it was to judge which vowel in a pair like /knɪk/ - /knɪkt/ they perceived as longer.

2.2 Compensatory vowel shorterning

Vowels were shown to be shorter in syllables with more segments, a phenomenon commonly known as *compensatory vowel shortening*. Following Katz (2008, 2012), we distinguish between compensatory shortening driven by adding segments to the onset from compensatory shortening driven by adding segments to the coda of a syllable, i.e. *onset* vs. *coda compensatory shortening*. We further differentiate – again following Katz (2012) – between *simplex* and *incremental compensatory shortening*: the former refers to the cross-linguistically widely attested fact that vowels are longer in CV or VC syllables than in CVC syllables (Farnetani and Kori, 1986; Lindblom and Rapp, 1973; Maddieson, 1985; Waals, 1992), and the latter makes reference to the observation that vowels are longer in CVC than in CCVC or CVCC syllables.

A large body of work on incremental shortening stems from the c-center literature (Browman and Goldstein, 1988, 2000; Marin and Pouplier, 2010), which provides evidence for shorter vowel durations in words with complex onsets (CCVC) as opposed to words with simplex onsets (CVC). Proponents of the c-center hypothesis argue that this is due to the global organization of onset clusters, which causes the vowel-adjacent consonant to move towards the vocalic nucleus. This shift towards the vowel then results in acoustic vowel shortening. Coda clusters, on the other hand, are supposed to be timed sequentially (Marin and Pouplier, 2010). Therefore, according to this theory, coda clusters should not cause incremental compensatory vowel shortening. There are, however, acoustic and articulatory studies showing evidence for coda-driven shortening (Katz, 2008, 2012; Marin and Pouplier, 2010; Munhall et al., 1992; Shaiman, 2001).

Concerning onsets, incremental shortening has been reported for various cluster types (Katz, 2012; Marin and Pouplier, 2010). However, there are mixed results on coda-driven shortening and manner of the coda consonant seems to affect the degree of compensatory shortening: there is more shortening of vowels preceding sonorants than of vowels preceding obstruents (Fowler, 1983; Katz, 2008, 2012; Waals, 1992). Similarly, the shortening effect in phonologically long vowels of Dutch is strongest in words with liquids in coda position and weakest in words with obstruents in coda position (Waals, 1992). For production, there is thus a tendency for longer segments (either phonologically or context-dependent as in sonorants) to be shortened to a greater extent (see also Turk and Shattuck-Hufnagel, 2000; White and Turk, 2010).

Because of this tendency also the influence of accentuation was included in the present study. In general, within an utterance a word is accented when a pitch accent is associated with the strongest syllable (Pierrehumbert and Beckman, 1988). This pitch accent causes changes in the fundamental frequency (f_0) (Fry, 1958) as well as differences in duration: accented syllables or words are produced with greater duration (de Jong, 1995b). Given that vowels in accented words are longer than in deaccented words (Cho, 2005; Lehiste, 1970), we investigate the influence of accentuation on incremental compensatory shortening in the present study to test the hypothesis that it is the longer segments to be shortened by coda clusters.

2.3 Compensation for compensatory shortening

It has been shown that listeners compensate for the effects of coarticulation. For example, when a /s-ʃ/ continuum is presented to listeners with two different vowel contexts /a, u/, then listeners give more /s/ than /ʃ/ answers in the context of /u/, because they factor out the coarticulatory rounding influence of the vowel (Mann and Repp, 1980). As far as the perception of compensatory shortening is concerned, only few work has been done so far. There is one study by Fowler and Thompson (2010) investigating listeners' perception of compensatory shortening in English. In this study another type of vowel shortening was investigated, namely *polysyllabic shortening*, in which the vowel is shorter in polysyllabic than in monosyllabic words. In polysyllabic shortening there is a distinction between anticipatory shortening (i.e. an additional syllable is added after the vowel: bay > baby) and backward shortening (i.e. an additional syllable is added before the vowel: bay > obey) (see Fowler and Thompson, 2010). These shortening effects have been extensively documented in Germanic languages such as Swedish (Lindblom and Rapp, 1973), English (e.g. Klatt, 1973; Lehiste, 1970; Port, 1981) and Dutch (Noteboom, 1972). Fowler and Thompson (2010) showed that English listeners compensate for compensatory shortening of polysyllabic versus monosyllabic words, i.e. listeners perceived vowels in polysyllabic contexts as longer than vowels with the same duration in monosyllabic contexts. Moreover, listeners compensated more for anticipatory shortening – which is also more pronounced in production (see Lehiste, 1972; Lindblom and Rapp, 1973) – than for backward shortening. For German the effect of polysyllabicity on vowel duration is less clear. Harrington et al. (2015), focusing on anticipatory shortening, found no effect of polysyllabicity on vowel duration, i.e. there were no durational vowel differences between /za:kt/ (monosyllable) and /za:ktɛ/ (disyllable). However, the duration ratio between the vowel and the following /kt/ cluster was less for disyllables than monosyllables. For perception, they found that listeners compensated for the effects of polysyllabicity, confirming Fowler and Thompson's (2010) results, and listeners compensated less for this effect in deaccented words. So far, it is unclear how listeners perceive incremental compensatory shortening and whether they compensate for it. If listeners compensate for compensatory shortening, they should perceive the vowel as being longer in words with complex codas.

To conclude, vowel shortening seems to be affected by a number of factors: there are differences depending on vowel length (see Waals, 1992), on coda manner (see Katz, 2012) and on language (see e.g. Fowler and Thompson, 2010; Harrington et al., 2015, reporting different results for English and German). The aims of the present study are to test (1) whether incremental coda-driven compensatory shortening can also be observed in Standard German, (2) whether there are differences in the degree of compensatory shortening depending on lengthening factors, such as the manner of the following cluster's consonant (obstruent vs. sonorant), the prosodic condition (accented vs. deaccented) as well as the vowel tensity (tense vs. lax), and (3) whether listeners perceive shortening effects (or whether they compensate for them).

2.4 Hypotheses

H1 Vowel duration is shorter in words with a complex coda than in words with a simplex coda.

H2a, H2b and H2c specifically test whether there is more shortening in segments that are contextually lengthened.

H2a Vowel duration is shortened to a greater extent in accented words.

H2b Vowel duration is shortened to a greater extent in words with a complex coda consisting of sonorants.

H2c Vowel duration is shortened to a greater extent in words containing tense vowels.

The last hypothesis is concerned with the listeners' perception of coda-driven compensatory vowel shortening.

H3 Listeners compensate for compensatory shortening and even more so in contexts that favor shortening.

2.5 Speech Production

2.5.1 Participants

We recorded 10 native speakers of Southern Standard German (6 females and 4 males, aged 20 to 35 years). None of the subjects reported any speaking or reading disorders. Some of them were undergraduate students of phonetics and thus had basic phonetic knowledge, but all participants were naïve as to the purpose of the experiment.

2.5.2 Method

2.5.3 Speech material

All test verbs were produced in sentence medial position in the carrier phrase "Maria hat [target word] vorgelesen" ("Maria read [target word] out"). Only real German monosyllabic words that contained either one (simplex) or two (complex) coda consonants were included.

The vowel was always preceded by one of two onset clusters (/kn/ vs. /kl/). These onsets were chosen because previous articulatory studies (see Bombien et al., 2013; Hoole et al., 2009) revealed different timing patterns within these consonant clusters, which may have an effect on vowel duration.

In order to test hypothesis **H2b** the coda consonant adjacent to a lax vowel was either a sonorant (/l/ or /ŋ/) or an obstruent (/k/). Here, the simplex set contained the imperatives *knall ("bang")*, *knick ("crease")*, *klick ("click")*, and *kling ("sound")*. Complex codas were derived by adding the third person singular *-t* to the stem. All words used in both the production and the perception experiment are listed in table 2.1.

	Simplex coda	**Complex** coda
obstruent coda	/knɪ**k**/ /klɪ**k**/	/knɪ**kt**/ /klɪ**kt**/
sonorant coda	/kna**l**/ /klɪ**ŋ**/	/kna**lt**/ /klɪ**ŋt**/

Table 2.1: *Speech material containing obstruent and sonorant codas.*

To account for hypothesis **H2c** also words containing tense vowels were included. Since we decided to use only real words containing the abovementioned onset clusters, the repertoir of words with an increasing number of coda consonants was strictly limited.[1]. Here, the simplex set contained the noun *Knut (the male name "Knut")* and the verb *kniet ("knees", 3rd P. Sg.)*. Complex codas were derived by, again, adding a consonant to the stem. To remain real words it was not possible to keep the stem constant. The complex set contained the imperative *knutsch ("kiss")* and the verb *kniest ("knee", 2nd, P. Sg.))*. In addition, the noun *Knie ("knee")* was included in order to investigate whether there is *simplex compensatory shortening* in /kniː/ vs. /kniːt/. All words used in both the production and the perception experiment are listed in table 2.2.

[1]Note that no direct comparison between tense and lax vowels was possible, as no real counterparts containing lax vowels exist in German.

Open Syllable	**Simplex** coda	**Complex** coda
/kniː/	/kniː**t**/	
	/kniː**t**/	/kniː**st**/
	/knuː**t**/	/knuː**tʃ**/

Table 2.2: *Speech material containing tense vowels.*

2.5.4 Experimental set-up

The recordings were made in a sound-attenuated booth at the Phonetics Institute in Munich. Seven repetitions of the sentences were presented isolated and in randomized order on a computer screen using the SpeechRecorder software version 2.6.6 (Draxler and Jänsch, 2004). In order to elicit the pitch accent either on *Maria* (i.e. the target word was deaccented) or on the target word (i.e. the target word was accented) each sentence was introduced with a question. “WER hat [target word] vorgelesen?” (“WHO read [target word] out?”) was used to evoke accentuation of the name (which is new information in this context) and the question “WAS hat Maria vorgelesen?” (“WHAT did Maria read out?”) was applied to the accented condition. In addition, capitals drew attention to words with a pitch accent. The experimenter ensured that incorrect prosodic patterns were repeated by the subjects.

2.5.5 Data analysis

Acoustic segmentation was first done automatically using MAUS (Schiel, 2004). The segment boundaries of the target words were then manually corrected in Praat (Boersma and Weenink, 1992) when necessary. Segmentation of C1 included closure, burst and aspiration. Duration measurements of both the vowel and C1 of the target word were carried out in EMU/R (Harrington, 2010). The durations of the vowel and C1 were normalized. As word duration is longer in words with complex codas, proportional durations were calculated by dividing vowel duration and C1 duration, respectively, by the duration of the word “hat” (*had*) in the carrier phrase (see the equations 2.1 and 2.2).

$$Prop.dur_V = \frac{duration_V}{duration_{\text{"}hat\text{"}}} \tag{2.1}$$

$$Prop.dur_{C1} = \frac{duration_{C1}}{duration_{\text{"}hat\text{"}}} \tag{2.2}$$

Since the word “hat” contains a short vowel and was always deaccented, there are some instances of target vowels – especially in the accented condition – that are longer than “hat” resulting in proportional vowel durations between 0 (zero) and 2. As we were interested in the degree of compensatory shortening, we calculated duration differences by subtracting the proportional V and C1 durations of words with a complex coda from the proportional vowel and C1 duration of words with a simplex coda (see equations 2.3 and 2.4), respectively.

$$V_{dur.diff} = (V_{\text{simplex coda}} - V_{\text{complex coda}}) \tag{2.3}$$

$$C1_{dur.diff} = (C1_{\text{simplex coda}} - C1_{\text{complex coda}}) \tag{2.4}$$

Thus values above 0 (zero) indicate V or C1 shortening. The durational difference of C1 (i.e. $C1_{dur.diff}$) was not measured for the words /kniː/ - /kniːt/ - /kniːst/ because of a missing constant C1 in these tokens.

2.5.6 Statistics

For the statistical analysis various mixed models were conducted . Prop.dur_V, Prop.dur_{C1}, $\text{V}_{dur.diff}$, and $\text{C1}_{dur.diff}$ each served as the dependent variable in one of the models. Coda complexity (two levels: simplex vs. complex), coda manner (two levels: sonorant vs. obstruent) and accentuation (two levels: accented vs. deaccented) were the fixed factors. Speaker and word were entered as random factors. The models were implemented using the lme4 package (Bates, 2008) in the statistical environment R.

2.5.7 Results

Obstruent vs. Sonorant

Firstly, durations did not differ with respect to the two different onset clusters /kn/ and /kl/. Secondly, absolute vowel durations were longer for accented tokens (mean = 64.6 ms) than for deaccented tokens (mean = 57 ms). Additionally, vowel durations were longer for words containing sonorant codas (mean = 76.4 ms) than obstruent codas (mean = 45.3 ms). On the basis of these contextual lengthening effects, we can test the influence of accentuation and coda manner on compensatory vowel shortening.

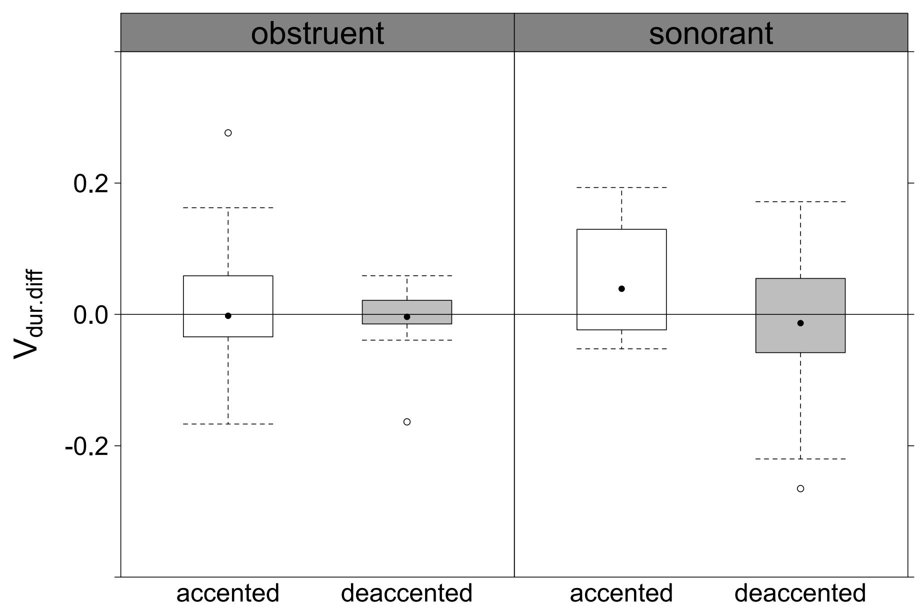


Figure 2.1: *Proportional vowel duration differences between simplex and complex codas separately for accented (white) and deaccented (gray) words containing obstruent and sonorant codas.*

Figure 2.1 shows the proportional vowel duration differences between simplex and complex codas separately for accented and deaccented words containing obstruent and sonorant codas. It is clearly observable that the differences in vowel duration between complex and simplex codas are around zero in all contexts, i.e. there seems to be no compensatory vowel shortening. However, a tendency towards vowel shortening can be observed in accented words containing a sonorant coda, as the median lies above zero.

A mixed model with Prop.dur$_V$ as the dependent variable and with fixed factors coda complexity, accentuation and coda manner revealed no significant effect of coda complexity. This means – commensurate with Figure 2.1 – coda clusters did not shorten the vowel acoustically, despite the observable trend in accented sonorant tokens. The model revealed significant effects of accentuation ($\chi^2[1] = 10.4$, $p<0.01$) and coda manner($\chi^2[1] = 5.9$, $p<0.05$) as well as a significant interaction between these two factors ($\chi^2[1] = 10.5$, $p<0.01$). Post-hoc Tukey tests revealed that accented sonorant tokens were longer than accented obstruent tokens ($p<0.001$) and that only accented sonorant tokens were longer than deaccented sonorant tokens ($p<0.001$).

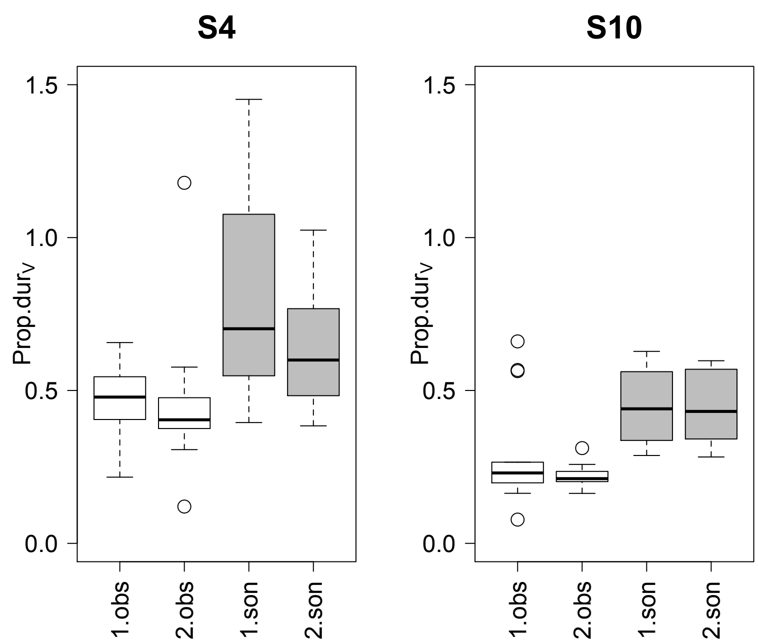


Figure 2.2: *Proportional vowel durations (only accented tokens) of two speakers separately for words with a simplex (1) or a complex (2) coda containing obstruent (obs, white) or sonorant (son, gray) codas. Speaker S4 (left) shows vowel shortening, speaker S10 (right) does not.*

Although in general there was no compensatory vowel shortening, further analysis showed that some of the speakers showed vowel shortening (e.g. speaker S4 in Figure 2.2) while others did not (e.g. speaker S10 in Figure 2.2). That is, vowel shortening seems to be speaker-dependent.

Figure 2.3 shows the proportional C1 duration differences between simplex and complex codas separately for accented and deaccented words containing obstruent and sonorant codas. There seems to be a tendency towards C1 shortening in complex coda tokens in all contexts, as values lie above zero.

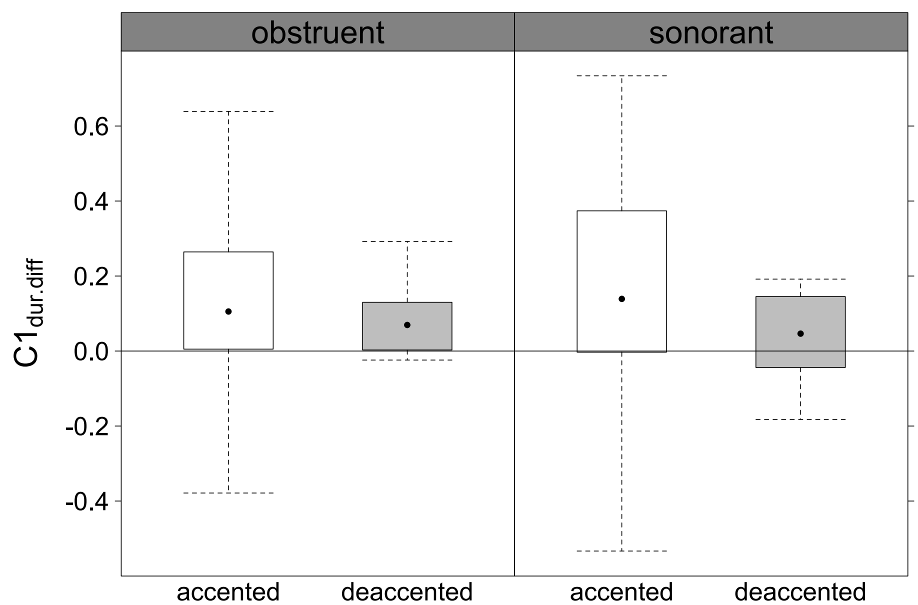


Figure 2.3: *Proportional C1 duration differences between simplex and complex codas separately for accented (white) and deaccented (gray) words containing obstruent and sonorant codas.*

A mixed model with Prop.dur$_{C1}$ as the dependent variable and the same fixed factors as above revealed only a significant effect of accentuation ($\chi^2[1] = 6.5$, $p<0.05$). This means that Prop.dur$_{C1}$ was longer in accented tokens than in deaccented tokens. There was no significant effect of coda complexity or coda manner. However, as mentioned above a shortening tendency of C1 can be observed, as the medians are clearly above zero (see Figure 2.3).

There were no significant effects on either $V_{dur.diff}$ or $C1_{dur.diff}$. This means that – commensurate with Figures 2.1 and 2.3 – neither accentuation nor coda manner had an effect on the degree of vowel or C1 shortening.

Tense vowels

In this section results of the influence of coda clusters and accentuation on the duration of tense vowels will be reported. First, results of *simplex compensatory shortening* are presented followed by the results of *incremental compensatory shortening*. Again, absolute vowel durations were longer in accented (mean = 124.2 ms) than in deaccented (mean = 95.7 ms) tokens.

Simplex compensatory shortening

As far as simplex compensatory shortening is concerned, Figure 2.4 shows the proportional vowel duration differences between words containing no coda consonant (/kniː/) and words containing a coda consonant (/kniːt/). Values lie clearly above zero, indicating that the vowels are generally shorter in words containing a coda consonant than in words containing no coda consonant. Thus there seems to be simplex compensatory shortening.

A mixed model with Prop.dur_V as the dependent variable and fixed factors accentuation and coda complexity revealed a significant effect of accentuation ($\chi^2[1] = 111.2$, $p<0.001$). This means that the proportional vowel duration was longer in accented than in deaccented tokens. Further, there was a significant effect of coda complexity ($\chi^2[1] = 42.1$, $p<0.001$) on Prop.dur_V. This means that the proportional vowel duration was significantly shorter in /kniːt/ than in /kniː/. This is in line with what is evident from Figure 2.4 showing values clearly above zero. Further, this finding is in line with the cross-linguistically attested fact that closed syllables contain shorter vowel durations than open syllables (Farnetani and Kori, 1986; Lindblom and Rapp, 1973; Maddieson, 1985; Waals, 1992).

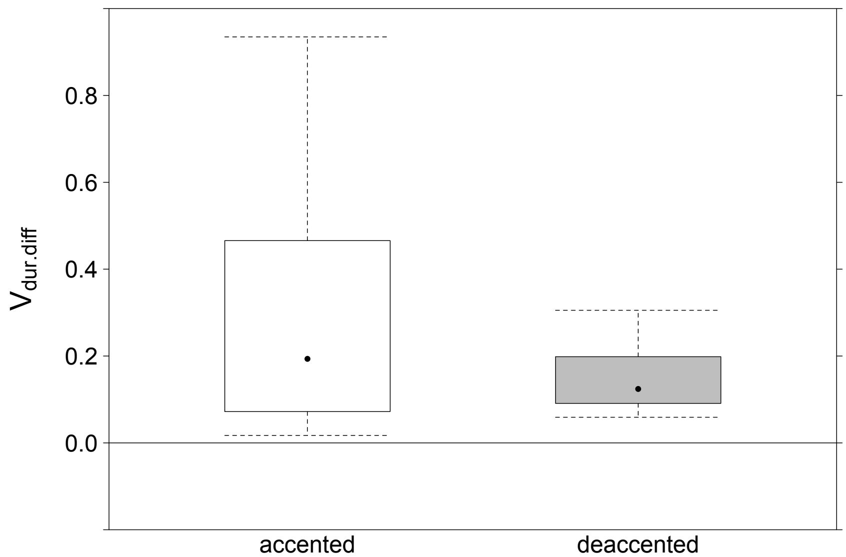


Figure 2.4: *Proportional vowel duration difference between /kniː/ and /kniːt/ separately for accented (white) and deaccented (gray) words.*

A mixed model with $V_{dur.diff}$ as the dependent variable revealed no significant effect of accentuation (cf. Figure 2.4), indicating that accentuation did not influence the degree of vowel shortening. This means that the vowel is shortened in both accented and deaccented words.

Incremental compensatory shortening

Concerning incremental compensatory shortening, Figure 2.5 shows the proportional vowel duration differences between simplex and complex codas separately for accented and deaccented words containing the tense vowels /uː/ and /iː/. $V_{dur.diff}$ seems to differ within the two word pairs (/kniːt/ vs. /kniːst/ (left) and /knuːt/ vs. /knuːtʃ/ (right)). There are negative values in words containing the tense vowel /iː/, indicating that here the vowel was lengthened due to an increasing number of coda consonants in both accented and deaccented tokens. The predominantly positive values in words containing the tense vowel /uː/, however, indicate slight vowel shortening due to an increasing number of coda consonants, especially in the accented condition.

A mixed model with Prop.dur$_V$ as the dependent variable and the fixed factors accentuation, vowel quality and coda complexity revealed a significant effect of accentuation ($\chi^2[1] = 224.0$, $p<0.001$), coda complexity ($\chi^2[1] = 5.8$, $p<0.05$) as well as a significant interaction between coda complexity and vowel quality ($\chi^2[1] = 18.2$, $p<0.01$). This is in line with what is evident from Figure 2.5. Post-hoc Tukey tests revealed that there were durational vowel differences only in /kniːt/ vs. /kniːst/ in both accented and deaccented words ($p<0.001$). This means, that the vowel was lengthened due to an increasing number of coda consonants in this particular word pair. The observed slight vowel shortening in /knuːt/ vs. /knuːtʃ/ turned out to be not statistically significant. The only significant effect on V$_{dur.diff}$ – commensurate with Figure 2.5 – was vowel quality ($\chi^2[1] = 25.8$, $p<0.001$).

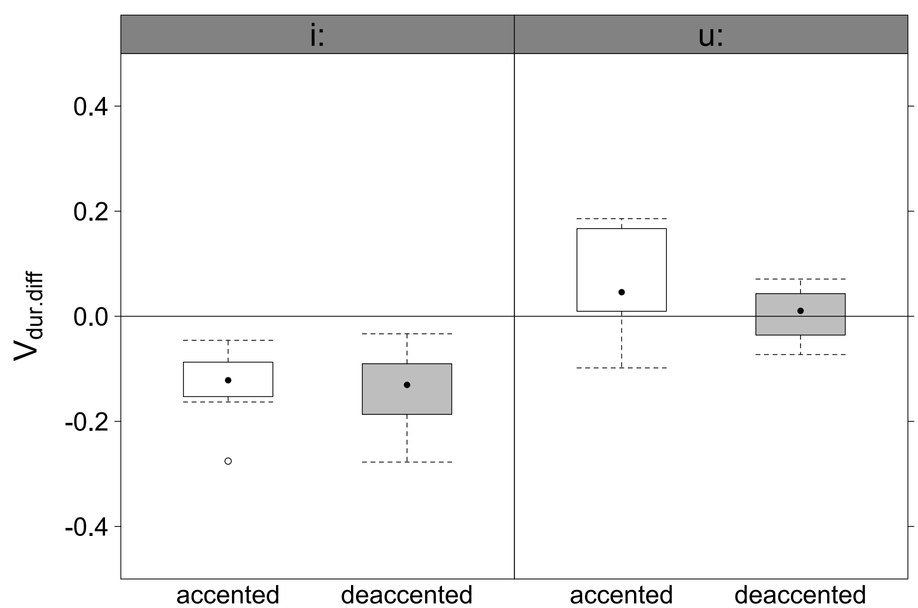


Figure 2.5: *Proportional vowel duration difference between simplex and complex codas separately for accented (white) and deaccented (gray) words containing the tense vowels /uː/ (right) and /iː/ (left).*

Figure 2.6 shows the proportional C1 duration differences (only for /knuːt/ vs. /knuːtʃ/) between simplex and complex codas separately for accented and deaccented words. It is obvious that C1 is shortened with an additional consonant. A mixed model with $C1_{dur.diff}$ as the dependent variable and fixed factors accentuation and coda complexity showed significant effects of accentuation ($\chi^2[1] = 48.3$, $p<0.001$) and coda complexity ($\chi^2[1] = 12.2$, $p<0.001$). This means that C1 was longer in accented tokens than in deaccented tokens and that C1 was significantly shorter in words containing a complex coda than in words containing a simplex coda.

There was no significant effect on $C1_{dur.diff}$. This means that – commensurate with Figure 2.6 – accentuation had no effect on the degree of C1 shortening.

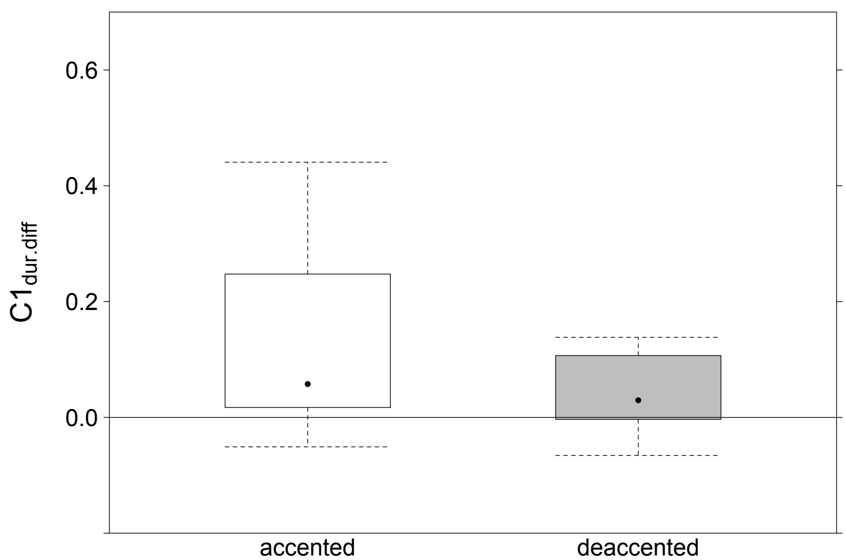


Figure 2.6: *Proportional C1 difference between simplex and complex codas separately for accented (white) and deaccented (gray) words.*

2.5.8 Interim summary: speech production

The results of acoustic vowel duration in syllables containing simplex and complex codas revealed that overall there is no incremental vowel shortening before complex codas in the production of German monosyllables. However, there are speakers showing vowel shortening, while others do not. This compression effect, thus, seems to be speaker-dependent. In addition there were tendencies toward vowel shortening in contexts favoring lengthening, such as in accented words, in words containing lax vowels preceding sonorants and in words containing tense vowels. Moreover, the analyses showed that there were tendencies toward C1 shortening in words containing coda clusters compared to words containing singletons for all speakers. This C1 shortening effect is independent of prosodic structure and coda manner and might be a reason for the less pronounced vowel shortening.

2.6 Speech Perception

A discrimination experiment was conducted to test whether listeners compensate for the shortening effects found in the productions of some speakers. Contrary to previous studies investigating listeners' compensation for compensatory shortening (Fowler and Thompson, 2010; Harrington et al., 2015), natural stimuli of five selected speakers from the production study instead of manipulated stimuli were used.

2.6.1 Participants

21 subjects (15 females and 6 males aged 20 to 55 years) with no hearing or reading disorders participated in the perception experiment. None of them were recorded for the production experiment described above. Some participants were undergraduate students of phonetics but again they were all naïve as to the purpose of the experiment.

2.6.2 Material

Realizations of all target words from the production experiment served as stimuli in the perception experiment, i.e. context sentences were not given. Productions from speakers – among them speaker S4 from Figure 2.2 – were chosen for whom compensatory shortening or no shortening was found. Despite the overall bias towards vowel lengthening in the word-pair /kniːt/ vs. /kniːst/, there were again speaker-dependent differences. For, example speaker S9 did not lengthen the vowel, so there were shortened or non-shortened word-pairs. Since C1 was also shortened in the production data, both vowel and consonant shortening were included in the perception experiment[2]. In order to test whether listeners are sensitive to the shortening of a particular segment or a combination of segments, productions matching the following combination criteria were selected:

sd.0 = no vowel shortening + no consonant shortening

sd.1 = vowel shortening + no consonant shortening

sd.2 = no vowel shortening + consonant shortening

sd.3 = vowel shortening + consonant shortening

Table 2.3: *Durational differences in V and C1 length between simplex and complex tokens of word pairs used for perception experiment.*

Degree	**Word-pairs**	**Durational differences (ms)**	
		accented	deaccented
sd.0	/knɪk/ - /knɪkt/	V: -3.65 ; C1: 3.05	V: 13.06 ; C1: 10.33
	/klɪk/ - /klɪkt/	V: 14.91 ; C1: -6.08	V: 6.4 ; C1: 3.1
	/knal/ - /knalt/	V: 8.96 ; C1: 14.90	V: 4.69 ; C1: 1.34
	/klɪŋ/ - /klɪŋt/	V: 7.70 ; C1: 2.65	V: 5.78 ; C1: 7.05
	/kniː/ - /kniːt/	V: -6.37	V: 3.1
	/kniːt/ - /kniːst/	V: 11.66	V: 0.31
	/knuːt/ - /knuːtsch/	V: 3.55 ; C1: -1.78	V: 11.56 ; C1: -10.44
sd.1	/knɪk/ - /knɪkt/	V: 28.74 ; C1: 7.54	V: 25.64 ; C1: 1.25
	/klɪk/ - /klɪkt/	25.02 ; C1: 7.97	/
	/knal/ - /knalt/	V: 27.74 ; C1: 11.55	V: 32.77 ; C1: 14.31
	/klɪŋ/ - /klɪŋt	V: 26.10 ; C1: 2.81	V: 28.05 ; C1: 13.59
	/kniː/ - /kniːt/	V: 104.95	V: 62.78
	/kniːt/ - /kniːst/	V: 67.64	V: 26.34
	/knuːt/ - /knuːtsch/	V: 57.58 ; C1: 3.64	/

[2]Again, this was not done for the following words /kniː/ - /kniːt/ - /kniːst/ because of the missing constant coda consonant. Therefore, in these tokens only sd.0 and sd.1 combinations were available.

sd.2	/knɪk/ - /knɪkt/	V: 16.42 ; C1: 50.71	V: 14.21 ; C1: 72.80
	/klɪk/ - /klɪkt/	V: 15.22 ; C1: 37.71	V: 1.62 ; C1: 28.61
	/knal/ - /knalt/	V: 3.66 ; C1: 30.17	V: 7.69 ; C1: 26.94
	/klɪŋ/ - /klɪŋt	V: 10.84 ; C1: 53.92	V: 7.91 ; C1: 33.63
	/knuːt/ - /knuːtsch/	V: 1.63 ; C1: 34.19	V: 20.17 ; C1: 36.38
sd.3	/knɪk/ - /knɪkt/	V: 31.52 ; C1: 26.52	/
	/klɪk/ - /klɪkt/	V: 48.34 ; C1: 95.85	V: 42.34 ; C1: 42.63
	/knal/ - /knalt/	V: 37.27 ; C1: 44.47	V: 39.50 ; C1: 44.59
	/klɪŋ/ - /klɪŋt	V: 52.43 ; C1: 68.85	V: 39.99 ; C1: 26.15
	/knuːt/ - /knuːtsch/	V: 47.24 ; C1: 26.00	V: 46.33 ; C1: 29.43

This resulted in five selected speakers (4 females, 1 male). Seven word pairs × two accentuation patterns × four combinations × six repetitions should result in 336 stimuli. However, not all combinations were found for some word pairs. Therefore, there were only 270 stimuli in total. Details for durational differences in vowel and C1 durations are given in table 2.3. Differences in vowel and C1 durations between simplex and complex words above 25 ms were treated as 'shortened', anything below 25 ms was treated as 'non-shortened'.

2.6.3 Experimental set-up

The discrimination task was conducted using the Praat ExperimentMFC-script. Words with simplex codas were paired with their complex coda counterparts, i.e. only *kling* and *klingt* were presented to the listener not *kling* and *knallt*. Only accented simplex words were paired with accented complex words. Likewise, only deaccented simplex words were paired with deaccented complex words. Each stimulus pair was repeated six times and presented counterbalanced in randomized order via headphones (Bayerdynamic DT770). Subjects were asked to identify the word that contains the longer vowel (i.e. the word with a simplex coda or the word with a complex coda). They had to click on the word, which was given in orthographic transcription, they perceived as containing the longer vowel. Once the participants had made their judgment, the next stimulus pair was presented automatically.

2.6.4 Statistics

As there were some empty cells in the data, generalized linear mixed models were conducted. The listeners' judgments (two levels: simplex vs. complex) always served as the dependent variable. Accentuation (two levels: accented vs. deaccented) and coda manner (two levels: sonorant vs. obstruent) were entered as fixed factors in the first analysis. For perception, stimuli containing tense and lax vowels were directly compared in the second analysis. Accentuation (two levels: accented vs. deaccented) and vowel tensity (two levels: lax vs. tense) were entered as fixed factors in the second analysis. In both analyses, the effect of the shortening combinations (four levels: sd.0, sd.1, sd.2, sd.3) were analyzed in separate mixed models. Speaker and listener were always treated as random effects. The models were implemented using the lme4 package (Bates, 2008) ('family=binomial') in the statistical environment R.

2.6.5 Results

Obstruent vs. Sonorant

Figure 2.7 shows the proportion of complex and simplex responses separately for accented and deaccented, obstruent and sonorant tokens. If listeners compensated for compensatory shortening, complex judgments would be above 50%. However, as is evident, listeners generally perceived the vowel to be longer in words with a simplex coda as opposed to the same tokens with a complex coda in 78.0% of all instances. A mixed model with the listeners' judgments as the dependent variable and fixed factors accentuation and coda manner revealed a significant effect of coda manner ($\chi^2[1] = 8.3$, $p<0.01$) and a significant interaction between coda manner and accentuation ($\chi^2[1] = 9.4$, $p<0.01$). Post-hoc Tukey tests revealed that – commensurate with Figure 2.7 – listeners judged the vowel to be longer significantly more often before simplex codas when the coda contained a sonorant ($p<0.05$) and when the tokens were originally produced in accented position ($p<0.05$).

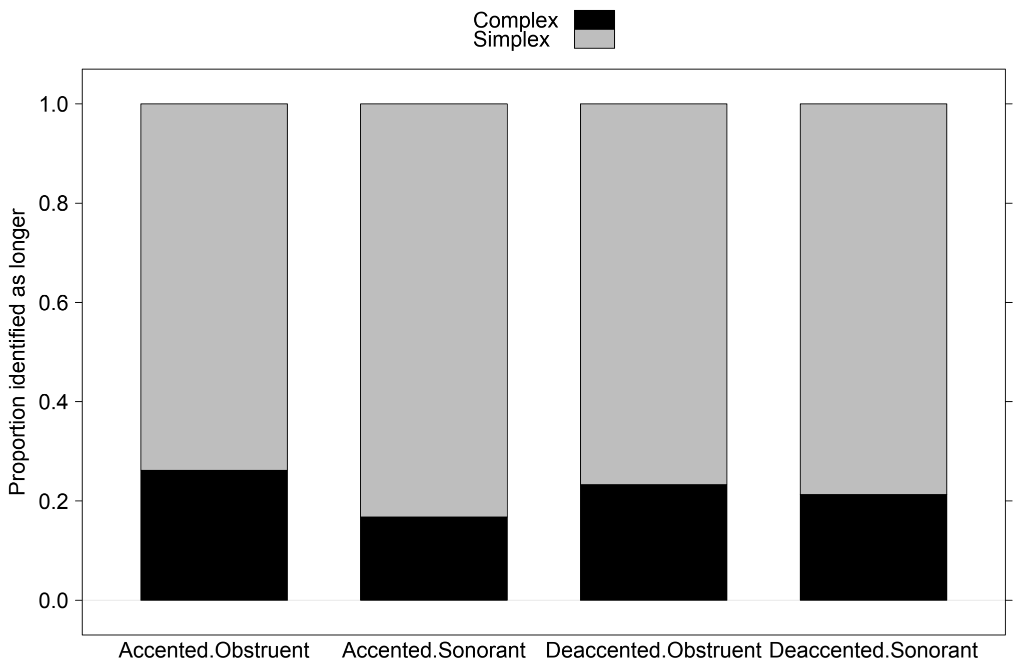


Figure 2.7: *Proportion of complex (black) and simplex (gray) responses separately for accented and deaccented, obstruent and sonorant tokens.*

In addition, the shortening combination influenced the proportion of responses ($\chi^2[3] = 17.0$, p<0.001), but overall listeners perceived vowels in about two thirds of all combinations to be longer before simplex than complex codas. As the combination effect was most pronounced for sd.0 versus sd.3 (and in these stimuli the absolute as well as the proportional vowel and C1 durations were either the same or shortened), only the statistics on this comparison will be reported: listeners perceived non-shortened vowels in sonorant pairs as longer significantly more often than in shortened vowels (accented: $\chi^2[1] = 26.4$, p<0.001, deaccented: $\chi^2[1] = 20.1$, p<0.001). This means that word pairs containing acoustically similar vowels (sd.0) were judged as having a longer vowel before a complex coda more often than word pairs containing shortened vowels (sd.3). This can be seen by means of higher black bars in sd.0 compared to sd.3 in the columns on the right-hand side of Figure 2.8. This finding suggests a tendency for compensation for compensatory shortening, although the overall percentage for as longer identified vowels before simplex codas is too high to speak of compensation for coarticulation.

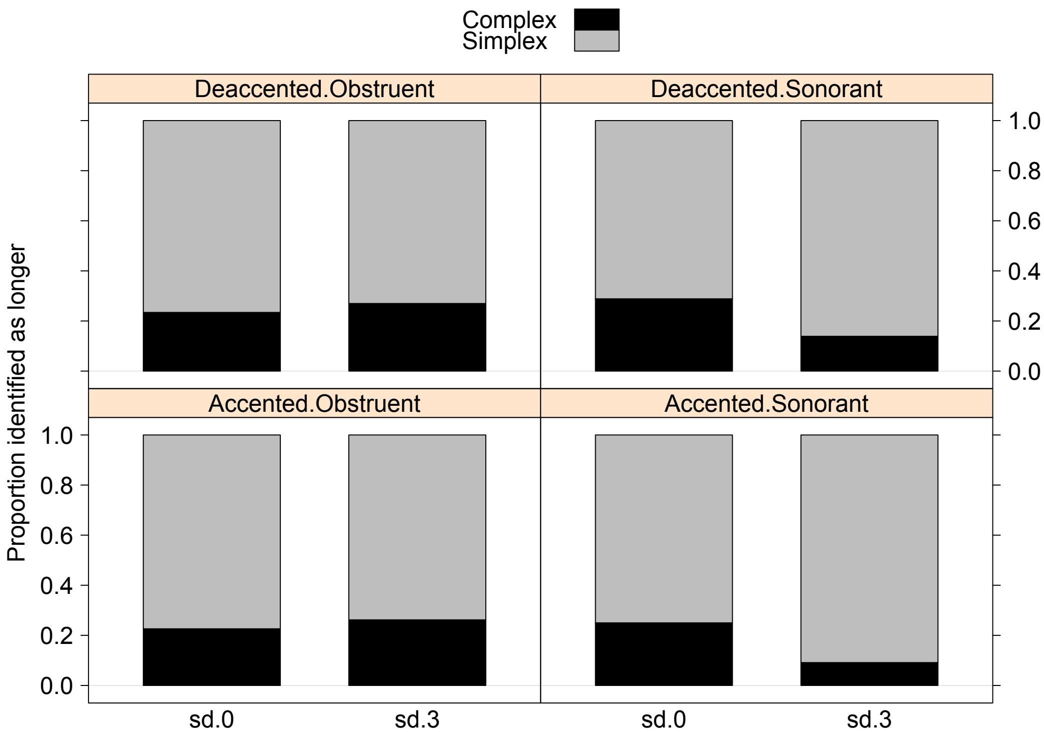


Figure 2.8: *Proportions of complex (black) and simplex (gray) responses in the two shortening combinations separately for accented (bottom row) and deaccented (top row), obstruent (left column) and sonorant (right column) tokens.*

Tense vs. lax vowels

Figure 2.9 shows the proportions of complex and simplex responses separately for accented and deaccented, tense and lax vowel tokens. Again, listeners perceived the vowel to be longer in words with a simplex coda as opposed to the same tokens with a complex coda in 74.8% of all instances. This indicates that listeners did not compensate for compensatory vowel shortening.

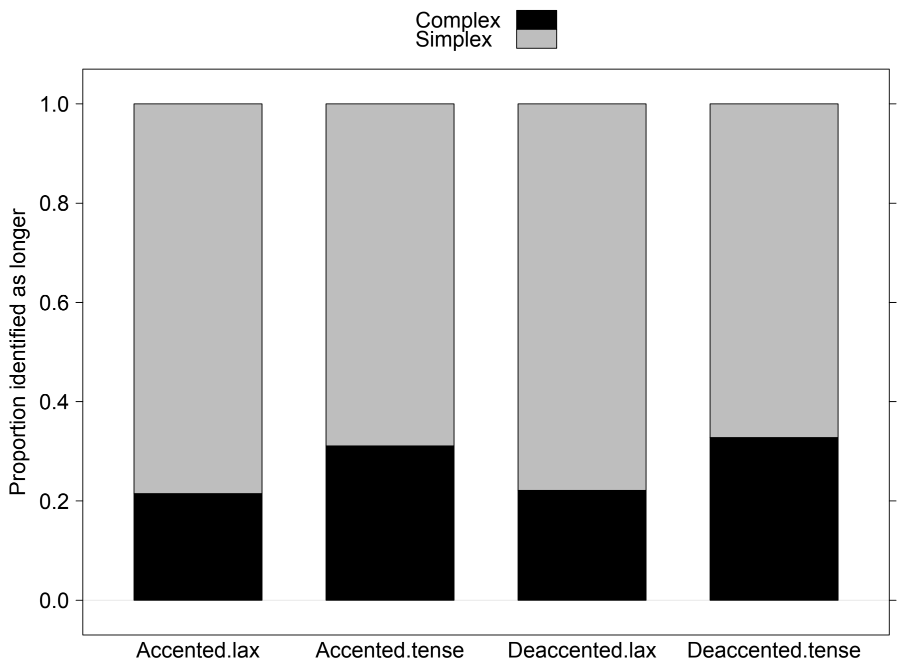


Figure 2.9: *Proportions of complex (black) and simplex (gray) responses separately for accented and deaccented, tense and lax vowel tokens.*

A mixed model with the listeners' judgements as the dependent variable and fixed factors accentuation, vowel tensity showed a significant effect of vowel tensity ($\chi^2[1] = 4.7$, $p<0.05$). This means that – commensurate with Figure 2.9 – listeners identified the vowel significantly more often as being longer before complex codas when the vowel was tense than when the vowel was lax, irrespective of whether the stimuli were originally produced in accented or deaccented position. Thus listeners do show some tendencies to compensate for coarticulation in words containing tense vowels.

A closer look into detail regarding the two most prone shortening combinations (i.e. sd.0 and sd.3) showed that listeners identified the vowels differently depending on vowel tensity. In word pairs with shortening (sd.3), listeners judged the vowel as longer in complex codas in deaccented tense than accented tense tokens ($\chi^2[1] = 10.8$, $p<0.001$) (see higher black bar in deaccented tense tokens of Figure 2.10). Further, in word pairs with no durational differences (sd.0), listeners perceived the vowel to be longer in complex coda tokens significantly more often in words containing tense vowels than in words containing lax vowels (accented: $\chi^2[1] = 14.9$, p<0.001; deaccented: $\chi^2[1] = 10.9$, p<0.001) (see higher black bars in accented and deaccented tense tokens of Figure 2.10).

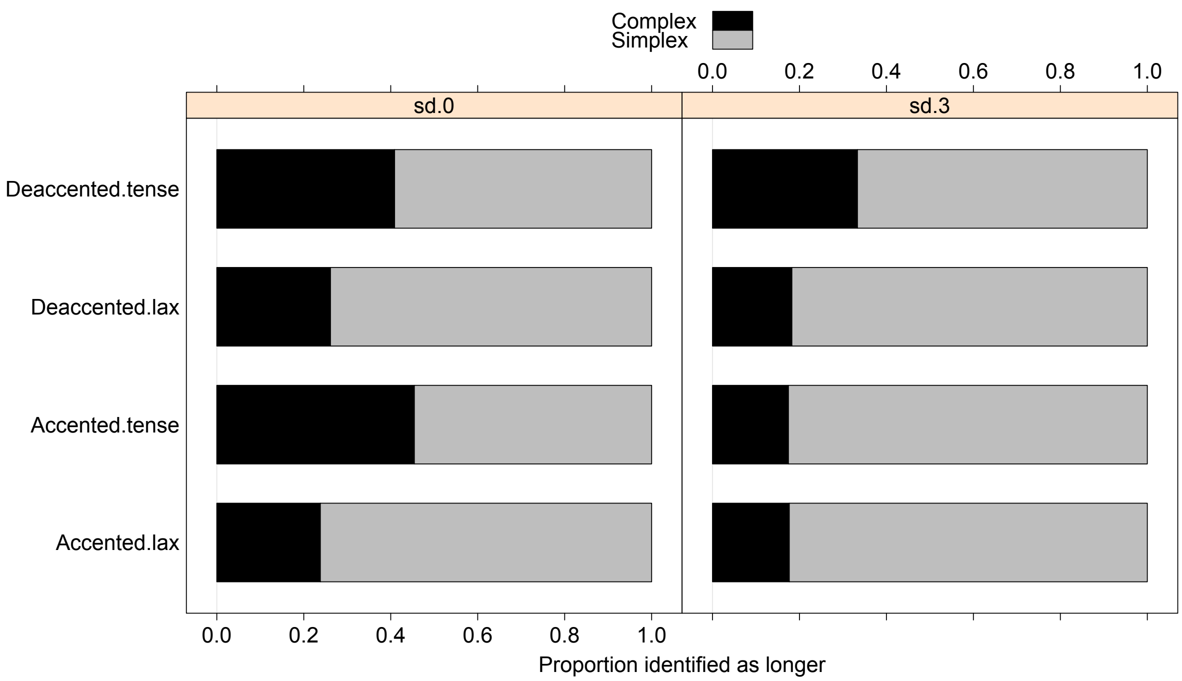


Figure 2.10: *Proportions of complex (black) and simplex (gray) responses in the two shortening combinations (sd.0 and sd.3) separately for deaccented tense, deaccented lax, accented tense and deaccented lax vowels.*

This means, that in ambiguous stimuli, where the absolute and proportional durations of V and C1 did not differ, listeners showed tendencies to compensate in words containing tense vowels. This finding suggests a tendency for compensation for compensatory shortening in words containing tense vowels, although the overall percentage for longer identified vowels before simplex codas is too high to speak of compensation for coarticulation.

2.6.6 Interim summary: speech perception

The investigation of listeners' perception of compensatory vowel shortening revealed only tendencies toward compensation for compression effects. In ambiguous stimuli, listeners' tended to compensate: they perceived the vowel with the same duration to be longer in words with complex codas more often than vowels differing in duration. However, the fact that listeners perceived longer vowels in words containing simplex codas in almost two-thirds of all instances indicates that listeners perceptually shorten the vowel duration in words containing complex codas.

2.7 Summary and discussion

The aim of the present study was to investigate incremental coda compensatory shortening in the production and perception of German monosyllables and the influence of accentuation, coda manner of articulation and vowel tensity on the degree of compensatory shortening. Recent studies on this field revealed quite different or even contradictory results (e.g. Crystal and House, 1988; Katz, 2012; Marin and Pouplier, 2010).

There were several findings from this study. Overall it can be concluded that there is no compensatory vowel shortening before complex codas in the production of German monosyllables, contradicting hypothesis **H1**. Further, this finding contradicts previous studies (e.g. Katz, 2012; Munhall et al., 1992; Shaiman, 2001) reporting vowel compression in words containing coda clusters. However, there were tendencies towards vowel shortening in words containing lax vowels preceding sonorant codas and in words containing tense vowels (only for /knuːt/ - /knuːtʃ/), especially in the accented condition. In addition, there were speaker-dependent differences: some speakers produced the vowel with a shorter duration as the number of coda consonants increased and this effect was even greater in contexts that favour shortening. These speaker-dependent differences are in line with Byrd (1995), who reported a c-center organization of coda clusters for some speakers and not for others.

Concerning the influence of (de)accentuation on the acoustic duration of lax vowels, results showed that only lax vowels preceding sonorant codas were shortened, while no such shortening was found for lax vowels preceding obstruent codas. At first, this prosody-dependent shortening effect contradicts findings from Mooshammer and Fuchs (2002) and Mooshammer and Geng (2008), showing that prosodic weakening (due to deaccentuation) shortens tense but not lax vowels. This *incompressibility* (Klatt, 1973) of lax vowels may be explained by their inherently short duration, which could be the reason why they cannot be compressed anymore. However, additional lengthening because of a sonorant coda may lengthen a lax vowel to be compressible in deaccented contexts.

2.7.1 Simplex compensatory shortening

This study investigated also the effect of simplex compensatory vowel shortening, albeit only in one word pair. The results showed that the vowel was shorter in closed CCVC than in open CCV syllables. This finding is in line with previous studies reporting simplex compensatory vowel shortening for a variety of languages (Farnetani and Kori, 1986; Lindblom and Rapp, 1973; Maddieson, 1985; Waals, 1992) and it provides further evidence that this kind of vowel shortening is a cross-linguistically universal pattern.

2.7.2 Incremental compensatory shortening

Lax vowels

Incremental compensatory vowel shortening was investigated in words containing lax vowels followed by either a sonorant or an obstruent coda. It was hypothesized that longer segments – either because of sonorant codas (**H2b**) or of prosody-induced lengthening (**H2a**) – are shortened to a greater extent. However, the lengthening of lax vowels by sonorant codas did not lead to compensatory vowel shortening. Likewise, prosody-induced lengthening did not cause compensatory vowel shortening of lax vowels. That is, contrary to hypotheses **H2a** and **H2b** there were no significant vowel shortening effects in contexts in which vowels are longer and which thus provide more segmental substance for shortening (see Turk and Shattuck-Hufnagel, 2000; White and Turk, 2010). Only a trend was observed in accented words containing a sonorant coda. This particular result is only to some extent in line with findings from Katz (2008, 2012) showing vowel shortening in words with sonorant codas. However, compensatory vowel shortening of lax vowels is speaker-dependent, and those speakers who exhibit vowel shortening effects show more shortening in accented tokens with sonorant codas than with obstruent codas and/or in deaccented tokens. In total, there is some evidence for more compensatory vowel shortening in lax vowels that are contextually lengthened.

In addition, not only the vowel duration was affected by an additional coda consonant but also the vowel following first coda consonant was shortened in complex coda tokens. All speakers – even those who did not shorten the vowel – showed a tendency for C1 shortening in all contexts, i.e. there are compression effects beyond vowel reduction. This finding is in line with the fact that consonants within a cluster are shorter compared to their duration as a singleton (e.g. Haggard, 1973; O'Shaughnessy, 1974; Sigurd, 1973). Further, it supports Harrington et al.'s (2015) finding on polysyllabic shortening that the duration ratio between the vowel and the following cluster (V/C ratio) is shortened in polysyllabic words as opposed to monosyllabic words. Thus not only the duration of the vowel may be compressed by an additional coda consonant but also the duration of the coda consonant or the entire V/C ratio. In the present data on lax vowels C1 shortening may have been a reason why vowel shortening was less marked.

Tense vowels

Concerning tense vowels, there was again only a tendency towards vowel shortening and only in the word pair */knu:t/* - */knu:tʃ/*. In these tokens, the vowel was somewhat shorter before a complex coda than before a simplex coda. This finding contradicts the results for shortening of phonologically long vowels in Dutch (Waals, 1992) and it also refutes hypothesis **H2c**. In addition, there was no shortening effect in the word pair */kni:t/* -

/kni:st/. To the contrary, there was vowel lengthening before a coda cluster. This may be explained by the different kind of coda manner: in simplex tokens the vowel is followed by an obstruent, while in complex tokens the vowel is followed by a sibilant. For English it has been shown that vowels are longer before fricatives than before obstruents (House and Fairbanks, 1953; Peterson and Lehiste, 1960) and this might be also the case in German. Again, there were no differences between accented and deaccented tokens in the degree of shortening/lengthening.

2.7.3 Listener's perception of compensatory vowel shortening

Perhaps the most important finding from this study comes from the perception experiment. Listeners perceived all vowels – irrespective of whether they were shortened or not – in words with a simplex coda to be longer than in words with a complex coda in about two-thirds of all stimuli. For word pairs with large duration differences, this result corresponds with Fowler and Thompson (2010), who found that longer vowels in monosyllables were perceived as such. This indicates that listeners pay attention to durational differences (see also Davis et al., 2002; Salverda et al., 2003).

Although vowel shortening was present in words containing sonorant codas at least for some speakers, vowels in accented tokens with sonorant simplex codas were significantly more often identified as being longer than their vowel counterparts before complex codas as opposed to vowel pairs in obstruent and deaccented contexts. This finding suggests that listeners anticipate a greater shortening effect in vowels preceding complex sonorant codas. This then implies that listeners do not compensate for compensatory shortening: if listeners compensate for coda-driven vowel shortening, then they should identify vowels with the same duration to be longer more often before complex codas than before simplex codas. By contrast, the results showed that listeners perceptually shorten the vowel in complex coda tokens.

However, listeners perceived vowels in sonorant word pairs with similar V and C1 durations to be longer before complex codas more often than in sonorant word pairs differing in V and C1 duration. Therefore, the finding for ambiguous stimuli is to some extent in line with Fowler and Thompson (2010), too: although the overall bias towards simplex responses indicates perceptual shortening of vowels before complex codas, listeners show a tendency to compensate for compensatory shortening in sonorant tokens. This further supports the hypothesis of greater compensatory shortening effects in more favorable contexts (**H3**), in this case sonorant codas.

Further, the results showed that listeners compensated to a greater extent in tokens containing tense vowels compared with tokens containing lax vowels when there were no durational differences in both prosodic conditions. This means, that in ambiguous stimuli – where the absolute and proportional durations of V and C1 did not differ – listeners showed tendencies to compensate in words containing tense vowels. This finding suggests a tendency for compensation for compensatory shortening in words containing tense vowels, although the overall percentage for longer identified vowels before simplex codas is too high to speak of compensation for coarticulation.

The observation that listeners did not compensate for compensatory vowel shortening may be explained by the fact that there was only little evidence that vowels are acoustically shortened before coda clusters. Therefore, listeners might not have much experience in dealing with compensatory vowel shortening in the particular contexts presented to them.

2.8 Conclusion

This study was a first attempt in investigating coda-driven incremental compensatory vowel shortening in German. The literature, so far, provided evidence for Dutch (Waals, 1992), for English (Katz, 2012; Munhall et al., 1992), for Romanian (Marin and Shigemori, 2014) and for Swedish (Lindblom and Rapp, 1973). However, the results presented in the literature are quite contradicting. For example, Marin and Shigemori (2014) and Katz (2012) reported coda-driven shortening only in words containing a lateral in coda position, but not for words containing an obstruent in coda position. Munhall et al. (1992) and Shaiman

(2001), on the other hand, indeed reported coda-driven shortening in words containing obstruents. The present study revealed that overall there was no compensatory vowel shortening caused by coda clusters. However, there was a trend for longer segments (e.g. lax vowels with sonorant codas; tense vowels; accented tokens) to be shortened to a greater extent than shorter segments (e.g. lax vowels with obstruent codas; deaccented tokens). In addition, this study revealed that shortening effects seem to be speaker-dependent. Further, there were tendencies for C1 shortening in all complex coda tokens compared with their simplex counterparts.

The production study presented here included real words containing complex onsets (either /kl/ or /kn/). These onset clusters were chosen, because they were shown to exhibit different intergestural timing patterns (Bombien et al., 2013; Hoole et al., 2009), which in turn may affect the acoustic vowel duration. It has to be considered, though, that the vowel in words with complex onsets should – according to the predictions made by Articulatory Phonology – be shorter than in words with simplex onsets. For example, the duration of the vowel should be compressed in /knɪk/ compared with /nɪk/. An additional coda consonant as in /knɪkt/ may, therefore, not compress the vowel anymore. This and the observed C1 shortening may account for the fact that vowel shortening was less pronounced.

The findings from the perception experiment suggests perceptual vowel shortening before coda clusters instead of compensation for compensatory shortening. However, in contexts that are more prone to shortening in production (i.e. accented words containing a sonorant coda and accented words containing tense vowels) there was also a trend towards compensation. The perception experiment conducted in the present study included natural productions from the production experiment. A perception experiment which includes synthetic stimuli is needed in order to eliminate possible other acoustic cues, such as acoustic shortening effects of the onset consonants in complex coda tokens.

Chapter 3

Articulatory mechanisms underlying incremental compensatory shortening

3.1 Introduction

This study examines whether compensatory shortening can be caused by both onset and coda clusters. The results of the previous chapter have revealed some evidence that lax vowels preceding sonorants and tense vowels are shortened to a greater extent in words containing complex codas than lax vowels preceding obstruents. For production, there is thus a tendency for longer segments to be shortened to a greater extent. Thus in the present study only words containing the tense vowel /aː/ were included. The overall aim of this study was to investigate the articulatory mechanisms leading to acoustic onset- and coda-driven vowel shortening. To do so, two temporal mechanisms proposed in the literature by Katz (2012) and three spatial mechanisms were taken into account. The proposed temporal mechanisms are vowel plateau compression and increased overlap between the vowel and its adjacent consonant in either onset (onset-driven shortening) or coda position (coda-driven shortening). The spatial parameters considered here are greater stiffness of the vowel gesture and coarticulatory influences on tongue back height (i.e. vowel target undershoot) as well as influences on jaw height.

3.1.1 Effects on vowel duration

The duration of vowels is influenced by a variety of segmental and suprasegmental factors. For example, it is well-documented that vowels are longer in accented contexts than in unaccented ones (Lehiste, 1970), before voiced than voiceless consonants (Peterson and Lehiste, 1960) and before sonorants than obstruents (Chen, 1970). The number of the surrounding consonants also affects the duration of a vowel. For example, vowels are longer in CV and VC syllables than in CVC syllables. This phenomenon is known as *simplex compensatory shortening* and is attested for various languages (Maddieson, 1985). In addition, vowels are longer in CVC than in CCVC (i.e. onset shortening) or CVCC (i.e. coda shortening) syllables. This vowel shortening due to an increasing number of consonants within a syllable is commonly known as *incremental compensatory shortening*. Compensatory shortening caused by clusters in onset position has been found in a variety of languages, e.g. for Swedish (Lindblom and Rapp, 1973), for English (Katz, 2012; Marin and Pouplier, 2010), for Italian (Farnetani and Kori, 1986) and for Romanian (Marin and Shigemori, 2014). Furthermore, also compensatory shortening due to an increasing number of consonants in coda position has been found in various languages, e.g. for Dutch (Waals, 1992), for English (Katz, 2012; Munhall et al., 1992), for Romanian (Marin and Shigemori, 2014) and for Swedish (Lindblom and Rapp, 1973).

3.1.2 Articulatory Phonology and the coupled oscillator model

Recent studies focusing on differences in the temporal coordination of articulatory gestures in onset and coda position have shown that consonant-to-vowel timing as well as consonant-to-consonant timing differs in onset and coda position. It has been proposed that this is due to the global timing of onset clusters with the following vowel (Browman and Goldstein, 1986, 1988; Pouplier, 2012). For onset clusters, the midpoint of the entire cluster – the so-called *c-center* – remains in a stable relationship with the vowel. This means that the midpoint of a single onset consonant is equal to the midpoint of an entire onset cluster (Marin and Pouplier, 2010). If an onset cluster shows c-center organization, the consonant adjacent to the vowel is expected to shift towards the vowel relative to its timing as a singleton, whereas the vowel-remote consonant shifts away from the vowel. This rightward shift of the vowel-adjacent onset consonant then results in greater overlap between the consonant and the vowel. Therefore, this articulatory shift is also predicted to influence acoustic vowel duration. Given that the c-center effect is hypothesized for onsets, complex onsets should cause acoustic compensatory shortening. Coda clusters, on the other hand, are proposed to be timed locally: the cluster's left edge is in a stable relationship with the preceding vowel irrespective of the number of consonants within the cluster (Browman and Goldstein, 1988, 2000; Pouplier, 2012). This sequential organization of coda clusters with the preceding vowel does not result in greater VC overlap and, therefore, complex codas should not cause acoustic shortening of the preceding vowel (Nam et al., 2009).

In general then onset consonants are supposed to be linked to the vowel in a different way than coda consonants. For onsets, the coupled oscillator model of Nam and Saltzman (2003), Nam et al. (2009) and Goldstein et al. (2009) assumes that each consonant on its own is in an in-phase timing relation with the vowel. In addition, the onset consonants are also in an anti-phase timing relation with each other. This additional timing relation causes the consonants to be produced one after the other, and thus to be perceptible. The resulting competition among these two timing relations leads to gestural overlap within the two onset consonants on the one hand, and between the vowel-adjacent consonant with the vowel. This leads then to the c-center effect and to a shift towards the vowel. For coda consonants, on the other hand, different timing patterns are assumed: only the vowel-adjacent consonant is timed anti-phase with the vowel and the consonants within a coda cluster are timed anti-phase with each other, resulting in a sequential order. However, findings from recent studies challenge the assumption of globally timed onset clusters and sequentially ordered coda clusters (see e.g. Byrd, 1995) as there is increasing acoustic and articulatory evidence for compensatory shortening caused by complex codas (Katz, 2012; Munhall et al., 1992). There are also some indications that global timing of onset clusters depends on consonant manner (Marin, 2013).

3.1.3 Results for coda-driven vowel shortening

In the literature, onset-driven shortening has been reported for various types of onset composition (Katz, 2012; Marin and Pouplier, 2010), which is in line with Articulatory Phonology predicting a c-center effect for onset clusters in general. For coda-driven vowel shortening, however, there are mixed or even contradictory results depending on the language, the manner of the following coda consonant and the vowel quality. For example, some researchers found coda-driven shortening in English (Katz, 2008, 2012) while others did not (Crystal and House, 1982, 1988, 1990). Furthermore, in some studies coda-driven shortening was observed only in vowels preceding sonorants (such as /rp/ and /lp/ coda clusters) but not preceding obstruents (Byrd, 1995; Katz, 2008, 2012; Marin and Pouplier, 2010), whereas other studies indeed found evidence for vowel shortening in words containing obstruent codas such as /ps/ and /sp/ (Munhall et al., 1992; Shaiman, 2001). In addition, long vowel shortening in Dutch has been observed for lateral codas, while only a small effect of obstruent codas could be found (Waals, 1992). These findings suggest that the manner of the following coda consonant and the vowel quality have some influence on whether the vowel is shortened or not. Based on these findings, the present study includes words containing both obstruent and sonorant codas. Any acoustic vowel shortening effect is expected to be greater in words containing sonorant codas compared with words containing obstruent codas. Further, chapter 2 has revealed some evidence that vowel shortening is more likely to be observed in tense vowels. Therefore, only tense vowels are included in the present study.

3.1.4 Influences of consonant manner on gestural overlap

If vowel shortening is a consequence of differences in articulation (i.e. increased overlap), then vowel compression should be observed – according to Katz (2012) – in all words with complex codas irrespective of the consonant's manner of articulation. Since shortening effects were found to depend on consonant manner, he concludes that a theory in which constraints on vowel duration are stated in auditory terms (e.g. Campbell and Isard, 1991;

Flemming, 2001; Myers, 1987) is much more appropriate, because the amount of coda-driven vowel shortening may depend on the perceptual relationship between the vowel and the following coda consonant. For instance, Katz (2012) observed that vowels are shortened preceding liquids but not preceding obstruents. He concludes that this is because liquids contain more information about the preceding vowel than obstruents do and this helps to satisfy the durational requirements of the preceding vowel (see also Katz, 2010). However, this perceptual approach does not account for Munhall et al.'s (1992) and Shaiman et al.'s (2001) finding that obstruent codas indeed cause vowel shortening. In addition, the vocalic dorsal gesture required for coda /l/ may be produced simultaneously with the vowel gesture (see Sproat and Fujimura, 1993), so that the vowel itself is not shortened at all.

Also in articulatory studies consonant manner has been shown to affect the timing between mainly onset clusters and their timing with the following vowel within a syllable. For example, Marin (2013) showed evidence for global organization of /sk/ onsets but not for /ps/ onsets. Goldstein et al. (2009) showed a significantly greater rightward shift, i.e. more overlap with the following vowel, in /sp/ onsets than in /pl/ onsets. Thus it may be the case that manner has also some influence on the timing between vowels and their adjacent coda consonants. The effect of manner on global timing of onset clusters is very likely a consequence of inter-consonantal timing, which can be measured in terms of overlap. For example, Bombien et al. (2013) and Hoole et al. (2009) reported more overlap between /k/ and /l/ than between /k/ and /n/ in German and French. Although they made no claims as to potential c-center organization of these clusters, this difference in intra-cluster timing may in turn affect the global organization proposed for onsets and the consonant's shift towards the vowel. Therefore, those two different types of onset consonants were included in the current study in order to further investigate those findings. Based on the results presented in Bombien (2011) and Hoole et al. (2009), more CC overlap between /k/ and /l/ than between /k/ and /n/ is predicted. Because of this stronger inter-consonantal coupling, less CV overlap resulting in less acoustic vowel shortening in /kl/ tokens is expected.

3.1.5 Articulatory mechanisms

Instances of acoustically measured incremental vowel reductions are, according to Katz (2012), caused by two possible articulatory mechanisms. They may either come about because of compression of the vowel gesture or by a consonantal shift towards the vowel, the latter increasing the CV or VC overlap, respectively. It is well attested, at least for onsets, that there is a shift towards the vowel in complex tokens (e.g. Marin, 2013; Marin and Pouplier, 2010; Pouplier, 2012). According to our knowledge, a combination of these two proposed mechanisms may also be conceivable. Therefore, the vowel plateau duration, which is considered the most stable part in articulatory measures (Oliveira et al., 2004), is expected to be shorter in words containing complex onsets or codas compared with their simplex counterparts. In addition, more CV overlap in words containing complex onsets and more VC overlap in words containing complex codas compared with simplex tokens is predicted.

Beyond durational measures, spatial parameters such as stiffness[1], tongue back and jaw position may broaden our understanding of the articulatory mechanism underlying acoustic compression. Stiffness is a measurement of the relative speed of the articulatory movement independent from the physical displacement of a particular gesture. Since velocity itself increases with physical displacement (Guenther, 1995; Ostry and Munhall, 1985), measuring only a gesture's velocity is not an appropriate measure. There is evidence that stiffness associated with tongue back movements is significantly lower (i.e. tongue back movements are slower) than stiffness associated with tongue tip or lips (Kuehn and Moll, 1976; Roon et al., 2007). This is because the tongue back has a greater mass and is therefore less flexible. Roon et al. (2007) suggest that differences in stiffness between consonants in a cluster containing different articulators may be responsible for differences in overlap. Byrd and Saltzman (1998) have shown that the stiffness of a gesture affects the duration of a movement, i.e. a higher stiffness corresponds to a faster gesture resulting in a shorter

[1] Although stiffness is not a spatial parameter, it is listed together with spatial parameters as it includes the velocity and the maximum displacement of a gesture.

segment (Ostry and Munhall, 1985). Stiffness also varies with speech rate in so far as faster speech rates are associated with higher stiffness (Gay, 1981; Kelso et al., 1985; Kuehn and Moll, 1976). In addition, Löfqvist (2004) found that movement duration is shorter for short consonants, which also have a higher stiffness than long consonants.[2] Therefore, greater stiffness for vowel gestures within complex onset and coda tokens, respectively, compared with simplex tokens is expected.

A further trigger for many kinds of acoustic shortening may be articulatory vowel target undershoot. Stevens and House (1963) noted that lax vowels have a greater tendency to undershoot their target values than tense vowels, which they attributed to the fact that lax vowels are acoustically shorter than tense vowels. Moon and Lindblom (1994) claim that vowel undershoot is a result of coarticulation between a vowel and its adjacent consonants. Especially in contexts favoring coarticulation, such as unstressed syllables or fast speech rate, vowel undershoot arises when a vowel gesture does not achieve its target position, because its vocal tract configuration has been postponed towards that of the adjacent consonant(s) (Moon and Lindblom, 1994). Guenther (1995) argues that in contexts that favor coarticulation the target region for vowels is extended so that a speaker uses more peripheral articulatory configurations. Thus there seems to be a relationship between coarticulation, vowel duration and target undershoot. Assuming that incremental vowel shortening is a result of coarticulation and extensive overlap of the vowel with its adjacent consonant, it can be hypothesized that the vowel gesture does not achieve its target position in complex tokens. Thus articulatory vowel target undershoot in complex onset and coda tokens, respectively, compared with their simplex counterparts is expected.

Munhall et al. (1992) explained their finding of acoustic coda-driven vowel shortening by a truncation of the offset of the jaw opening gesture by the closing gesture in syllables containing coda clusters. In addition, the shorter acoustic vowel durations in unaccented speech compared with accented speech was attributed to the fact that the opening gesture is truncated by the closing gesture (Harrington et al., 1995). Beckman et al. (1992) also di-

[2] However, it has to be considered, that stiffness is displacement proportional to peak velocity. Further, the duration of a segment is a byproduct of its displacement and stiffness.

rectly attributed their observed smaller jaw displacement in unaccented vowels as opposed to accented vowels to truncation. This is in line with the assumption that accented vowels are produced with a greater mouth opening and thus with a longer duration (de Jong, 1995b). Further, Siddins et al. (2014) showed that there was undershoot of the first formant frequency (F1) in the low vowel /aː/ in words containing complex codas compared with words containing simplex codas. F1 is linked to tongue height, i.e. the higher the tongue, the lower F1. For /aː/ a higher tongue position would result in F1 undershoot (Lindblom and Sundberg, 1971). Such a higher tongue position may be due to the tongue itself in complex coda tokens. However, given that the jaw is correlated to the tongue body height during the production of vowels (e.g. Gracco, 1993), a higher jaw position will also lead to F1 undershoot. Therefore, differences in jaw height may account for compensatory vowel shortening. As this study includes also the vowel /aː/ a higher jaw position in complex onset and coda tokens compared with simplex tokens is thus expected.

3.2 Hypotheses

As there are many different hypotheses included in this chapter, they are listed again.
The first two hypotheses relate to acoustic measures of incremental vowel shortening:
H1 Vowels are shorter in words with complex onsets or codas compared with their simplex counterparts.
H2 If vowels are shortened, then this effect will be more pronounced in words containing sonorant codas than in words containing obstruent codas.
H3 and H4 are specifically based on previous findings by Bombien (2011) and Hoole et al. (2009) on inter-consonantal timing and its further implications:
H3 There is more CC overlap between /k/ and /l/ than between /k/ and /n/.
H4 If there is more CC overlap within /kl/ than within /kn/, then CV overlap will be less in /kl/ resulting in less acoustic vowel shortening in /kl/ tokens.
The following two hypotheses specifically test the two articulatory mechanisms in syllable timing that possibly lead to incremental compensatory shortening suggested by Katz

(2012):

H5 The vowel plateau duration will be shorter in words containing complex onsets or codas compared with their simplex counterparts.

H6 There is more CV overlap in words containing complex onsets and more VC overlap in words containing complex codas compared with simplex tokens.

Finally, the last three hypotheses are concerned with spatial mechanisms:

H7 Vowel gestures within complex onset and coda syllables, respectively, show a greater stiffness compared with vowel gestures within simplex syllables.

H8 There is articulatory vowel target undershoot in complex onset and coda tokens, respectively, but not in their simplex counterparts.

H9 The jaw is lowered to a lesser extent in complex tokens than in simplex tokens.

3.3 Method

3.3.1 Subjects

Data from five speakers of Southern Standard German (3 female, 2 male; aged between 19 and 27) were recorded in a sound attenuated booth at the Phonetics Institute in Munich. None of the subjects reported any speech or reading disorders. Three of the five participants were undergraduate students of phonetics, but they were naïve as to the purpose of the experiment. One participant was the author of this thesis.

3.3.2 Data acquisition

Articulatory movement data were collected using 3D Electromagnetic Articulography (EMA) (AG501, Carstens Medizinelektronik) (Hoole et al., 2003). The system allows positions and movement data during speech to be measured by tracking — using an electromagnetic field — the three-dimensional positions of sensors glued to various points on the subject's vocal tract. The articulatory data were sampled at a frequency of 250 Hz. Audio data were simultaneously collected with a shotgun microphone at a sampling rate of 25.6 kHz.

3.3.3 Recordings

In total, ten sensors were used for the recordings. Three sensors were placed on the tongue, spaced at relatively equidistant points: on the tongue tip (attached approximately 1 cm behind the actual tongue tip), on the tongue mid and on the tongue back (attached at the approximate velar constriction region). Additional sensors were placed on the upper and lower lips in order to estimate lip aperture and below the lower teeth to measure jaw movement. Reference sensors were placed on the nose bridge, upper incisor (maxilla) and behind the ears (on the right and left mastoid process). These reference sensors were used for head movement correction. All sensors, except those on the right and left mastoid process, were fixed midsagittaly.

Speakers were familiarized with examples of the utterances before data collection. Seven repetitions of the sentences were presented in randomized order on a computer screen. In total there were 315 stimuli. Speakers were visually cued when to speak by a green box framing the utterance.

3.3.4 Speech material

The test items were non-existent words containing simplex and complex onsets and codas, respectively. To address our research questions, the test items in this study differed from the words usually used in c-center studies. The consonants surrounding the tense vowel /aː/ were kept identical for onset and coda analyses, thus preventing any other influences on vowel duration. This means that for onset and coda analyses, different cluster compositions were exploited. All target words were accented and embedded in a specific carrier phrase. Since the tongue back and jaw position is low for /aː/, for the environment in the carrier phrases the high vowel /i/ was used. For syntactical reasons, two different carrier phrases had to be used for various types of target words. Since German words with complex codas and word-final -t usually indicate a verb, carrier phrase (1) was used only for items ending in -t: thus all items embedded in the following carrier phrase contained complex codas.

Carrier phrase (1):

"Melanies Omi [**target word**] ihm einmal."
("Melanie's grandma [**target word**] him once.")

The remaining words were placed in way as to refer to the name of Melanie's grandma in carrier phrase (2). Thus items embedded in the following carrier phrase all contained simplex codas as well as simplex and complex onsets.

Carrier phrase (2):
"Melanies Omi [**target word**] imitiert ein Lied."
("Melanie's grandma [**target word**] imitates a song.")

Onset tokens

Onset clusters contained either /kn/ or /kl/. Correspondingly, simplex onsets consisted of /n/ and /l/. These clusters were chosen because of their different timing behavior: /k/ and /l/ overlap to a greater extent than /k/ and /n/ (Bombien, 2011; Hoole et al., 2009). A second motivation for choosing them was because of their different articulators (i.e. tongue back and tongue tip), which facilitates their measurement and analysis. The vowel /aː/ was chosen as it requires an active tongue back movement, which is ideal for measuring following the alveolars requiring tongue tip movement. The target words used for onset analyses are listed in table 3.1.

Simplex onset	**Complex onset**
/**n**aːp/	/**kn**aːp/
/**l**aːp/	/**kl**aːp/
/**l**aːm/	/**kl**aːm/

Table 3.1: *Speech material for onset analyses containing simplex and complex onsets.*

Coda tokens

Coda clusters contained either /pt/ or /mt/. Correspondingly, simplex codas consisted of /p/ and /m/. These clusters were chosen because the labials are articulated independently of the vowel gesture. Due to different articulators (i.e. upper/lower lips and tongue tip) the clusters are easy to measure and distinguish. The target words for coda analyses are listed in table 3.2.

Simplex coda	**Complex coda**
/naː**p**/	/naː**pt**/
/laː**p**/	/laː**pt**/
/laː**m**/	/laː**mt**/

Table 3.2: *Speech material for coda analyses containing simplex and complex codas.*

A minimal pair such as /knaːm/ - /naːm/ - /naːmt/ was not included, because /knVm/-sequences are rare in German.

3.3.5 Data segmentation

The segmentation and labeling of the speech signals was first done automatically using MAUS (Schiel, 2004). If necessary, the segment boundaries of the target words were then hand-corrected in Praat (Boersma and Weenink, 1992). This was especially the case in determining the beginning of the vowel. As onsets contained /n/ and /l/, the vowel onset was set at the point of spectral change in F2. For codas containing /m/, the vowels were measured up to the point of spectral change in F2 from vowel to nasal. Vowels preceding stops were measured until the end of periodic structure or vibration in the spectrogram. Stops were divided into silence period and aspiration.

The articulatory segmentation was done in EMU (Harrington, 2010). In order to address our research questions, the following articulators were labeled: tongue back (**TB**; for /k/ and /aː/), tongue tip (**TT**; for /n/, /l/ and /t/), lip aperture (**LA**; for /p/ and /m/) and the jaw (**JAW**). Time points of articulatory landmarks were computed with semi-automatic algorithms using EMA-coil trajectories and their corresponding velocity signals. Different velocity signals were used for different articulators:

TB: Only the vertical (i.e. up-down) dimension was used for the analysis. /k/ is produced by lifting the tongue back to the soft palate and the vowel /aː/ is produced by lowering the tongue back to the pharynx.

TT: The movement of the tongue tip was measured for the alveolars /n/, /l/ and /t/. Alveolar articulation involves both lifting and fronting of the tongue tip. Therefore the tangential velocity of the tongue tip was used for the analysis. Tangential velocity is defined as the square root of the sum of the squared first derivatives of the trajectory's vertical (V_X) and anterior-posterior (V_Y) dimensions: TV = $\sqrt{V_X^2 + V_Y^2}$.
LA: For bilabials /p/ and /m/ two articulators (the upper and lower lips) are involved in the vertical as well as in the anterior-posterior dimension. Therefore, the Euclidean distance (d) between the respective EMA coils (u, l) was calculated as a measure of lip aperture (LA): d = $\sqrt{(u_x - l_x)^2 + (u_y - l_y)^2 + (u_z - l_z)^2}$.
JAW: The jaw is lowered for the back vowel /aː/ and raised for the following high vowel /i/. Therefore, the vertical component was used to measure jaw movement.

The following landmarks were placed to capture a gesture: $\mathbf{G}_{on}$ (gesture onset), $\mathbf{V}_{on}$ (maximum velocity at onset), $\mathbf{P}_{on}$ (beginning of plateau), **max** (maximum), $\mathbf{P}_{off}$ (end of plateau), $\mathbf{V}_{off}$ (maximum velocity at offset) and $\mathbf{G}_{off}$ (gesture offset). Figure 3.1 schematically displays the positioning of all articulatory landmarks for a gesture by the example of tongue tip.

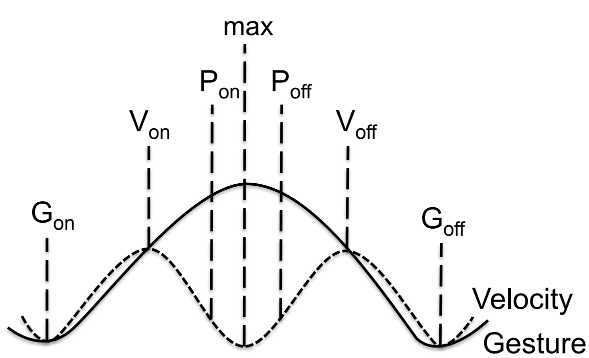


Figure 3.1: *Schematic illustration of landmark positioning for a tongue tip gesture and its corresponding tangential velocity profile.*

The onset and offset of a gesture ($\mathbf{G}_{on}$, $\mathbf{G}_{off}$) and beginning and end of the plateau ($\mathbf{P}_{on}$, $\mathbf{P}_{off}$), are interpolated values representing the 20% threshold of the difference between the two nearby peaks in the velocity signal. For example, the beginning of the plateau ($\mathbf{P}_{on}$) is positioned at the 20% threshold between the two peaks in velocity, i.e. where velocity is at its maximum, and **max** where velocity is at its minimum. In comparison to other static and dynamic thresholds like zero crossing this method seems to be the one yielding the most stable results (Oliveira et al., 2004).

3.3.6 Data analysis

Acoustic measurements

The acoustic vowel duration was measured in order to see whether the vowel was shortened in words containing complex onset and codas, respectively, compared with their simplex counterparts. Here, absolute vowel durations were measured because no appropriate normalization technique could be found.

Articulatory measurements

Concerning the measurement of articulatory timing (e.g. CV_{lag} and VC_{lag}), prior analyses have shown that vowels vary acoustically depending on the presence or absence of a pitch accent on the target word and also on the number of consonants within a cluster (see e.g. Marin, 2013; Marin and Pouplier, 2010; Peters and Kleber, 2013). Because of this (potential) variability, methods applied in other c-center studies where the measurement is based on anchor points either in the following coda consonant (see Marin and Pouplier, 2010) or the acoustic vowel midpoint (see e.g. Pouplier, 2012) were not used, because the measured duration may be variable and an observed shift towards the vowel may be an artefact of this procedure. Aside from that, it was possible to articulatorily label the vowel gesture in the present corpus and therefore to directly measure CV and VC overlap. Thus the normalization method described in Bombien (2011) was applied instead. Following this method, first the lag between two neighboring segments' plateaus was determined and then the lag was normalized over the entire duration of the two gestures. This measurement procedure was applied to CC_{lag}, CV_{lag} and VC_{lag}. The duration of the vowel gesture's plateau ($V_{plateau}$) was calculated by subtracting the plateau offset from plateau onset. Using the start and end of the gesture's plateau was found to be the most stable timing measure in articulatory analyses, i.e. the one with the lowest variation coefficient (Oliveira et al., 2004). All temporal measurements are listed in Table 3.3.6. Further, Figure 3.2 schematically illustrates the specific landmarks used for temporal measurements.

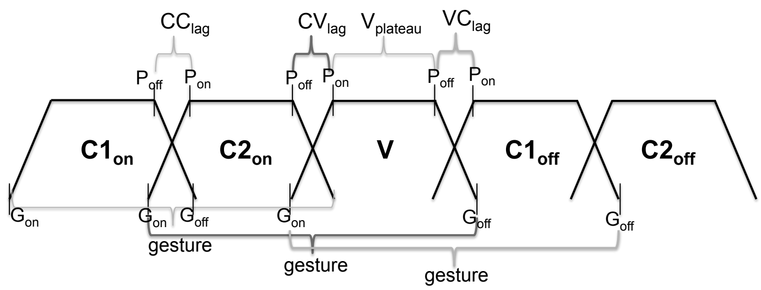


Figure 3.2: *Schematic illustration of gestures and landmarks used in the present analysis.*

Measure	Calculation	Note
CC_{lag}	$= \frac{P_{on}[C2_{on}]-P_{off}[C1_{on}]}{G_{off}[C2_{on}]-G_{on}[C1_{on}]}$	Negative values signify overlapping plateaus; positive values signify a lag between two neighboring segments.
CV_{lag}	$= \frac{P_{on}[V]-P_{off}[C2_{on}]}{G_{off}[V]-G_{on}[C2_{on}]}$	Negative values signify overlapping plateaus; positive values signify a lag between two neighboring segments.
VC_{lag}	$= \frac{P_{on}[C1_{off}]-P_{off}[V]}{G_{off}[C1_{off}]-G_{on}[V]}$	Negative values signify overlapping plateaus; positive values signify a lag between two neighboring segments.
$V_{plateau}$	$= P_{off}[V] - P_{on}[V]$	Higher values signify longer plateaus; lower values signify shorter plateaus.

Table 3.3: *Measurements of articulatory timing.*

In addition to temporal measurements, spatial measurements were included such as the vowel gesture's stiffness and the vertical position of the vowel and jaw gestures. Stiffness is a measurement of articulator movement that characterizes speed independent from its physical displacement as described in equation 3.1.:

$$Stiffness = \frac{\text{peak velocity (mm/ms)}}{\text{maximum displacement (mm)}} \tag{3.1}$$

Figure 3.3 schematically illustrates the stiffness measurement. Maximum displacement is the distance from $\mathbf{G}_{on}$ to **max**, i.e. the extent to which the gesture is lowered. This procedure was applied to all vowel gestures.

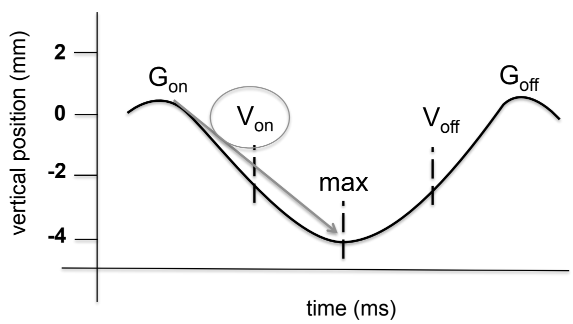


Figure 3.3: *Schematic illustration of stiffness measure.*

Furthermore, the vertical position of both the tongue back and the jaw during the production of tense vowel /aː/ between V_{on} and V_{off} of each gesture was measured as illustrated by the gray circle in Figure 3.4.

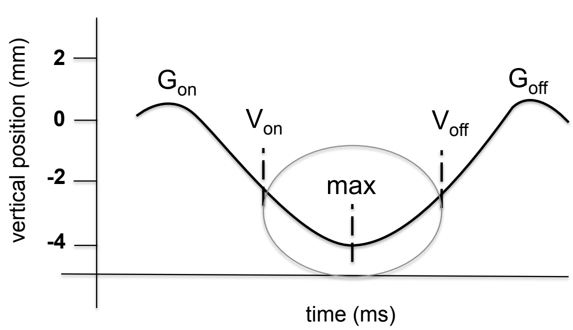


Figure 3.4: *Schematic illustration of vertical tongue back and jaw position measure.*

To minimize speaker-dependent anatomical influences, stiffness, tongue back and jaw position data were Lobanov-normalized (Lobanov, 1971). The measurements on tongue back and jaw height were then time-normalized because of different lengths from V_{on} to V_{off}.

3.3.7 Statistics

For the statistical analysis various mixed models were conducted. Vowel duration, $V_{plateau}$, CC_{lag}, CV_{lag}, VC_{lag}, stiffness, vertical tongue back position and vertical jaw position each served as the dependent variable in one of the models. Onset manner (two levels: nasal vs. lateral), coda manner (two levels: plosive vs. nasal), onset complexity (two levels: simplex vs. complex) and coda complexity (two levels: simplex vs. complex) were the fixed factors. The speaker was entered as a random factor. The models were implemented using the lme4 package (Bates, 2008) in the statistical environment R. Mixed models were preferred because repeated measures ANOVAs are calculated on cell means, which requires data manipulation and entails loss of data.

3.4 Results

The results section is structured into four parts. First, the acoustic vowel duration will be presented depending on onset and coda complexity in section 3.4.1. Then, the results for syllable timing will be presented in section 3.4.2 for onsets first, including CC_{lag}, vowel plateau duration and CV_{lag}, followed by the results for codas, including vowel plateau duration and VC_{lag}. Section 3.4.3 shows results for stiffness separately for onsets and codas. Finally, the results for tongue back and jaw height will be displayed separately for onsets and codas in section 3.4.4.

3.4.1 Acoustic vowel duration

Onset

Prior to the articulatory analyses, it was tested whether an increasing number of consonants in onset position influences the acoustic vowel duration. Words with complex onsets and simplex codas and their counterparts with simplex onsets were included in this analysis (cf. Table 3.1). Figure 3.5 shows the acoustic vowel durations in words containing either simplex (/n/ vs. /l/) or complex onsets (/kn/ vs. /kl/) and the simplex coda (/m/ vs. /p/). It is evident that the acoustic vowel duration is shorter in tokens containing complex onsets than in tokens containing simplex onsets. This effect seems to be independent of onset and coda manner.

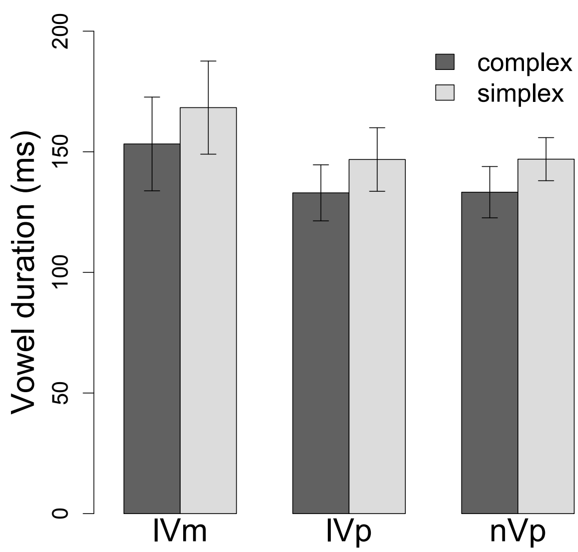


Figure 3.5: *Acoustic vowel durations separately for different sets (lVm, lVp, nVp) containing simplex (light gray) and complex (dark gray) onsets.*

A mixed model with the acoustic vowel duration as the dependent variable and with fixed factors onset complexity (simplex vs. complex), onset manner (/n/ vs. /l/) and coda manner (/m/ vs. /p/) showed a significant effect of onset complexity ($\chi^2[1] = 16.8$, p<0.001). This means that the acoustic vowel duration was shorter in complex than in simplex tokens (cf. Figure 3.5). Furthermore, neither onset, nor coda manner significantly affected the acoustic duration of the tense vowel /aː/ and there was no significant interaction. This indicates that the degree of acoustic vowel shortening remains the same irrespective of /l/ and /n/ onsets and /p/ and /m/ codas.

Coda

In order to test whether an increasing number of consonants in coda position influences the acoustic vowel duration, words with complex codas and simplex onsets and their counterparts with simplex codas (cf. Table 3.2) were included.

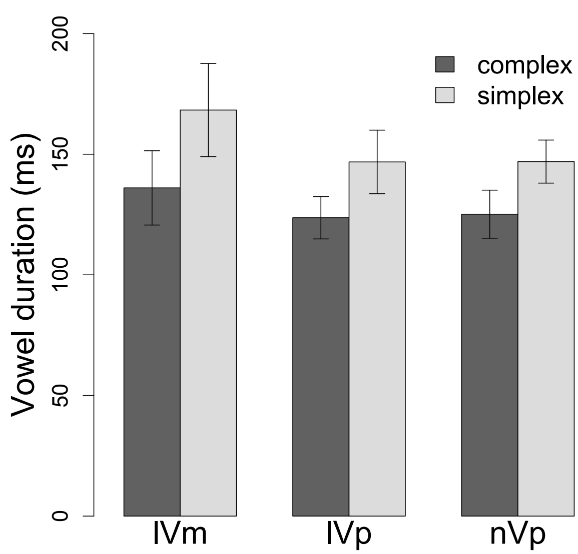


Figure 3.6: *Acoustic vowel durations separately for different sets (lVm, lVp, nVp) containing simplex (light gray) and complex (dark gray) codas.*

Figure 3.6 shows acoustic vowel durations containing either simplex (/p/ vs. /m/) or complex codas (/pt/ vs. /mt/) and the simplex onsets (/n/ vs. /l/). Again, vowel durations seem to be shorter in tokens containing complex codas than in tokens containing simplex codas. This effect seems to be independent of onset and coda manner. A mixed model with the acoustic vowel duration as the dependent variable and with fixed factors coda complexity (simplex vs. complex), onset manner (/n/ vs. /l/) and coda manner (/m/ vs. /p/) showed a significant effect of coda complexity ($\chi^2[1] = 64.9$, $p<0.001$) but no effect of onset or coda manner and no interaction between these factors. This means that there was acoustic vowel shortening driven by an increasing number of coda consonants irrespective of onset and coda manner.

These findings for onset- and coda-driven compensatory vowel shortening are in line with previous studies (Katz, 2008, 2012; Marin and Pouplier, 2010; Munhall et al., 1992; Shaiman, 2001; Waals, 1992).

3.4.2 Syllable timing

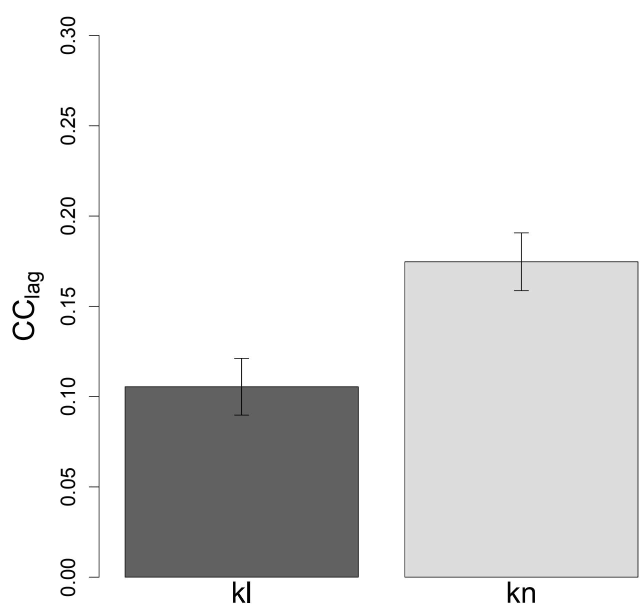


Figure 3.7: *Consonant-to-consonant timing within the two onset clusters /kl/ (left, dark grey) and /kn/ (right; light gray).*

Before reporting the results for CV and VC timing in complex versus simplex tokens, interconsonantal timing differences within the clusters /kl/ and /kn/ was examined, as these differences are important for the further hypotheses. Figure 3.7 shows the CC_{lag} of the two different onset clusters. It can clearly be observed that there was more overlap between /k/ and /l/ than between /k/ and /n/, indicated by lower lag values. A mixed model with CC_{lag} as the dependent variable and with onset manner (/l/ vs. /n/) as fixed factor showed a significant effect ($\chi^2[1] = 64.5$, $p<0.001$). This significant effect supports what is evident from Figure 3.7 that there is more overlap between /k/ and /l/ than between /k/ and /n/ and further supports findings by Bombien et al. (2013) and (Hoole et al., 2009).

As mentioned in section 3.2, two temporal articulatory mechanisms are taken into account leading to compensatory shortening driven by consonant clusters. The first is that shorter plateau durations of the vowel gesture are a possible factor resulting in acoustic vowel shortening. The idea is that the vowel gesture may be compressed in words containing complex onsets and codas, respectively, resulting in shorter vowel durations compared with their simplex counterparts. Therefore, the plateau duration of the vowel gestures was measured. Higher values signify longer vowel plateau durations, lower values signify shorter vowel plateau durations. The second is that there is a shift of the vowel-adjacent consonant towards the vowel in complex tokens compared with simplex tokens. This then leads to a greater CV or VC overlap resulting in shorter acoustic vowel durations. Therefore, plateau lags between the vowel and its adjacent consonant in onset as well as in coda position were compared in order to investigate possible differences due to an additional consonant. Thus lower lag values in complex onset and coda tokens, respectively, signify greater overlap and a shift towards the vowel.

Onset

Concerning the temporal mechanisms causing the observed onset-driven compensatory shortening, first results for vowel plateau duration will be presented followed by the results for CV_{lag}.

Vowel plateau duration

Figure 3.8 shows that – despite a tendency towards vowel plateau compression in words containing a sonorant coda (i.e. lVm) – vowel plateau durations seem not to differ with respect to onset complexity. Further, plateau durations seem to be longer in words containing the sonorant coda /m/ than in words containing the obstruent coda /p/. These observations were confirmed statistically: a mixed model with vowel plateau duration as the dependent variable and with fixed factors onset complexity (simplex vs. complex), onset manner (/l/ vs. /n/) and coda manner (/m/ vs. /p/) revealed that only coda manner had a significant influence on vowel plateau duration ($\chi^2[1] = 4.7$, $p<0.05$).

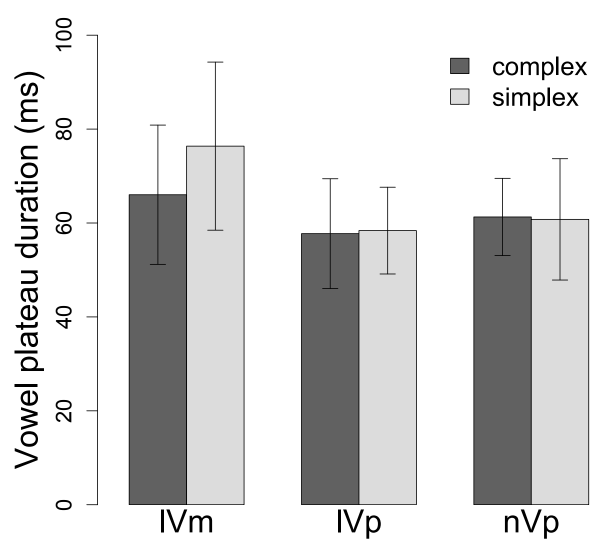


Figure 3.8: *Vowel plateau duration depending on onset complexity separately for word pairs included (lVm, lVp, nVp) containing simplex (light gray) and complex (dark gray) onsets.*

This effect is in line with Figure 3.8, showing greater vowel plateau durations in words containing /m/ in coda position. Further, the finding that there was no significant effect of onset complexity suggests that a compression of the vowel gesture's plateau does not seem to be an underlying articulatory mechanism leading to compensatory vowel shortening.

CV_{lag}

Concerning the timing between the vowel-adjacent onset consonant ($C2_{on}$) and the vowel, Figure 3.9 shows that onset consonants seem to differ in their CV timing: there seems to be a shift towards the vowel only in /n/ words, because it is only in these tokens that the CV_{lag} is markedly smaller in complex onset tokens compared with simplex onset tokens. A mixed model with CV_{lag} as the dependent variable and with the same fixed factors as above revealed a significant effect of onset complexity ($\chi^2[1] = 5.3$, $p<0.05$) and a significant interaction between onset complexity and onset manner ($\chi^2[1] = 12.6$, $p<0.001$).

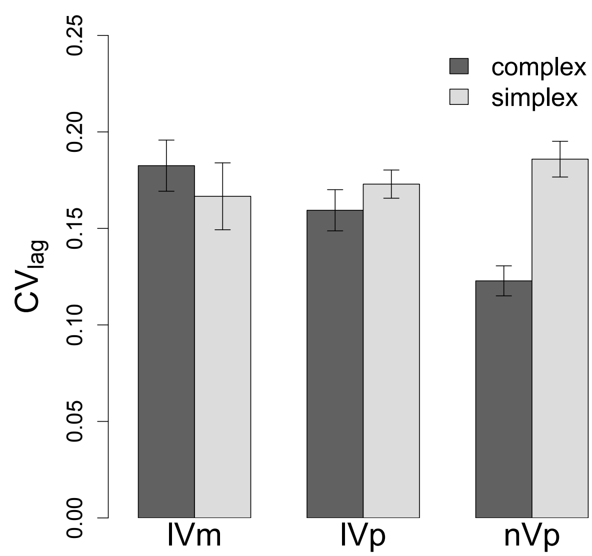


Figure 3.9: *CV_{lag} depending on onset complexity separately for word pairs included (lVm, lVp, nVp) containing simplex (light gray) and complex (dark gray) onsets.*

This interaction effect is in line with what is evident from Figure 3.9: in words containing the onset /l/ and the coda /m/ (i.e. lVm), the onset consonant in complex tokens seems to shift slightly away from the vowel, as the lag values are somewhat higher in complex than in simplex tokens. In words containing the onset /l/ and the coda /p/ (i.e. lVp), on the other hand, the consonant seems to shift towards the vowel in complex tokens, indicated by somewhat lower lag values in complex compared with simplex tokens. This shift seems to be greater in words containing the onset /n/ and the coda /p/ (i.e. nVp). Post-hoc Turkey tests revealed that there was a significant difference between simplex and complex tokens only in nVp tokens (p<0.001). This means that – commensurate with Figure 3.9 – only in /kn/ onsets there was a shift towards the vowel, which was not the case in /kl/ onsets. That finding implies that only the complex onset /kn/ shows a global organization with the following vowel and presumably a c-center organization, whereas the onset cluster /kl/ does not.

Coda

Concerning the temporal mechanisms causing coda-driven compensatory shortening, the results for vowel plateau duration will be reported first, followed by the results for VC_{lag}.

Vowel plateau duration

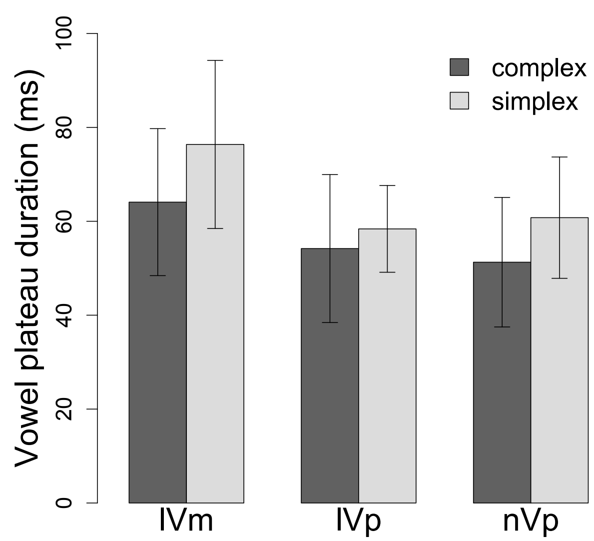


Figure 3.10: *Vowel plateau duration depending on coda complexity separately for word pairs included (lVm, lVp, nVp) containing simplex (light gray) and complex (dark gray) codas.*

Figure 3.10 shows a tendency towards vowel plateau compression in complex coda tokens. However, a mixed model with vowel plateau duration as the dependent variable and with fixed factors coda complexity (simplex vs. complex), onset manner (/l/ vs. /n/) and coda manner (/m/ vs. /p/) revealed no significant effect of coda complexity. Coda manner significantly influenced vowel plateau duration ($\chi^2[1] = 4.8$, $p<0.05$). This is because the vowel plateau is generally longer before /m/ than before /p/ (cf. Figure 3.10). Further, the mixed model revealed no effect of onset manner and there was no interaction. This result further suggests that compression of the vowel gesture's plateau does not seem to be an underlying articulatory mechanism leading to compensatory shortening, although there seems to be a tendency toward vowel plateau compression in all complex coda tokens (cf. Figure 3.10).

VC_{lag}

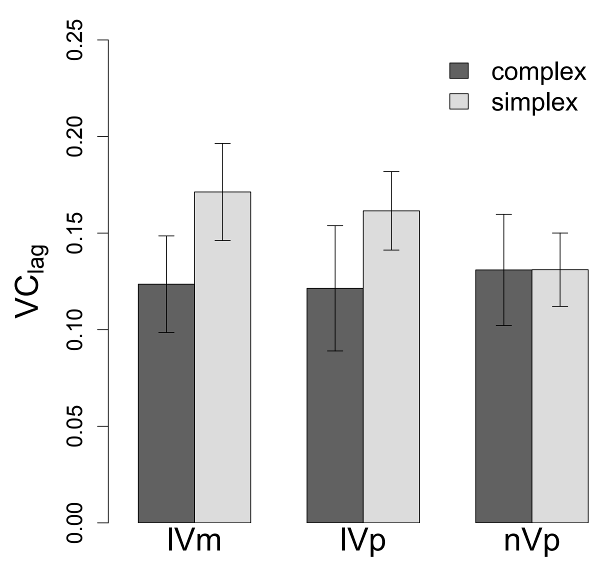


Figure 3.11: *VC_{lag} depending on coda complexity separately for word pairs included (lVm, lVp, nVp) containing simplex (light gray) and complex (dark gray) codas.*

Concerning the timing between the vowel and its adjacent coda consonant ($C1_{off}$),

Figure 3.11 shows tendencies for the consonant to shift towards the vowel in /l/ words as indicated by lower lag values in complex coda tokens. A mixed model with VC_{lag} as the dependent variable and with the same fixed factors as above revealed a significant main effect of coda complexity ($\chi^2[1] = 6.0$, p<0.01) and a significant interaction between onset manner and coda complexity ($\chi^2[2] = 4.1$, p<0.05). This interaction effect is in line with Figure 3.11, showing differences in the timing of the vowel and $C1_{off}$ only for /l/ words. Post-hoc Tukey test revealed – commensurate with Figure 3.11 – a significant difference between simplex and complex tokens in words with /l/ in onset position (p< 0.01), i.e. there was a shift of $C1_{off}$ towards the vowel in complex coda tokens. This finding contradicts Articulatory Phonology, proposing a sequential organization of coda clusters.

To conclude, the results for temporal parameters showed that vowel plateau compression seems not to be a trigger for compensatory shortening. The vowel plateau duration was not compressed by adding a consonant to the onset of a syllable and the vowel plateau duration was compressed only by trend by adding a consonant to the coda of a syllable. Further, there was a shift towards the vowel only in /kn/ onsets but not in /kl/ onsets. For coda clusters, the results showed an increased overlap with the preceding vowel in /l/ words but not in /n/ words. This leads to the conclusion that VC(C) timing depends on the consonant in onset position. Further, there seems to be a relationship between onsets and codas: while there was no shift towards the vowel in complex onsets in /l/ words, there was a shift towards the vowel in complex codas in these tokens. Thus the proposals by Katz (2012) could not be consistently confirmed in our data. Maybe differences in stiffness or tongue back and jaw height may account for compensatory shortening. The results for these three proposed mechanisms will be reported in the following sections.

3.4.3 Stiffness

Differences in stiffness were considered as another mechanism underlying acoustic vowel shortening. A greater stiffness indicates a faster gesture resulting in a shorter segment.

Therefore, the vowel stiffness was presumed to be greater in complex tokens than in simplex tokens. The results for stiffness will be reported first for complex vs. simplex onsets and subsequently for complex vs. simplex codas.

Onset

Figure 3.12 shows that there is a clear difference between complex and simplex tokens in all word pairs: the vowel gestures' stiffness is greater in complex onset tokens than in simplex onset tokens. A mixed model with stiffness as the dependent variable and with fixed factors onset complexity (simplex vs. complex), onset manner (/l/ vs. /n/) and coda manner (/m/ vs. /p/) revealed a significant effect of onset complexity ($\chi^2[1] = 45.1$, $p<0.01$), which is commensurate with the observed greater stiffness in vowels following complex onsets. There was no significant main effect of onset or coda manner and no interaction effect. This means that the vowel gesture was produced with a greater stiffness as the number of onset consonants increased, which may be attributed to causing acoustic compensatory shortening.

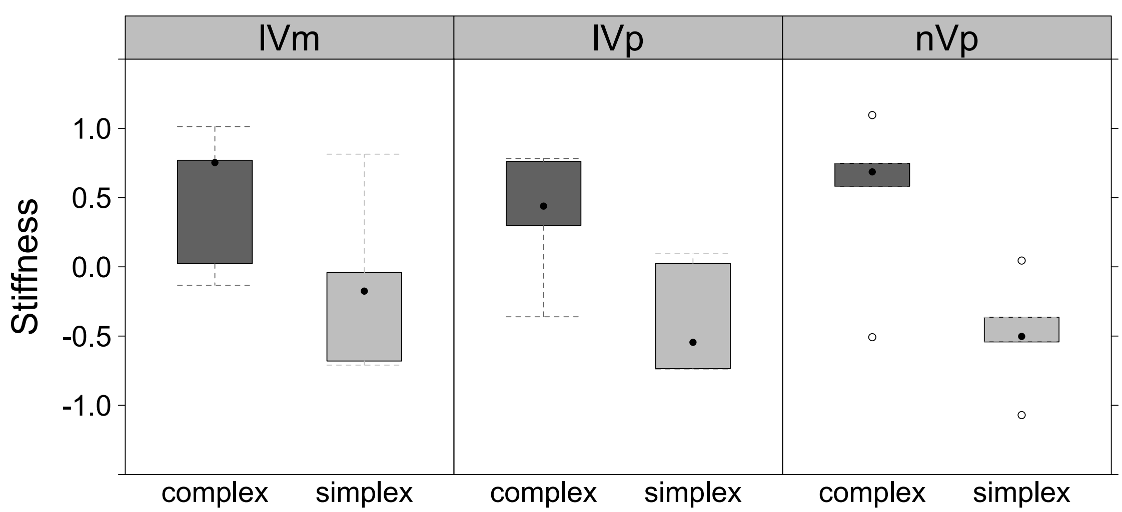


Figure 3.12: *Normalized stiffness of the vowel gesture depending on onset complexity (simplex: light gray; complex: dark gray) separately for the different word pairs.*

Coda

As far as coda consonants are concerned, Figure 3.13 shows that there seem to be no differences between complex and simplex tokens. However, a greater stiffness in complex than in simplex coda tokens for words with /n/ in onset position and /p/ in coda position (nVp-tokens) can be observed. A mixed model with stiffness as the dependent variable and with fixed factors coda complexity (simplex vs. complex), onset manner (/l/ vs. /n/) and coda manner (/m/ vs. /p/) revealed a significant interaction between coda complexity and onset manner ($\chi^2[1] = 4.3$, $p<0.05$). Post-hoc Tukey tests showed a slightly significant effect of coda complexity only for nVp tokens ($p<0.05$), which is in line with the observations from Figure 3.13. It can be concluded from these results that coda-driven acoustic vowel shortening appears not to be entirely triggered by increased stiffness in complex coda tokens.

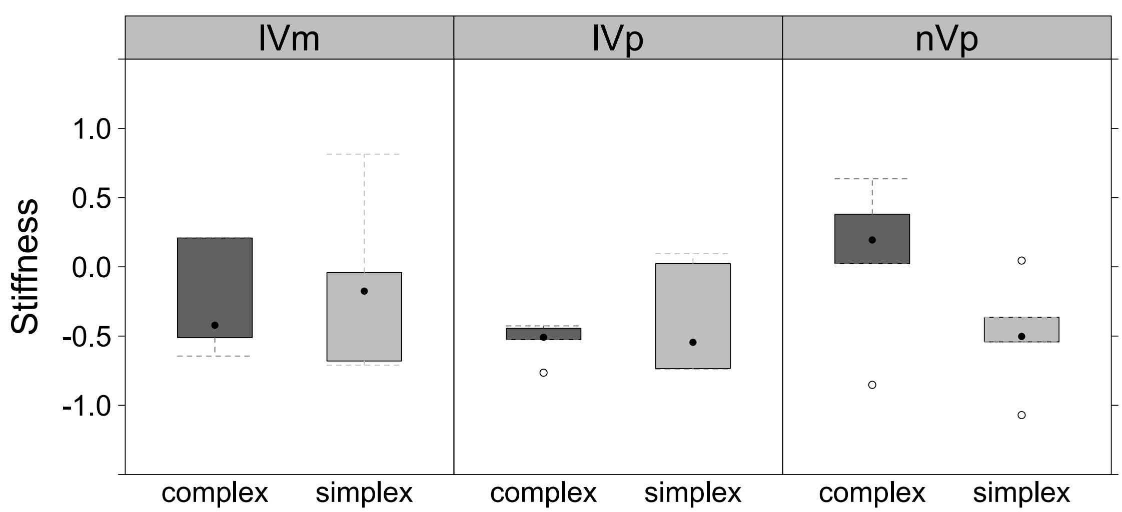


Figure 3.13: *Normalized stiffness of the vowel gesture depending on coda complexity (simplex: light gray; complex: dark gray) separately for the different word pairs.*

3.4.4 Tongue back and jaw position

The last parameters to be considered are the vertical positions of the tongue back and the jaw depending on onset and coda complexity, respectively. It was assumed that complex onsets and codas induce vowel undershoot, i.e. the tongue back gesture to produce the vowel was assumed to be less lowered in complex onset and coda tokens than in simplex tokens. Further, it was assumed that the jaw is lowered to a lesser extent during the production of the tense vowel /aː/ in complex onset and coda tokens than in simplex tokens.

The following section reports results separately for onsets and codas. First, results for vertical tongue back position depending on onset complexity will be shown, followed by results for vertical jaw position depending on the same factor. Afterwards, results of vertical tongue back position depending on coda complexity will be shown followed by results for vertical jaw position depending on the same factor.

Onset

Vertical tongue back position

Figure 3.14 shows the vertical position of the tongue back (time normalized) during the production of the tense vowel /aː/ separately for all five speakers. As can be seen, there are differences in tongue position between simplex and complex tokens. In words containing complex onsets, the tongue back seems to be lowered to a lesser extent than in words containing simplex onsets. A mixed model with tongue back position at the temporal midpoint (0.5) as the dependent variable and onset complexity (simplex vs. complex) as fixed factor revealed that onset complexity significantly affected the vertical position of the tongue back producing the vowel ($\chi^2[1] = 40.7$, $p<0.001$). This significant effect of onset complexity confirms what is evident from Figure 3.14 that the tongue is lowered to a greater extent in words containing simplex onsets as opposed to words containing complex onsets. This finding of a higher tongue position in complex tokens can be related to coarticulatory-induced vowel target undershoot in these tokens and it may account for onset-driven compensatory vowel shortening found in the acoustics.

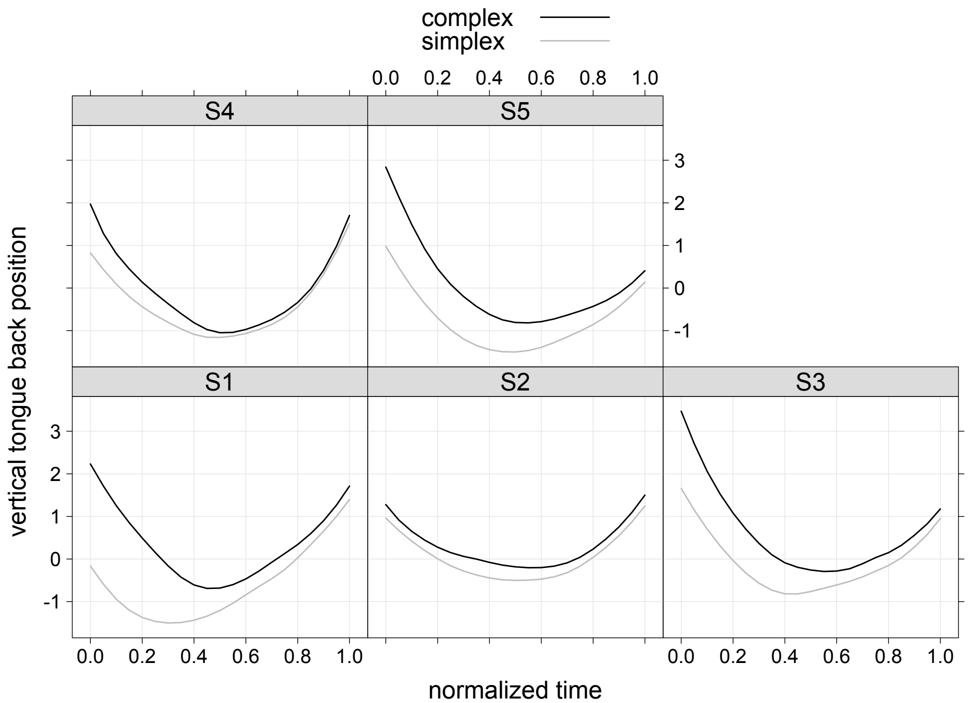


Figure 3.14: *Aggregated vertical tongue back position during the production of /aː/ depending on onset complexity (simplex: gray; complex: black) separately for each speaker.*

Vertical jaw position

Further, the vertical position of the jaw – which is lowered during the production of the tense vowel /aː/ – was measured in order to see whether onset clusters influence the jaw position. Figure 3.15 shows the vertical jaw positions in words containing complex and simplex onsets separately for each speaker. It is quite obvious that there are no differences regarding onset complexity, i.e. the jaw is lowered to the same extent in tokens containing simplex onsets as in tokens containing complex onsets. This observation was also confirmed statistically: a mixed model with jaw position at the temporal midpoint (0.5) as the dependent variable and onset complexity (simplex vs. complex) as fixed factor showed that there was no significant effect on the vertical position of the jaw. This means that the vertical jaw position did not differ with respect to onset complexity, which is in line with what is evident from Figure 3.15.

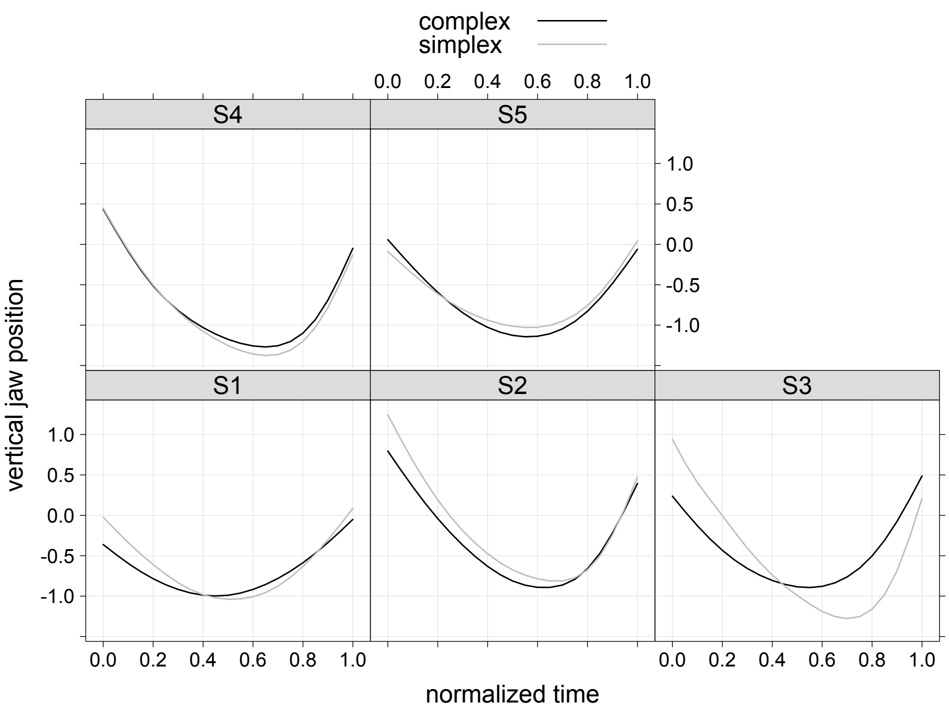


Figure 3.15: *Aggregated vertical jaw position during the production of /aː/ depending on onset complexity (simplex: gray; complex: black) separately for each speaker.*

To conclude, the results for the vertical tongue back position showed that onset complexity had a significant influence: there was coarticulatory induced vowel target undershoot in complex onset tokens as opposed to simplex onset tokens. Thus vowel target undershoot in complex onset tokens may explain compensatory vowel shortening. Concerning the vertical position of the jaw, on the other hand, the results showed that there were no differences between complex and simplex onset tokens. Thus an additional consonant in onset position affects the position of the vowel gesture but not the position of the jaw.

Coda

In this section, the same analyses as described above for onsets were also conducted for codas. First, the results for vertical tongue back position depending on coda complexity will be reported, followed by the results for vertical jaw position depending on coda complexity.

Vertical tongue back position

Figure 3.16 shows the vertical tongue back position in words containing simplex and complex codas, respectively, separately for all five speakers. Except for speaker S4, there are differences in tongue back position depending on coda complexity: the tongue back was lowered to a lesser extent to produce the tense vowel /aː/ in words containing complex codas than in words containing simplex codas.

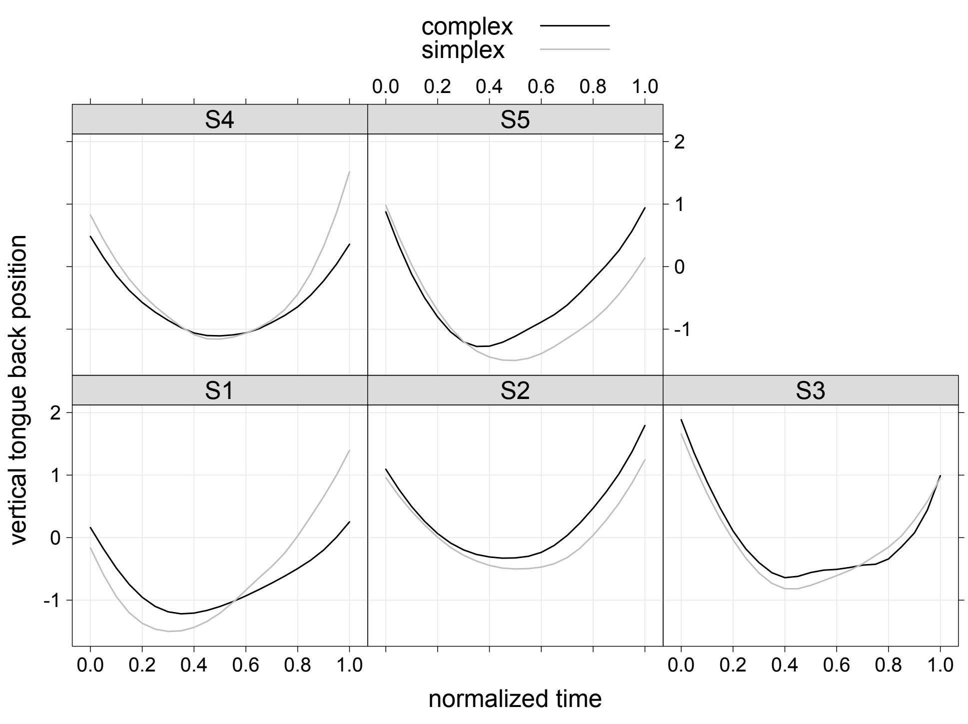


Figure 3.16: *Aggregated vertical tongue back during the production of /aː/ position depending on coda complexity (simplex: gray; complex: black) separately for each speaker.*

A mixed model with tongue back position at the temporal midpoint (0.5) as the dependent variable and coda complexity (simplex vs. complex) as fixed factor revealed a significant effect on tongue back height ($\chi^2[1] = 7.2$, p<0.01). This effect is in line with what is evident from Figure 3.16 that the tongue back is lowered to a lesser extent in complex coda tokens than in simplex coda tokens. This means that an increasing number of coda consonants did affect the vertical position of the preceding vowel gesture: there was coarticulatory-induced vowel target undershoot. This result may account for compensatory vowel shortening found in the acoustics.

Vertical jaw positon

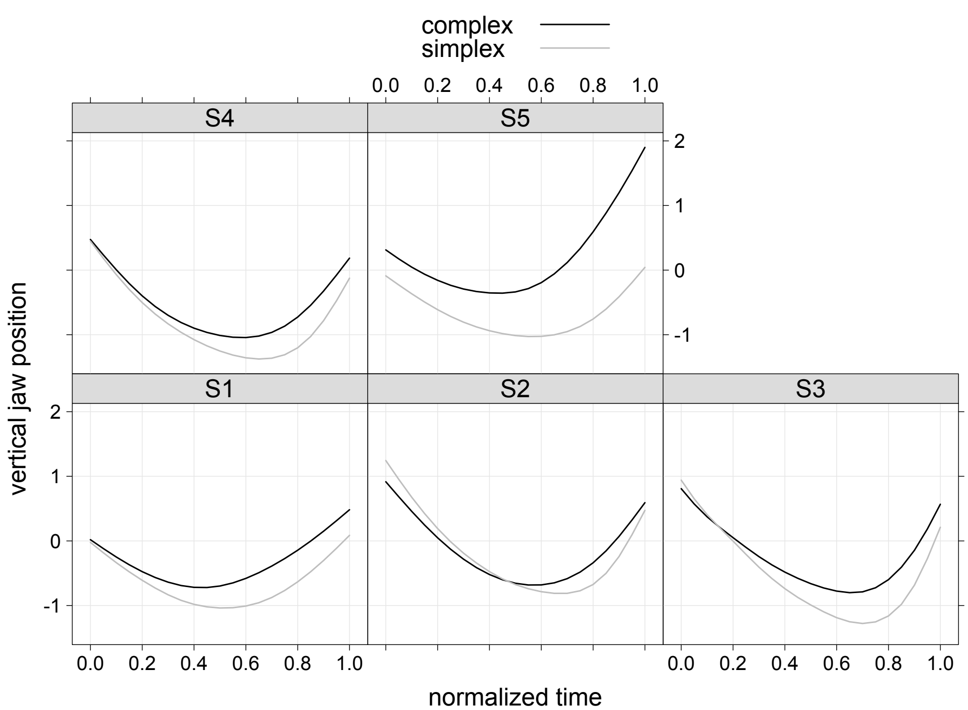


Figure 3.17: *Aggregated vertical jaw position during the production of /aː/ depending on coda complexity (simplex: gray; complex: black) separately for each speaker.*

In addition, for the vertical jaw position, there seem to be differences due to an increasing number of coda consonants. Figure 3.17 clearly shows that the jaw was lowered to a lesser extent in words containing complex codas than in words containing simplex codas. A mixed model with jaw position at the temporal midpoint (0.5) as the dependent variable and coda complexity (simplex vs. complex) as fixed factor revealed a significant effect ($\chi^2[1] = 8.9$, $p<0.01$). That is, the jaw was lowered to a lesser extent in complex coda tokens compared with simplex coda tokens. This finding indicates that coda clusters are linked to the jaw and is in line with previous findings showing evidence that the jaw plays an active role in vowel shortening (see e.g. Harrington et al., 1995; Munhall et al., 1992; Siddins et al., 2014).

The conclude, the results showed that coda complexity significantly influenced the vertical position of the tongue: there was coarticulatory induced vowel target undershoot in complex coda tokens as opposed to simplex coda tokens. Thus vowel target undershoot in complex coda tokens may be a trigger for compensatory vowel shortening. Concerning the vertical position of the jaw, the results showed that there were differences between complex and simplex coda tokens: the jaw was lowered to a lesser extent in complex coda tokens than in simplex coda tokens. Thus an additional consonant in coda position affects both the position of the vowel gesture and the position of the jaw. It has to be considered, though, that the differences in tongue back position might be an artefact due to the differences in jaw height, since the tongue and the jaw are depending on each other.

3.5 Summary and discussion

This study presented a first attempt at explaining incremental compensatory shortening from an articulatory point of view. Furthermore, this study was – according to our knowledge – the first one in which the same consonants flanked the vowel in onset and coda analyses, thus preventing any other influences on acoustic vowel duration. This method contradicts the usually applied procedure in other c-center studies (e.g. Marin, 2013; Marin and Pouplier, 2010; Pouplier, 2012) using for example /kl/ in onset position and /lk/ in coda position in order to compare the vowel's timing with either a preceding or a following lateral and whether there are timing differences attributed to the additional velar consonant. It is a point against this method that onset /kl/ is not the same as coda /lk/ and the timing of /k/ to /l/ to the vowel differs from the timing of the vowel to /l/ to /k/ (see Brunner et al., 2014, for a discussion). In addition, different coda consonants used in previous studies (see e.g. Marin and Pouplier, 2010; Pouplier, 2012, comparing singleton voiced sonorants with clusters involving voiceless stops) might have influenced the acoustic vowel durations. Moreover, this study aimed to measure CV and VC overlap directly rather than indirectly using anchor points (such as the coda consonant in CV analyses or the onset consonant in VC analyses). The present study has revealed several results. First, the results for acoustic vowel shortening will be summarized and discussed, followed by a summary and discussion of the temporal and spatial mechanisms included in the present study, each separately for onset and coda clusters.

3.5.1 Incremental compensatory shortening

This study has shown that there is incremental compensatory shortening caused by both onset and coda clusters in German monosyllables, confirming hypothesis **H1**. This effect does not depend on other factors such as onset or coda manner, contrary to hypotheses **H2** and **H4** proposing greater shortening in words containing /kn/ onsets and in words containing sonorant coda tokens. The present results thus confirm previous findings on acoustic incremental vowel shortening (e.g. Katz, 2012; Munhall et al., 1992; Pouplier, 2012; Waals, 1992). In this chapter only tense vowels were included, supporting the findings for phonologically long vowels in Dutch (Waals, 1992). However, as there were no differences in vowel shortening depending on coda manner – i.e. whether the syllable contained a sonorant or an obstruent in coda position – the present findings contradict those reported by Katz (2012) for English. The finding that there were no differences depending on onset cluster composition in onset-driven vowel shortening is in line with other studies focusing on vowel shortening caused by onset clusters (e.g. Katz, 2012; Marin and Pouplier, 2010). To conclude, the evidence for onset-driven acoustic vowel shortening is in line with the predictions made by Articulatory Phonology (Browman and Goldstein, 1986, 1988, 2000) but the evidence for coda-driven shortening contradicts these predictions.

3.5.2 Consonant-to-consonant timing

The results for consonant-to-consonant timing showed that there was a difference between the two onset clusters /kl/ and /kn/: there was more overlap between /k/ and /l/ than between /k/ and /n/, confirming hypothesis **H3**. This finding is in line with previous reported timing differences within these two onset clusters (Bombien, 2011; Bombien et al., 2013; Hoole et al., 2009). Explanations for these differences may be a manner based model, the DAC (degree of articulatory constraint) model (Recasens, 1999, 2007; Recasens and Espinosa, 2006; Recasens et al., 1995; Recasens and Pallarés, 1999, 2001; Recasens et al., 1997) and perceptual recoverability (Byrd, 1996; Chitoran et al., 2002; Hoole et al., 2009; Silverman, 1997). Manner based ranking is similar to the sonority hierarchies proposed by Sievers (1901) or Selkirk (1984) (see also chapter 1): stops, fricatives, nasals, liquids, glides, vowels (see e.g. Wright, 2004). Gestural overlap is supposed to be greater when the two sounds are further apart within this rank order (e.g. stop+vowel sequences). Thus in a manner-based model it is assumed that there are timing differences: more overlap is expected in /kl/ (as stops and liquids are listed further apart) than in /kn/. However, manner based ranking is more descriptive than explanatory and does not explain the observed overlap differences. According to the DAC model, speech sounds are classified by their coarticulatory influence on neighboring sounds by allocating DAC values. the higher the difference between DAC values, the higher the amount of coarticulation between neighboring gestures. The lateral /l/ is assumed to exert more overlap than /n/ because of laterality requirements including the tongue back (Recasens, 2007). Consonant sequences following the perceptual recoverability principle are timed in such a way that essential perceptual cues of one consonant will not be hidden by another consonant. In order to perceive the cluster /kn/ as such, the velar lowering for the nasal is not allowed to overlap with the release of the plosive, because otherwise the plosive is not perceivable anymore (see also Hoole et al., 2009, for this account). Therefore, the degree of overlap may be reduced, because it would otherwise tend to mask the perception of the plosive. This perceptual or articulatory account as proposed by Hoole et al. (2009) seems to be the most plausible model explaining those observed differences, because it states that the onset cluster /kn/ is not allowed to overlap to the same extent as /kl/, the earlier velum lowering would otherwise tend to mask the perception of the plosive (see also Kühnert et al., 2006). Further, this account could explain the rarity of /kn/ in the world's languages (Ladefoged and Maddieson, 1996) and the loss of velar consonants in English, because it might be difficult for speakers to maintain the exact timing of gestures under certain circumstances.

3.5.3 Vowel plateau compression

The acoustic shortening effect was considered to be possibly caused by a compression of the vowel gesture's plateau, following Katz's (2012) proposal. However, results have shown that the duration of the vowel plateau remains the same, irrespective of an increasing number of consonants in onset and coda position, respectively. This finding therefore contradicts hypothesis **H7** proposing vowel plateau compression in words containing complex onsets and codas, respectively, compared with words containing simplex onsets and codas. It instead suggests that compensatory shortening found in the acoustic analysis was not caused by compression of the vowel plateau.

3.5.4 Consonant-to-vowel timing: CV_{lag} and VC_{lag}

According to Katz (2012), another mechanism leading to compensatory shortening was increased overlap between the vowel and its adjacent consonants. If this was the case, more CV overlap and more VC overlap – and consequently smaller CV_{lag} and smaller VC_{lag} – should be observed in complex tokens compared with simplex tokens.

Concerning CV_{lag}, the results were quite different for the two onset clusters that were included in this study. In /n/ words, there was a notable shift towards the vowel, supporting a c-center organization of /kn/ onsets. For /l/ words, on the other hand, there was no evidence that /l/ shifts towards the vowel in complex onsets. This may be due to the observed greater overlap between /k/ and /l/ than between /k/ and /n/. Different coordination patterns between consonant gestures seem to affect the degree of consonant-to-vowel timing: because there is more overlap between /l/ and the preceding /k/, the /l/ is not shifted towards the vowel, resulting in less CV overlap. Hypothesis **H4** proposed a shift towards the vowel in complex onset tokens leading to greater CV overlap. The findings for CV_{lag} partly confirm this hypothesis, as there was a shift towards the vowel only in /kn/ onsets but no shift in /kl/ onsets presumably due to the greater consonant-to-consonant overlap in /kl/ onsets. But this difference in both consonant-to-consonant and consonant-to-vowel timing does not appear to account for shortening effect in acoustic vowel duration, as the amount of incremental vowel shortening caused by an increasing number of onset consonants was the same in both contexts (i.e. in /kl/ and /kn/). This leads to the conclusion that consonant-to-consonant timing affects consonant-to-vowel timing but not necessarily the acoustic vowel duration. The different timing behavior of onset clusters /kn/ and /kl/ supports recent findings that whether or not an onset cluster shows a c-center effect depends on cluster composition. For example, Marin (2013) reported a c-center organization for /sp/, /sk/ and /sm/ onsets but not for /ps/, /ks/, /kt/ and /kn/ onsets. Our results for /kl/ support the evidence from other studies on German consonant+lateral onsets that found no shift towards the vowel in German /pl/ and /gl/ onsets (Brunner et al., 2014; Pouplier, 2012) and suggests that German consonant+lateral onsets may be a special case. As mentioned above, according to the DAC model of lingual coarticulation, /l/ is assumed to exert more overlap than /n/ because it requires laterality including the tongue back (Recasens, 2007). Therefore, also see more CV overlap in /l/ words compared with /n/ words should be observed. One explanation for why /l/ exhibits less coarticulation and thus less overlap with the following vowel may be coarticulatory directionality. The lateral may exhibit more anticipatory coarticulation (i.e. right-to-left) than carry-over coarticulation (i.e. left-to-right). Therefore, it might overlap to a greater extent with the preceding consonant and to a lesser extent with the vowel. This could be explained by both a vocalic dorsal gesture as well as a consonantal apical gesture required for laterals (Krakow, 1999; Sproat and Fujimura, 1993). Sproat and Fujimura (1993) further suggest that the vocalic gesture has a strong affinity for the nucleus of the syllable. That is, the tongue dorsum gesture occurs much earlier with respect to its associated tongue tip closure when /l/ is syllable final than when syllable initial (Krakow, 1999). The two gestures of /l/ are thus synchronously with the vocalic gesture in syllable final /l/s. In addition, also Goldstein et al. (2009) observed an overall less shift for /l/ in /pl/ than in other onset clusters. They hypothesized that this was due to the multiple gestures required for /l/. Specifically, they claimed that, if both the tongue tip and the tongue body constrictions for /l/ were coupled with the vowel, this could result in an overall tighter coupling of /l/ with the vowel. Because of this tighter coupling, there is less shift of /l/ toward the vowel in complex onset tokens.

For VC_{lag}, there was no timing difference concerning the two coda clusters included (i.e. /pt/ and /mt/). However, onset manner had an influence on VC timing, although the coda was kept constant in all tokens: /l/ in onset position caused more overlap between the vowel and $C1_{off}$ in complex codas than in simplex codas (and thus a shift towards the vowel), whereas no differences in VC overlap between simplex and complex codas have been observed in words containing /n/ in onset position. This finding partially confirms hypothesis **H4** and it contradicts the predictions made by Articulatory Phonology that coda clusters are sequentially organized. Furthermore, this finding reveals evidence that onsets and codas do not behave independently of each other within a syllable (see also Hawkins and Nguyen, 2000).

Figure 3.18 schematically illustrates the results for syllable timing depending on onset complexity (left) and coda complexity (right). Since coda manner (/p/ vs. /m/) never had a significant effect on syllable timing, Figure 3.18 displays results schematically only for words containing /p/ in coda position for simplification.

Concerning onsets, the results can be summarized as follows: (1) in complex onset tokens there is a greater overlap within /kl/ than within /kn/, (2) in complex onset tokens, relative to simplex onset tokens, there is a shift towards the vowel only in /kn/ onsets but not in /kl/ onsets, (3) in complex onset tokens, relative to simplex onset tokens, the vowel plateau durations remain the same.

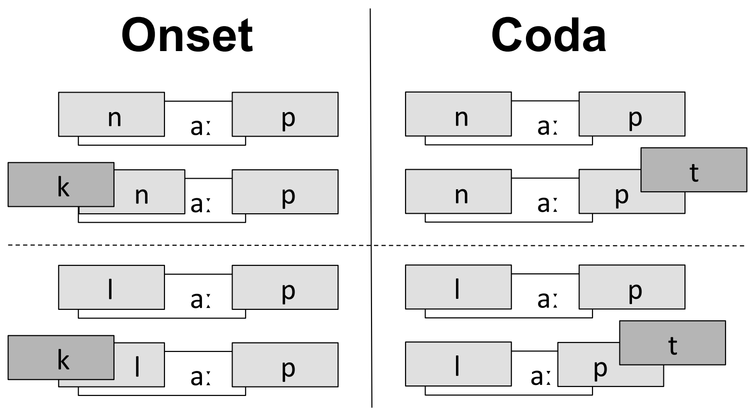


Figure 3.18: *Schematic illustration of results concerning syllable timing. Results are displayed separately for onsets (left) and codas (right) depending on complexity and on the two different onset consonants.*

Concerning codas, the results can be summarized as follows: (1) in complex coda tokens, relative to simplex coda tokens, there is a shift of $C1_{off}$ towards the vowel only in words containing /l/ in onset position but not in words containing /n/ in onset position, (2) in complex coda tokens, relative to simplex coda tokens, the vowel plateau durations remain unchanged.

To conclude, the temporal organization of the onset cluster /kn/ is in line with the predictions made by Articulatory Phonology. However, the observed lesser (C)CV and greater VC(C) overlap in /l/ words are not predicted by Articulatory Phonology. This finding and the acoustically shorter vowel durations before complex codas suggest that coda clusters are not entirely sequentially organized. Instances of acoustic vowel shortening do not seem to be (only) a result of a shift towards the vowel in complex onset and coda tokens, respectively: there was no shift towards the vowel in /kl/ onsets and in complex coda tokens with /n/ in onset position, but there was compensatory shortening in all tokens. This is to some extent in line with the study by Byrd (1995) showing that at least some speakers produced coda clusters with c-center organization although she did not find vowel shortening in her acoustic analyses. This leads to the conclusion that compensatory shortening is not necessarily caused by increased overlap (i.e. a c-center organization) and vice versa.

3.5.5 Stiffness

Concerning stiffness, it was suggested in hypothesis **H7** that vowels are produced with a greater stiffness in words with complex onsets and codas. A greater stiffness in complex tokens could account for the durational differences found in the acoustics. The results of the present study showed that onsets and codas behave differently concerning their influence on vowel gesture stiffness: vowel gestures were produced with a greater stiffness in words with complex onsets than in words with simplex onsets. Coda clusters, on the other hand, did not influence the stiffness of the preceding vowel gesture (only in nVp tokens). Thus hypothesis **H7**, predicting greater stiffness in both complex onsets and codas, can be partly confirmed. These findings provide further evidence that onset clusters are linked to the vowel gesture, whereas coda clusters are not. In addition, this linking seems to be independent of whether or not an onset consonant within a cluster shows a shift towards the vowel compared with its timing as a singleton. Remember that a shift towards the vowel was only observed for /kn/ onsets but not for /kl/ onsets. The vowel gesture's stiffness, however, was influenced by both onset cluster types. This result further supports Articulatory Phonology and the coupled oscillator model predicting that onsets are linked wo the vowel gesture whereas codas are not. Thus for the two onset clusters /kn/ and /kl/ there are marked differences in consonant-to-consonant timing as well as in consonant-to-vowel timing. Concerning their effect on vowel stiffness, however, these two onset clusters do not behave differently as both of them cause an increase in vowel's stiffness. An unresolved issue was so far whether it is the velocity itself that increases in complex onsets or whether the observed greater stiffness is caused by a greater displacement in simplex onset tokens. Since stiffness is peak velocity divided by displacement, a greater displacement in simplex tokens would result in lower stiffness. Figure 3.19 shows the two components of the stiffness measurement – peak velocity and maximum displacement – aggregated per speaker for complex (C; red) and simplex (S; black) tokens separately for onsets (left) and codas (right).

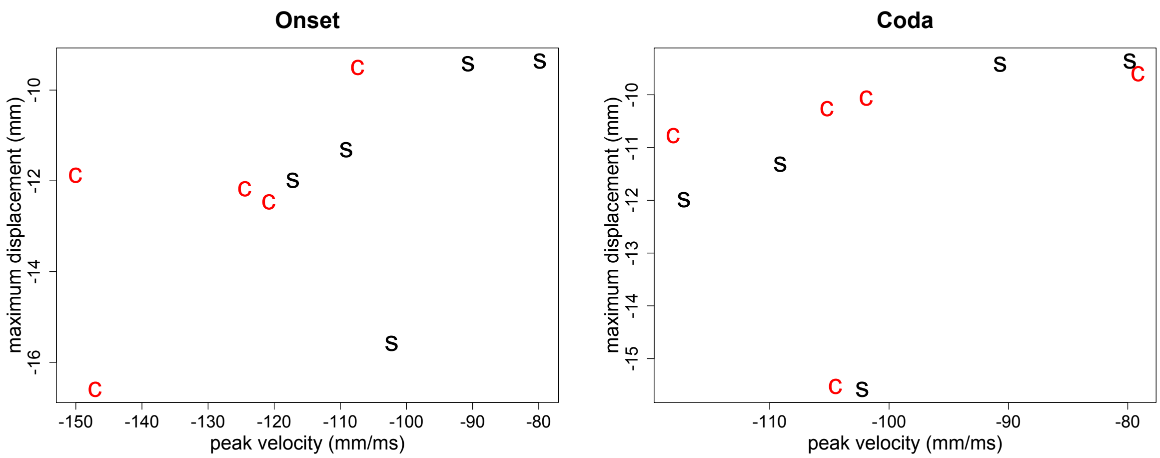


Figure 3.19: *Aggregated peak velocity and maximum displacement separately for words containing simplex (S) and complex (C) onsets (left) and codas (right), respectively.*

Concerning onsets, there seems to be a clear separation between simplex and complex tokens such that complex onset tokens were produced with nearly the same displacement as simplex tokens but with a greater velocity. This is evident from Figure 3.19 as values for simplex are on the right side (i.e. lower velocity), whereas values for complex are on the left side (i.e. higher velocity). For codas, however, there seem to be no differences either in displacement or in peak velocity as simplex and complex values are placed randomly in the two-dimensional space. Therefore, it can be concluded that the vowel gesture was indeed produced faster in complex onset tokens than in simplex onset tokens. One explanation might be that speakers hyperarticulate in complex onset tokens since peak velocity represents a plausible measure of "biomechanical/articulatory effort" (Nelson, 1984; Nelson et al., 1984; see also Moon and Lindblom, 1994). This might be related to the predictions made by the coupled oscillator model (Goldstein et al., 2007a; Nam et al., 2009; Nam and Saltzman, 2003). In this model, complex onsets are in a competitive coupling relationship:

anti-phase with each other and in-phase with the following vowel. It has been observed that consonants in onsets start roughly synchronously with the vowel (de Jong, 2003; Löfqvist and Gracco, 1999). Thus with an additional consonant in onset position, speakers might have less time in coupling both onset consonants with the vowel and therefore the vowel gesture is produced faster in order to produce hyperarticulated forms perceptible for listeners. This is not the case in coda clusters, since there is no competitive coupling.

3.5.6 Tongue back position

As a spatial parameter, the vertical position of the tongue back was taken into account. It was assumed in hypothesis **H8** that there will be vowel undershoot in complex onsets and codas, respectively, compared with simplex tokens. Results have shown that complex onsets caused the vowel gesture to undershoot its target position as the tongue back was lowered to a lesser extent than in words with simplex onsets. There was also a difference in tongue back position depending on coda complexity: the vowel gesture was lowered to a lesser extent in complex coda tokens compared with simplex coda tokens. That is, also in complex coda tokens there was coarticulatory induced vowel target undershoot, which may lead to acoustic vowel shortening. These results confirm hypothesis **H8** and may be a trigger for onset- and coda-driven compensatory vowel shortening.

3.5.7 Jaw position

Jaw positions during speech production received great attention in articulatory studies as this is non-invasive measurement. So far studies focused on the influence of different consonants (Edwards, 1985; Lindblom, 1983; Mooshammer et al., 2007) and prosodic weakening (Harrington et al., 1995) on jaw position and movement, on coarticulation (Fletcher and Harrington, 1999) and on the relationship between jaw and lip positions in perturbation experiments (Löfqvist and Gracco, 1997). In the present study the influence of onset and coda complexity on jaw position was examined. Hypothesis **H9** proposed that the jaw is lowered to a lesser extent in complex onset and coda tokens, respectively, than in simplex tokens. Concerning the vertical jaw position during the production of the tense vowel /aː/, the results have shown marked differences depending on whether the cluster occurred in onset or coda position. For onsets, there was no difference in jaw position as the number of consonants increased. That is, adding a consonant to the onset of a syllable does not influence the position of the jaw during the vowel. Having shown less tongue back lowering in complex onset tokens, no difference in jaw position was somewhat surprising, leading to the conclusion that onsets are linked to the tongue gesture of the following vowel and this linking is independent of the jaw position. For codas, on the other hand, there was a significant difference in jaw height: the jaw was lowered to a lesser extent in complex coda tokens than in simplex coda tokens. Thus hypothesis **H9** can be confirmed for complex codas but not for complex onsets. Accordingly, coda-driven compensatory shortening may be caused by a reduction of the opening gesture for the vowel. This finding is in line with Siddins et al. (2014) who showed a lower F1 in complex coda tokens compared with simplex coda tokens. But it contradicts Munhall et al. (1992) who revealed no differences in jaw position. Instead, they reported that the jaw lowering gestures of all three talkers were significantly shorter in duration before a cluster than before a singleton consonant. This particular measurement was not included in the present study, but further analyses are possible to investigate whether it is the case that either jaw height or jaw lowering duration causes vowel shortening or whether a combination is also possible.

Different jaw heights have been reported in the literature depending on manner of articulation. Despite the supportive function of the jaw for all oral consonants to achieve their place of articulation (see e.g. Browman and Goldstein, 1990; Saltzman and Munhall, 1989), consonants were found to differ in jaw position. Lindblom (1983) showed that sibilants had the highest jaw position, consonants showed a somewhat reduced jaw height, followed by nasals. Liquids were produced with the lowest jaw position. However, most studies focusing on jaw height found evidence that this classification as described by Lindblom (1983) needs to be subdivided: there is evidence that /t/ requires a higher jaw position than for example /k/ (Elgendy, 1999; Keating et al., 1994; Lee et al., 1994; Perkell, 1969; Tuller et al., 1981). Therefore, the jaw position also depends on place of articulation with an increasing influence of the jaw on the tongue going from the back to the front (Mooshammer et al., 2007). Mooshammer et al. (2007) explained the lower jaw position for velar stops as opposed to the high jaw position for alveolar stops by anatomical factors: the tongue-tip is more distant from the mandibular joint and is thus affected to a greater extent. Additionally, they explained the high jaw position for the alveolar voiceless stop /t/ by acoustic enhancements: they assume that the higher position of the lower teeth enhances the acoustical outcome of the burst salience. Therefore, the higher jaw position during the production of the vowel in complex coda tokens may be triggered by the additional /t/ in these tokens. That is, the upward lifting of the jaw for the production of the alveolar stop even influences the jaw position during the vowel. This leads to the conclusion that the anticipatory coarticulated vowel in complex coda tokens is linked to the jaw movement in coda position. However, differences in jaw height should also have been observed in the data included in the study by Munhall et al. (1992): for example, they compared jaw height during the vowel in simplex /p/ with complex /ps/ codas. As mentioned above, the highest jaw position is required for sibilants. Therefore, one might have expected differences between simplex and complex codas also in Munhall et al.'s (1992) study. One of the reasons for why they did not find significant differences in jaw height depending on coda complexity may be that they included the open mid vowel /ɛ/ and the near open vowel /æ/ in their study, whereas in the present study the open vowel /a/ was included, which requires more jaw lowering and which may thus be more susceptible for jaw height influences.

3.6 Conclusion

To conclude, onset-driven compensatory vowel shortening is most likely to be caused by differences in vowel stiffness and vowel target undershoot. For coda-driven compensatory vowel shortening, on the other hand, the most plausible articulatory mechanism is the observed difference in jaw height[3].

Furthermore, this study reveals evidence for differences between onsets and codas within a syllable, not only in their timing relation but also in their spatial relation with the vowel. Firstly, only for the complex onset /kn/ a shift towards the vowel was observable but not for /kl/. For coda, a shift was observable for /l/ words but not for /n/ words. Secondly, only onset clusters affected the vowel stiffness, whereas complex codas did not. Thirdly, only coda consonants affected the jaw position required for the vowel, whereas no such effect was found for onset consonants.

To conclude, additional onset consonants seem to affect especially the tongue movement of the following vowel: the vowel gesture was produced with a greater stiffness (and also faster, see Figure 3.19) as the number of onset consonants increased and the tongue back was lowered to a lesser extent in complex tokens, i.e. there was coarticulatory induced vowel target undershoot. These differences in speed and position may account for the differences in vowel duration found in the acoustics. Additional coda consonants, on the other hand, seem to be linked especially to the jaw movement required for vowel production. That is, the jaw seems to play an active role in coda-driven vowel shortening. The observed difference in jaw height may account best for the observed acoustic vowel compression caused by complex codas. These results show further evidence that presumably only onset clusters are coupled with the vowel. It is especially the active role of the jaw in complex codas that needs to be considered in the coupled oscillator model. In addition to temporal aspects, spatial aspects depending on syllable position need to be considered in Articulatory Phonology and the coupled oscillator model.

[3]There were also differences in tongue back height but it is unclear how far this finding is just an artefact due to differences in jaw height.

Chapter 4

Stability of gestural timing within syllables

4.1 Introduction

This study investigates the influence of prosodic weakening, onset complexity and vowel tensity on the gestural timing of all syllable constituents (i.e. CC, CV and VC sequences). Further, it reports according to the author's knowledge for the first time effects of prosodic weakening on (C)CV overlap and thus on the global organization of onset clusters. The literature so far does not reveal whether the presumed shift towards the vowel in complex onsets compared with simplex onsets will be diminished or enhanced under prosodic influences. Chapter 3 showed that there was a shift of the vowel-adjacent consonant towards the vowel only for /kn/ onsets. Therefore, this particular onset cluster is included in the present study. For reasons related to the articulatory mechanisms underlying gestural timing (cf. chapter 3), not only the timing between gestures but also the plateau durations of each syllable constituent and its stiffness are included . This was further motivated by recent studies showing either articulatory strengthening and less overlap at higher prosodic boundaries (Bombien et al., 2006; Byrd and Choi, 2010; Byrd et al., 2006), or less overlap at lower prosodic boundaries and plateau shortening instead (Mücke et al., 2008).

4.1.1 Prosodic influences on acoustic and articulation

Stress is the phonetic prominence given to a particular syllable in a word, whereas accent is the relative emphasis of words in a phrase or sentence (e.g. Cutler, 1984; Gussenhoven, 2004; Hayes, 1995; Pierrehumbert and Beckman, 1988). Stress and accent are not necessarily independent within an utterance. For instance, accent is always realized on a syllable which is marked for stress (Cutler, 1984). Accented words or syllables are associated with a higher fundamental frequency (f0) compared to unaccented words or syllables (Fry, 1958) and the duration of the syllable constituents differs with respect to (de)accentuation (see e.g. Cho, 2005; Lehiste, 1970). Generally, segments tend to be longer in prosodically stronger contexts (i.e. stressed syllables, accented words, at higher prosodic boundaries) than in prosodically weaker contexts (i.e. unstressed syllables, deaccented words, at lower prosodic boundaries). Especially vowels were shown to differ in duration with respect to lexical stress and phrasal accent (Lehiste, 1970): they are longer in lexically stressed and accented syllables than in lexically unstressed and deaccented syllables. Also onset consonants were shown to be lengthened in stressed and accented syllables (Crystal and House, 1988; Ingrisano and Weismer, 1979; Klatt, 1974; Stathopoulos and Weismer, 1983; Umeda, 1977) and some lengthening was observed in coda consonants as well (Turk and Sawusch, 1997). However, stress and accent are generally found to affect vowels to a greater extent than consonants (see Cho and Keating, 2007; Pierrehumbert and Talkin, 1992). According to Pierrehumbert and Talkin (1992) the nucleus of a syllable is affected by accentuation, whereas the syllable onset is affected by prosodic boundaries.

Articulatory studies have shown that prosodic prominence in general induces a change in the temporal and spatial characteristics of gestures adjacent to the boundary such that these gestures are produced longer, faster and with a stronger linguo-palatal contact (Beckman et al., 1992; Byrd and Choi, 2010; Byrd et al., 2006, 2005; Byrd and Saltzman, 2003; Cho and Keating, 2009; Cho and McQueen, 2005; Fougeron and Keating, 1997; Keating, 2006; Krivokapić, 2007; Kuzla et al., 2007; Lee et al., 2006). Domain initial and domain final lengthening have been observed to increase cumulatively for larger prosodic boundaries (Byrd, 2000; Byrd and Saltzman, 1998; Cho, 2006; Cho and Keating, 2001; Fougeron, 2001; Keating et al., 2003; Tabain, 2003). Further, less temporal overlap between articulators adjacent to a prosodic boundary was observed (see e.g. Byrd, 2000; Byrd et al., 2000; Byrd and Saltzman, 1998) and there is less overlap across stronger boundaries (Byrd, 2000; Cho, 2004). Byrd and Choi (2010) reported boundary effects on preboundary and postboundary clusters. They found that intergestural timing in heterosyllabic clusters (i.e. clusters across a word boundary) is highly affected by prosodic variation such that there is less overlap at strong boundaries. This effect was also found in coda clusters but to a lesser degree, whereas the timing between onset clusters was nearly unaffected by variation of the prosodic boundary.

Another type of prosodic strengthening is caused by accent/stress, which is often assumed to be linked with distinctiveness, inducing lexical or phonemic contrasts (de Jong, 2004; de Jong and Zawaydeh, 2002). Such a distinctiveness can be reached through extreme articulation like longer durations and larger magnitudes (e.g. Beckman and Edwards, 1994; Cho, 2006; de Jong, 1995a,b). Based on this, de Jong (2004) argued that the finding of less coarticulation was due to hyperarticulation (see also Lindblom's (1990) Hyper- and Hypospeech theory (H&H-Theory)) in accented words, resulting in an enhancement of distinctiveness. The influence of phrasal accent on articulation was investigated by a number of studies (Cambier-Langeveld, 2000; Cambier-Langeveld and Turk, 1999; Cho and McQueen, 2005; Dilley et al., 1996; Eefting, 1991; Fougeron, 2001; Harrington et al., 1995; Menadier et al., 1998; Mooshammer, 2010; Turk and Sawusch, 1997), primarily investigating lengthening of segments associated with a phrasal accent. In an EPG study, Cho and Keating (2009) found a cumulative effect of prosodic positions, stress and also a weak effect of accent on the linguo-palatal contact duration of word initial consonants. Edwards et al. (1991) analyzed jaw movements in [pap] sequences and lower lip movements in [pe] sequences. They found that the final closing gestures were longer and slower, but not more displaced in accented syllables. In addition, Harrington et al. (1995) and Beckman et al. (1992) showed that the acoustic differences between accented and unaccented words can be explained by lower jaw positions and truncation in unaccented words. Cho (2004) reported less vowel-to-vowel (carryover and anticipatory) coarticulation when words were accented than when words were deaccented. Generally, accentuation was found to cause a larger, longer and faster articulatory movement (see e.g. Avesani et al., 2007; Cho, 2006; Fowler, 1995).

To summarize, lengthening in the vicinity of high prosodic boundaries was found for consonants (Bombien et al., 2007; Byrd and Choi, 2010). With respect to CC overlap, Bombien et al. (2007) found that consonantal overlap decreased at higher prosodic boundaries, but there was no effect of lexical stress. Byrd and Choi (2010), on the other hand, found no effect of phrase boundaries on gestural overlap of onset clusters. Further, Mücke et al. (2008), investigating nasal place assimilation in Vn#g sequences, reported more gestural overlap when the target words were uttered in focus position than when produced in unaccented position. They explained their findings with a shortening of the gestures' plateaus in prosodically weaker contexts instead of indirect shortening due to more overlap. Thus the effect of accent on gestural overlap is not quite clear and there might be a relationship between plateau duration and gestural overlap.

Since Jun (2004) proposed that a slower gesture is longer and therefore overlaps to a greater extent with neighboring gestures, also the gestures' stiffness, which is considered to be the relative speed of an articulator independent of its displacement was included. Roon et al. (2007) concluded that differences in stiffness of neighboring segments may broaden our understanding of overlapping patterns. In addition, chapter 3 showed that stiffness is an important factor in gestural timing.

4.1.2 π-gesture approach and Articulatory Phonology

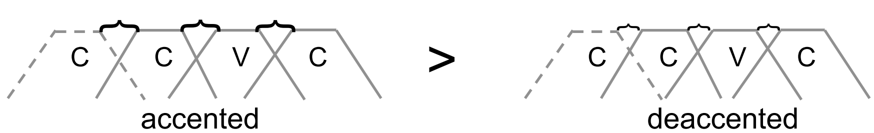
Within the framework of Articulatory Phonology (e.g. Browman and Goldstein, 1986, 1988, 2000), prosodic effects on articulation are explained by the abstract π-gesture (Byrd and Saltzman, 2003). This π-gesture affected in earlier versions (e.g. Byrd, 2000; Byrd et al., 2000; Byrd and Saltzman, 1998) the stiffness of the boundary adjacent gestures. In more recent versions (Byrd and Saltzman, 2003), the stiffness approach is replaced by local clock slowing. While the π-gesture is active, this clock slowing causes lengthening of the gestural movements and less overlap. As a consequence of lesser degree of overlap, the π-gesture causes gestures to become larger in spatial magnitude (i.e. there is articulatory strengthening). The activation level of a π-gesture waxes continuously towards its peak – which is located at the position of the boundary – and then wanes again. Thus the π-gesture's influence on speech gestures should increase until the π-gesture's peak activation is reached and then to decrease (Byrd et al., 2006; Byrd and Saltzman, 2003). Therefore, the prosodic effects on gestures are strongest at the π-gesture's peak activation and decrease with distance from this peak. Given that accent is generally found to affect vowels to a greater degree than consonants (see e.g. Cho and Keating, 2007; Pierrehumbert and Talkin, 1992), the peak activation should be situated in the nucleus of the accented syllable. Therefore, in a (C)CVC syllable, the vowel is expected to be affected most. The vowel-adjacent consonants – $C2_{on}$ and $C1_{off}$ – are expected to be affected as well but to a lesser extent. $C1_{on}$ (in the cluster condition) is the consonant most apart from the vowel and thus it is expected to be least affected by accentuation. Saltzman et al. (2008) presented yet a newer version, in which the π-gesture is replaced by the more general μ-gestures. These μ-gestures control μ_T-gestures and μ_S-gestures. μ_T-gestures modulate the temporal behavior of gestures, whereas μ_S-gestures modulate articulatory strengthening effects.

Further, within the framework of Articulatory Phonology, onset clusters are claimed to exhibit a c-center organization, which is primarily expressed by a shift of the vowel-adjacent consonant towards the vowel in complex onset tokens compared with simplex onset tokens (see chapter 1 for more details). In a recent study on German onset clusters, Brunner et al. (2014) showed that the 'c-center stability' is more frequently found in words containing tense vowels than in words containing lax vowels. Thus in addition to cluster composition (e.g. Marin, 2013), vowel tensity seems to be a further factor influencing the organization of onset clusters. Additionally, it has been shown for German that deaccented/unstressed tense vowels have considerably shorter acoustic durations than accented/stressed tense vowels, whereas unstressed vs. stressed lax vowels differ primarily in quantity (Hoole and Mooshammer, 2003; Mooshammer and Fuchs, 2002). In particular, Mooshammer and Geng (2008) showed that lax as opposed to tense vowels were spatially but not temporally reduced and thus not acoustically shortened in unstressed position. The present study therefore includes lax and tense vowels to see whether there are not only differences in the acoustics but also in articulatory timing patterns, especially in the timing between the vowels and their surrounding consonants in accented versus deaccented words.

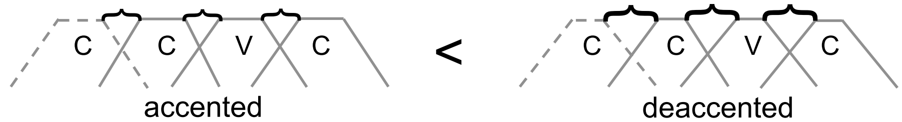
The aim of this study is to investigate the interplay between accentuation, vowel tensity and articulatory timing patterns with respect to all syllable constituents. The study is extended to prosodic effects on CC, CV and VC overlap, plateau durations and stiffness. Of particular interest are the effects of (de)accentuation on syllable coordination proposed by Articulatory Phonology. Therefore, this study includes syllables with simplex and complex onsets. Both the H&H-Theory (Lindblom, 1990) and the π-gesture approach (Byrd and Saltzman, 2003) would predict less overlap in accented words in order to enhance distinctiveness or as a consequence of local clock slowing. However, also less overlap in deaccented words may occur (see Mücke et al., 2008). The question that arises is how speakers coordinate speech gestures in order to mark phrasal accent.

4.2 Hypotheses

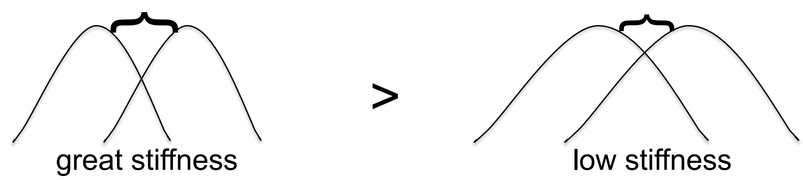
H1 Consonants and vowels are shorter in their acoustic durations when uttered in deaccented positions, but lax vowels are less effected by deaccentuation than tense vowels.
H2 There is more overlap between all syllable constituents (CC, CV and VC sequences) in the deaccented than in the accented condition.



H3 If there is, however, less overlap in deaccented tokens then plateau durations should be shortened.



H4 If it is the case that shorter segments overlap to a lesser degree, we should also observe less overlap in words containing lax vowels than in words with tense vowels.
H5 Shorter segments (such as lax vowels and deaccented words) are produced with greater stiffness and exhibit less overlap.



4.3 Method

4.3.1 Subjects and recordings

The acoustic and articulatory data used for analyses in the present study were collected during the the recording session for the experiment described in chapter 3. Thus speakers and the recording procedure was the same. Further details can be found in section 3.3.

4.3.2 Speech material

The test items were non-existent words containing simplex and complex onsets. The speech materials contained part of the materials used in chapter 3 plus words containing lax vowels and deaccented tokens. In German, lax vowels are phonologically short while tense vowels are phonologically long, for example [mɪtə] ('center') vs. [miːtə] ('rent'). The difference between short /a/ and long /aː/ is primarily in duration (Jessen, 1993; Mooshammer and

Fuchs, 2002).

All target words were embedded in a specific carrier phrase with only one intermediate (and thus one intonation) phrase. The test sentences contained one pitch accent that was either on *Melanie's* (i.e. the target word was deaccented) or on the target word. *Melanie* was used, because of its sonorant sound sequences and its accent on the first syllable. This means, that in the deaccented condition the pitch accent is quite distant from the target word. The target words were used indicating the name of Melanie's grandma in the following carrier phrase:

"Melanies Omi [**target word**] imitiert ein Lied."
("Melanie's grandma [**target word**] imitates a song.")

"WELCHE Omi von Melanie imitiert ein Lied?" ("WHICH grandma of Melanie imitates a song?") was used to trigger a pitch accent on the target word in the accented condition. The question "WESSEN Omi [target word] imitiert ein Lied?" ("WHOSE grandma [target word] imitates a song?") was used to evoke the pitch accent on Melanie's in the deaccented condition. Thus in the deaccented condition the target word was already present in the corresponding question and therefore was more likely to be deaccented as it was old information for the speakers. In addition, during the recordings, blue coloured letters drew attention to the words that had to be pronounced with a pitch accent. Thus the word to be accented was graphically highlighted on the computer screen. The experimenter ensured that subjects repeated incorrect prosodic patterns.

Simplex onsets consisted of /n/, the onset cluster included was /kn/. This onset cluster was chosen because only in words containing this onset cluster there was a clear shift towards the vowel due to complexity (see chapter 3). Thus it can be tested whether deaccentuation affects this shift towards the vowel or not. The target words used are listed in table 4.1 and were produced either in accented or deaccented position.

	Simplex onsets	**Complex onsets**
lax vowels	/**n**ap/	/**kn**ap/
tense vowels	/**n**aːp/	/**kn**aːp/

Table 4.1: *Speech material including tense and lax vowels uttered in accented in deaccented sentences.*

4.3.3 Acoustic segmentation and measurements

The segmentation procedure was the same as in chapter 3. In addition, the target word's accentuation pattern (i.e. accented vs. deaccented) was manually labelled. For all acoustic analyses absolute durations were measured because no appropriate normalization technique could be found. Acoustic durations of all syllable constituents and the fundamental frequency (f0) were analyzed in order to verify that speakers correctly produced deaccented patterns. This verification is necessary prior to further analyses.

4.3.4 Articulatory segmentation and measurements

The articulatory segmentation was done in EMU (Harrington, 2010) and the same articulators as in chapter 3 were labeled. For timing measurements as CC_{lag}, CV_{lag} and VC_{lag}, the normalization method described in Bombien (2011) was applied as it was also done in the previous chapter. Following this method, first the lag between two neighbouring segments was determined and this lag then was normalized on the entire duration of the two gestures (see equations 4.1 - 4.3). Thus negative values signify overlapping plateaus, positive values signify a lag between two neighboring segments.

$$CC_{lag} = \frac{P_{on}[C2_{on}] - P_{off}[C1_{on}]}{G_{off}[C2_{on}] - G_{on}[C1_{on}]} \tag{4.1}$$

$$CV_{lag} = \frac{P_{on}[V] - P_{off}[C2_{on}]}{G_{off}[V] - G_{on}[C2_{on}]} \tag{4.2}$$

$$VC_{lag} = \frac{P_{on}[C1_{off}] - P_{off}[V]}{G_{off}[C1_{off}] - G_{on}[V]} \tag{4.3}$$

Plateau durations were calculated by subtracting the plateau's onset by the plateau's offset. Thus higher values signify longer plateaus and lower values signify shorter plateaus.
In addition, also the stiffness of each gesture (i.e. for /k/, /n/, /aː/, / a/ and /p/) was calculated using the formula described in equation 4.4.

$$Stiffness = \frac{\text{peak velocity (mm/ms)}}{\text{maximum displacement (mm)}} \tag{4.4}$$

To minimize speaker-dependent influences, Lobanov-Normalization (Lobanov, 1971) was conducted on stiffness data.

4.3.5 Statistics

For the statistical analysis various mixed models were conducted. Fundamental frequency, acoustic vowel and consonant durations, CC_{lag}, CV_{lag}, VC_{lag}, plateau durations and the gestures' stiffness each served as the dependent variable in one of the models. Tensity (two levels: lax vs. tense), accentuation (two levels: accented vs. deaccented) and onset complexity (two levels: simplex vs. complex) were the fixed factors. The speaker was entered as random factor. The models were implemented using the lme4 package (Bates, 2008) in the statistical environment R.

4.4 Results

The results will be presented in five parts: section 4.4.1 reports results on acoustic analyses including the fundamental frequency (f0) during the vowels and the acoustic durations of the syllable constituents. Section 4.4.2 presents results on CC_{lag} as well as on the corresponding plateau durations of $C1_{on}$ and $C2_{on}$. Section 4.4.3 reports results on CV_{lag} as well as on the corresponding plateau durations of $C2_{on}$ and the vowel (V). Consequently, section 4.4.4 presents results on VC_{lag} as well as on the corresponding plateau durations of the vowel (V) and C1$_{off}$. Finally, in section 4.4.5 results on the gestures' stiffness are presented.

4.4.1 Acoustic

Fundamental frequency

Prior to articulatory analyses, the fundamental frequency (f0) during both tense and lax vowels – tense and lax vowels were shown to have considerably the same f0 (Fischer-Jørgensen, 1990) – was measured in order to see whether speakers successfully differentiated accented tokens from deaccented tokens.

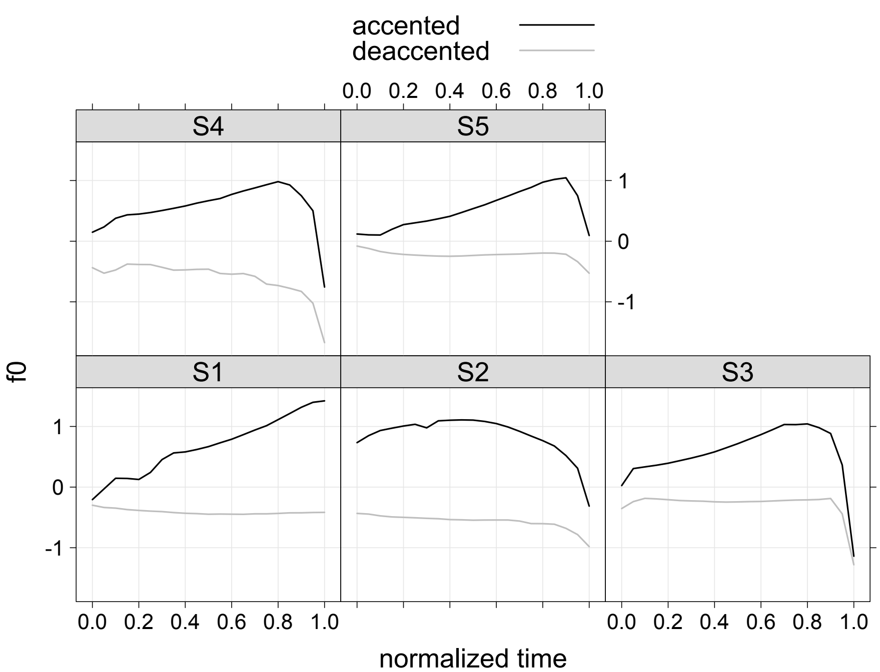


Figure 4.1: *Aggregated fundamental frequency during tense and lax vowels in accented (black) and deaccented (gray) words separately for each speaker.*

Figure 4.1 shows clear differences between words uttered in accented and deaccented positions for each speaker: f0 is higher in vowels produced in accented monosyllables than in deaccented monosyllables. This observation was confirmed statistically: a mixed model with f0 at the temporal midpoint (0.5) as the dependent variable and with accentuation (accented vs. deaccented) as fixed factor revealed a significant effect ($\chi^2[1] = 302.2$, p<0.001).

Thus commensurate with the predictions from the literature (see section 4.1.1), speakers distinguished accented and deaccented tokens by means of f0 differences. This observation is fundamental for further acoustical and articulatory analyses dealing with the consequences of prosodic weakening.

Acoustic durations of syllable components

The next step was to determine to which extent the syllable constituents were acoustically affected by prosodic weakening. Figure 4.2 shows the acoustic duration of $C1_{on}$ (left) and $C2_{on}$ (right) in accented and deaccented words.

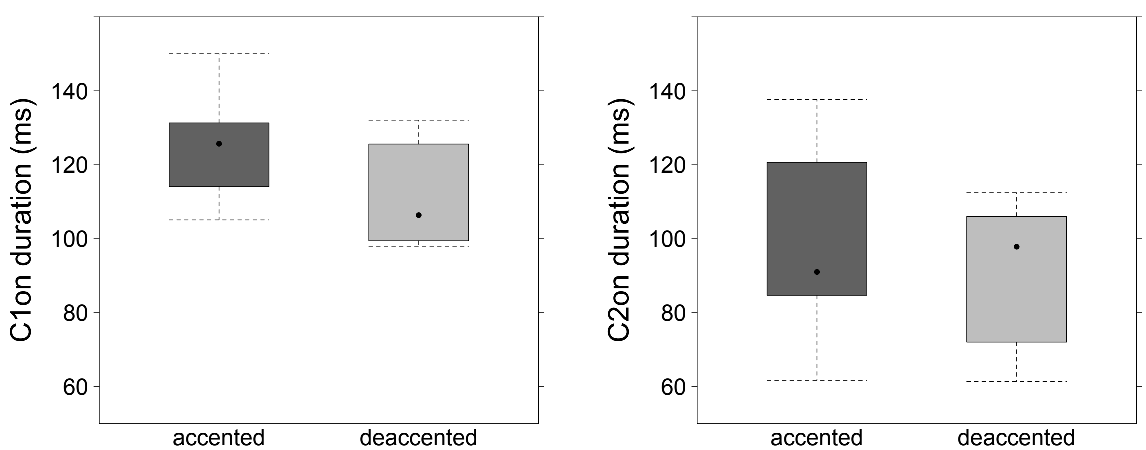


Figure 4.2: *Acoustic durations of* $C1_{on}$ *(left) and* $C2_{on}$ *(right) in accented (dark gray) and deaccented (light gray) words.*

A mixed model with the acoustic $C1_{on}$ duration as the dependent variable and with accentuation (accented vs. deaccented) as the fixed factor revealed that – commensurate with Figure 4.2 – there was a significant effect ($\chi^2[1] = 15.2$, p<0.001). Thus the duration of $C1_{on}$ was significantly shorter in deaccented words than in accented words.

As far as the acoustic duration of $C2_{on}$ is concerned, Figure 4.2 shows only slight differences between accented and deaccented tokens. However, a mixed model with the acoustic $C2_{on}$ duration as the dependent variable and with accentuation (accented vs. deaccented) as the fixed factor revealed a significant effect ($\chi^2[1] = 11.0$, p<0.001). This means, that the duration of $C2_{on}$ was shorter in deaccented words than in accented words.

Further, the acoustic vowel durations were measured separately for lax and tense vowels. This separate analysis was motivated by findings from previous studies showing prosody-induced acoustic shortening only for tense but not for lax vowels (Harrington et al., 2015; Mooshammer and Fuchs, 2002; Mooshammer and Geng, 2008).

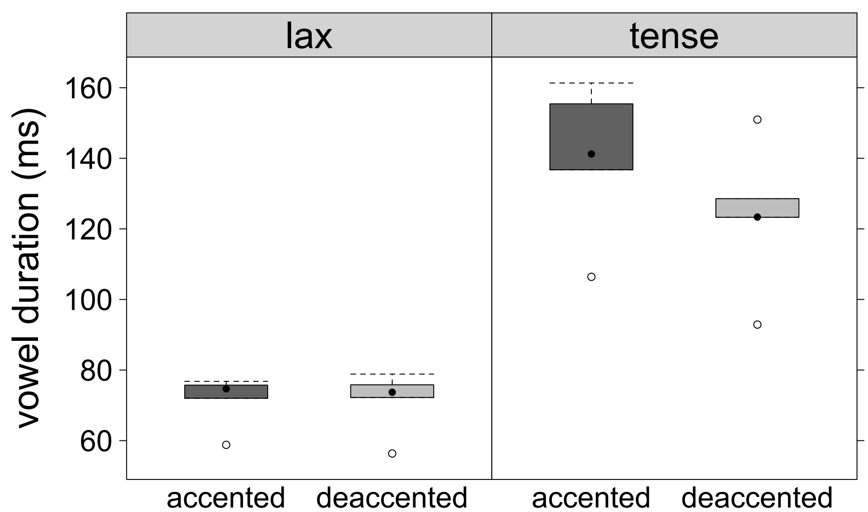


Figure 4.3: *Acoustic durations of lax (reft) and tense (right) vowels in accented (dark gray) and deaccented (light gray) words.*

As is evident from Figure 4.3, tense vowels are generally longer in acoustic duration than lax vowels. Further, only tense vowels seem to be acoustically shorter in deaccented words than in accented words.

A mixed model with the acoustic vowel duration as the dependent variable and with fixed factors accentuation (accented vs. deaccented) and tensity (tense vs. lax) revealed that there was a significant effect of accentuation ($\chi^2[1] = 17.4$, p<0.001), tensity ($\chi^2[1] = 404.1$, p<0.001), as well as a significant interaction between accentuation and tensity ($\chi^2[1] = 17.2$, p<0.001). Post-hoc Tukey tests revealed that – commensurate with Figure 4.3 – lax vowels were always shorter in their duration than tense vowels, irrespective of accentuation (accented: p<0.01; deaccented: p<0.01). Further, only tense vowels were acoustically shorter in deaccented words than in accented words (p<0.001). This result is in line with previous findings of prosodic influences on tense and lax vowels by Mooshammer and Fuchs (2002), Mooshammer and Geng (2008) and Harrington et al. (2015) and provides further evidence that lax vowels are not acoustically affected by prosodic weakening. It further substantiates the suggestion that lax vowels are incompressible because of their inherently short acoustic duration (Klatt, 1973). In spite of this result, there may be articulatory shortening or reduction. For example, Mooshammer and Fuchs (2002) showed that lax vowels were significantly shortened in unstressed positions, at least by some speakers, although the lax vowel space area was maintained.

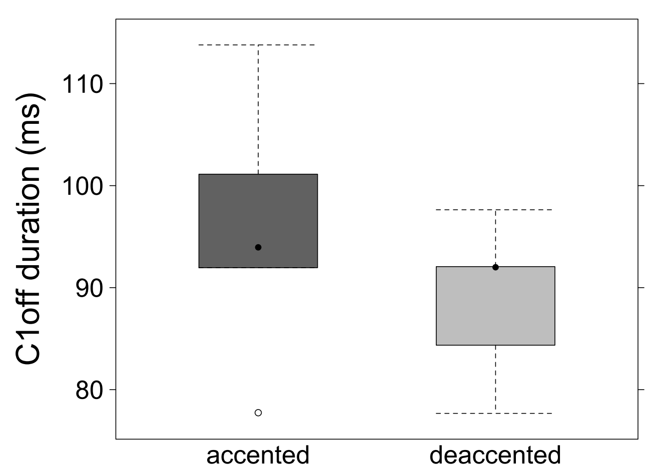


Figure 4.4: *Acoustic C1$_{off}$ durations in accented (dark gray) and deaccented (light gray) words.*

Concerning the acoustic duration of C1$_{off}$, Figure 4.4 clearly shows shorter durations in deaccented words than in accented words. This observation was confirmed statistically: a mixed model with the acoustic $C1_{off}$ duration as the dependent variable and with accentuation (accented vs. deaccented) as the fixed factor revealed a significant effect ($\chi^2[1] = 6.2$, $p<0.05$). In the following sections, the results on syllable timing including CC_{lag}, CV_{lag} and VC_{lag} are reported. For each of those three lags, the plateau durations of the corresponding two gestures are additionally presented in order to investigate the relationship between overlap and plateau duration.

4.4.2 CC_{lag}

In order to account for the influence of accentuation on the temporal organization within onset clusters and on the corresponding plateau durations, the CC_{lag} between /k/ and /n/ and the corresponding plateau durations in accented and deaccented words was analyzed (i.e. only words containing complex onsets were analyzed). Figure 5.3 shows the CC_{lag} (upper panel) and the plateau durations of $C1_{on}$ and $C2_{on}$ (lower two panels).

There seems to be a trend towards more overlap in the deaccented condition in words containing lax vowels (see upper panel of Figure 5.3). A mixed model with CC_{lag} as the dependent variable and with accentuation (accented vs. deaccented) and tensity (tense vs. lax) as fixed factors revealed a significant main effect of accentuation ($\chi^2[1] = 4.5$, $p<0.05$), confirming this observation. There was no effect of tensity and no interaction between accentuation and tensity. As is evident from Figure 5.3, the effect of accentuation may only hold for lax vowel tokens. The non-significant interaction between tensity and accentuation may be explained by the fact that there is a trend towards the same direction also in tense vowel tokens.

Further, the plateau durations of $C1_{on}$ appear not to differ with respect to accentuation or tensity. That is, the duration of /k/ seems to remain the same, irrespective of vowel tensity and accentuation (see Figure 5.3 a). This was confirmed by a mixed model with $C1_{on}$ plateau duration as the dependent variable and with the same fixed factors as above, revealing no significant main or interaction effect.

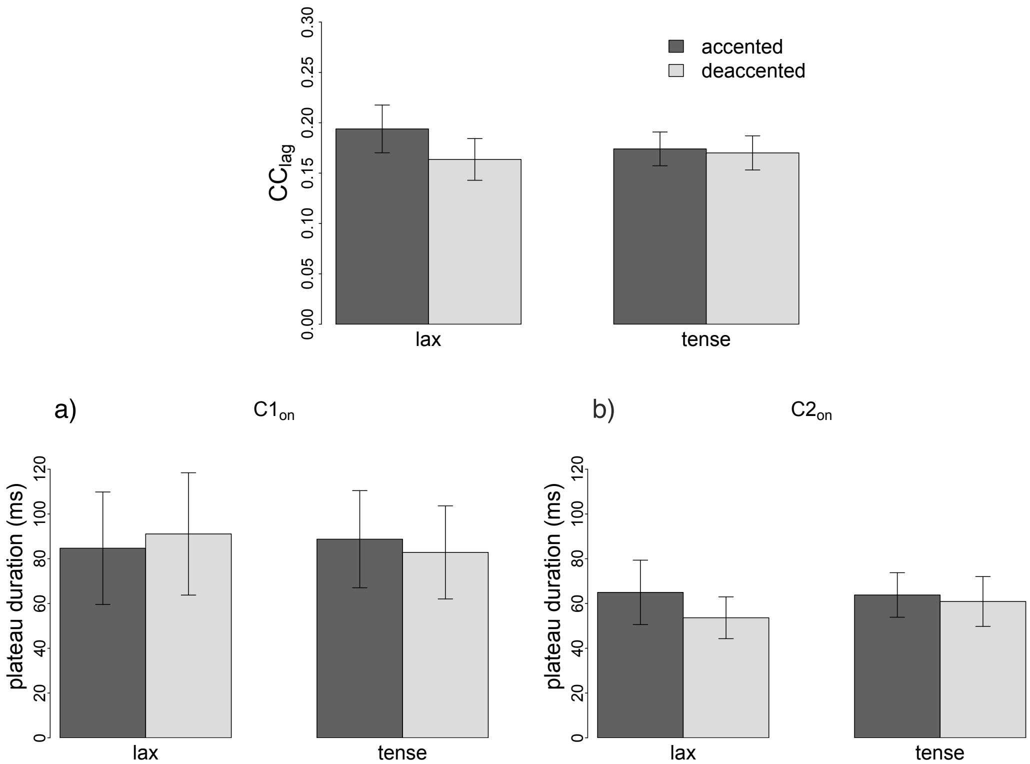


Figure 4.5: *CC_{lag} in accented (dark gray) vs. deaccented (light gray) tokens containing lax (left) and tense (right) vowels (upper panel). Plateau durations of $C1_{on}$ (lower left panel) and of $C2_{on}$ (lower right panel) in accented (dark gray) vs. deaccented (light gray) tokens containing lax (left) and tense (right) vowels.*

Concerning $C2_{on}$, there seems to be a tendency towards shorter plateau durations in deaccented words, especially in words containing lax vowels (see Figure 5.3 b). A mixed model with $C2_{on}$ plateau duration as the dependent variable and with the same fixed factors as above revealed a significant effect of accentuation ($\chi^2[1] = 7.9$, p<0.01). As is evident from Figure 5.3 b, this effect may only hold for lax vowel tokens. However, tensity did not influence the plateau duration of $C2_{on}$ and there was no significant interaction between the two factors. The non-significant interaction between tensity and accentuation may be explained by the fact that there is a trend towards the same direction also in tense vowel tokens.

To conclude, CC_{lag} was affected by accentuation such that there was more overlap in deaccented words and the corresponding plateau of $C2_{on}$ was shortened, too. Thus greater overlap can also occur with plateau shortening. This result supports findings reported by Byrd and Choi (2010), Byrd et al. (2006) and Bombien et al. (2006), showing less overlap at higher prosodic boundaries, but it contradicts findings reported by Mücke et al. (2008).

4.4.3 CV_{lag}

In order to investigate the influence of onset complexity and accentuation both on the timing between the vowel-adjacent onset consonant ($C2_{on}$) and the vowel and on the corresponding plateau durations, words containing simplex (/n/) as well as complex (/kn/) onsets were included. This gives the opportunity to investigate the effect of prosodic weakening on the vowel-adjacent consonant's shift towards the vowel in complex onset tokens. Figure 4.6 shows the CV_{lag} (upper panel) and the plateau durations of $C2_{on}$ and the vowel (lower two panels). The upper panel of Figure 4.6 shows higher lag values in the deaccented condition. This means that, in general, there was less CV overlap in deaccented words compared with accented words. Further, there seems to be more overlap in terms of lower lag values in complex onsets compared with their simplex counterparts, indicating a shift towards the vowel.

These observations were confirmed by a mixed model with CV_{lag} as the dependent variable and with fixed factors accentuation (accented vs. deaccented), tensity (tense vs. lax) and onset complexity (simplex vs. complex), showing a significant main effect of accentuation ($\chi^2[1] = 43.9$, $p<0.001$) and onset complexity ($\chi^2[1] = 27.8$, $p<0.001$). Further, there was a significant three-way interaction between onset complexity, accentuation and tensity ($\chi^2[4] = 11.0$, $p<0.05$). Figure 4.6 shows, that the shift towards the vowel is less in deaccented tokens and even more so in deaccented lax vowel tokens. Post-hoc Tukey tests showed significant differences between simplex and complex accented lax tokens ($p<0.001$), between simplex and complex accented tense tokens (p<0.001) and a trend between simplex and complex deaccented tense tokens (p<0.10), but there was no difference between simplex and complex deaccented lax tokens. Thus the shift towards the vowel – and a potential c-center organization – is reduced or even non-existent in prosodically weak words. Further, c-center organization seems to be more stable and robust to prosodic influences in words containing tense vowels. This finding supports to some extent the results reported in Brunner et al. (2014), who showed that a cluster's c-center is more stable in words containing tense vowels as opposed to lax vowels and diphthongs.

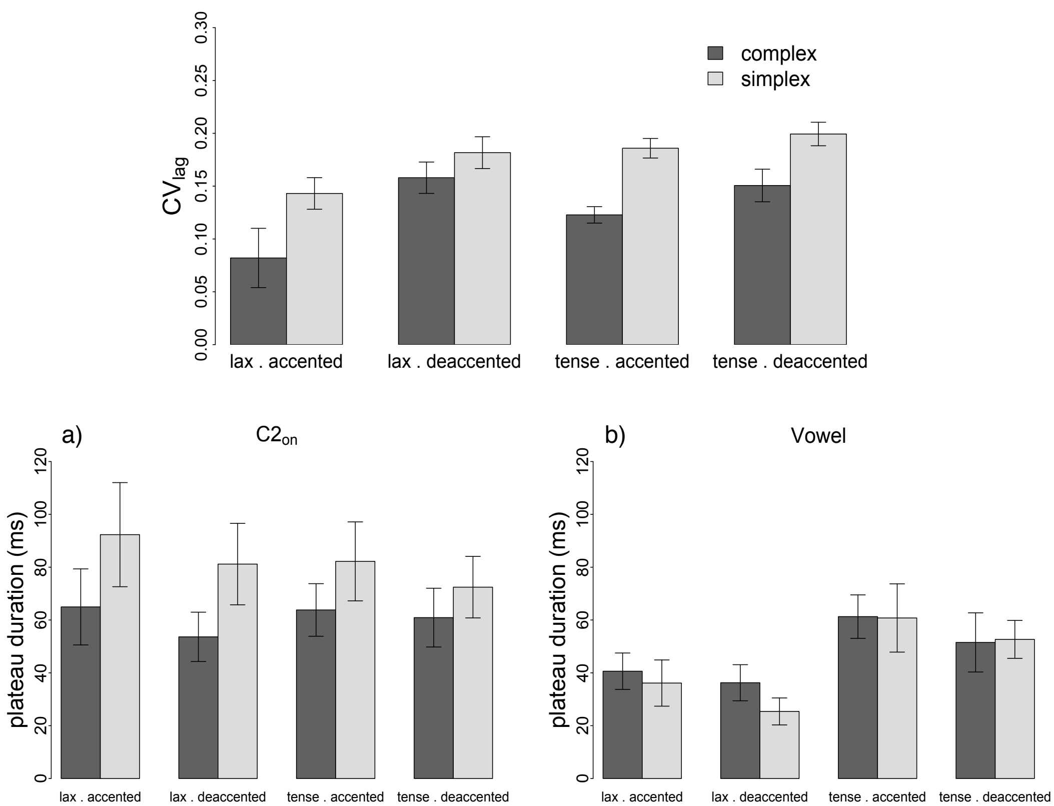


Figure 4.6: *CV_{lag} in accented vs. deaccented tokens containing lax and tense vowels with simplex (light gray) or complex (dark gray) onsets (upper panel). Plateau durations of $C2_{on}$ (lower left panel) and of the vowel (lower right panel) in accented vs. deaccented tokens containing lax and tense vowels with simplex (light gray) or complex (dark gray) onsets.*

Concerning the plateau durations of $C2_{on}$, additional consonants in onset position seem to compress the plateau durations (see Figure 4.6 a). A mixed model with $C2_{on}$ plateau duration as the dependent variable and with with the same fixed factors as above showed a significant effect of onset complexity ($\chi^2[1] = 79.3$, $p<0.001$), accentuation ($\chi^2[1] = 15.1$, $p<0.001$) and a significant interaction between onset complexity, accentuation and tensity ($\chi^2[4] = 10.4$, $p<0.05$). Post-hoc Tukey tests revealed significant differences between simplex and complex accented lax tokens ($p<0.001$), between simplex and complex deaccented lax tokens ($p<0.001$) and between simplex and complex accented tense tokens ($p<0.001$), but not between simplex and complex deaccented tense tokens. This means that the plateau duration of $C2_{on}$ was not shortened due to an increasing number of onset consonants in deaccented tense tokens. Further, there was a significant interaction between onset complexity and tensity ($\chi^2[1] = 7.5$, $p<0.01$). Post-hoc Tukey test revealed that the plateau duration of $C2_{on}$ was significantly shorter in complex onsets in both vowel contexts (lax: $p<0.001$; tense: $p<0.001$), but it differed only between complex lax and complex tense tokens ($p<0.001$), not between simplex ones.

Concerning the vowel plateau durations, Figure 4.6 b shows clear differences between lax and tense vowels: there are shorter plateau durations in lax vowels compared with tense vowels. Further, vowel plateau durations seem to be shorter in deaccented than in accented words. A mixed model with vowel plateau duration as the dependent variable and with the same fixed factors as above revealed a significant main effect of tensity ($\chi^2[1] = 29.4$, $p<0.001$), indicating that tense and lax vowels can indeed be distinguished articulatorily in their plateau durations: acoustically shorter lax vowels have articulatorily shorter plateaus whereas acoustically longer tense vowels exhibit longer plateau durations. Further, there was a significant main effect of accentuation ($\chi^2[1] = 4.6$, $p<0.05$). That is, vowel plateau durations were shorter in deaccented than in accented tokens. Onset complexity did not affect the plateau durations of lax and tense vowels and there was no significant interaction.

To conclude, CV_{lag} was greater in deaccented words, i.e. there was more overlap between $C2_{on}$ and the vowel in accented than in deaccented words. Additionally, the plateau durations of the vowel and $C2_{on}$ were shorter in deaccented words. Thus less CV overlap in deaccented words as opposed to accented words may be caused by vowel plateau compression combined with $C2_{on}$ plateau compression. These findings for less overlap in deaccented words and shorter plateau durations as opposed to accented words are in line with the results presented by Mücke et al. (2008). However, it has to be considered that the plateau durations of $C2_{on}$ were shorter in words containing complex onsets than in words containing simplex onsets, although there were shorter CV_{lags} in complex onset tokens. This indicates, that more overlap can emerge together with plateau shortening.

4.4.4 VC_{lag}

Concerning VC sequences, the upper panel of Figure 4.7 shows clear differences in overlap between tense and lax vowels. The lower lag values in words containing lax vowels in both accented and deaccented tokens indicate more overlap between the vowel and $C1_{off}$. Further, lower VC_{lag} values can be observed in deaccented words, indicating more overlap than in accented words. Finally, there seem to be no differences in VC_{lag} depending on onset complexity.

A mixed model with VC_{lag} as the dependent variable and with fixed factors accentuation (accented vs. deaccented), tensity (tense vs. lax) and onset complexity (simplex vs. complex) showed that – commensurate with Figure 4.7 – accentuation had a significant influence on the timing between the vowel and $C1_{off}$ ($\chi^2[1] = 5.1$, $p<0.05$). That is, there was more overlap in words produced without a pitch accent compared with words produced with a pitch accent.

In line with the observation mentioned above, there was also a significant effect of tensity ($\chi^2[1] = 32.3$, $p<0.001$). This means that there was more VC overlap indicated by lower lag values in words containing lax vowels compared with words containing tense vowels. Further, there was no effect of onset complexity and no interaction.

As described earlier in section 4.4.3, vowel plateau durations were shorter in deaccented words and for lax vowels (see Figure 4.7 a).

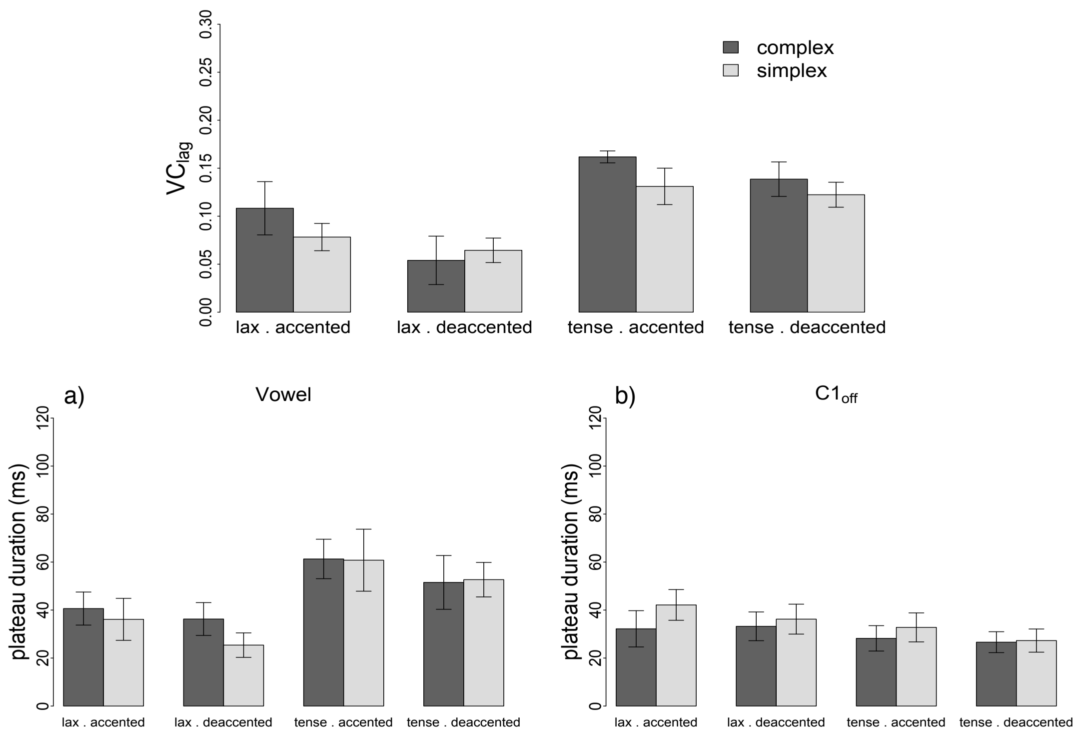


Figure 4.7: *VC_{lag} in accented vs. deaccented tokens containing lax and tense vowels with simplex (light gray) or complex (dark gray) onsets (upper panel). Plateau durations of the vowel (lower left panel) and of $C1_{off}$ (lower right panel) in accented vs. deaccented tokens containing lax and tense vowels with simplex (light gray) or complex (dark gray) onsets.*

Concerning the plateau durations of $C1_{off}$, Figure 4.7 b shows no differences with respect to accentuation. However, the plateau durations seem to be shorter in words containing tense vowels than in words containing lax vowels and there are some tendencies towards shorter plateau durations in complex onset tokens than in simplex onset tokens. A mixed model with $C1_{off}$ plateau duration as the dependent variable and with the same fixed factors as above revealed a significant main effect of tensity ($\chi^2[1] = 18.5$, $p<0.001$). The duration of $C1_{off}$ is longer in words containing lax vowels than in words containing tense vowels. Further, there was a significant effect of onset complexity ($\chi^2[1] = 8.4$, $p<0.01$). Thus the plateau duration of $C1_{off}$ is generally shorter in words with an additional onset consonant. However, this effect might only hold for accented lax tokens. Accentuation had no significant influence on the plateau duration of $C1_{off}$ and there was no interaction. The non-significant interaction may be explained by the same trends in all tokens.

To conclude, VC_{lag} was affected by accentuation such that there was more overlap in deaccented words and the vowel plateau was shortened, too. Thus greater overlap can also occur with plateau shortening. Therefore, this result contradicts Mücke et al. (2008) who explained less overlap by plateau shortening. On the other hand, this result supports previous studies showing less overlap at higher prosodic boundaries (Bombien et al., 2006; Byrd and Choi, 2010; Byrd et al., 2006).

4.4.5 Stiffness

Stiffness of all syllable constituents was analyzed to test whether there are differences between accented and deaccented tokens as well as between lax and tense vowels. A mixed model with stiffness of $C1_{on}$ as the dependent variable and with fixed factors accentuation (accented vs. deaccented) and tensity (tense vs. lax) revealed no significant main or interaction effect. The same result was found for the stiffness of $C2_{on}$ depending on the same factors. Thus $C1_{on}$ and $C2_{on}$ were produced with the same velocity-to-displacement ratio in both prosodic conditions, although both segments were acoustically shorter in deaccented words and $C2_{on}$ was also articulatorily shortened. The vowel gestures' stiffness was calculated separately for lax and tense vowels.

Figure 4.8 shows that lax vowels seem to be produced with a greater stiffness in accented words. Further, accented lax vowels seem to be produced with a greater stiffness than accented tense vowels. A mixed model with the vowel stiffness as the dependent variable and with the same fixed factors as above revealed a significant main effect of tensity ($\chi^2[1] = 6.5$, $p<0.05$) and a significant interaction between accentuation and tensity ($\chi^2[1] = 4.7$, $p<0.05$). Post-hoc Tukey tests revealed significant differences in stiffness only between accented lax and tense vowels ($p<0.001$). This means that accented lax vowels were produced with a greater velocity-to-displacement ratio than accented tense vowels.

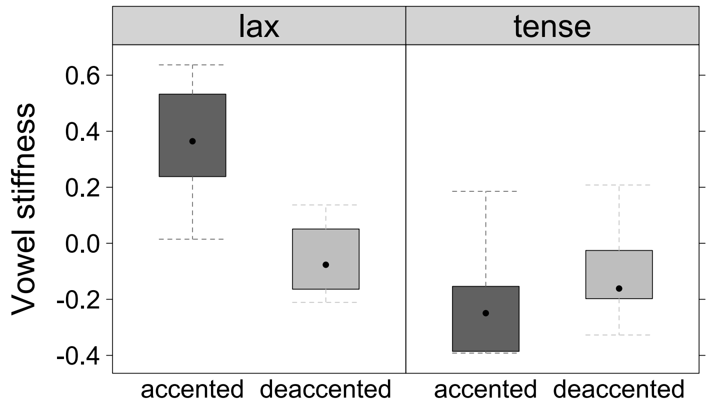


Figure 4.8: *Normalized stiffness of lax (left) and tense (right) vowels in accented (dark gray) and deaccented words (light gray).*

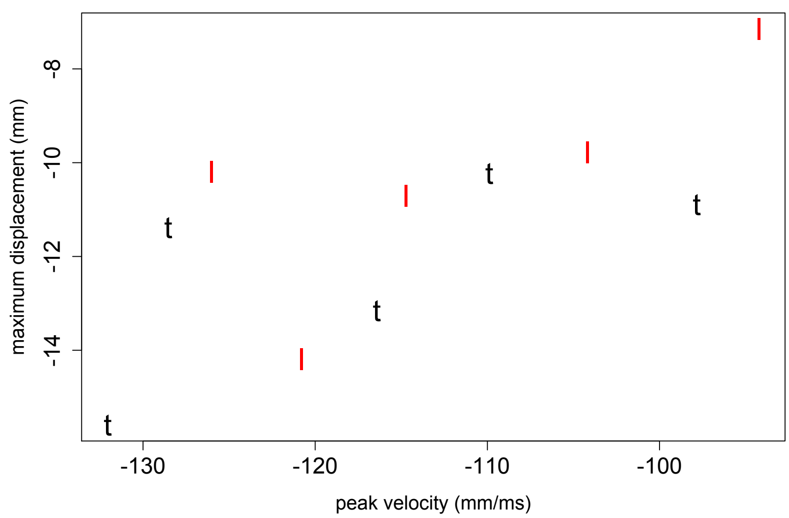


Figure 4.9: *Aggregated peak velocity and maximum displacement of lax (l; red) and tense (t; black) vowels in accented words.*

This difference in stiffness may be due to a greater velocity in lax vowels or due to a smaller displacement. Figure 4.9 illustrates both parameters included in the stiffness analysis: maximum displacement and peak velocity for tense (t; black) and lax (l; red) vowels in accented words. It can be observed that the difference between accented tense and lax vowels in stiffness is most likely caused by a generally smaller displacement of lax vowels. This means that the vowel gesture producing the tense vowel /aː/ is more displaced than the vowel gesture producing the lax vowel /a/.

As far as $C1_{off}$ is concerned, there seems to be a greater stiffness in words containing lax vowels and in deaccented words (see Figure 4.10). Commensurate with this observation, a mixed model with the stiffness of $C1_{off}$ as the dependent variable and with the same fixed factors as above revealed that there was a significant main effect of accentuation ($\chi^2[1] = 11.3$, $p<0.001$) and tensity ($\chi^2[1] = 22.8$, $p<0.001$). This means that $C1_{off}$ was produced with a greater velocity-to-displacement ratio in words containing lax vowels than in words containing tense vowels as well as in deaccented words than in accented words.

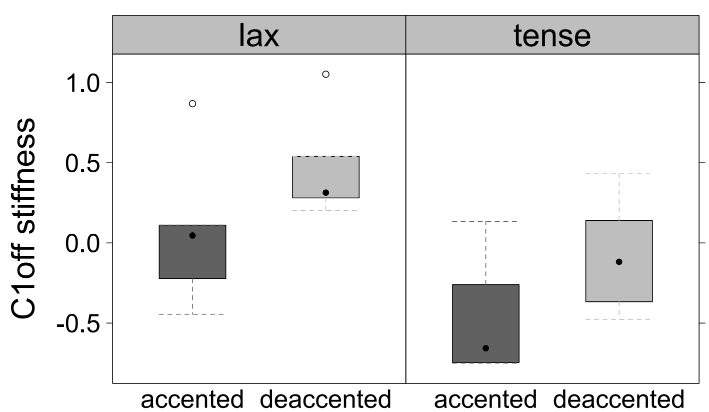


Figure 4.10: *Normalized stiffness of $C1_{off}$ in accented (dark gray) and deaccented (light gray) words containing lax and tense vowels.*

Again, these differences in stiffness may be due to a greater velocity or due to a smaller displacement.

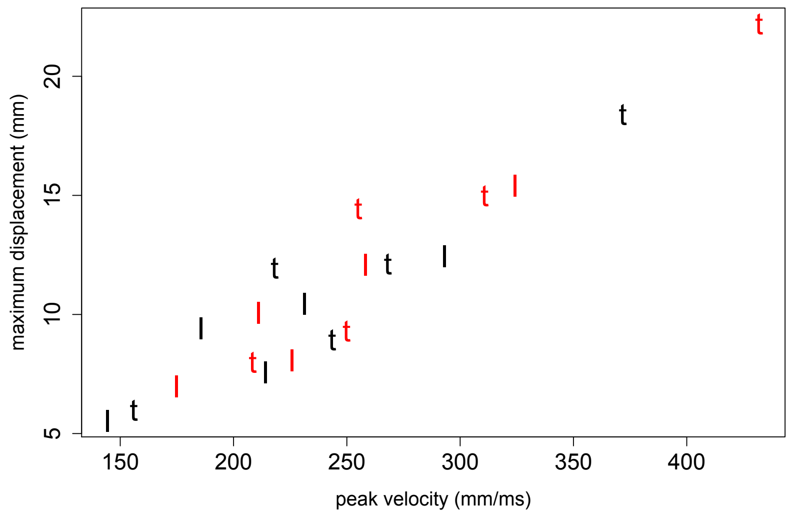


Figure 4.11: *Aggregated peak velocity and maximum displacement of $C1_{off}$ in words containing lax (l) and tense (t) vowels in accented (red) and deaccented (black) positions.*

Both parameters included in the stiffness analysis (maximum displacement and peak velocity) are illustrated in Figure 4.11 separately for words containing tense (t) and lax (l) vowels in accented (red) and deaccented (black) conditions. It can be observed that the greater stiffness in words containing lax vowels is most likely caused by a smaller displacement of $C1_{off}$ in this particular context: the maximum displacement of $C1_{off}$ in words containing lax vowels (l) is smaller than in words containing tense vowels (t).

The difference in stiffness between deaccented words seems also most likely to be caused by a smaller displacement in this particular context: the maximum displacement of $C1_{off}$ in deaccented tokens (black) is smaller than in accented tokens (red).

To conclude, stiffness of $C1_{on}$ and $C2_{on}$ did not differ with respect to accentuation and vowel tensity. Vowel stiffness was greater in accented lax than in accented tense vowels, because of a smaller maximum displacement in lax vowel tokens. Stiffness of $C1_{off}$ differed with respect to both vowel tensity and accentuation. These differences were explained by a smaller maximum displacement of $C1_{off}$ in words containing lax vowels and in deaccented words.

4.5 Summary and dicsussion

The aim of the present study was to further investigate the influence of prosodic weakening on the timing between gestures within a syllable (e.g. CC, CV and VC sequences) and to investigate possible differences between lax and tense vowels. Further, not only the timing between gestures was considered but also the plateau durations of each syllable constituent and its stiffness. The motivation behind this was that recent studies showed articulatory lengthening and less overlap at higher prosodic boundaries (Bombien et al., 2006; Byrd and Choi, 2010; Byrd et al., 2006). These observations are in line with Lindblom's (1990) H&H-Theory in which speech is hyperarticulated at higher and hypoarticulated at lower prosodic boundaries. Further, these observations are in accordance with the local slowing down of the clock during the activation of the π-gesture (Byrd and Saltzman, 2003). Mücke et al. (2008), on the other hand, reported less overlap at lower prosodic boundaries and explained their findings by plateau shortening in these particular contexts. Because of these contradicting results, further research was needed on the effects of (de)accentuation on syllable timing and the present study investigated possible relationships between timing differences and plateau durations.

4.5.1 Fundamental frequency and acoustic durations

The acoustic analyses showed that prosodic weakening shortens the duration of $C1_{on}$, $C2_{on}$, $C1_{off}$ and tense vowels, whereas lax vowels did not differ in their durations, confirming hypothesis **H1**. This difference in acoustic vowel shortening depending on vowel tensity is in line with previous studies by Mooshammer and Fuchs (2002), Mooshammer and Geng (2008) and Harrington et al. (2015). Thus the present study provides further evidence that lax vowels are incompressible which may be attributed to their generally short duration (Klatt, 1973). Further, the prosody-dependent acoustic shortening of the consonants supports findings reported in previous studies (e.g. Crystal and House, 1988; Ingrisano and Weismer, 1979; Klatt, 1974; Stathopoulos and Weismer, 1983; Turk and Sawusch, 1997; Umeda, 1977). In addition, speakers produced deaccented vowels with a lower fundamental frequency than accented vowels. All these findings are in line with previous studies showing that deaccented words are produced with a lower fundamental frequency and with shorter durations (see e.g. Cho, 2004; de Jong, 1995a; Fry, 1958; Lehiste, 1970).

4.5.2 Gestural coordination

$\mathbf{CC}_{lag}$

Articulatory results showed more CC overlap, as well as plateau shortening of $C2_{on}$ in the deaccented condition than in the accented condition. This finding supports hypothesis **H2** and is in line with Byrd and Choi (2010) who found gestural lengthening of this consonant in lexically stressed position and with Bombien et al. (2007), showing less overlap within consonant clusters at lower prosodic boundaries. It further indicates, that shorter plateau durations can also co-occur with increased overlap.

$\mathbf{CV}_{lag}$

Three major findings regarding CV sequences arise from this study. The first one is that the CV_{lag} is greater in deaccented than in accented tokens, which can be explained by shorter plateau durations of both $C2_{on}$ and the vowel. This result contradicts hypothesis **H2** predicting more overlap in deaccented words than in accented words. However, it confirms hypothesis **H3** predicting less overlap in deaccented words than in accented words caused by plateau shortening in the prosodically weaker context. Thus this result in line with the findings described in Mücke et al. (2008).

The second finding concerning CV sequences is that the vowel-adjacent consonant shifted towards the vowel in complex onsets confirming the predictions for globally organized onsets in Articulatory Phonology.

The third major finding is that this cluster dependent shift is diminished in deaccented tokens. There was no significant shift towards the vowel in deaccented tokens containing lax vowels and only a trend for $C2_{on}$ to shift towards the vowel in deaccented tokens containing tense vowels. This indicates that an onset cluster's c-center organization is less stable in lax vowels, supporting Brunner et al.'s (2014) finding of a more frequent c-center stability in words containing tense vowels than lax vowels. Further, this finding supports to some extent hypothesis **H4**, predicting less overlap in words containing lax vowels. The less marked shift towards the vowel in deaccented tokens may be explained by an additional lengthening of the plateau durations in accented as opposed to deaccented tokens, which may then result in a greater shift towards the vowel. This is not to say that articulatory lengthening in principal results in more overlap as articulatory strengthening should cause less overlap (see Cho, 2004) (e.g. there was less overlap in accented CC and VC sequences, as well as generally less overlap in tense vowels). In this particular case, however, the shift towards the vowel may be greater in the accented condition, because the temporal overlap between longer plateau gestures (due to accentuation) is reached faster. To conclude, the predictions for complex onsets made by the Articulatory Phonology are depending not only on the cluster's composition (see Marin, 2013), but also on vowel tensity (see also Brunner et al., 2014) and prosodic structures.

$\mathbf{VC}_{lag}$

Concerning VC sequences, there was more overlap in deaccented words, which is again in line with Byrd and Choi (2010) and Bombien et al. (2007). In addition, it supports hypothesis **H2**. Despite this result, there were shorter plateau durations of the vowel in the deaccented condition. This again indicates that shorter plateau durations can also co-occur with an increased overlap.

Further, onset composition affected the plateau duration of the coda consonant: there were shorter plateau durations of $C1_{off}$ as the number of onset consonants increased. This indicates that onsets and codas do not behave independently of each other within a syllable (see also Hawkins and Nguyen, 2000) and supports the findings presented in Peters and

Kleber (2014).

To conclude, this study shed further light on the interplay between plateau duration and overlap. The results showed, that plateau compression might be accompanied by an increased overlap (as in CC_{lag} and VC_{lag}), or a decreased overlap (as in CV_{lag}). However, what differed was that only in CV_{lag} plateau durations of both corresponding gestures (i.e. that of $C2_{on}$ and the vowel) were shortened, whereas in CC_{lag} and VC_{lag} only the plateau of one corresponding gesture was shortened. Thus only a compression of both plateaus leads to less overlap. Note, however, that this study does not allow for a direct comparison of plateau shortening between segments of different articulators, since the tongue tip is flexible resulting in short gestures and long plateau durations for nasals, whereas the tongue back moves slowly resulting in long vowel gestures and short plateau durations for vowels (compare for example Figure 4.6 a and Figure 4.6 b). For this reason, the results concerning plateau durations have to be extended by taking into account also gesture durations (see e.g. Byrd et al., 2005, measuring gesture durations).

Figure 4.12 schematically illustrates the results on syllable timing depending on prosodic prominence and onset complexity. Since vowel tensity had only a marginal effect, the results are displayed without referring to this factor. The most important findings regarding syllable timing in accented versus deaccented speech can be summarized as follows: (1) in deaccented words there is increased CC and VC overlap, (2) in deaccented words there is decreased CV overlap caused by a shortening of the corresponding plateaus and (3) in deaccented words the shift of the vowel adjacent consonant towards the vowel in complex onset tokens is diminished.

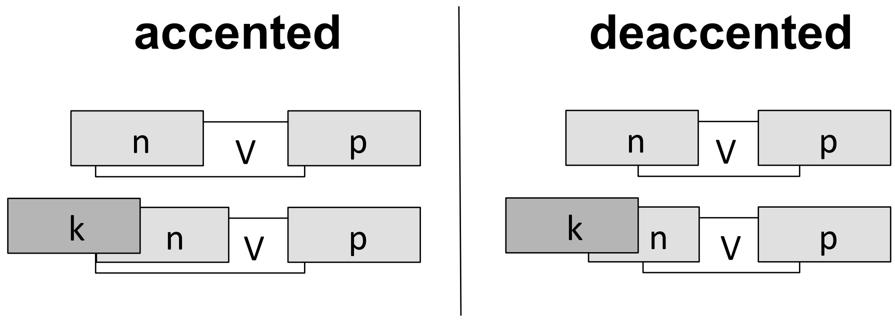


Figure 4.12: *Schematic illustration of results on syllable timing depending on onset complexity in the accented and deaccented condition.*

Tense vs. lax vowels

Less overlap in words containing lax vowels was expected (**H4**). This hypothesis, however, has to be refuted. In fact, the results of the present study showed more VC overlap in words containing lax vowels than in words containing tense vowels.

Stiffness

Stiffness, was taken into account to better understand the observed differences in overlap. It was assumed that greater stiffness of a gesture results in less overlap as predicted in hypothesis **H5**, based on suggestions made by Jun (2004) and Roon et al. (2007). However, the results of the present study cannot account for this hypothesis. For example, the stiffness of $C1_{on}$ and $C2_{on}$ was not affected by prosody or tensity, although there was less CC overlap in accented words. Concerning the vowel gestures' stiffness, results showed that accented lax vowels were produced with a greater stiffness than accented tense vowels. This contradicts the result for generally more VC overlap in words containing lax vowels. Further, for $C1_{off}$ the stiffness was greater in deaccented words and in words containing lax vowels. However, these findings contradict the result for more VC overlap in words containing lax vowels and the result for more VC overlap in deaccented words.

The findings on stiffness support suggestions made by Saltzman and Byrd (2000a), Saltzman and Byrd (2000b) and Cho (2006) that stiffness is insufficient to explain the entire scope of articulatory changes caused by prosodic influences. Furthermore, previous findings (c.f. Beckman and Edwards, 1994; Mooshammer and Harrington, 2005) indicate that speakers distinguish accented from deaccented words mainly by temporal differences as opposed to spatial differences. Nonetheless, it might be the case that gestures with a higher stiffness (e.g. lax vowels and $C1_{off}$) reach their target value earlier and this in turn leads to more overlap.

The observed difference in vowel stiffness between accented tense and lax vowels was related to a lower maximum displacement in lax vowels. Further, also the differences in $C1_{off}$ stiffness between accented and deaccented tokens as well as between lax and tense vowels were attributed to lower maximum displacement in deaccented and lax vowel tokens. These findings may be explained by different jaw positions in these particular contexts. For example, it is assumed that accented vowels are produced with a greater mouth opening and thus with a longer duration than deaccented vowels (de Jong, 1995b). Therefore, the maximum displacement of the lower lip is greater in accented tokens compared with deaccented tokens, because the jaw is lowered to a greater extent. However, the jaw height was found to be the same in German lax and tense vowels (Hoole and Mooshammer, 2003). Thus the lower maximum displacement of $C1_{off}$ with a preceding lax vowel as well as the lower maximum displacement of accented lax vowels cannot entirely be explained by differences in jaw height. However, Hoole and Mooshammer (2003) did not include the tense-lax vowel pair /aː, a/, thus it has to be proved whether there might be jaw height differences between these particular vowels.

4.6 Conclusion

Three major conclusions arise from this study. The first one is that lax vowels may indeed be articulatorily compressed due to deaccentuation without showing acoustic shortening.

The second conclusion is that syllable constituents are affected differently by prosodic conditions: there was more CC and VC overlap, but less CV overlap in deaccented words. Such differences cannot be explained by the π-gesture approach. However, only the plateau durations of $C2_{on}$ and the vowel were affected by prosodic weakening. This particular finding can indeed be explained by the π-gesture approach which predicts a waxing of π-gesture towards the nucleus (Byrd et al., 2006; Byrd and Saltzman, 2003), thus affecting $C2_{on}$ and the vowel. In addition, these prosody-dependent plateau shortening effects may account for the observed lesser CV overlap. Further, the fact that the plateau duration of $C1_{on}$ was not affected by prosodic weakening supports the assumption that the effect of the π-gesture decreases with distance from its peak activation. $C1_{off}$, on the other hand, was not shortened in deaccented tokens but rather its displacement was reduced. This observed difference in displacement between accented and deaccented words is in line with the π-gesture approach, too, predicting articulatory strengthening at higher prosodic boundaries. Thus it appears that speakers use both less overlap and temporally lengthened segments to mark phrasal accent. According to the author's knowledge, the π-gesture approach has only been used for modeling the effects of prosodic boundaries. However, Saltzman et al. (2007) have already presented a concept for the implementation of the π-gesture on syllable level investigating the effects of stress. The results from the present study indicate that this approach needs to consider that accentuation does not necessarily cause gestural lengthening and less overlap. In fact, gestural shortening combined with less overlap can occur. Further, this approach needs to consider that not all syllable constituents are affected in the same way by (de)accentuation.

The third conclusion is that the shift towards the vowel – and thus the potential c-center organization of an onset cluster – was diminished in deaccented tokens. Thus the effects of prosodic weakening need to be included in theories modeling syllable structure such as Articulatory Phonology. It might be the case that the vowel-adjacent consonant $C2_{on}$ does not shift towards the vowel in deaccented words because otherwise (due to articulatory masking) segments would be lost in perceptual terms.

As far as gestural stability is concerned, the present study showed that all syllable constituents were affected by prosodic weakening. Especially the fact that the onset cluster exhibited more overlap in deaccented words than in accented words may be related to the attested diachronic loss of the onset cluster /kn/ in English (Luick, 1964) and the fact that /kn/ is rare in the world's languages (Ladefoged and Maddieson, 1996). That is, if /k/ and /n/ overlap to a greater extent in deaccented words, then it might be the case that a potentially earlier velum lowering acoustically hides the release of the preceding plosive.

This study showed that hypoarticulated speech, as in deaccented words, does not necessarily lead to more overlap or shortening of segments. In fact, the relationship between overlap and stiffness is not entirely clear: lower stiffness may co-occur with increased or decreased overlap. It may be the case that articulatory undershoot due to greater coarticulation at lower prosodic boundaries leads to hypoarticulated speech. Therefore, spatial parameters, such as tongue and jaw height might shed further light on articulatory differences in accented versus deaccented speech. For example, the tongue back might be lowered to a lesser extent during the production of the vowel in deaccented words. This will then result in articulatory vowel target undershoot. In addition, Harrington et al. (1995) and Beckman et al. (1992) attributed differences between accented and deaccented words to lesser jaw lowering or truncation.

Chapter 5

Gestural timing in first language acquisition

5.1 Introduction

This study is concerned with syllable organization in (pre-)school children using EMA. It investigates consonant-to-consonant timing as well as consonant-to-vowel timing with some emphasis on acoustic vowel duration in both syllable onsets and syllable codas. To do so, three different cluster types were included in onset position: /kn/, /kl/ and /pl/. Those clusters were shown to differ in their interconsonantal timing (Bombien et al., 2013). The clusters /kn/ and /kl/ were already included in chapter 3, revealing evidence for a global organization of /kn/ but not of /kl/. This finding was attributed to the fact that interconsonantal overlap was greater within /kl/ and this is why the vowel-adjacent consonant does not shift towards the vowel in the cluster condition. Also in coda position three different cluster types were included: /kn/, /kl/ and /ks/. These clusters were chosen because (1) two of them are comparable with onset clusters, (2) they involve the tongue-back and the tongue tip, i.e. they are all velar+alveolar clusters and (3) they are linked differently to the syllable: /kn/ and /kl/ are extrasyllabic whereas /ks/ is a syllabic coda cluster. Until now it was unclear how children coordinate complex onsets and codas with the vowel compared to simplex onsets and codas.

5.1.1 Cluster simplification and reduction

It has been shown that children are able to produce consonant clusters as early as the age of two (Lleó and Prinz, 1996) but there is a protracted development until adult-like mastery (Smit et al., 1990). For instance, their early attempts typically result in cluster reduction [bɛd] for "bread" (McLeod et al., 2001b) or in cluster simplification [bwɛd] for "bread" (McLeod et al., 2001b). With advancing age cluster simplification occurs more often, while cluster reduction occurs less frequent (Smit, 1993; Watson and Scukanec, 1997). A particular type of cluster reduction is based on the principles of Sonority Theory (Clements, 1990; Steriade, 1982). Children tend to reduce word-initial clusters to the segment of lower sonority, irrespective of its position within the cluster (see e.g. Barlow, 1997; Chin, 1996; Ohala, 1998, 1999; Pater and Barlow, 2003). Following a general sonority scale, an onset cluster such as /kl/ is typically reduced to /k/, rather than /l/. However, reduction to the segment of lower sonority is not obligatory: plosive-nasal clusters (/kn/, /gn/) were shown to be more likely reduced to /n/, whereas plosive-liquid clusters (/kl/, /gl/) are more likely to be reduced to /k/ (Fox, 2003; Lleó and Prinz, 1996; Ohala, 1999; Ott et al., 2006; Romonath, 1991). Other nonadult realizations of consonant clusters reported in the literature on cluster acquisition include, for example, epenthesis (e.g. [gəriːn] for "green") (Bortolini and Leonard, 1991; Dyson and Paden, 1983; McLeod et al., 2001a) and coalescence (e.g. [fɪm] for "swim") (Dyson and Paden, 1983).

5.1.2 Asymmetries in acquisition

Studies focusing on the acquisition of consonant clusters have shown that there are asymmetries in consonant cluster acquisition depending on syllable position, i.e. there is a difference in the acquisition of onset and coda clusters. Lleó and Prinz (1996) showed that German-speaking children aged 0;9-2;1 had acquired coda clusters several months before onset clusters and they produced generally more correct coda clusters than onset clusters. The same effect was found for Dutch (Levelt et al., 2000) and English (Templin, 1957) speaking children. These findings are surprising as singleton codas tend to be acquired after singleton onsets (e.g. Fikkert, 1994; Levelt et al., 2000; Salidis and Johnson, 1997; Vihman and Ferguson, 1987) and the CV syllable is the only universal syllable type in languages (Bell and Hooper, 1978). This is also reinforced by the *frame/content theory* (e.g. Davis and MacNeilage, 2004) showing evidence that the CV pattern is closely related to the closing-opening mouth movement during babbling.

Kirk and Demuth (2005) propose structural, morphological, frequency-based and articulatory reasons that may account for the differences in consonant cluster acquisition, which are briefly addressed in the following. One explanation they propose for the earlier acquisition of coda clusters is that these clusters may have a simpler structure and are thus easier to acquire. Consonant clusters in onset position are generally analyzed as a branching onset, whereas clusters in coda position are non-branching (Selkirk, 1982) and might according to Kirk and Demuth (2005) be easier to acquire. However, this account is only descriptive and not explanatory. It is not clear why onsets, which are just described to be 'branching', should be harder to acquire than 'non-branching' codas. Another explanation Kirk and Demuth (2005) propose is that only word-final clusters may contain important morphological information and this may draw children's focus to the ends of words. They assume that this attention may then lead to more accurate productions in children's speech. As the abovementioned asymmetries were observed for children acquiring English, German and Dutch, this assumption is plausible. However, the presence of morphemes in either coda or onset position is language-dependent. If this hypothesis is true, the opposite effect in cluster acquisition should emerge in children acquiring languages with morphemes in word-initial position. Furthermore, Kirk and Demuth (2005) claim that frequency may also affect the acquisition of consonant clusters: children acquire frequent syllable types used more often in child directed speech earlier (Levelt et al., 2000) and highly frequent words are produced with a greater accuracy (Berry and Eisenson, 1956; Leonard and Ritterman, 1971). However, according to our knowledge, an accurate production depends to a large extent on neighborhood density (cf. Wright, 2003): if neighborhood density is low, then – following Lindblom's (1990) H&H-Theory – highly frequent words do not have to be produced accurately. Concerning the earlier acquisition of frequent syllable types, Kirk and Demuth (2003) attributed the earlier acquisition of coda clusters in English-learning children to the fact that words containing coda clusters are much more frequent in English than words containing onset clusters. The last explanation proposed in Kirk and Demuth (2005) concerning the asymmetry in cluster acquisition is articulatory based. They assume that some sequences are easier to produce than others. For example, concerning onset clusters, some consonantal sequences are acquired before others (Smit et al., 1990; Templin, 1957). Thus it may be more difficult for a child to accurately produce the cluster /kn/ in onset position, as it requires a lowering of the velum to produce the nasal, compared with the cluster /kl/. Kirk and Demuth (2005) tested all these hypotheses with 1;5-2;7 year old children and concluded that an articulatory attempt best explains the asymmetries found in children's cluster acquisition as some clusters are easier to produce and thus to acquire. They came to this conclusion by comparing children's performance in /s/+stop sequences in onset position with children's performance in stop+/s/ and /s/+stop sequences in coda position. Their results have shown that acquisition patterns were quite different in /s/+stop sequences in onsets compared to stop+/s/ sequences in coda position, but they were similar in /s/+stop sequences in both onset and coda position, leading to the conclusion that articulatory factors, irrespective of phonetic context, play a role in explaining cluster acquisition asymmetries. Unfortunately, /s/+stop sequences were the only sequences included in both onset and coda position.

5.1.3 Physiological studies with children

Auditorily based transcription studies, which predominate in language acquisition research due to the subjects' age, can only capture stages of language acquisition and development that are perceptible to the adult transcriber (Scobbie, 1998). There is, however, accumulating evidence that articulatory gestures, which are present in speech production, can be perceptually masked at the surface due to the degree of overlap (cf. Goldstein et al., 2007b; Pouplier and Hardcastle, 2005, on speech errors). Thus it is important not to analyze only acoustic data but to include also articulatory data in language acquisition research. The children in the present study are older compared with those in the above-referenced studies focusing on syllable acquisition (i.e. they have already acquired both onset and coda clusters), but they still exhibit the greater variability in speech motor control and timing patterns compared with adults (Clark et al., 2001; Nittrouer et al., 2005; Riely and Smith, 2003) necessary for testing the predictions. These differences in speech motor control between children and adults were shown to last even until adolescence (Cheng et al., 2007; Walsh and Smith, 2002), which may be attributed to immaturity in the cortical networks involved in the planning and execution of speech (Smith, 2010) as well as to the fact that the motor control system must adjust to the craniofacial growth (Vorperian et al., 2005). All of these findings suggest that the development of speech motor control in (pre-)school children is not completed, although they produce words with consonant sequences correctly identified by adult listeners. That is, even if adults perceive younger children's clusters correctly, children nevertheless diverge from an adult-like realization of articulatory forms (see e.g. Hawkins, 1973, 1979).

5.1.4 Clusters in Articulatory Phonology

The abovementioned asymmetries in cluster acquisition suggest that the syllable plays an important role in language acquisition. A model which has a precise articulatory definition of the syllable is Articulatory Phonology (Browman and Goldstein, 1986, 1988, 1992). In this dynamic model, the timing of consonants with the adjacent vowel differs as a function of syllable position. Onset clusters are hypothesized to be globally timed with the following vowel, while coda clusters are said to be timed sequentially with the preceding vowel (Browman and Goldstein, 1988, 2000; Marin and Pouplier, 2010). As a consequence, the global organization of onset clusters is assumed to acoustically shorten the duration of the following vowel, whereas the local timing of coda clusters should not cause shortening of the preceding vowel (see section 3.1 for more details).

Further, adults' consonant-to-consonant timing also differs as a function of syllable position. Recent studies have shown that there is less gestural overlap within clusters in onset position compared to clusters in coda position (e.g. Chitoran et al., 2002; Pouplier, 2012). One explanation is perceptual recoverability requirements (Byrd, 1996; Chitoran et al., 2002): greater overlap between consonants in coda position is possible because there is recoverability on the basis of acoustic cues like formant transitions or release of the word-final consonant. However, onset clusters are not allowed to overlap to a greater extent because otherwise they will not be identified correctly, as no other acoustic cues are available (Marin, 2013) and word-initial phonetic detail is important for lexical access (Marslen-Wilson, 1987). Further, a greater gestural overlap between two adjacent consonants can lead to assimilation, or to consonantal loss in perception when one gesture hides another (Browman and Goldstein, 1990).

Proposed differences in timing patterns between onset and coda clusters both in their consonant-to-consonant timing and in their consonant-to-vowel timing may be a reason for the observed asymmetries in children's consonant cluster acquisition. In the coupled oscillator model (Nam and Saltzman, 2003), two modes are available in coupling of consonantal and vocalic gestures: in-phase and anti-phase. In-phase coupling is considered more dominant and more stable than anti-phase coupling (Goldstein et al., 2006; Nam et al., 2009), with which Goldstein et al. (2006) explain the fact that CV is a universal syllable and also acquired before the VC syllable. In complex onsets, such as CCVC, there is competitive coupling between the onset consonants. This is because these consonants are timed in-phase with the following vowel and anti-phase with each other. In complex codas, on the other hand, there is no such competitive coupling between the consonants, as only the vowel-adjacent coda consonant is timed anti-phase with the vowel and also anti-phase with the second coda consonant. Nam et al. (2009) demonstrate in their computer simulation that children may well produce the two consonants of an onset cluster, but that they are both initially coupled solely in-phase with the following vowel, because this coupling pattern is more dominant. As a consequence of the dominant in-phase coupling and a not yet stabilized anti-phase coupling between the two onset consonants, one gesture of the cluster will be perceptually masked. According to Nam et al. (2009) the sequential anti-phase consonant-to-consonant coupling in onset clusters needs to be acquired first so that both consonants are perceivable for adults. That is, the asymmetrical pattern observed in acquisition may arise from a perceptual masking due to different coupling mechanisms and increased gestural overlap (see also Goldstein et al., 2007b; Pouplier and Hardcastle, 2005). However, there is no indication at which age competitive coupling will be acquired. The present study specifically tests the assumption that children exhibit a tendency towards non-competitive coupling in complex onsets.

In addition to differences in consonant-to-consonant as well as in consonant-to-vowel timing in both onset and coda position, it has been shown that there are substantial differences in consonant-to-consonant timing depending on cluster composition in adults speech productions. For example, Bombien (2011), Bombien et al. (2013) and Hoole et al. (2009) have shown more overlap within the onset cluster /kl/ than within /kn/ in adult speakers of German and French. Also the onset cluster /pl/, despite allowing for articulatory independence, is somewhat less overlapped than /kl/. Bombien et al. (2013) assume that /kl/ is an articulatorily easy cluster, because only slight tongue tip rising is required to produce the lateral constriction for /l/ so that the stop release of /k/ is nonetheless recoverable. This is not the case in /kn/: an earlier velum lowering may hide the acoustic outcome of the plosive (Bombien et al., 2013; Hoole et al., 2009). In the present study, the clusters /kn, kl, pl/ produced by children are examined in order to investigate whether it can be assumed that /kl/ is a good and articulatorily easy cluster. If it is the case that /k/ and /l/ are well suited – as assumed by Bombien et al. (2013) – then children should show a similar timing behavior as adults.

To summarize, during the course of speech development children have to acquire the precise use of several articulatory parameters: timing differences depending on syllable position, including consonant-to-consonant timing as well as consonant-to-vowel timing, and timing differences depending on cluster composition. Those differences in timing patterns may be a reason for the observed asymmetries in children's consonant cluster acquisition.

The aim of this study is to investigate consonant-to-consonant and consonant-to-vowel timing including various cluster types in syllable onset and coda position, respectively, in children's productions whose speech development is not yet completed. Articulatory data of four (pre-)school children acquiring German are reported.

5.2 Hypotheses

Based on the findings described in Bombien et al. (2013) and on the fact that (pre-)school children produce perceivable onset clusters, the following prediction concerning specific cluster types is made:

H1 There is less overlap within the two onset clusters /kn/ and /pl/ than within the onset cluster /kl/.

Based on the results from a computer simulation by Nam et al. (2009), the following prediction about consonant-to-vowel timing in onset position is made:

H2 Children should exhibit a tendency towards non-competitive coupling in onset clusters.

H3 Onset clusters overlap to a lesser degree than coda clusters.

H4 Children produce coda clusters with a non-competitive coupling.

5.3 Method

5.3.1 Speech material and experimental set-up

The words included in this experiment were real and picturable words containing simplex and complex onsets and codas, respectively. We did not use the mirror-image of onset clusters in coda position as it is usually the case in c-center studies focusing on syllable organization in speech production. The goal of this chapter was instead to further investigate Kirk and Demuth's (2005) conclusion of an articulatory explanation. Therefore the attempt was made to include the same clusters in syllable onset position as in syllable coda position. Therefore, as far as it was possible, the same clusters in syllable-initial as in syllable-final position were included. A further issue was to focus on consonant-to-consonant timing in onset and coda position and this is also only possible by comparing the same clusters. In addition, clusters found to be coordinated differently in adult productions were included to see whether children might show similar coordination patterns. However, this method limits the eligible clusters. The onset clusters used in this study were /kn/, /kl/, /pl/. Simplex onsets consisted of the alveolars /n/ and /l/. The coda clusters contained /kn/, /kl/ and /ks/, simplex codas therefore contained /k/. It has to be noted, however, that the German clusters /kn/ and /kl/ in word-final position emerge from schwa-deletion (e.g. "Dackel" /dakəl/ → /dakl/). These consonant combinations, thus, may not be attributed to the coda clusters but instead to be extra-syllabic (/dakl̩/). However, it was not possible to use other clusters in syllable onset as well as in syllable coda because of three reasons: (1) there are no other clusters occurring in both onset and coda position in German; (2) articulatory measurements, such as overlap within clusters, are only possible when different places of articulation are present; (3) no information – except for /kl/ and /kn/ – on good vs. bad clusters. Further, whether a cluster shows a global or sequential timing with the vowel does not seem to be an appropriate measure for determining its syllable affiliation, because of different implications for onsets und codas as well as for some onsets (see Marin, 2013; Pouplier and Beňuš, 2011). Using the clusters /kn, kl/ in word-final position provides the possibility for examining the timing behavior of possibly extra-syllabic codas. Therefore, the cluster /ks/, which is considered to be syllabic, was additionally included in coda position in order to investigate whether there are timing differences. The cluster /ks/ was not included in onset position and the cluster /pl/ was not included in coda position, as they are rare and uncommon at these positions in German phonotactics (Wiese, 1996). Three repetitions of each target word listed in Table 5.1 were obtained using a picture naming task.

Onset		Coda	
simplex	**complex**	**simplex**	**complex**
/**no**ːte/	/**kno**ːtn/	/lɔ**k**/	/lɔ**kŋ**/
/**l**ak/	/**kl**ak/	/sa**k**/	/da**kl**/
/	/	/gəbɛ**k**/	/gəvɛ**ks**/
/**l**ast/	/**pl**ats/	/	/

Table 5.1: *Target words containing either complex (left column each) or simplex (right column each) separately for Onset (left) and Coda (right).*

All target words were embedded between the words /pɪtsa/ (*pizza*) and /apfəl/ (*apple*) and were always accented. This /a/-context allows for optimal tracking of the articulators' movements because of the active tongue back movement. Prior to the recording, children were first familiarized with the recording procedure and with the pictures they had to name. In order to keep it interesting for children, different pictures across repetition were used. The pictures remained the same during one block but were different across repetition. This means that during one block only picture of the target word changed, so that the children had only to focus on this picture. This procedure is illustrated in Figure 5.1. After every repetition, the children saw a picture of a treasure map and a pirate getting closer until he reached the treasure at the end of the experiment. Thus the children knew how much of the experiment they already passed.

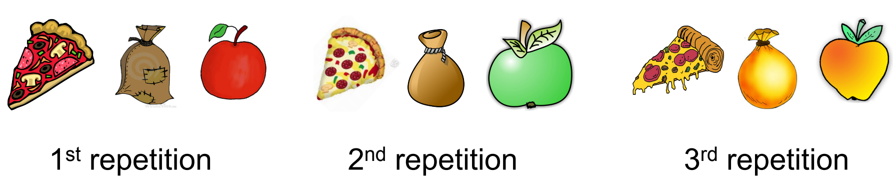


Figure 5.1: *Examples of picture naming task used in the present study including the target word "Sack".*

5.3.2 Acoustic and articulatory recordings

Articulatory movement data were collected in a sound-attenuated booth using the 3D Electromagnetic Articulography (EMA) system (AG501, Carstens Medizinelektronik) at the Munich Institute of Phonetics and Speech Processing (Hoole et al., 2003) in order to measure positions from sensors glued to various points on the subject's vocal tract. The articulatory data were sampled at a frequency of 250 Hz. Audio-data was simultaneously collected with a shotgun microphone at a sampling rate of 25.6 kHz.

In total, only seven sensors were attached to various parts on the children's heads keeping the gluing procedure as short as possible for the children's comfort. Because children's tongues are smaller than adults' tongues (Vorperian et al., 2005), only two sensors were placed on the tongue: on the tongue tip (attached approximately 1 cm behind the actual tongue tip) and on the tongue back (attached on the estimated velar constriction region). Additional sensors were placed on the upper and lower lips in order to measure lip aperture (LA). Reference sensors were placed on the nose bridge and behind the ears (on the right and left mastoid process) in order to correct for head movements. All sensors, except those on the right and left mastoid process were fixed midsagittally.

5.3.3 Subjects

Data from eight typically developing (pre-)school children aged 5 years 10 months to 7 years 5 months were recorded. None of the children reported any speaking or hearing disorders and they were all L1 speakers of Standard Southern German. Since only one layer of glue was used, some of the sensors did not last the whole recording session. Thus it was able to collect articulatory data for only four (aged 5 years and 11 months to 7 years and 1 month) out of those eight children. Therefore, in the present study only articulatory and acoustic results from those four children will be reported.

5.3.4 Acoustic and articulatory segmentation

The segmentation of the speech signals was first done automatically using MAUS (Schiel, 2004). If necessary, the segment boundaries of the target words were then hand-corrected in Praat (Boersma and Weenink, 1992). Stops were acoustically divided into closure phase and aspiration. As the onsets always contained the lateral /l/ or the nasal /n/, the beginning of the vowel was set at the point of spectral change in F2. The articulatory segmentation was done in EMU (Harrington, 2010). The sensors included in the analysis were those attached to the tongue tip (for the alveolars /n/, /l/ and /s/), to the tongue back (for the velar /k/) and to the upper and lower lips (for the labial /p/). To label articulatory gestures, the vertical movement for velars and the tangential velocity (i.e. the sum of the vertical and the horizontal tongue movement) for alveolars was used. For the labial, the lip aperture was calculated as the Euclidean distance between upper and lower lips (see chapter 3 for more details).

5.3.5 Measurements of articulatory timing

The present study focuses on temporal articulatory measures, i.e. vertical and horizontal movements of the relevant sensors at specific moments in time. Children's timing in syllable onset and coda position was analyzed in two ways:

1. consonant-to-consonant timing (CC_{lag}) in complex onsets and codas
2. consonant-to-vowel timing in complex vs. simplex onsets (CV_{lag}) and codas (VC_{lag})

Although /kl/ and /kn/ were included in both word-initial and word-final position, the cluster /kn/ had to be excluded from the analysis on consonant-to-consonant timing in word-final position because /kn/ was almost always assimilated to /kŋ/. To investigate children's coordination within a cluster in both onset and coda position, the lag between the two consonantal gestures' plateaus was calculated. Plateau onsets (P_{on}) and offsets (P_{off}) were defined on the basis of changes in the articulator's velocity (see chapter 3 for more information). This CC_{lag} was calculated by subtracting the initial consonant's plateau offset by the second consonant's plateau onset, as described in the following equation 5.1:

$$CC_{lag} = P_{on}[C2] - P_{off}[C1] \tag{5.1}$$

In order to investigate children's consonant-to-vowel timing depending on onset and coda complexity, the lag between the vowel-adjacent consonant and the vowel was calculated. CV_{lag} was measured between the consonant's plateau offset and the acoustic vowel midpoint for each token (see e.g. Pouplier, 2012), as described in equation 5.2.

$$CV_{lag} = Mid_{vowel} - P_{off}[C2_{on}] \tag{5.2}$$

The same procedure was applied to the VC_{lag} measure with the exception of subtracting the coda consonants plateau onset by the acoustic vowel midpoint. Figure 5.2 schematically illustrates the measurement applied for CV_{lag} and VC_{lag}.

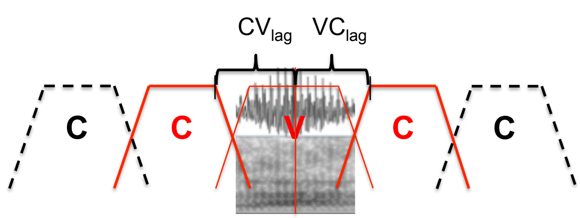


Figure 5.2: *Schematic illustration of measurement procedure used for CV_{lag} and VC_{lag}.*

In a second step, the singleton lag was divided by the corresponding cluster lag for each speaker and item resulting in lag ratios. Ratios above 1 indicate a shift towards the vowel, whereas values below 1 indicate a shift away from the vowel. As these lag ratios contain information about simplex and complex tokens, the terms (k)l-, (k)n- and (p)l-onsets will be used in the following when referring to CV_{lag} and the terms k(l)-, k(n)- and k(s)-codas in the following when referring to VC_{lag}. Further, the hypothesis for greater interconsonantal overlap in word-final position compared with word-initial position had to be tested. Although /kl/ and /kn/ were elicited in both word-initial and word-final position, the cluster /kn/ had to be excluded from the analysis because /kn/ was always assimilated to /kŋ/, which is a common assimilation process in German (Jun, 2004; Kohler, 1990). In order to establish the difference in consonant-to-consonant timing between onset and coda /kl/, first the CC_{lag} between /k/ and /l/ was calculated as described in equation (3) in both syllable positions. In a second step, mean values per speaker and syllable position were calculated. The mean values of onset /kl/ were then subtracted by the mean values of coda /kl/ to see whether this cluster is timed differently depending on syllable position. The same procedure was applied to calculate the overlap difference between /kl/ and /ks/ in coda position, with subtracting the mean values of /kl/ by the mean values of /ks/.

5.4 Statistics

For the statistical analyses mixed models were conducted. CC_{lag}, CV_{lag}, VC_{lag} and the acoustic vowel duration each served as the dependent variable. Onset or coda complexity (two levels each: simplex vs. complex) as well as cluster set were the fixed factors. The speaker was entered as random factor. The models were implemented using the lme4 package (Bates, 2008) in the statistical environment R. For differences in consonant-to-consonant timing depending on syllable position a one-sample t-test was conducted.

5.5 Results

This section will report both consonant-to-consonant and consonant-to-vowel timing in syllable onset and syllable coda position. First, the results for onset clusters will be reported followed by the results for coda clusters.

5.5.1 Onset

Consonant-to-consonant timing

As far as CC_{lag} in onset clusters is concerned, Figure 5.3 shows a clear difference between the clusters /kl/ and /kn/ and between /kl/ and /pl/. The predominantly negative values in the cluster /kl/ suggest that plateau offset of $C1_{on}$ and plateau onset of $C2_{on}$ did overlap, whereas the predominantly positive values in the clusters /kn/ and /pl/ indicate that there was a lag between the two plateaus.

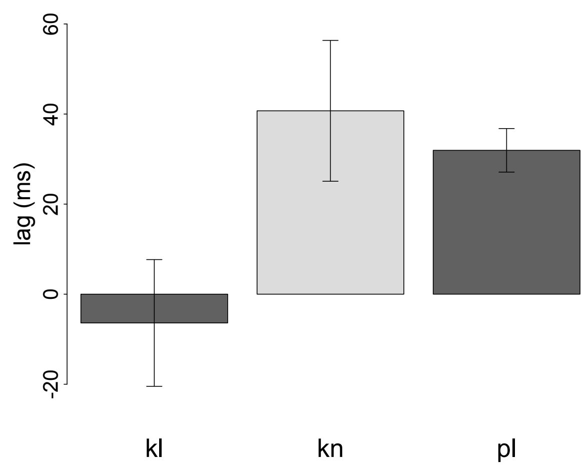


Figure 5.3: *Consonant-to-consonant timing within the three different onset clusters /kl/, /kn/ and /pl/.*

The results of a mixed model revealed a significant effect of onset cluster set ($\chi^2[1] = 8.5$, $p<0.05$). As the fixed factors contained three levels, post-hoc Tukey test were conducted, revealing a significant difference between /kn/ and /kl/ ($p<0.05$). This greater overlap within the cluster /kl/ compared to the cluster /kn/ is in line with previous findings in adult speech productions (Bombien, 2011; Hoole et al., 2009). Further, the model revealed only a tendency towards a difference between /kl/ and /pl/ ($p<0.10$), although the onset cluster /kl/ exhibits a greater amount of overlap, whereas the onset cluster /pl/ overlaps to a lesser degree (see Figure 5.3). Commensurate with Figure 5.3, the clusters /kn/ and /pl/ did not significally differ in their interconsonantal timing.

Consonant-to-vowel timing

In addition to consonant-to-consonant timing in onset clusters, children's consonant-to-vowel timing depending on onset complexity was analyzed.

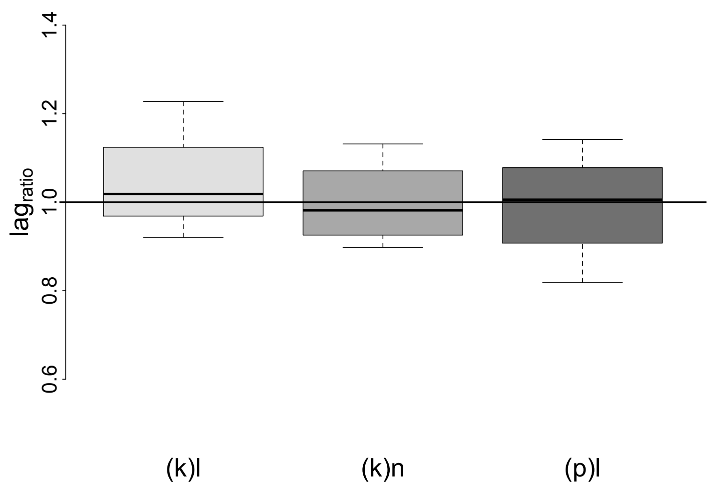


Figure 5.4: *Lag ratios separately for the three different onset cluster sets (k)l, (k)n and (p)l.*

Recall that the predictions made within the coupled oscillator imply a competitive coupling between the onset consonants and the vowel in syllables containing complex onsets. If this is the case, calculated lag ratios between simplex and complex onsets should lie above 1. It was expected that children should exhibit a tendency towards non-competitive coupling. Therefore, only small differences between /kn/, /pl/ and /kl/ concerning their coupling with the vowel should be observable in children's productions, and the calculated lag ratios between simplex and complex tokens should be around 1.

As far as this particular timing between the vowel-adjacent onset consonant with the following vowel (CV_{lag}) in words containing simplex and complex onsets is concerned, Figure 5.4 shows the lag ratios calculated for simplex and complex onset tokens separately for the three different onset cluster sets. Lag ratios lie around 1, indicating that there is no shift of the vowel-adjacent consonant towards the vowel as the number of onset consonants increased. There seem to be no timing differences between the three different onset clusters.

In line with this observation, a mixed model with CV_{lag} as the dependent variable and fixed factors onset complexity (simplex vs. complex) and onset cluster set (/kl/ vs. /kn/ vs. /pl/) revealed no significant main effects and no interaction. This indicates that onset clusters in children's productions are not globally timed with the vowel. There was therefore non-competitive coupling in all onset clusters produced by pre-school children. There was a significant effect of onset cluster set ($\chi^2[2] = 27.0$, $p < 0.001$). This can be explained by the fact that k(n) tokens contained the tense vowel /oː/, while the other tokens the contained lax vowel /a/. Therefore, the time interval from $C2_{on}$ to the acoustic vowel midpoint was longer on all k(n) tokens than in k(l) or p(l) tokens. Post-hoc Tukey tests revealed significant differences between complex k(n) and complex (kl) tokens ($p < 0.5$) as well as between simplex k(n) and simplex k(l) tokens ($p < 0.05$).

Acoustic vowel duration

Having shown that complex onsets exhibit the same timing behavior with the following vowel as simplex onsets (i.e. there was no shift towards the vowel in complex onset tokens compared with simplex onset tokens), an increasing number of onset consonants should, therefore, not influence the acoustic vowel duration. Figure 5.5 shows that there seem to be no differences in acoustic vowel duration between simplex and complex onset tokens, despite a trend towards shorter vowels in complex onset tokens compared with simplex onset tokens in (k)l-onsets. Further, it shows longer vowel durations in (k)n-onsets because this word pair contained the tense vowel /oː/.

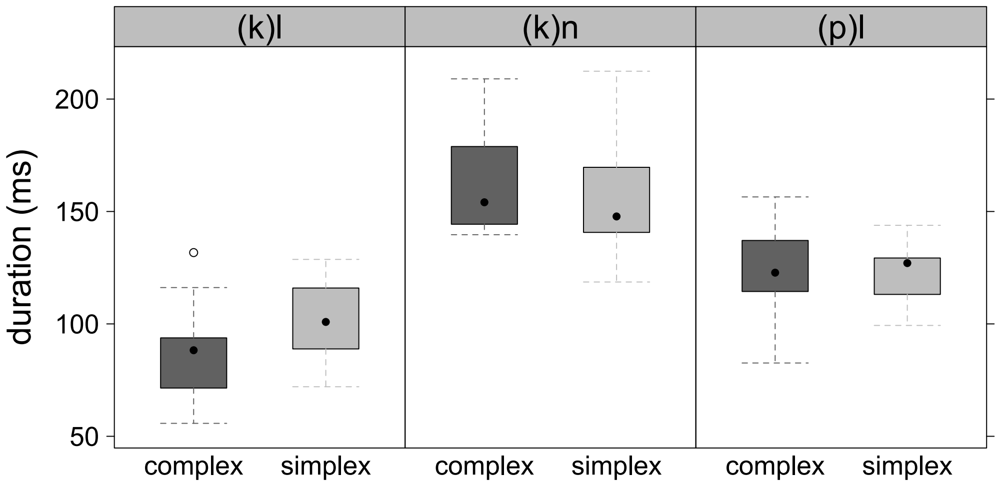


Figure 5.5: *Acoustic vowel durations in words containing complex onsets (dark gray) and simplex onsets (light gray) separately for the three different onset cluster sets (k)l, (k)n and (p)l.*

This observation was confirmed by a mixed model with the same fixed factors as above: there was no significant effect of onset complexity on the acoustic vowel duration. That is, in line with the previous finding that complex onsets show the same timing with the vowel as simplex onsets, there were also no differences in acoustic vowel duration. There was a significant effect of onset cluster set ($\chi^2[2] = 77.3$, $p<0.001$). Post-hoc Tukey tests showed longer vowel durations in (k)n onsets than in p(l) and k(l) onsets ($p< 0.01$), irrespective of onset complexity, as well as longer vowel durations in complex p(l) onsets than in complex k(l) onsets ($p< 0.01$).

5.5.2 Coda

Consonant-to-consonant timing

A further issue was to compare the degree of interconsonantal overlap within a cluster in either onset or coda position. It was assumed that a allowed greater degree of overlap within clusters in syllable-final position, which are nonetheless perceivable by the listeners, may be a reason for the reported earlier acquisition of coda clusters compared with onset clusters. The only cluster set accounting for different overlap patterns in syllable onset position compared with syllable coda position is /kl/. Figure 5.6 shows that there seem to be slight differences in interconsonantal overlap between syllable positions: /kl/ in coda position seems to overlap less than /kl/ in onset position as the difference is mainly positive. However, a one-sample t-test revealed that the duration difference is not significant, which can be explained by the fact that the median lies below zero.

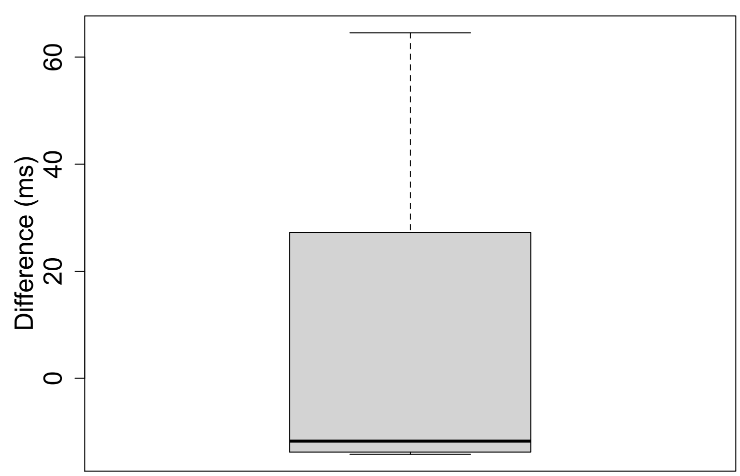


Figure 5.6: *Duration difference in* CC_{lag} *between onset and coda /kl/.*

Consonant-to-vowel timing

A further assumption of the coupled oscillator model is that coda clusters are generally organized in an anti-phase alignment with the preceding vowel and thus timed with a non-competitive coupling. There are no reasons to assume anything different for children, as there is no dominating in-phase coupling. This then implies that lag ratios should lie around 1, because no timing differenced between simplex and complex codas are expected within this model. Concerning this timing relation between the vowel-adjacent coda consonant and the preceding vowel (VC_{lag}) in words containing simplex and complex codas, Figure 5.7 shows that lag ratios for all three cluster sets lie slightly below 1. This indicates that the vowel-adjacent coda consonant shifts somewhat away from the vowel as the number of consonants in syllable-final position increases. For the two potentially extrasyllabic clusters /kl/ and /kn/ this effect seems to be more pronounced than for the syllabic coda cluster /ks/, but there seem to be no great differences.

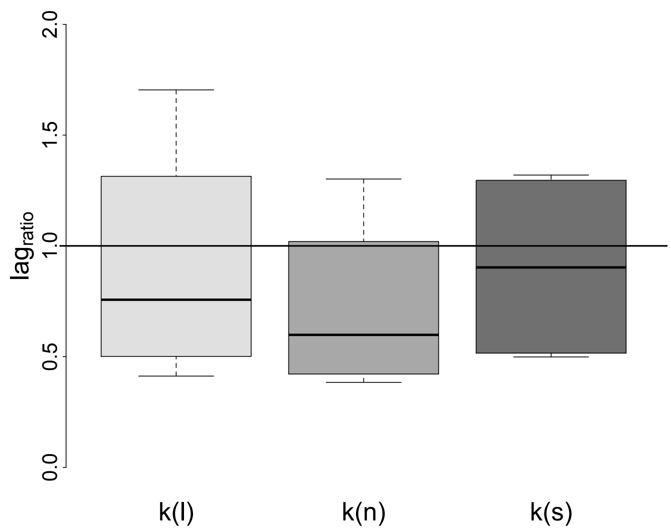


Figure 5.7: *Lag ratios separately for the three different coda cluster sets k(l), k(n) and k(s).*

Consistently with the observation in Figure 5.7, a mixed model with VC_{lag} as dependent variable and fixed factors coda complexity (simplex vs. complex) and coda cluster set (/kl/ vs. /kn/ vs. /ks/) revealed a significant effect of coda complexity ($\chi^2[1] = 10.2$, p<0.01) and of coda cluster set ($\chi^2[2] = 21.2$, p<0.01). Post-hoc Tukey tests revealed significant differences between simplex and complex codas only for k(n) tokens (p<0.05). There was thus a shift away from the vowel in k(n) tokens. Despite the above-mentioned trend, there was no significant difference in VC_{lag} between simplex and complex codas for k(l) and k(s) tokens. From this it can be deduced that the cluster sets k(l) and k(s) are sequentially organized, whereas in the cluster set k(n) the vowel-adjacent coda consonant shifts away from the vowel.

Acoustic vowel duration

As mentioned in section 5.1 coda clusters are not supposed to influence the duration of the vowel by the predictions of Articulatory Phonology. This is because no shift towards the vowel is expected in complex codas compared with simplex codas. As there was no shift towards the vowel in k(l) and k(s) codas, one would therefore expect no vowel shortening in these word pairs. For k(n) codas on the other hand there was a shift away from the vowel. Having extrasyllabic codas in k(l) and k(n), monosyllabic words become disyllabic with an additional coda consonant. Thus there may be polysyllabic shortening, in which the vowel is shorter in polysyllabic than in monosyllabic words (see chapter 2 for more details on polysyllabic shortening) (Klatt, 1973; Lehiste, 1970; Lindblom and Rapp, 1973; Noteboom, 1972; Port, 1981).

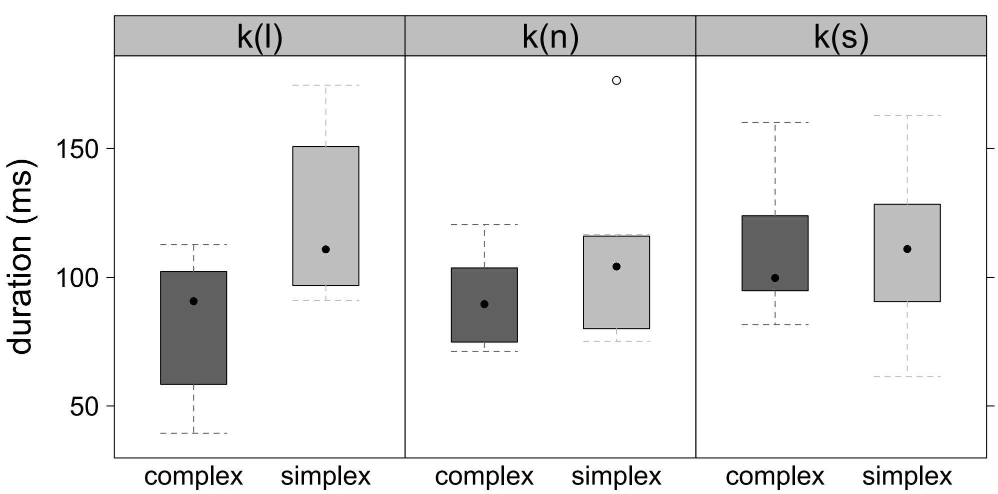


Figure 5.8: *Acoustic vowel durations in words containing complex (dark gray) and simplex (light gray) codas separately for the three different coda cluster sets k(l), k(n) and k(s).*

It is clearly observable from Figure 5.8 that in k(l) codas there are marked differences in vowel duration in that way that vowels are shorter in complex codas than in simplex codas. Such a tendency is also observable in k(n) codas but there seems to be no difference in vowel duration depending on coda complexity in k(s) codas.

Consistently with this observation, a mixed model with the acoustic vowel duration as dependent variable and fixed factors coda complexity (simplex vs. complex) and coda cluster set (/kl/ vs. /kn/ vs. /ks/) revealed a significant interaction between these two factors ($\chi^2[2] = 10.7$, $p<0.01$). Post-hoc Turkey tests showed significant vowel shortening only in k(l) codas ($p<0.01$). There was no vowel shortening in k(n) codas and k(s) codas. This means, that there was polysyllabic shortening in k(l) codas, whereas no such shortening effect could be found in k(n) codas, although both clusters are considered to be extrasyllabic.

5.6 Summary and discussion

There were several results from this study. First the results concerning children's articulatory timing in syllable onsets will be reported in section 5.6.1. This involves consonant-to-consonant timing and consonant-to-vowel timing as well as acoustic vowel duration. In section 5.6.1 the results of timing patterns in syllable codas will be reported. This includes again consonant-to-consonant timing and consonant-to-vowel timing as well as acoustic vowel duration.

5.6.1 Onset

Consonant-to-consonant timing

The analysis on consonant-to-consonant timing in syllable onset position has shown that there were marked differences within the onset clusters included: children produced more overlap in /kl/ compared to /kn/ and /pl/, confirming hypothesis **H1**. This is in line with results reported in Bombien (2011), Bombien et al. (2013) and Hoole et al. (2009) who reported these timing differences for adult speakers of German and French. Thus this pattern seems not only to be language independent, but it also emerges early during language acquisition. Perceptual recoverability (Byrd, 1996; Chitoran et al., 2002; Silverman, 1997) best accounts for the observed differences between /kl/ and /kn/. These two clusters seem to be very similar as both include the tongue back followed by the tongue tip, except for the lowering of the velum producing the nasal. In order to perceive the cluster /kn/ as such, the velar lowering for the nasal is not allowed to overlap with the release of the plosive, because otherwise the plosive is not perceivable anymore (see also Hoole et al., 2009). Therefore, the degree of overlap may be reduced. An articulatory account may explain the observed differences between /kl/ and /pl/ best. The greater articulatory independence for /pl/ would allow for more overlap compared to /kl/ because the tongue tip is freely moveable (see Bombien, 2011). However, this was not the case. In fact, the cluster /pl/ overlapped to a lesser extent than /kl/. Bombien (2011) argued that the observed greater overlap in /kl/ is due to the strong articulatory interdependence causing a very narrow window of overlap.

Consonant-to-vowel timing

The results on consonant-to-vowel timing in words containing simplex and complex onsets revealed that there was no shift towards the vowel in complex onset tokens. Therefore, it can be concluded that the clusters included in the present analysis did not show a global organization and thus a potential c-center, confirming hypothesis **H2**. This finding might explain why onset clusters are said to be acquired after coda clusters.

In onset clusters in children's productions only the vowel-adjacent consonant seems to be in an in-phase coupling with the vowel and the two onset consonants are coupled anti-phase with each other. In onset clusters as proposed by the coupled oscillator model, however, both onset consonants are in an in-phase relationship with the vowel and in an anti-phase relationship with each other. The absence of in-phase relationship of C1 with the vowel – as observed in children's productions – results in a right-edge alignment of an onset cluster with the following vowel rather than a c-center alignment. Such a temporal pattern, i.e. right-edge alignment instead of c-center alignment, is similar to that reported for "impure-s" clusters in Italian (Hermes et al., 2008) as well as for onset clusters in Moroccan Arabic (Shaw et al., 2009, 2011) and Tashlhiyt Berber (Goldstein et al., 2007a; Hermes et al., 2011). In these languages, the syllable affiliation of onsets is controversial and whether or not an onset cluster shows a competitive coupling relationship with the vowel has been used to predict its syllable affiliation. In all of these studies, an observed non-competitive coupling of onset clusters has been taken as empirical evidence that these consonant sequences are a series of simplex consonants rather than complex onsets or an additional syllable with a consonantal nucleus. However, unlike the consonant sequences in these languages, the clusters included in the present study are described as complex onsets without any doubt. In addition, chapter 3 revealed evidence for a global organization of the onset cluster /kn/ in adult productions. Further, it has to be considered that non-competitive has also been reported for complex onsets in Slovak (Pouplier and Beňuš, 2011) and for some onset clusters of Romanian (Marin, 2013), two languages where word-initial consonant sequences are indisputably assumed to form complex onsets. In the study by Marin (2013) manner seems to affect the timing of onset consonants with the following vowel. For example, she revealed evidence for a global organization of /sp/, /sk/, /sm/ onsets but not for /ps/, /ks/, /kt/ and /kn/ onsets. Therefore, it was suggested by Marin (2013), Pouplier (2012) and Brunner et al. (2014) that interconsonantal timing might affect the shift towards the vowel. Also in chapter 3 it was assumed that the greater overlap between /k/ and /l/ leads to less overlap between /l/ and the following vowel. However, in the present study, the onset clusters /kl/ and /kn/ differed in their intragestural timing (just like for adults reported in chapter 3), but there were no differences in consonant-to-vowel timing. Thus – unlike in adults' productions – differences in consonant-to-consonant timing seem not to affect consonant-to-vowel timing in children's productions.

An assumption made by Articulatory Phonology is that a shift towards the vowel would shorten it acoustically. Thus because no such a shift was present, no vowel shortening was expected in children's productions. Acoustic analyses confirmed this as there was no vowel shortening caused by an increasing number of onset consonants in none of the onset clusters included.

5.6.2 Coda

Consonant-to-consonant timing

Consonant clusters in coda position should show a greater amount of overlap than consonant clusters in onset position. Less overlap in onset clusters was expected because of perceptual recoverability requirements (Chitoran et al., 2002). Therefore, more overlap within the cluster /kl/ was hypothesized in coda position than in onset position (**H3**). However, this was not the case as there was nearly the same amount of overlap in both syllable positions. This finding is in line with Marin (2013), reporting no differences in interconsonantal timing depending on syllable position. We agree with her conclusion that gestural overlap may depend more on the consonants' identification rather than on syllable position, because to correctly identify the cluster /kl/ as such, the amount of overlap is limited. Given the anyway great amount of overlap within this cluster, it might be the case that even more overlap (in any syllable position) might prevent a correct identification of both cluster components (see also Byrd and Tan, 1996).

Consonant-to-vowel timing

Concerning consonant-to-vowel timing in syllable coda position, this study has shown that in k(n) codas the vowel-adjacent coda consonant shifts away from the vowel as the number of coda consonants increased. In the cluster sets k(l) and k(s), on the other hand, there was no difference concerning consonant-to-vowel timing in clusters compared to singletons. Thus complex codas are sequentially organized in pre-school children's productions. This finding is in line with the predictions made by Articulatory Phonology predicting sequential organization of coda clusters and confirms hypothesis **H4**. It is interesting that k(l) an k(n) codas differ in their timing behavior with the preceding vowel. Both clusters (/kl/ and /kn/) are expected to be extrasyllabic codas (Kohler, 1990) but this cannot explain why there is a shift away from the vowel only in k(n). It might be the case that the assimilation process of syllable-final /kn/ to /kŋ/ (Kohler, 1990) leads to a different alignment and thus to the observed shift away from the vowel.

As far as the acoustic vowel duration is concerned, no vowel shortening in k(s) codas was expected because of the observed sequential timing. Results confirmed this: there were no differences in vowel duration depending on coda complexity. This finding is in line with the predictions made by Articulatory Phonology, but it contradicts studies reporting coda-driven vowel shortening in adult productions (see also chapter 3 for evidences on coda-driven vowel shortening; Katz, 2012; Munhall et al., 1992; Shaiman, 2001; Waals, 1992). For k(l) and k(n) codas another type of vowel shortening, namely polysyllabic shortening, needed to be considered. The acoustic results showed that there was indeed vowel shortening but only in k(l) codas and only by trend in k(n) codas. This cannot be explained by syllable timing: the vowel-adjacent coda consonant was sequentially organized with the preceding vowel in k(l), but there was a shift away from the vowel in k(n). Therefore, from an articulatory point of view, no polysyllabic vowel shortening should have been observed. This finding rather implies that also vowel shortening in children's productions is independent of syllable organization (see also Byrd, 1995). It further supports the idea that vowel shortening can occur in syllables showing a sequential alignment, like in chapter 3. The finding that children show polysyllabic shortening only in one word-pair is not in line with results reported in studies focusing on polysyllabic shortening in adult productions. This effect has been well documented for various languages and for various word-pairs (Klatt, 1973, 1976; Lehiste, 1970; Lindblom and Rapp, 1973; Noteboom, 1972; Port, 1981). Thus the findings of the present study imply to some extent that children have less precise control on vowel durations than adults.

5.7 Conclusion

In the present study articulatory data was successfully acquired for children by means of electromagnetic articulography (EMA). The recordings monitored the upper and lower lips, the tongue-tip and the tongue-back movements. We investigated the timing patterns during the production of complex and simplex onsets and codas, respectively.

The goal of this study was to investigate children's coordination patterns including different types of clusters with respect to syllable position. This was based on findings by Kirk and Demuth (2005) supporting evidence that an articulatory attempt best explains the reported asymmetries in consonant cluster acquisition. To do so, previously documented differences in consonant-to-consonant timing as well as in consonant-to-vowel timing on both syllable onset and syllable coda position were taken into account as being possible reasons leading to the observed asymmetries in consonant cluster acquisition. As mentioned in section 5.1, Nam et al. (2009) assume that in the production of a CV syllable, the in-phase coupling is stronger or easier to produce than the anti-phase coupled VC syllable. They contribute this to the fact that CV syllables are acquired before VC syllables. They showed in their study that CC coupling occurs earlier in the VCC simulation than in the CCV simulation. From this they assume that children produce consonants in onset position synchronously because there is dominant in-phase coupling of the CV syllable leading to the fact that consonant sequences will not be perceptible. Therefore, the sequential consonant-to-consonant coupling needs to be acquired (Nam et al., 2009). It was suggested that the competitive coupling in complex onsets – according to the coupled oscillator model – may be a reason for the later acquisition of onset consonants. And indeed, there was no evidence in our data, that children produce onset clusters as hypothesized by a competitive coupling model or by Articulatory Phonology. Instead, the present study showed that there is sequential timing on both onset and coda position. The finding for sequential timing of onset clusters contradicts Articulatory Phonology and differs from adult productions (see e.g. Pouplier, 2012, and also chapter 3 showing a global organization of the onset cluster /kn/). This leads to the conclusion that a global organization of onset clusters may be acquired later during adolescence.

Further, the present study showed that children seem to acquire adult like consonant-to-consonant timing before they acquire adult like consonant-to-vowel timing in onset position. This supports Nam et al.'s (2009) conclusion that children need to stabilize the sequential consonant-to-consonant coupling in order to produce clusters that are perceived as such. If children coordinate consonants and vowels with an in-phase coupling without having stabilized anti-phase consonant-to-consonant timing, this then will result in synchronous productions leading to perceptual loss of one onset consonant (see Nam et al., 2009). Thus children during language acquisition may stretch the whole CCV sequence into separate constituents in order to produce perceivable onset clusters, resulting in the sequential timing patterns observed in the present study. This may be the reason why coda clusters in general are acquired first, because in coda clusters there is no additional coupling and consonants are allowed to overlap to a greater extent being nonetheless perceivable. In addition, such a stretching of the CCV sequence may explain the fact that young children do show nonadult realisations of onset clusters, such as epenthesis (Bortolini and Leonard, 1991; Dyson and Paden, 1983; McLeod et al., 2001a). When children get older and the motor control system reached adult sizes (Fitch and Giedd, 1999; Kent and Vorperian, 1995; Lieberman et al., 2001; Vorperian, 2000; Vorperian et al., 1999), then they might be able to couple both onset consonants in-phase with the following vowel.

To conclude, the results have shown that children seem to prefer sequential or anti-phase timing in both onset and coda position. However, this conclusion needs to be substantiated with adult data producing the same words embedded in the same experimental set-up used here. Recent studies have shown that the proposed c-center organization is not a universal pattern holding for all types of onset clusters cross-linguistically. To the contrary, the question whether a onset cluster shows a c-center organization or not depends on cluster composition (see e.g. Marin, 2013). It might be interesting to see, whether clusters that generally do not exhibit c-center organization are also reported to be acquired before clusters that exhibit c-center organization.

Chapter 6

Summary and conclusion

Previous studies focusing on articulatory timing patterns reported cluster specific differences concerning both consonant-to-consonant timing as well as consonant-to-vowel timing. For example, Bombien (2011), Bombien et al. (2013) and Hoole et al. (2009) reported differences for German and French in consonant-to-consonant timing for the onset clusters /kl/ and /kn/: there was more overlap between /k/ and /l/ than between /k/ and /n/. Studies focusing on consonant-to-vowel timing typically aim to revise the predictions made by Articulatory Phonology (e.g. Browman and Goldstein, 1988, 1990, 2000). According to this theory, onset clusters are organized around their c-center, i.e. the vowel adjacent consonant shifts towards the vowel resulting in increased CV overlap and shorter acoustic vowel durations as opposed to simplex onsets (Marin and Pouplier, 2010; Pouplier, 2012). However, recent studies showed that some onset clusters do exhibit a c-center organization while others do not (see e.g. Brunner et al., 2014; Marin, 2013; Marin and Pouplier, 2010). The composition of the cluster as well as its consonant-to-consonant timing was considered to influence the global organization of onset clusters with the following vowel (Marin, 2013). For coda clusters, Articulatory Phonology predicts a sequential alignment with the preceding vowel. Thus in complex codas the same amount of VC overlap as in simplex codas is expected. Consequently, the acoustic vowel duration should remain unaffected. However, the proposed sequential organization also might be influenced by the composition of the cluster. For example, Marin and Pouplier (2010) showed c-center organization of /lp/ and /lk/ codas in English and Byrd (1995) showed that at least some speakers produced coda clusters with c-center organization. In addition, evidence for acoustic vowel shortening caused by complex codas (e.g. Katz, 2012; Munhall et al., 1992; Shaiman, 2001) is not in line with the predictions made by Articulatory Phonology.

The aim of this thesis was to investigate the effects of syllable structure on consonantal timing and vowel compression. Chapter 2 focused on compensatory vowel shortening caused by coda clusters in both production and perception. Using EMA, chapter 3 focused on the articulatory mechanisms leading to compensatory vowel shortening driven by both onset (/kn/ and /kl/) and coda (/pt/ and /mt/) clusters. Chapter 4 investigated gestural timing stability within syllables under the influence of prosodic weakening, including complex and simplex (/kn/ and /n/) onsets as well as tense and lax vowels (/aː/ and /a/). Finally, chapter 5 investigated syllable organization during language acquisition focusing on consonant-to-consonant as well as on consonant-to-vowel timing in onset (/kl/, /kn/, /pl/) and coda position (/kl/, /kn/, /ks/).

In this chapter the experiments described in chapters 2-5 will be briefly summarized and the influence of syllable structure on consonantal timing and vowel compression will be discussed.

6.1 Incremental compensatory vowel shortening

In chapters 2 and 3 there was a special focus on *incremental vowel shortening*, i.e. whether the vowel becomes shorter in CCVC (onset-driven shortening) or CVCC (coda-driven shortening) syllables than in CVC syllables. Chapter 2 focused only on compensatory vowel shortening driven by complex codas including real words with a complex onset, whereas chapter 3 focused on compensatory vowel shortening driven by both complex onsets and codas including non-words with the tense vowel /aː/.

The results of chapter 2 showed that overall there was no compensatory vowel shortening before complex codas in the production of German monosyllables. However, vowel shortening seems to be speaker-dependent: some speakers did shorten the vowel, while others did not. Further, there were tendencies for longer segments (i.e. tense vowels or lax vowels preceding sonorants in accented contexts) to be shortened to a greater extent than shorter segments (i.e. lax vowels preceding obstruents or vowels in deaccented contexts). For the word pair /kni:t/ - /kniːst/, however, vowel lengthening occurred. This particular word pair differed form other in our study, as it was the only one, in which the vowel following coda consonant (C1) was not kept constant. That is, the C1 in simplex coda tokens was a plosive, whereas in complex coda tokens it was a sibilant. This may be the reason for why vowel lengthening occurred. In addition, there was acoustic shortening of the first coda consonant (C1) in the cluster condition.

The next question was how listeners perceive compensatory vowel shortening and whether they compensate for it. The results showed that listeners were very sensitive to the vowel durations presented to them and they perceived the vowel as being shorter in complex coda tokens in about two-thirds of all instances – even in word pairs containing no durational differences of the vowel or C1. This finding suggests perceptual vowel shortening in words with coda clusters. For accented words containing sonorant codas with no duration differences, however, there was a trend towards compensation for compensatory shortening. Further, listeners compensated more in words containing tense vowels in both accented and deaccented contexts than in words containing lax vowels. Thus in contexts that are more prone to shortening in production (i.e. accented words containing a sonorant coda and accented words containing tense vowels) there was also a trend towards compensation in perception, when the presented signal was ambiguous. As the perception experiment contained natural productions, a perception experiment including synthetic stimuli is needed in the future for two reasons: firstly, other acoustic cues, such as possible acoustic shortening of the syllable onset in complex coda tokens, can be eliminated in this way. Secondly, a continuum between shortened and non-shortened vowels would be preferable to measure compensation for compensatory shortening more accurately in terms of spectral slope and cross-over points.

The study presented in chapter 2 included real words containing complex onsets (either

/kl/ or /kn/). Complex onsets, however, should – according to Articulatory Phonology – compress the vowel acoustically, e.g. the vowel is longer in /nɪk/ than in /knɪk/. Thus an additional coda consonant as in /knɪkt/ may not compress the vowel anymore, especially not a lax vowel (see Klatt, 1973). This and the observed C1 shortening may be a reason for why vowel shortening was less pronounced.

Chapter 3, therefore, included only the tense vowel /aː/ in accented non-words containing simplex onsets and simplex vs. complex codas in order to examine coda-driven vowel shortening. In addition, onset-driven vowel shortening was investigated in non-words containing simplex codas and simplex vs. complex onsets. The acoustic analyses showed that there was significant vowel shortening driven by both onset and coda clusters. This leads to the conclusions that (1) tense vowels are more prone to shortening effects (see also Mooshammer and Geng, 2008, Mooshammer and Fuchs, 2002 and Harrington et al., 2015, for differences in vowel compression between lax and tense vowels), and that (2) the syllable has to be kept as simple and constant as possible.

6.2 Articulatory mechanisms underlying vowel shortening

Chapter 3 investigated the articulatory mechanisms leading to the observed compensatory vowel shortening of the tense vowel /aː/ driven by both onset and coda clusters. Different kinds of articulatory mechanisms were taken into account leading to these shortening effects: increased CV or VC overlap, shorter vowel plateau durations, greater stiffness of the vowel gesture, vowel undershoot and less jaw lowering in complex tokens as opposed to simplex tokens were considered.

Concerning CV and VC overlap it was expected that there would be increased overlap (and thus a shift towards the vowel) in complex tokens leading to shorter acoustic vowel durations. For onsets, results showed that there was a shift towards the vowel in the cluster /kn/ but not in the cluster /kl/. We concluded that the observed greater CC overlap within /kl/ onsets leads to lesser (C)CV overlap. One explanation for why /l/ exhibits less coarticulation and thus less overlap with the following vowel may be that the lateral exhibits more anticipatory coarticulation (i.e. right-to-left). Further, there was a shift towards the vowel in complex coda tokens only in words containing /l/ in onset position.

Concerning vowel plateau duration, it was expected that the vowel plateau will be compressed in complex syllables leading to shorter acoustic vowel durations. However, there were no differences either due to an increasing number of consonants in onset position or due to an increasing number of consonants in coda position.

As far as stiffness is concerned, it was expected that vowel gestures in complex tokens have a greater stiffness (and are thus produced faster), which would result in shorter acoustic vowel durations. The results showed that only complex onsets affected the stiffness of the vowel gesture: the vowel was produced with a greater stiffness in words containing complex onsets as opposed to simplex onsets. For codas, on the other hand, there were no differences in stiffness as the number of consonants increased. That is, the stiffness of a vowel gesture was affected by syllable onset clusters but not by syllable coda clusters. Further analyses showed that this greater stiffness was due to greater peak velocity in complex onset tokens as opposed to simplex onset tokens. One explanation may be that speakers might have less time to couple both onset consonants with the vowel (Goldstein et al., 2007a; Nam et al., 2009; Nam and Saltzman, 2003) and therefore the vowel gesture is produced faster. This is not the case or not necessary in coda clusters, since there is no competitive coupling.

Further, we considered whether vowel undershoot and differences in jaw height could lead to vowel shortening. To do so, the vertical tongue back and jaw positions during the vowel were measured. Concerning onsets, results showed that the tongue back was lowered to a lesser extent producing the vowel /aː/ in complex tokens compared with simplex tokens. That is, there was vowel undershoot. Despite less tongue back lowering, there were no differences in jaw position depending on onset complexity. However, differences concerning both tongue back and jaw position in complex coda tokens could have been observed. The tongue back was lowered to a lesser extent during the production of /aː/ in words containing complex codas than in words containing simplex codas, i.e. there was vowel undershoot. In addition, the jaw was lowered to a lesser extent during the production of the vowel in complex coda tokens than in simplex coda tokens. It has to be considered, though, that the observed tongue back differences in coda tokens could be an artefact caused by differences in jaw height, since the jaw is coupled with the tongue. The observed differences in jaw height might be attributed to the high jaw position required for the additional /t/ in the cluster condition.

To conclude, onset-driven compensatory vowel shortening is most likely to be caused by differences in vowel stiffness and vowel target undershoot. For coda-driven compensatory vowel shortening, on the other hand, the most plausible articulatory mechanism is the observed difference in jaw height. Furthermore, the study presented in chapter 3 revealed evidence that onsets and codas behave different not only in their presumed timing relation but also in their spatial relation with the vowel.

1. A shift towards the vowel was observed for the complex onset /kn/. This was not the case for the onset cluster /kl/. In addition, a shift towards the vowel in complex tokens was observed only for /l/ words not for /n/ words, although the coda was kept constant.
2. Only onset clusters affected the vowel stiffness, whereas complex codas did not.
3. The jaw position was affected by complex codas, whereas no such effect was found for onset consonants.

6.3 Stability of gestural timing within syllables

The study presented in chapter 4 investigated the effects of prosodic weakening and vowel tensity on the timing between all syllable constituents (CC, CV and VC sequences). The speech material included complex (/kn/) and simplex (/n/) onsets in order to investigate the effect of prosodic weakening on the global organization of /kn/ (see chapter 3). Further, possible differences in timing patterns between lax and tense vowels were examined. In particular, Brunner et al. (2014) recently showed that the c-center organization of onset clusters in German is more stable in words containing tense vowels than lax vowels.

To account for the influence of prosodic weakening on articulatory timing patterns within a syllable, not only the timing between the syllable constituents was considered but also the plateau duration of each syllable constituent and its stiffness. The motivation behind this was that less overlap at higher prosodic boundaries (Bombien et al., 2006; Byrd and Choi, 2010; Byrd et al., 2006) but also less overlap caused by plateau shortening in unaccented contextsMücke et al. (2008) can occur.

The results showed that the syllable constituents are affected differently by prosodic conditions: there was more CC and VC overlap, but less CV overlap in deaccented words. Further, only the plateau durations of $C2_{on}$ and the vowel were shortened by prosodic weakening. This particular finding can be explained by the fact that the π-gesture associated with the syllable nucleus already waxes continuously towards the π-gesture's peak activation (Byrd et al., 2006; Byrd and Saltzman, 2003). In addition, these prosody-dependent plateau shortening effects may account for the observed lesser CV overlap. Thus it appears that speakers use both less overlap and temporally lengthened segments to mark phrasal accent.

Further, the study presented in chapter 4 showed that prosodic weakening prevents the cluster's shift towards the vowel. Thus prosodic weakening decouples onset clusters from the vowel via plateau shortening. It was concluded that the vowel adjacent consonant $C2_{on}$ does not shift towards the vowel in deaccented words because otherwise segments could be lost in perceptual terms because of articulatory masking. This leads to the conclusion that the predictions made by the Articulatory Phonology depend not only on the cluster's composition (see Marin, 2013), but also on prosodic structures.

To conclude, this study has shown that prosodic weakening may cause more overlap between gestures or less overlap between gestures accompanied by gestural shortening. These different prosodic effects on syllable timing and gestural duration need to be included in the π-gesture approach. Further, the particular and quite different effects of prosodic weakening reported in this study need to be included in theories modeling syllable structure such as Articulatory Phonology.

6.4 Timing patterns in first language acquisition

The last chapter was concerned with syllable timing in children whose language development is not yet completed. The motivation behind this was that asymmetries in cluster acquisition have been reported (Levelt et al., 2000; Lleó and Prinz, 1996; Templin, 1957). We attributed these asymmetries to articulatory timing differences depending on syllable position. It was assumed that onset clusters are acquired later than coda clusters because of two articulatory based reasons: (1) onset clusters in adult productions are globally organized (e.g. Goldstein et al., 2007a; Marin and Pouplier, 2010; Pouplier, 2012), i.e. they are timed anti-phase with each other and in-phase with the following vowel (competitive coupling; see Nam and Saltzman (2003)) and (2) onset clusters in adult productions are timed more rigidly than coda clusters to ensure their perceptual recoverability (e.g. Byrd, 1996; Byrd and Choi, 2010; Chitoran et al., 2002; Pouplier, 2012). Thus it might be easier for children to produce sequentially organized coda clusters, which are also allowed to overlap to a greater extent and nonetheless perceivable.

Our results showed that the onset clusters /kl/ and /kn/ differed in consonant-to-consonant timing: there was greater overlap between /k/ and /l/ than between /k/ and /n/. Also the onset clusters /pl/ and /kl/ differed slightly in consonant-to-consonant overlap: there was less overlap between /p/ and /l/ than between /k/ and /l/. This confirmed previous studies on adult productions (Bombien, 2011; Bombien et al., 2013; Hoole et al., 2009) and suggests that the cluster specific consonant-to-consonant timing pattern (which is necessary for perceptual recoverability) is acquired early in childhood. Further, these onset clusters did not shift towards the vowel, i.e. they were sequentially organized with the following vowel. This finding differs from adult productions (see e.g. Pouplier, 2012). In addition, it was suggested that onset clusters overlap to a lesser extent in their consonant-to-consonant timing than coda clusters. However, there was the same amount of overlap in

/kl/ clusters in onset as in coda position. This suggests that the degree of overlap within a cluster depends more on whether the consonants can be both identified (cf. Marin, 2013). Concerning coda clusters, /kl/ and /ks/ are timed with a sequential alignment with the preceding vowel, whereas in /kn/ the vowel adjacent coda consonant shifted away from the vowel. In general then, the results on consonant-to-vowel timing suggest that children prefer a sequential timing relationship with the vowel, regardless of syllable position.

The present study showed that children acquire consonant-to-consonant timing before consonant-to-vowel timing in onset position. Additionally, a decoupling of C1 in onset position can explain the occurrence of epenthesis in early cluster productions (Bortolini and Leonard, 1991; Dyson and Paden, 1983; McLeod et al., 2001a). Fikkert (1994) showed for two children the following pattern: first there is cluster reduction, followed by the production of both cluster components separated from an intervening vowel, and finally a correct cluster production. This supports Nam et al.'s (2009) idea that the sequential consonant-to-consonant coupling has to stabilize. This may explain why coda clusters in general are acquired first, because in coda clusters there is no additional dominant in-phase coupling and consequently the anti-phase timing needs not to stabilize.

6.5 Conclusion and implications for future work

Within the framework of Articulatory Phonology, onset clusters are generally predicted to exhibit c-center organization. This thesis supported evidence that not all onset clusters exhibit such global timing with the following vowel. While there was more CC overlap in /kl/ than in /kn/ clusters, only the latter seems to exhibit c-center organization (see chapter 3). Further, coda clusters are not entirely sequentially organized. In fact, the onset consonant seems to have an effect on VC(C) timing: only in words with /l/ in onset position did $C1_{off}$ shift towards the vowel. This was not the case in words with /n/ in onset position, although $C1_{off}$ was always a labial. The study presented in chapter 3 also showed that consonant-to-consonant timing affects consonant-to-vowel timing, but there were no differences in their influence on the vowel gesture's stiffness. Thus the observed greater vowel stiffness (which is related to greater velocity) in complex onsets is independent of any presumed c-center organization. We can therefore conclude that onset clusters that do not shift towards the vowel can nonetheless be linked to the following vowel. Further, the results presented in chapter 3 showed that onset and coda clusters behave quite differently, not only in their proposed temporal organization, but also in their spatial organization with the following/preceding vowel.

Further, the study presented in chapter 4 revealed evidence that speakers use different kinds of articulatory strategies to mark phrasal accent: less overlap or more overlap accompanied by gestural lengthening. Further, the coordination patterns within CCVC syllables differed with respect to syllable position: there was more CC and VC overlap but less CV overlap in deaccented words. The lesser CV overlap in deaccented words may be a consequence of plateau shortening of the two corresponding gestures. In addition, the global organization of onset clusters diminishes in deaccented position.

Finally, chapter 5 revealed evidence that children coordinate onset clusters in a sequential alignment with the following vowel. During language acquisition, therefore, onset clusters are in the same timing relationship with the vowel as coda clusters.

The c-center hypothesis has been used to articulatorily explain/investigate the phonological status of onset clusters in languages where the syllabic affiliation of onsets is considered to be controversial. Generally, it has been claimed that an absent c-center organization indicates that the consonants do not form a complex onset in the particular language but a series of simplex onsets or are considered to be extrasyllabic with a consonantal nucleus (see chapter 1 for more details) (e.g. Hermes et al., 2013, 2011; Shaw et al., 2009). The results obtained in this dissertation, however, show – like studies on Slovak (Pouplier and Beňuš, 2011) and Romanian (Marin, 2013) – that the c-center hypothesis cannot be used as a tool for clarifying the syllable affiliation of onsets. This is because otherwise the onset cluster /kl/ would be affiliated in a different way that the onset cluster /kn/. Further, this would indicate that the syllabic affiliation of a particular cluster would change in prosodically weak contexts. In addition, children would exhibit a different phonological syllable structure than adults do.

For future work, an investigation of the following questions should be considered:

- It would be advisable to conduct a perception experiment using synthetic stimuli with manipulated (i.e. lengthened and shortened) vowel durations like those presented in Fowler and Thompson (2010). Complex vs. simplex onsets (≈ anticipatory shortening) and complex vs. simplex codas (≈ backward shortening) pairs would need to be included. We would expect – commensurate with the findings presented in Fowler and Thompson (2010) and Harrington et al. (2015) – that listeners compensate for compensatory shortening and even more so for word pairs in accented positions.

- The results presented in chapter 3 showed that there were differences between /n/ and /l/ words not only in CC and CV overlap in onset tokens, but also in VC overlap in coda tokens: in /l/ words there was no shift towards the vowel in complex onsets (e.g. /klaːp/ vs. /laːp/) but there was a shift in complex codas (e.g. /laːp/ vs. /laːpt/) and the reverse for /n/ in onset position. Therefore, it would be interesting to look at overlap patterns in /klaːp/ - /laːp/ - /laːpt/ - /klaːpt/. The question to be investigated is: is it the case that CCVCC coordination can be predicted by CCVCC coordination, and vice versa?

- The results of chapter 3 showed that onset clusters are associated with a greater vowel gesture stiffness. It would be interesting to see whether children show the same pattern. It might be the case that they cannot yet coordinate their articulators precisely enough. This may further explain differences in consonant cluster acquisition.

- There is a disagreement on whether there is a correspondence between diachronic changes and children's mispronunciations during language acquisition (Grammont, 1933; Jakobson, 1941; Paul, 1880; Stampe, 1979) or not (cf. Drachman, 1978; Foulkes and Vihman, 2013; Greenlee and Ohala, 1980). According to Ohala (1993), sound change occurs when a listener misperceives speakers' productions. Thus children as inexperienced listeners may play a role (Ohala, 1981, 1993). Bombien et al. (2013) and Hoole et al. (2009) state that if there is more overlap between /k/ and /n/, the earlier velum lowering would hide the release of the consonant. This might be an explanation or an underlying mechanism for observed sound change patterns from /kn/ to /n/ in English (Vennemann, 2000). TaDA (Nam et al., 2007) or VocalTract-Lab (Birkholz, 2004; Birkholz et al., 2007; Birkholz and Kröger, 2006) can be used to synthesize a continuum with different overlap patterns creating acoustic stimuli. Thus a continuum with increasing CC overlap between /kn/ and /kl/ for minimal pairs such as *Knick* vs. *Nick* and *Klack* vs *Lack* can be synthesized. The resulting stimuli could then be presented to adults and children in a procedure in which they have to name the corresponding word or picture after each presented stimuli. If earlier velum lowering masks the closure release of /k/ in *Knick*, then subjects should hear *Nick*. However, the same amount of overlap (e.g. between the velar gesture for /k/ and the tongue dorsum gesture for /l/) should not lead to perceptual masking of the plosive. Accordingly, children may be 'worse' than adults in performing the task and hear *Knick* later (i.e. in stimuli with less CC overlap) than adults.

- The study presented in chapter 5 has revealed evidence that children aged 5-6 years produce onset clusters with a sequential alignment with the following vowel. So far it is not clear at which age children would acquire the proposed global alignment with the following vowel. Therefore, a follow-up study with older children (e.g. aged 8-9 years) or a longitudinal study would be desirable to gain more insight into motor control development in children.

- It might be interesting to see how syllables are timed in languages with no syllable onsets like Arrernte. Could it be the case that coda clusters are more stable and exhibit competitive coupling with the preceding vowel?

To conclude, consonant-to-consonant timing influences consonant-to-vowel timing. Further, increased consonant-to-vowel timing in complex onsets influences the plateau duration of $C2_{on}$ and $C1_{off}$; i.e. while the vowel plateau duration remains nearly unaffected, the plateaus of the vowel surrounding consonants are compressed due to an additional onset consonant. Further, it is not only the consonantal timing with the vowel that acoustically compresses the vowel duration (as proposed by Articulatory Phonology), it is also the consonantal influence on spatial parameters. Generally, this thesis showed that syllable timing is influenced by prosodic structure, segmental make-up and language acquisition.

This thesis was a first attempt to measure consonant-to-vowel overlap directly rather than indirectly as it was usually the case in c-center studies using anchorpoints (e.g. Browman and Goldstein, 1988, 2000; Byrd, 1995; Marin, 2013; Marin and Pouplier, 2010; Pouplier, 2012). The advantage of directly measuring overlap patterns is that confounding factors (such as shortening of segments) can be disregarded. The disadvantage of this direct measurement is that the items included are limited to those in which articulatory movement patterns are observable (i.e. different articulators have to be used) and which contain a vowel requiring an active tongue movement. The clusters included for articulatory analyses in this thesis were restricted to those exhibiting different types of consonant-to-consonant timing in German, including therefore only /kl, kn/ in chapter 3 and /kl, kn, pl/ in chapter 5. Moreover, only one cluster (/kn/) was used for analyses in chapter 4. Therefore, the results reported in the present thesis need to be verified for other languages and for other clusters preceded/followed by various vowels (such as /i/, /o/ or /u/).

Appendix A

Speech Material

A.1 Complete speech material of Chapter 3 and 4

	CVC	**CCVC**	**CVCC**	**CCVCC**
	Nahp	Knahp	nahpt	knahpt
	Lahp	Klahp	lahpt	klahpt
tense vowels	Lahm	Klahm	lahmt	klahmt
	Kahp	/	kahpt	/
	Kahm	/	kahmt	/
	Pahp	/	pahpt	/
	Napp	Knapp	nappt	knappt
	Lapp	Klapp	lappt	klappt
lax vowels	Lamm	Klamm	lammt	klammt
	Kapp	/	kappt	/
	Kamm	/	kammt	/
	Papp	/	pappt	/

	lateral codas	
tense vowels	Pahl	Pahlm
	Kahl	Kahlm
lax vowels	Pall	Palm
	Kall	Kalm

A.2 Complete speech material of Chapter 5

Onset		**Coda**	
simplex	*complex*	*simplex*	*complex*
Note	Knoten	Lok	Locken
Lack	Klack	Sack	Dackel
Last	Platz	Gebäck	Gewächs
Laster	Pflaster	Platz	platzen
Lied	Kleid	Blut	bluten
Napf	Knopf	Ball	Hulk
Kette	Skelett	Wolle	Wolke

additional words

Naht	Knauf	Socken	Pumuckl
Kopf	Skateboard	Schaukel	Dackel
gelb	Eigelb	Wal	Wale
Suppe	Schuppe	Sippe	Schippe

Appendix B

Previous works

The author of this thesis has previously been among the authors of the following works:

Ruch, H. & Peters, S. (2012). Pre-aspiration and sound change in two varieties of Spanish. In *Proceedings of the 13th Conference on Laboratory Phonology, Stuttgart.*

Peters, S. & Kleber, F. (2013). Compensatory vowel shortening before complex coda clusters in the production and perception of German monosyllables. *Journal of the Acoustical Society of America*, 134, p. 4202.

Peters, S. & Kleber, F. (2014). Articulatory mechanisms underlying incremental compensatory vowel shortening in German. In *Proceedings of the 10th International Seminar on Speech Production, Cologne, Germany.*

Peters, S., & Kleber, F. (2014). The influence of accentuation and onset complexity on gestural timing within syllables. In *Proceedings of the 7th Speech Prosody Conference, Dublin, Ireland.*

Kleber, F., Peters, S. (2014). Children's imitation of coarticulatory patterns in different prosodic contexts. In *Proceedings of the 14th Conference on Laboratory Phonology, Tokyo, Japan.*

Ruch, H. & Peters, S. (2016). On the Origin of Post-Aspirated Stops: Production and Perception of /s/ + Voiceless Stop Sequences in Andalusian Spanish. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 7(1), 2.
DOI: http://doi.org/10.5334/labphon.2

Bibliography

Avesani, C., Vayra, M., and Zmarich, C. (2007). On the articulatory bases of prominence in Italian. In *Proceedings of the 16th International Congress of Phonetic Sciences. Saarbrücken*, pages 981–984.

Barlow, J. (1997). *A constraint-based account of syllable onsets: Evidence from developing systems*. PhD thesis, Indiana University.

Barton, D., Miller, R., and Macken, M. (1980). Do children treat clusters as one unit or two? *Papers and Reports on Cild Language Development*, 18:93–137.

Bates, D. (2008). lme4: An R package for fitting and analyzing linear, nonlinear and generalized linear mixed models. Software application.

Beckman, M. and Edwards, J. (1994). Articulatory evidence for differentiating stress categories. In Keating, P., editor, *Papers in Laboratory Phonology 3: Phonological Structure and Phonetic form*, pages 7–33. Cambridge University Press.

Beckman, M., Edwards, J., and Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory analysis. In Docherty, G. and Ladd, D., editors, *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, pages 68–86. Cambridge: Cambridge University Press.

Bell, A. and Hooper, J. (1978). *Syllables and segments*. North-Holland, Amsterdam.

Bernhardt, B. and Stoel-Gammon, C. (1998). *Handbook of phonological development. From the perspective of constraint-based nonlinear phonology*. San Diego: Academic Press.

Berry, M. and Eisenson, J. (1956). *Speech disorders: principles and practices of therapy*. New York: Appleton.

Birkholz, P. (2004). Control of an articulatory speech synthesizer based on dynamic approximation of spatial articulatory targets. In *Proceedings of the Interspeech 2004 - ICSLP, Jeju, Korea*, pages 1125–1128.

Birkholz, P., Jackél, D., and Kröger, B. (2007). Simulation of losses due to turbulence in the time-varying vocal system. *IEEE Transactions on Audio, Speech and Language Processing*, 15(4):1218–1226.

Birkholz, P. and Kröger, B. (2006). Vocal tract model adaptation using magnetic resonance imaging. In *Proceedings of the 7th International Seminar on Speech Production, Ubatuba, Brazil*, pages 493–500.

Blevins, J. (1995). The Syllable in Phonological Theory. In Goldsmith, J., editor, *The Handbook of Phonological Theory*. Blackwell: Oxford.

Boersma, P. and Weenink, D. (1992). Praat: Doing phonetics by computer. Software program (Website).

Bombien, L. (2011). *Segmental and prosodic aspects in the production of consonant clusters - On the goodness of clusters*. PhD thesis, Ludwig-Maximilans-University, Munich.

Bombien, L., Mooshammer, C., and Hoole, P. (2013). Articulatory coordination in word-initial clusters of German. *Journal of Phonetics*, 41:546–561.

Bombien, L., Mooshammer, C., Hoole, P., Kühnert, B., and Schneeberg, J. (2006). An EPG study of initial /kl/ clusters in varying prosodic conditions in german. In *Proceedings of the 7th International Seminar on Speech Production, Ubatuba, Brazil*, pages 457–460.

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

Bortolini, U. and Leonard, L. (1991). The speech of phonologically disordered children acquiring italian. *Clinical Linguistics and Phonetics*, 5:1–12.

Breen, G. and Pensalfini, R. (1999). Arrernte: A language with no syllable onsets. *Linguistic Inquiry*, 30:1–26.

Browman, C. and Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3:219–252.

Browman, C. and Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, 45:140–155.

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

Browman, C. and Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In Kingston, J. and Beckman, M., editors, *Papers in Laboratory Phonology I: Between the grammar and physics of speech*, pages 341–376. Cambridge: Cambridge University Press.

Browman, C. and Goldstein, L. (1992). Articulatory Phonology: An Overview. *Haskins Laboratories Status Report on Speech Research*.

Browman, C. and Goldstein, L. (1995). Gestural syllable position effects in American English. In Bell-Berti, F. and Raphael, L., editors, *Producing speech: Contemporary Issues for Katherine Safford Harris*. New York: American Institute of Physics.

Browman, C. and Goldstein, L. (2000). Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la Communication Parlèe*, 5:25–34.

Brunner, J., Geng, C., Sotiropoulou, S., and Gafos, A. (2014). Timing of German onset and word boundary clusters. *Papers in Laboratory Phonology 14. Mouton de Gruyter*, 5(4):403–454.

Byrd, D. (1992). Perception of Assimilation in Consonant Clusters: A Gestural Model. *Phonetica*, 49:1–24.

Byrd, D. (1995). C-centers revisted. *Phonetica*, 52:263–282.

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

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

Byrd, D. and 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. *Papers in Laboratory Phonology 10. Mouton de Gruyter*.

Byrd, D., Kaun, A., Narayanan, S., and Saltzman, E. (2000). Phrasal signatures in articulation. In Broe, M. and Pierrehumbert, J., editors, *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, pages 70–87. Cambridge University Press.

Byrd, D., Krivokapić, J., and Lee, S. (2006). How far, how long: on the temporal scope of prosodic boundary effects. *Journal of the Acoustical Society of America*, 120(3):1589–1599.

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

Byrd, D. and Saltzman, E. (1998). Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics*, 26:173–199.

Byrd, D. and Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary adjacent lengthening. *Journal of Phonetics*, 31(2):149–180.

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

Cambier-Langeveld, T. (2000). *Temporal marking of accent and boundaries*. PhD thesis, University of Amsterdam.

Cambier-Langeveld, T. and Turk, A. (1999). A cross-linguistic study of accentual lengthening: Dutch vs. English. *Journal of Phonetics*, 27(3):255–280.

Campbell, N. and Isard, S. (1991). Segment durations in syllable frame. *Journal of Phonetics*, 19(1):37–48.

Chen, M. (1970). Vowel Length Variation as a Function of the Voicing of the Consonant Environment. *Phonetica*, 22:129–159.

Cheng, H., Murdoch, B., Goozee, J., and Scott, D. (2007). Physiological development of tongue-jaw coordination from childhood to adulthood. *Journal of Speech, Language, and Hearing Research*, 50:352–360.

Chin, S. (1996). The role of the sonority hierarchy in delayed phonological systems. In Powell, T., editor, *Pathologies of speech and language: Contributions of clinical phonetics and linguistics*, pages 109–117. New Orleans: International Clinical Phonetics and Linguistics Association.

Chitoran, I., Goldstein, L., and Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian. In Gussenhoven, C. and Warner, editors, *Laboratory Phonology 7*, pages 419–448. Berlin: Mouton de Gruyter.

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

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*, 117:3867–3878.

Cho, T. (2006). Manifestation of prosodic structure in articulation: Evidence from lip kinematics in English. In Goldstein, L., Whalen, D., and Best, C., editors, *Laboratory Phonology 8: Varieties of phonological competence*. Mouton de Gruyter, Berlin/New York.

Cho, T. and Keating, P. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics*, 29:155–190.

Cho, T. and Keating, P. (2007). Effects of initial position versus prominence in English. *UCLA Working Papers in Phonetics*, 106:1–33.

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

Cho, T. and McQueen, J. (2005). Prosodic influences on consonant production in Dutch: effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, 33(2):121–157.

Clark, H., Robin, D., McCullagh, G., and Schmidt, R. (2001). Motor control in children and adults during a non-speech oral task. *Journal of Speech, Language, and Hearing Research*, 44:1015–1025.

Clements, G. (1990). The role of the sonority cycle in core syllabification. In Kingston, J. and Beckman, M., editors, *Papers in Laboratory Phonology I: Between the grammar and physics of speech*, pages 283–333. Cambridge, U.K.: Cambridge University Press.

Clements, G. and Keyser, S. (1983). *CV phonology: a generative theory of the syllable.* (Linguistic Inquiry Monograph 9). Cambridge MA: MIT Press.

Cooke, J. (1980). The organization of simple, skilled movements. In Stelmach, G. and Requin, J., editors, *Tutorials in motor behavior*. North-Holland, Amsterdam.

Crystal, T. and House, A. (1982). Segmental durations in connected speech signals: Preliminary results. *Journal of the Acoustical Society of America*, 72(3):705–716.

Crystal, T. and House, A. (1988). Segmental durations in connected speech signals: Current results. *Journal of the Acoustical Society of America*, 83(4):1553–1573.

Crystal, T. and House, A. (1990). Articulation rate and the duration of syllables and stress groups in connected speech. *Journal of the Acoustical Society of America*, 88(1):101–112.

Cutler, A. (1984). Stress and accent in language production and understanding. In Gibbon, D. and Richter, H., editors, *Intonation, accent, and rhythm: Studies in discourse phonology*, pages 77–90. Berlin/New York: Mouton de Gruyter.

Davis, B. and MacNeilage, P. (2004). The frame/content theory of speech evolution: from lip smacks to syllables. *Primatologie*, 6:305–328.

Davis, M., Marslen-Wilson, M., and Gaskell, M. (2002). Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 28:218–244.

Davis, S. (1990). Italian Onset Structure and the Distribution of il and lo. *Linguistics*, 28:43–55.

de Jong, K. (1995a). On the status of redundant features: The case of backing and rounding in American English. In Connell, B. and Arvaniti, A., editors, *Phonology and phonetic evidence: Papers in laboratory phonology IV*, pages 68–86. Cambridge: Cambridge University Press.

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

de Jong, K. (2003). Temporal constraints and characterizing syllable structure. In Local, J., Ogden, R., and Temple, R., editors, *Papers in Laboratory Phonology VI*, pages 253–268. Cambridge: Cambridge University Press.

de Jong, K. (2004). Stress, lexical focus, and segmental focus in English: patterns of variation in vowel duration. *Journal of Phonetics*, 32(4):493–516.

de Jong, K. and Zawaydeh, B. (2002). Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration. *Journal of Phonetics*, 30:53–75.

Demuth, K. and Fee, E. (1995). *Minimal words in eraly phonological development*. MS. Brown University & Dalhouse University.

Dilley, L., Shattuck-Hufnagel, S., and Ostendorf, M. (1996). Glottalization of word-initial vowels as function of prosodic structure. *Journal of Phonetics*, 24(4):423–444.

Drachman, G. (1978). Child language and change: a conjecture and some refutations. In Fisiak, J., editor, *Recent developments in historical phonology*. Mouton, The Hague.

Draxler, C. and Jänsch, K. (2004). Speechrecorder - a universal platform independent multi-channel audio recording software. In *Proc. LREC*, pages 559–562.

Duanmu, S. (2002). *The phonology of standard Chinese*. Oxford: Oxford University Press, Usa.

Dyson, A. and Paden, E. (1983). Some phonological acquisition strategies used by two-year-olds. *Journal of Childhood Communication Disorders*, 7:6–18.

Edwards, J. (1985). Contextual effects on lingual-mandibular coordination. *Journal of the Acoustical Society of America*, 78:1944–1948.

Edwards, J., Beckman, M., and Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research*, 47:421–436.

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

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

Elbert, S. and Pukui, M. (1979). *Hawaiian Grammar*. HOnolulu: University Press of Hawai'i.

Elgendy, A. (1999). Jaw contribution to the timing control of pharyngeal consonant production. In *Proceedings of the 14th International Congress of Phonetic Sciences*, pages 2415–2418.

Farnetani, E. and Kori, S. (1986). Effects of syllable and word structure on segmental durations in spoken Italian. *Speech Communication*, 5(1):17–34.

Fikkert, P. (1994). *On the acquisition of prosodic structure*. PhD thesis, University of Leiden, The Netherlands.

Fischer-Jørgensen, E. (1990). Intrinsic F_0 in Tense and Lax Vowels with Special Reference to German. *Phonetica*, 47(3-4):99–140.

Fitch, T. and Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America*, 106(3):1511–1522.

Flemming, E. (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology*, 18:7–44.

Fletcher, J. and Harrington, J. (1999). Lip and jaw coariculation. In Hardcastle, W. and Hewlett, N., editors, *Coarticulation: Theory, Data and Techniques*. Cambridge University Press.

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

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

Foulkes, P. and Vihman, M. (2013). First language acquisition and phonological change. In Honeybone, P. and Salmones, J., editors, *Oxford Handbook of Historical Phonology*. Oxford: Oxford University Press.

Fowler, C. (1983). Converging sources of evidence on spoken and perceived rhythms of spech: Cyclic production of vowels in monosyllabic stress feet. *Journal of Experimental Psychology*, 112(3):386–412.

Fowler, C. (1995). Acoustic and kinematic correlated of contrastive stress accent in spoken English. In Bell-Berti, F. and Raphael, J., editors, *Producing speech: Contemporary issues: For Katherine Safford Harris*, pages 355–373. Melville, Long Island, New York: AIP Publishing.

Fowler, C. and Thompson, J. (2010). Listeners preception of compensatory shortening. *Atten Percept Psychophys*, 72(2):481–491.

Fox, A. (2003). *Kindliche Aussprachestörungen*. Idstein: Schulz-Kirchner Verlag.

Fox-Boyer, A. (2014). Phonologieerwerb. In Fox-Boyer, A., editor, *Handbuch der Sprachentwicklung und Sprachstörungen*. Band II. München: Elsevier.

Fry, D. (1958). Experiments in the perception of stress. *Language and Speech*, 1:120–152.

Gafos, A. (1999). *The Articulatory Basis of Locality in Phonology*. New York: Garland.

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

Goldstein, L., Byrd, D., and Saltzman, E. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. In *Action to Language via the Mirror Neuron System*, pages 215–249. New York: Cambridge University Press.

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

Goldstein, L., Nam, H., Saltzman, E., and Chitoran, I. (2009). Coupled oscillator model of speech timing and syllable structure. In Fant, G., Fujisaki, H., and Shen, J., editors, *Frontiers in Phonetics and Speech Science*, pages 239–250. Festschrift for Wu Zongji.

Goldstein, L., Pouplier, M., Chen, L., Saltzman, E., and Byrd, D. (2007b). Dynamic action units slip in speech production errors. *Cognition*, 103:386–412.

Gracco, V. (1993). Some organisational characteristics of speech movement control. *Journal of Speech and Hearing Research*, 37:4–27.

Grammont, M. (1933). *Traité de phonétique*. Paris: Delagrave.

Greenberg, J. (1966). Synchronic and diachronic universals in phonology. *Language*, 42:508–517.

Greenlee, M. and Ohala, J. (1980). Phonetically motivated parallels between child phonology and historical sound change. *Language Sciences*, 2(1):283–308.

Guenther, F. H. (1995). Speech Sound Acquisition, Coarticulation, and Rate Effects in a Neural Network Model of Speech Production. *Phsychological Review*, 102(3):594–621.

Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.

Haggard, M. (1973). Abbreviation of consonants in English pre- and post-vocalic clusters. *Journal of Phonetics*, 1(1):9–24.

Hardcastle, W. and Barry, W. (1989). Articulatory and perceptual factors in /l/ vocalisations in English. *Journal of the International Phonetic Association*, 15(2):3–17.

Hardcastle, W. and Roach, P. (1979). An instrumental investigation of coarticulation in stop consonant sequences. In Hollien, H. and Hollien, P., editors, *Current issues in the phonetic sciences*, volume 9, pages 531–540. Current Issues in Linguistic Theory. Amsterdam: John Benjamins.

Harrington, J. (2010). *Phonetic Analysis of Speech Corpora*. Chichester: Wiley-Blackwell.

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

Harrington, J., Kleber, F., Siddins, J., and Reubold, U. (2015). The relationship between prosodic weakening and sound change: evidence from the German tense/lax vowel contrase. *Journal of Laboratory Phonology*, 6(1):87–117.

Haugen, E. (1956). The syllable in linguistic description. In Halle, M., Lunt, H., and McLean, H., editors, *For Roman Jacobson*. The Hague: Mouton.

Hawkins, S. (1973). Temporal coordination of consonants in the speech of children: Preliminary data. *Journal of Phonetics*, 1:181–217.

Hawkins, S. (1979). Temporal coordination of consonants in the speech of children: Further data. *Journal of Phonetics*, 7:234–267.

Hawkins, S. and Nguyen, N. (2000). Predicting syllable-coda voicing from the acoustic properties of syllable onsets. In *SWAP-2000*, pages 167–170.

Hayes, B. (1989). Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry*, 20:253–306.

Hayes, B. (1995). *Metrical Theory*. Chicago: Chicago University Press.

Hermes, A., Grice, M., Mücke, D., and Niemann, H. (2008). Articulatory indicators of syllables affiliation in word initial consonant clusters in Italian. In *Proceedings of the 8th International Seminar on Speech Production, Strasbourg, France*, pages 433–436.

Hermes, A., Grice, M., and Niemann, H. (2012). Articulatory coordination and the syllabification of word-initial consonant clusters in Italian. In Hoole, P., Pouplier, M., Bombien, L., Mooshammer, C., and Kühnert, B., editors, *Consonant clusters and Structural Complexity*, chapter 157-176. Mouton de Gruyter [Interface Exploration Series].

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

Hermes, A., Ridouane, R., Mücke, D., and Grice, M. (2011). Kinematics of syllable structure in Tashlhiyt Berber: The case of vocalic and consonantal nuclei. In Baum, S., Menard, L., and Gracco, V., editors, *Proceedings of the 9th International Seminar on Speech Production, Montreal, Canada*.

Hock, H. (1986). *Principles of Historical Linguistics.* Mouton, The Hague.

Hockett, C. (1955). *A manual of phonology*, volume 21, memoir 11. International Journal of Linguistics Monograph Series.

Honorof, D. and Browman, C. (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, Stockholm*, volume 3, pages 552–555.

Hoole, P., Bombien, L., Kühnert, B., and Mooshammer, C. (2009). Intrinsic and Prosodic Effects on Articulatory Coordination in Initial Consonant Clusters. In G. Fant, H. F. and Shen, J., editors, *Frontiers in Phonetics and Speech Science. Festschridt for Wu Zongji*, pages 275–286. Commercial Press, Beijing.

Hoole, P. and Mooshammer, C. (2003). Articulatory analysis of the German vowel system. In Auer, P., Gilles, P., and Spiekermann, H., editors, *Silbenschnitt und Tonakzente*, pages 129–152. Tübingen: Niemeyer.

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

House, A. and Fairbanks, G. (1953). The Influence of Consonant Environment upon the Secondary Acoustical Characteristics of Vowels. *The Journal of The Acoustical Society of America*, 25:105–113.

Ingrisano, D. and Weismer, G. (1979). s-duration: methodological and linguistic factors. *Phonetica*, 36:32–43.

Jakobson, R. (1941). *Child language, aphasia, and phonological universals.* The Hague: Mouton.

Jakobson, R. and Halle, M. (1956). *Fundamentals of Language.* Mouton, The Hague.

Jessen, M. (1993). Stress conditions on vowel quality and quantity in German. *Working Papers of the Cornell Phonetics Laboratory*, 8:1–27.

Jun, J. (2004). Place assimilation. In Hayes, B., Kirchner, B., and Steriade, D., editors, *Phonetically based phonology.* Cambridge: Cambridge University Press.

Katz, J. (2008). English compensatory shortening and phonetic representations. Presented at *Consonant Clusters and Structural Complexity.* Institute of Phonetics and Speech Processing, Munich.

Katz, J. (2010). *Compression effects, perceptual asymmetries, and the grammar of timing.* PhD thesis, Massachusetts Institute of Technology.

Katz, J. (2012). Compression effects in English. *Journal of Phonetics*, 40(3):390–402.

Kauschke, C. (2012). *Kindlicher Spracherwerb im Deutschen*. Walter de Gruyter.

Kavitskaya, D. (2001). *Compensatory Lengthening: Phonetics, Phonology, Diachrony*. PhD thesis, University of California, Berkeley.

Kawasaki, H. (1986). Phonetic explanation for phonological universals: The case for distinctive vowel nasalization. In Ohala, J. and Jaeger, J., editors, *Experimental Phonology*, pages 81–103. Cambridge, MA: MIT Press.

Keating, P. (1985). CV phonology, experimental phonetics, and coarticulation. *UCLA Working Papers in Phonetics*, 62:1–13.

Keating, P. (2006). Phonetic encoding of prosodic structure. In *Speech production: models, phonetic processes and techniques*, pages 167–186. New York: Psychology Press.

Keating, P., Cho, T., Fougeron, C., and Hsu, C. (2003). Domain-initial articulatory strengthening in four languages. In Local, J., Ogden, R., and Temple, R., editors, *Laboratory Phonology 6*, pages 143–161. Cambridge: Cambridge University Press.

Keating, P., Lindblom, B., Lubker, J., and Kreiman, J. (1994). Variability in jaw height for segments in English and Swedish VCVs. *Journal of Phonetics*, 22:407–422.

Kehoe, M. and Lleó, C. (2003). The acquisition of nuclei: a longitufÂ´dinal analysis of phonological vowel length in three German-speaking children. *Journal of Child Language*, 30:527–556.

Kehoe, M. and Stoel-Gammon, C. (2001). Development of syllable structure in English-speaking children with particular reference to rhymes. *Journal of Child Language*, 28:393–432.

Kelso, J., Vatikiotis-Bateson, E., Saltzman, E., and Kay, B. (1985). A qualitative dynamic analysis of reterant speech production: Phase portraits, kinematics, and dynamic modeling. *Journal of the Acoustical Society of America*, 77(1):266–280.

Kent, R. and Vorperian, H. (1995). Anatomic development of the craniofacial-oral-laryngeal systems: A review. *Journal of medical speech-language pathology*, 3:145–190.

Kirk, C. and Demuth, K. (2003). Onset/coda asymmetries in the acquisition of clusters. In Beachley, B., Brown, A., and Conclin, F., editors, *Proceedings of the 27th Annual Boston University Conference on Language Development*. Sommerbille, MA: Cascadilla Press.

Kirk, C. and Demuth, K. (2005). Asymmetries in the acquisition of word-initial and word-final consonant clusters. *Journal of Child Language*, 32:709–734.

Klatt, D. (1973). Interaction between two factors that influence vowel duration. *Journal of the Acoustical Society of America*, 54(4):1102–1104.

Klatt, D. (1974). The duration of [s] in English words. *Journal of Speech and Hearing Research*, 17:51–63.

Klatt, D. (1976). Linguistic uses of segmental duration in English: acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59:1208–1221.

Kochetov, A. (2006). Syllable position effects and gestural organization: Evidence from Russian. In Goldstein, L., Whalen, D., and Best, C., editors, *Papers in Laboratory Phonology VIII*, pages 565–588. Berlin/New York: Mouton de Gruyter.

Kohler, K. (1977). *Einführung in die Phonetik des Deutschen*. Erich Schmidt Verlag.

Kohler, K. (1990). Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In Hardcastle, W. and Marchal, A., editors, *Speech Production and Speech Modelling*, pages 60–92. Dordrecht: Kluwer.

König, W., Rein, K., Wagner, E., and Zehetner, L. (1991). *Bayerns Mundarten: Dialektproben mit Kommentaren und einer Einführung in die Verbreitung und Verwendung des Dialekts in Bayern*. TR-Verlagsunion GmbH, München.

Krakow, R. (1989). *The articulatory organization of syllables: A kinematic analysis of labial and velar gestures*. PhD thesis, Yale University.

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

Krivokapić, J. (2007). Prosodic planning: effects of phrasal length and complexity on pause duration. *Journal of Phonetics*, 35.

Kuehn, D. and Moll, K. (1976). A cineradiographic study of VC and CV articulatory velocities. *Journal of Phonetics*, 4:303–320.

Kühnert, B., Hoole, P., and Mooshammer, C. (2006). Gestural overlap and c-center in selected French consonant clusters. In Yehia, H., Demolin, D., and Laboissiére, R., editors, *Proceedings of the 7th International Seminar on Speech Production, UFMG, Belo Horizonte*, pages 40–48.

Kuzla, C., Cho, T., and Ernestus, M. (2007). Prosodic strengthening of German fricatives in duration and assimilatory devoicing. *Journal of Phonetics*, 35(3):301–320.

Ladefoged, P. and Maddieson, I. (1996). *The Sounds of the World's Languages*. Blackwell: Oxford.

Lee, S., Beckman, M., and Jackson, M. (1994). Jaw targets for strident fricatives. In *Proceedings of ICSLP, Yokohama*, pages 37–40.

Lee, S., Byrd, D., and Krivokapić, J. (2006). Functional data analysis of prosodic effects in articulatory timing. *Journal of the Acoustical Society of America*, 119(3):1666–1671.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, Mass: M.I.T. Press.

Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America*, 51:2018–2024.

Leonard, L. and Ritterman, S. (1971). Articulation of /s/ as a function of cluster and word frequency of occurrence. *Journal of Speech and Hearing Research*, 14:476–485.

Levelt, C., Schiller, N., and Levelt, W. (2000). The acquisition of syllable types. *Language Acquisition*, 8:237–264.

Lieberman, D., McCarthy, R., Hiiemae, K., and Palmer, J. (2001). Ontogeny of postnatal hyoid and larynx descent in humans. *Archives or Oral Biology*, 46:117–128.

Lindblom, B. (1983). Economy of speech gestures. In MacNeilage, P., editor, *The production of speech*. New York: Springer.

Lindblom, B. (1990). Explaining Phonetic Variation: A Sketch of H and H-Theory. In Hardcastle, W. and Marchal, A., editors, *Speech Production and Speech Modelling*, pages 403–439. Kluwer: Dordrecht.

Lindblom, B. and Rapp, K. (1973). Some temporal regularities of spoken Swedish. In *Papers in linguistics, Publication 21*. The University of Stockholm (PILUS).

Lindblom, B. and Sundberg, J. (1971). Acoustical consequences of lip, tongue, jaw, and larynx movement. *Journal of the Acoustical Society of America*, 50:1166–1179.

Lleó, C. and Prinz, M. (1996). Consonant clusters in child phonology and the directionality of syllable structure assignment. *Journal of Child Language*, 23:31–56.

Lobanov, B. (1971). Classification of Russian Vowels by Different Speakers. *Journal of the Acoustical Society of America*, 49:606–608.

Löfqvist, A. (2004). Movement duration and movement stiffness in long and short consonants. *Journal of the Acoustical Society of America*, 116:2630.

Löfqvist, A. and Gracco, V. (1997). Lip and Jaw Kinematics in Bilabial Stop Consonant Production. *Journal of Speech, Language, and Hearing Research*, 40:877–893.

Löfqvist, A. and Gracco, V. (1999). Interarticulator programming in VCV sequences: Lip and tongue movements. *Journal of the Acoustical Society of America*, 105:1854–1876.

Luick, K. (1964). *Historische Grammatik der englischen Sprache*. Blackwell: Oxford.

MacNeilage, P. (1998). The frame/content theory of speech evolution of speech production. *Behavioral and brain sciences*, 21:499–546.

Maddieson, I. (1985). Phonetic cues to syllabification. In Fromkin, V., editor, *Phonetic linguistics: Essays in honor of Peter Ladefoged*, pages 203–221. New York, NY: Academic Press.

Mann, V. and Repp, B. (1980). Influence of vocalic context on perception of the [ʃ]-[s] distinction. *Perception and Psychophysics*, 28:213–228.

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

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

Marin, S. and Shigemori, L. S. B. (2014). Vowel Compensatory Shortening in Romanian. In *Proceedings 10th International Seminar on Speech Production, Cologne, Germany.*

Marslen-Wilson, W. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1-2):71–102.

Mattingly, I. (1981). Phonetic representation and speech synthesis by rule. In Myers, T., Laver, J., and Anderson, J., editors, *The cognitive representation of speech*, pages 415–420. Amsterdam: North Holland.

McLeod, S., Doorn, J. V., and Reed, V. (2001a). Consonant Cluster Development in Two-Year-Olds: General Trends and Individual Difference. *Journal of Speech, Language, and Hearing Research*, 44:1144–1171.

McLeod, S., Doorn, J. V., and Reed, V. (2001b). Normal acquisition of consonant clusters. *American Journal of Speech- Language Pathology*, 10(2):99–110.

Menadier, Y., Pitermann, M., and Marchal, A. (1998). Effects of contrastive focal accent on linguopalatal coarticulation in the French [kskl] clusters. In *Proceedings of th 5th international conference on spoken language processing*, volume 5, pages 1871–1874.

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

Mooshammer, C. (2010). Acoustic and laryngographic measures of the laryngeal reflexes of linguistic prominence and vocal effort in german. *Journal of the Acoustical Society of America*, 127(2):1047–1058.

Mooshammer, C. and Fuchs, S. (2002). Stress distiction in German: simulating kinematic parameters of tongue-tip gestures. *Journal of Phonetics*, 30(3):337–355.

Mooshammer, C. and Geng, C. (2008). Acoustic and articulatory manifestations of vowel reduction in German. *Journal of the International Phonetic Association*, 28(2):117–136.

Mooshammer, C. and Harrington, J. (2005). Linguistic prominence and loudness: a systematic comparison between lexical word stress, sentence accent and vocal effort. Abstract of Talk presented at BeST, Leiden.

Mooshammer, C., Hoole, P., and Geumann, A. (2007). Jaw and order. *Language and Speech*, 50:145–176.

Morin, Y. (1992). Phonological interpretations of historical lengthening. In Dressler, W., Prinzhorn, M., and Rennison, J., editors, *Proceedings of the 7th International Phonology Meeting*.

Mücke, D., Grice, M., and Kirst, R. (2008). Prosodic and lexical effects on German place assimilation. In *Poster Presentation at the 8th International Seminar on Speech Production, Strasbourg, France*.

Munhall, K., Fowler, C., Hawkins, S., and Saltzman, E. (1992). "Compensatory shortening" in monosyllables of spoken English. *Journal of Phonetics*, 20:225–239.

Myers, S. (1987). Vowel shortening in English. *Natural language and linguistic theory*, 5:485–518.

Nam, H., Goldstein, L., and Proctor, M. (2007). TADA (TAsk Dynamics Application). available at http://www.haskins.yale.edu/tada_download/index.php.

Nam, H., Goldstein, L., and Saltzman, E. (2009). Self-organization of syllable structure: a coupled oscillator model. In Pellegrino, F., Marisco, E., and Chitoran, I., editors, *Approaches to phonological complexity*, pages 299–328. Berlin: Mouton de Gruyter de Gruyter.

Nam, H. and Saltzman, E. (2003). A Competitive, Coupled Oscillator Model of Syllable Structure. In *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona*, pages 2253–2256.

Nelson, W. (1984). Physical principles for economics of skilled movements. *Biological Cybernetics*, 46:135–147.

Nelson, W., Perkell, J., and Westbury, J. (1984). Mandible movements during increasingly rapid articulations of single syllables: Preliminary observations. *Journal of the Acoustical Society of America*, 75:945–951.

Nittrouer, S., Estee, S., Lowenstein, J., and Smith, J. (2005). The emergence of mature gestural patterns in the production of voiceless and voiced word-final stops. *Journal of the Acoustical Society of America*, 117:351–364.

Noteboom, S. (1972). *Production and perception of vowel duration: a study of durational properties of vowels in Dutch*. PhD thesis, University of Utrecht.

Ohala, D. (1998). Medial cluster reduction in eraly child speech. In Clark, E., editor, *Proceedings of the 30th Annual Child Language Research Forum*, pages 111–120. Stanford: CSLI.

Ohala, D. (1999). The influence of sonority on children's cluster reductions. *Journal of Communication Disorders*, 32:397–422.

Ohala, J. (1981). The listener as a source of sound change. In Masek, C., Hendrick, R., and Miller, M., editors, *Papers from the Parasession on Language and Behavior*, pages 178–203. Chicago Linguistic Society: Chicagp, IL.

Ohala, J. (1990). The phonetics and phonology of assimilation. In Kingston, J. and Beckman, M., editors, *Papers in Laboratory Phonology I: Between the grammar and physics of speech*. Cambridge: Cambridge University Press.

Ohala, J. (1992). Alternatives to the sonority hierarchy for explaining the shape of morphemens. In *Papers from the Parasession on the syllable*, pages 319–383. Chicago: Chicago Linguistic Society.

Ohala, J. (1993). Sound change as nature's speech perception experiment. *Speech Communication*, 13:155–161.

Ohala, J. and Kawasaki, H. (1984). Prosodic phonology and phonetics. *Phonology*, 1:113–127.

Ohala, J. and Kawasaki-Fukumori, H. (1997). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In Elaisson, S. and Jahr, E., editors, *Language and its ecology: essays in memory of Einar Haugen*, pages 343–365. Trends in Linguistics. Studies and Monographs. Berlin: Mouton de Gruyter.

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

Oliveira, L., Yanagawa, M., Goldstein, L., and Chitoran, I. (2004). Towards standard measures of articulatory timing. *Journal of the Acoustical Society of America*, 116(4):2643–2644.

O'Shaughnessy, D. (1974). Consonant durations in clusters. *IEEE Transactions on Audio, Speech and Language Processing*, 22(4):282–295.

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

Ott, S., van de Vijver, R., and Höhle, B. (2006). The effect of phonotactic constraints in German-speaking children with delayed phonological acquisition: Evidence from Production of word-initial consonant clusters. *International Journal of Speech-Language Pathology*, 4:323–334.

Pater, J. and Barlow, J. (2003). Constraint conflict in cluster reduction. *Journal of Child Language*, 30:487–526.

Paul, H. (1880). *Prinzipien der Sprachgeschichte*. Halle: Niemeyer.

Perkell, J. (1969). *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study*. York, Pennsylvania: Maple Press.

Peters, S. and Kleber, F. (2013). Compensatory vowel shortening before complex coda clusters in the production and perception of German monosyllables. *Journal of the Acoustical Society of America*, 134(5):4202.

Peters, S. and Kleber, F. (2014). Articulatory mechanisms underlying incremental compensatory shortening in German. In *Proceedings of the 10th International Seminar on Speech Production, Cologne, Germany.*

Peterson, G. and Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America*, 32:693–703.

Philipp, M. (1974). *Phonologie des Deutschen*. Stuttgart: Kohlhammer.

Pierrehumbert, J. and Beckman, M. (1988). *Japanese Tone Structure*, volume Linguistic Inquiry Monograph 15. Cambridge, MIT Press.

Pierrehumbert, J. and Talkin, D. (1992). Lenition of /h/ and glottal stop. In Docherty, G. and Ladd, D., editors, *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, pages 90–127. Cambridge: Cambridge University Press.

Pike, K. and Pike, E. (1947). Immediate Constituents of Mazateco Syllables. *International Journal of American Linguistics*, 13:78–91.

Port, R. (1981). Linguistic timing factors in combination. *Journal of the Acoustical Society of America*, 69:262–274.

Pouplier, M. (2012). The gestural approach to syllable structure: Universal, language- and cluster-specific aspects. In Fuchs, S., Weirich, M., Pape, D., and Perrier, P., editors, *Speech planning and dynamics*. New York, Oxford, Wien: Peter Lang.

Pouplier, M. and Beňuš, S. (2011). On the phonetic status of syllabic consonants: Evidence from Slovak. *Laboratory Phonology*, 2(2):243–273.

Pouplier, M. and Hardcastle, W. (2005). A Re-Evaluation of the Nature of Speech Errors in Normal and Disordered Speakers. *Phonetica*, 62:227–243.

Recasens, D. (1999). Lingual coarticulation. In Hardcastle, W. and Hewlett, N., editors, *Coarticulation: Theory, Data and Techniques*, pages 80–104. Cambridge: Cambridge University Press.

Recasens, D. (2007). Patterns of CVC coarticulatory direction according to the DAC model. In Prieto, P., Mascaró, J., and Solé, M.-J., editors, *Segmental and prosodic issues in Romance phonology*, volume 282, pages 25–40. Current Issues in Linguistic Theory. Amsterdam/Philadelphia: John Benjamins.

Recasens, D. and Espinosa, A. (2006). Articulatory, positional and contextual characteristics of palatal consonants: evidence from Majorcan Catalan. *Journal of Phonetics*, 34:295–318.

Recasens, D., Fontdevila, J., and Pallarés, M. (1995). Velarization degree and coarticulatory resistance for /i/ in Catalan and German. *Journal of Phonetics*, 23(1-2):37–52.

Recasens, D. and Pallarés, M. (1999). A study of /r/ and /rr/ in the light of the 'DAC' coarticulation model. *Journal of Phonetics*, 27(2):143–170.

Recasens, D. and Pallarés, M. (2001). Coarticulation, blending and assimilation in Catalan consonant clusters. *Journal of Phonetics*, 29(2):273–301.

Recasens, D., Pallarés, M., and Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America*, 102(1):544–561.

Redford, M. and Diehl, R. (1999). The relative perceptual distinctiveness of initial and final consonants in cvc syllables. *Journal of the Acoustical Society of America*, 106:1555–1565.

Repp, B. (1978). Perceptual integration and differentiation of spectral cues for intervocalic stop consonants. *Perception and Psychophysics*, 24(5):471–485.

Riely, R. and Smith, A. (2003). Speech movements do not scale by orofacial structure size. *Journal of Applied Psychology*, 94:2119–2126.

Romonath, R. (1991). *Phonologische Prozesse an sprachauffälligen Kindern*. Berlin: Edition Marhold im Wissenschaftsverlag Spiess.

Roon, K., Gafos, A., Hoole, P., and Zeroual, C. (2007). Influence of articulator and manner on stiffness. In *Proceedings of the 16th International Congress of Phonetic Sciences. Saarbrücken*, pages 409–412.

Salidis, J. and Johnson, J. (1997). The production of minimal words: a longitudinal case study of phonological development. *Language Acquisition*, 6:1–36.

Saltzman, E. and Byrd, D. (2000a). Demonstrating effects of parameter dynamics on gestural timing. *Journal of the Acoustical Society of America*, 107(5):2904.

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

Saltzman, E., Goldstein, L., Holt, K., Kluzik, J., and Nam, H. (2007). Gait wheels and foot cycles: a parallel between the dynamics of locomotion and speech. Poster presented at the 11th Conference on Cognitive and Neural Systems, Boston.

Saltzman, E. and Munhall, K. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, 1(4):333–382.

Saltzman, E., Nam, H., Krivokapić, J., and Goldstein, L. (2008). A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. In Barbosa, P., Madureira, S., and Reis, C., editors, *Proceedings of the 4th International Conference on Speech Prosody, Campinas, Brazil*.

Salverda, A., Dahan, D., and McQueen, J. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*, 90:51–89.

Schiel, F. (2004). MAUS Goes Iterative. In *Proceedings of the IVth International Conference on Language Resources and Evaluation*, pages 1015–1018.

Scobbie, J. (1998). Interactions between the acquisition of phonetics and phonology. In M Gruber, D. Higgins, K. O. and Wysocki, T., editors, *Papers from the 34th Annual Regional Meeting of the Chicago Linguistic Society*, pages 343–358. Chicago.

Selkirk, E. (1982). The syllable. In van der Hulst, H. and Smith, N., editors, *The structure of phonological representations - Part II*. Dordrecht, The Nederlands: Foris.

Selkirk, E. (1984). On major class features and syllable theory. In Aronoff, M. and Oehrle, R., editors, *Language sound and structure*, pages 107–136. Cambridge: MIT Press.

Shaiman, S. (2001). Kinematics of compensatory shortening: the effect of speaking rate and coda composition on intra- and inter-articulatory timing. *Journal of Phonetics*, 29:89–107.

Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In MacNeilage, P., editor, *The production of speech*. New York: Springer-Verlag.

Shattuck-Hufnagel, S. (1992). The role of word structure in segmental serial ordering. *Cognition*, 42:213–259.

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

Shaw, J., Gafos, A., Hoole, P., and Zeroual, C. (2011). Dynamic invariance in the phonetic expression of syllable structure: A case study of Moroccan Arabic consonant clusters. *Phonology*, 28:455–490.

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

Sievers, E. (1901). *Grundzüge der Phonetik*. Leipzig: Breitkopf und Härtel.

Sigurd, B. (1973). Maximum rate and minimal durations of repeated syllables. *Language and Speech*, 16:373–395.

Silverman, D. (1997). *Phasing and recoverability*. Outstanding dissertations in linguistics. New York: Garland.

Smit, A. (1993). Phonologic error distributions in the Iowa-Nebraska articulation norms project: Word-initial consonant clusters. *Journal of Speech and Hearing Research*, 36:931–947.

Smit, A., Hand, L., Freilinger, J., Bernthal, J., and Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders*, 55:779–798.

Smith, A. (2010). Development of neural control of orofacial movements of speech. In Hardcastle, W. and J.Laver, editors, *The Handbook of Phonetic Sciences*. Blackwell: Oxford.

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

Stampe, D. (1979). *A Dissertation on Natural Phonology*. PhD thesis, University of Chicago.

Stathopoulos, E. and Weismer, G. (1983). Closure durations of stop consonants. *Journal of Phonetics*, 11:395–400.

Steriade, D. (1982). *Greek prosodies and the nature of syllabification*. PhD thesis, MIT, Cambridge.

Stevens, K. and House, A. (1963). Perturbation of vowel articulations by consonantal context: An acoustical study. *Journal of Speech and Hearing Research*, 6(2):111–128.

Tabain, M. (2003). Effects of prosodic boundary on /aC/ sequences: articulatory results. *Journal of the Acoustical Society of America*, 113:2834–2849.

Templin, M. (1957). *Certain language skills in children: their development and interrelationships*. (Monograph Series No. 26). Minneapolis: University of Minnesota, The Institute of Child Welfare.

Treiman, R. (1985). Onsets and Rimes as Units of Spoken Syllables: Evidence from Children. *Journal of Experimental Psychology*, 39:161–181.

Tuller, B., Harris, K., and Gross, B. (1981). Electromyographic study of the jaw muscles during speech. *Journal of Phonetics*, 9:175–188.

Tuller, B. and Kelso, J. (1991). The production and perception of syllable structure. *Journal of Speech and Hearing Research*, 34:501–508.

Turk, A. and Sawusch, J. (1997). The domain of accentual lengthening in American English. *Journal of Phonetics*, 25(1):25–41.

Turk, E. and Shattuck-Hufnagel, S. (2000). Word boundary-related duration patterns in English. *Journal of Phonetics*, 28(4):397–440.

Umeda, N. (1977). Consonant duration in American English. *Journal of the Acoustical Society of America*, 60:846–858.

Vennemann, T. (1988). *Preference laws for syllable structure and the explanation of sound change: with special reference to German, Germanic, Italian and Latin*. Berlin: Mouton de Gruyter.

Vennemann, T. (2000). Triple-cluster reduction in Germanic: Etymology without sound laws? *Historische Sprachforschung (Historical Linguistics)*, 113:239–258.

Vihman, M. and Ferguson, C. (1987). The acquisition of final consonants. In Viks, U., editor, *Proceedings of the 11th International Congress of Phonetic Sciences, Tallinn, Estonia*, pages 381–384.

Vorperian, H. (2000). *Anatomic Development of the Vocal Tract Structures as Visualized by MRI*. PhD thesis, University of Wisconsin-Madison, Madison, WI.

Vorperian, H., Kent, R., Gentry, L., and Yandell, B. (1999). Magnetic resonance imaging procedures to study the concurrent anatomic development of the vocal tract structures: Preliminary results. *International Journal of Pediatric Otorhinolaryngology*, 49:197–206.

Vorperian, H., Kent, R., Lindstrom, M., Kalina, C., Gentry, L., and Yandell, B. (2005). Development of vocal tract length during early childhood: A magnetic resonance imaging study. *Journal of the Acoustical Society of America*, 117:338–350.

Waals, J. (1992). Contextual effects on vowel duration. *Speech Communication*, 11:513–546.

Walsh, B. and Smith, A. (2002). Articulatory movements in adolescents: evidence for protracted development of speech motor control processes. *Journal of Speech, Language, and Hearing Research*, 45:1119–1133.

Watson, M. and Scukanec, G. (1997). Phonological changes in the speech of two-year-olds: A longitudinal investigation. *Infant-Toddler Intervention*, 7:67–77.

White, L. and Turk, E. (2010). English words in the Procrustean bed: Polysyllabic shortening reconsidered. *Journal of Phonetics*, 38:450–471.

Wiese, R. (1996). *The phonology of German*. Oxford: Clarendon Press.

Wright, R. (2003). Factors of lexical competition in vowel articulation. In Local, J., Ogden, R., and Temple, R., editors, *Laboratory Phonology 6*. Cambridge University Press. Cambridge.

Wright, R. (2004). A review of perceptual cues and cue robustness. In Hayes, B., Kirchner, R., and Steriade, D., editors, *Phonetically based phonology*, pages 34–57. Cambridge: Cambridge University Press.

Acknowledgements

It is a pleasure to thank all those people who have contributed to this thesis.

First of all, I would like to thank Felicitas Kleber for all her support, reading, helpful comments and reflections over the last three years. All this has been indispensable for me and this work.

I thank my supervisor Jonathan Harrington for his advice and support. His critical reading was invaluable.

Quite a number of people were involved in acoustic and articulatory labeling or acquisition of the data analyzed here. For that I wish to express my gratitude to Susanne Waltl, Carolin Sabath and of course Rebecca Stegmaier for her kind way in dealing with children. Her help was inalienable for EMA recordings with children.

A big “thank you” to Jessica Siddins for proof-reading.

I am grateful to Véronique Bukmaier and Manfred Pastätter for so many helpful comments and discussions.

I would also like to thank the *Deutsche Forschungsgemeinschaft* and the *Europäische Forschungsrat* for facilitating this thesis through financial support within the German Research Council Grant No. HA3512/9-1 “The relationship between coarticulation, prosodic weakening, and sound change” and the European Research Council grant number 295573 “Sound change and the acquisition of speech”.