Segmental and prosodic aspects in the production of consonant clusters — On the goodness of clusters

Lasse Bombien

München 2011

Segmental and prosodic aspects in the production of consonant clusters — On the goodness of clusters

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

vorgelegt von
Lasse Bombien

Erstgutachter: Dr. phil. habil. Phil Hoole
Zweitgutachter: Prof. Dr. Jonathan Harrington
Tag der mündlichen Prüfung: 27. Januar 2011

Meinen Eltern, Edith und Michael Bombien

Contents

List of Figures

List of Tables

Zusammenfassung auf Deutsch

Einleitung

Die vorliegende Arbeit ist eine experimentelle Untersuchung, die sich mit Einflüssen prosodischer sowie segmenteller Art auf die Artikulation von Konsonantsequenzen (Clustern) am Wortanfang im Deutschen beschäftigt.[1] Eine grundsätzliche Fragestellung dieser Arbeit betrifft die Tatsache, dass einige Konsonantsequenzen sich in vielen Sprachen der Welt, in denen Cluster erlaubt sind, behaupten, während andere äußerst selten oder gar nicht vorkommen. So kann man die Folge /kl/ als häufig betrachten, z.B. lat. ‚clavis' „Schlüssel": französich ‚clé/clef' [kleː/klɛf], spanisch (musik.) ‚clavo' [klavo], polnisch ‚klucz' [klutʂ̥], russisch ‚ключ' [klʉt͡ɕ], aber italienisch ‚chiave' [kjave], spanisch ‚llave' [ʎaβe]; deutsch „Klaue" [klaʊə]: englisch ‚claw' [klɔː], schwedisch ‚klöv' [kløv]. /kn/ tritt anscheinend eher selten auf, z.B. deutsch „Knie" [kniː]: germanisch ‚*knewa-' , lateinisch ‚genu' , schwedisch ‚knä' [knæ], isländisch ‚hné' [n̥jeː], englisch ‚knee' [niː], französisch ‚genou' [ʒenu]; deutsch „Knoten" [knoːtən]: schwedisch ‚knut' [knʉːt], isländisch ‚hnútur' [n̥uːtʏr̥], englisch ‚knot' [nɒt], lateinisch ‚nodus' , französisch ‚nœud' [nø].

Es stellt sich also die Frage, ob verschiedene Cluster unterschiedlich gut geeignet sind, sich in der Sprachen der Welt diachronisch stabil zu behaupten. Dazu ist es notwendig, die Eignung bzw. die Güte der Cluster an messbaren Parametern festzumachen. Dies soll in dieser Arbeit auf der Grundlage von zwei fundamentalen Annahmen getan werden. Da ist zum einen das Prinzip der parallelen Übertragung (parallel transmission) (Mattingly, 1981). Dies beruht auf der Annahme, dass Koartikulation, also die Überlappung von artikulatorischen *Gesten*, der Übertragung von sprachlichem Inhalt zuträglich ist, indem Information über mehrere Gesten gleichzeitig verfügbar ist. Durch diese parallele Übertragung wird der Informationsfluss beschleunigt, da durch überlappende Gesten mehrere Sprachlaute innerhalb desselben Zeitfensters geäußert werden können. In Konsonant-Vokal-Folgen (CV) kann diese Überlappung maximal sein, da die entsprechenden Gesten simultan einsetzen (Öhman, 1967). In Konsonantsequenzen gilt jedoch die Einschränkung, dass die Entstehung akustischer Korrelate der

[1]Cluster in anderen Positionen verhalten sich anders, finden hier jedoch keine Beachtung. Es kann daher davon ausgegangen werden, dass, wann immer die rede von Clustern ist, wortinitiale gemeint sind.

konsonantischen Gesten nicht durch gestische Überlappung verhindert werden darf. Mit anderen Worten: Gestische Wiederherstellbarkeit (*gestural recoverability*, z.B. Chitoran, Goldstein & Byrd, 2002) muss trotz Überlappung gewährleistet sein. In dieser Arbeit wird angenommen, dass ein *gutes* Cluster einen Kompromiss zwischen diesen beiden Annahmen darstellt. Die Überlappung von Gesten spiegelt die artikulatorische Koordination wider. Zum Beispiel in CV Strukturen setzen Konsonant und Vokal, wie oben erwähnt, gleichzeitig ein. Konsonant-Konsonant (CC) Strukturen hingegen müssen eine andere Koordination aufweisen, damit die Wiederherstellbarkeit der einzelnen Laute sowie auch der Abfolge gewährleistet bleibt. In dieser Studie werden die Koordinationsmuster von Konsonantenclustern untersucht und durch systematische Variation segmenteller und prosodischer Faktoren auf ihre Stabilität hin geprüft.

Es gibt verschiedene Konzepte die mit Hilfe derer mögliche Koordinationsmuster in Konsonantenclustern vorhergesagt werden können. Diese Arbeit diskutiert drei verschiedene Modelle, die in ihren Vorhersagen nicht vollständig mit einander übereinstimmen. Eines dieser Modelle lässt sich aus dem oben genannten Prinzip der parallelen Übertragung herleiten. Nach Mattingly hängt der Grad der Überlappung, den aufeinanderfolgende Sprachlaute erlauben, davon ab, wie stark der für die Artikulation der Laute notwendige Konstriktionsgrad ist, beziehungsweise wie groß die Differenz zwischen den Konstriktionsgraden der beiden Laute ist. In Plosiv-Vokal Folgen beispielsweise ist der Konstriktionsgrad für den einen Laut maximal, für den zweiten minimal. Diese Kombination erlaubt maximale Überlappung. Viel weniger Überlappung würde demzufolge in einer Plosiv-Frikativ oder gar Plosiv-Plosiv Folge erwartet werden, da in diesen Fällen beide Laute eine starke Konstriktion aufweisen. Wird dieser Gedankengang fortgeführt ergibt sich, dass Lautfolgen zyklisch zwischen hoher und niedriger Konstriktion pendeln. Diese Annahme findet darin Bestätigung, dass die gängigste Silbenstruktur in den Sprachen der Welt aus CV Folgen besteht. Des weiteren erinnert diese Beschreibung der häufig beobachteten und beschriebenen sogenannten Sonoritätshierarchie (z.B. Sievers, 1901; Selkirk, 1984): Silben tendieren zu steigender Sonorität im Kopf, maximaler Sonorität im Kern und sinkender Sonorität in der Coda. Ein solcher Erklärungsansatz, der maßgeblich auf der Artikulationsart basiert, versagt jedoch in Hinblick auf Phänomene wie den sogenannten *place-order-effect* (z.B. Chitoran, 2002). Dieser Begriff beschreibt die Beobachtung, dass *front-back* Cluster (z.B. /tk/) mit größerer Überlappung produziert werden als *back-front* Cluster (z.B. /kt/). Hierfür wurde das Prinzip der gesturalen Wiederherstellbarkeit verantwortlich gemacht. Bei zu großer Überlappung in *back-front* Clustern maskiert der alveolare Verschluss die Lösung des vorangehenden velaren Verschlusslautes und verhindert die Entstehung eines Lösungsgeräusches, das eine wesentliche Rolle bei der korrekten Perzeption von Plosiven spielt. Als drittes wird das DAC (Degree of Articulatory Constraint) Modell für linguale Koartikulation diskutiert, das anhand von mechanischen Trägheitseigenschaften der Zunge Sprachlaute danach klassifiziert, wieviel Koartikulation benachbarter

Segmente sie erlauben und wieviel Koartikulation sie auf benachbarte Segmente ausüben.

Prosodische Variation wird verwendet, um zu ermitteln, welche der beobachteten Koordinationsmuster stabil sind. Ein zentrales Anliegen dabei ist, ob beobachtete stabile Muster mit diachronischer Stabilität von Clustern in den Sprachen der Welt übereinstimmt. Außerdem soll herausgefunden werden, ob das bekannte Phänomen der prosodischen Stärkung (z.B. Fougeron & Keating, 1997) auch bei deutschen Clustern auftritt. Es handelt sich dabei um die Beobachtung, dass Sprachlaute an prosodischen Grenzen länger und stärker artikuliert werden als in prosodisch unmarkierten Positionen. Ähnliches wurde ebenso für prosodische Prominenz (Wortbetonung, Phrasenakzent) gezeigt (z.B. Cho & Keating, 2009). Prosodische Effekte werden innerhalb des π-Gesten Modells (Byrd & Saltzman, 2003), einer Erweiterung der Artikulatorischen Phonologie (Browman & Goldstein, 1986, 2000) diskutiert. Dieses Modell sieht vor, dass die Ausführung von artikulatorischen Bewegungen/Gesten an prosodischen Grenzen in Abhängigkeit von deren Stärke verlangsamt und vergrößert werden. Diese Entschleunigung ist am stärksten an der Grenze und wird mit Abstand von diesem Punkt graduierlich schwächer.

Diese komplexen Sachverhalte werden in dieser Arbeit in drei Experimenten von verschiedenen Seiten und auf verschiedene Art beleuchtet.

Elektropalatographie von /sk/, /ks/, /kl/ und /kn/ unter prosodischer Variierung

Sieben Sprecher wurden mit Hilfe der Elektopalatographie (EPG) aufgenommen. Für EPG wird mithilfe einer künstlichen Gaumenplatte, die anhand eines Oberkieferabdrucks erstellt wird und mit elektrischen Kontakten versehen ist, der Kontakt zwischen Zunge und hartem Gaumen gemessen. Daraus folgt, dass nur Sprachlaute mit linguopalatalem Kontakt gemessen werden können. Für andere, z.B. Labiale ist das Verfahren nicht geeignet. Das Sprachmaterial enthielt zweisilbige Zielwörter mit den initialen Clustern /sk/, /ks/, /kl/ und /kn/: zwei Zielwörter pro Cluster, eines mit Betonung auf der ersten, das andere mit Betonung auf der zweiten Silbe (/ks/ bildet die Ausnahme dann kein angemessenes Testwort mit Betonung auf der zweiten Silbe gefunden wurde). Die Zielwörter waren in verschiedene Trägersätze eingebettet, um unterschiedliche starke prosodische Grenzen unmittelbar vor dem Zielwort zu elizitieren.

Zunächst wurden nur die prosodisch unmarkierte Fälle im Hinblick auf Koordination im Sinne von Überlappung betrachtet. Dabei ergaben sich sprecherübergreifend für /kl/ und /kn/ stabile Muster. /kl/ wurde in allen Fällen mit sehr viel Überlappung produziert, während /kn/ mit sehr wenig Überlappung produziert wurde. Für /sk/ und /ks/ wurde ebenfalls wenig Überlappung beobachtet, allerdings waren die Ergebnisse für beide Cluster sehr variabel. Die Ergebnisse sind nicht vollständig mit den Vorhersagen im Einklang, aber zum Teil. Bemerkenswert ist der große Unterschied zwischen /kl/ und /kn/, der sich am ehesten durch das Prinzip der gesturalen Wiederherstellbarkeit erklären lässt: Starke Überlappung in /kn/ hätte zur Voraussetzung, dass sich das Velum senkt, was jedoch schwer möglich ist, da der velare Plosiv das Velum an Ort und Stelle hält. Ein zu frühes Absenken des Velums würde aber auch dazu führen, dass die für die Perzeption des Plosivs wesentlichen Lösungsgeräusche verloren gingen (siehe auch /pl/ vs. /pn/ in Kühnert, Hoole & Mooshammer, 2006).

Prosodische Variation wurde eingesetzt, um die Stabilität der beobachteten Muster zu überprüfen. Da bei /ks/ und /sk/ keine konsistenten Muster auftraten, beschränkte sich die weitere Analyse auf die Cluster /kl/ und /kn/. Drei prosodische Grenzkategorien wurden definiert, die (absteigend nach ihrer Stärke geordnet) folgendermaßen bezeichnet werden können: äußerungsinitial, phraseninitial und wortinitial. Es wird im folgenden die Dichotomie betont–unbetont in Bezug auf das Cluster verwendet. Betont bezeichnet die Fälle, in denen die erste Silbe betont ist, unbetont hingegen die, in denen die zweite Silbe, die nicht das Cluster enthält, betont ist. Es zeigte sich, dass die Koordination in /kn/ durch prosodische Variation beeinflusst wurde, indem die Überlappung an schwachen Grenzen größer war als an starken Grenzen. In /kl/ hingegen erwies sich die Überlappung als recht stabil. In beiden Clustern wurde der Plosiv konsistent gelängt, wenn dem Cluster eine starke Grenze voranging. In /kn/ ist die apikale Verschlussdauer im Nasal in betonten Fällen länger als in unbetonten. Dies ist nicht der Fall für den Lateral in /kl/.

Die Ergebnisse bestätigen, dass artikulatorische Stärkung auch in Clustern auftritt. Insbesondere wird das Modell der π-Geste unterstützt, da der Stärkungseffekt graduierlich abnimmt und nur den ersten Konsonanten betrifft. Die Tatsache, dass die Koordination in /kl/ weniger von prosodischer Variation beeinflusst wird als in /kn/, kann eventuell erklären, warum /kl/ in den Sprachen der Welt erfolgreich ist und warum beispielsweise im Englischen der Plosiv in /kn/ verloren gegangen ist.

Elektromagnetische Artikulographie von /kn/, /kl/, /ks/, /pl/ und /ps/ unter prosodischer Variierung

Vier Sprecher wurden unter Verwendung von elektromagnetischer Artikulographie (EMA) aufgenommen. EMA verwendet Spulen, die an den Artikulationsorganen angeklebt werden können. Verschiedene Magnetfeldgeneratoren induzieren elektrische Spannungen in die Spulen, anhand derer die Position der Spule im Feld bestimmt werden kann. Ein offensichtlicher Vorteil gegenüber EPG liegt darin, dass die Beschränkung auf linguo-palatale Artikulationen hier nicht gilt. Im ersten Schritt sollen die EPG Ergebnisse für /kl/ und /kn/ der vorangehenden Studie repliziert werden. Der zweite Schritt soll diese erweitern: In der vorangehenden Untersuchung zeigte sich, dass /kl/ eine bemerkenswerte Stabilität aufwies. Hier soll überprüft werden, ob eventuell /l/ besonders geeignet für gestische Überlappung ist. Daher wird der Vergleich mit dem einzigen anderen Cluster im Deutschen bestehend aus stimmlosem Plosiv+/l/ betrieben: /pl/. Zum weiteren Vergleich werden die Kombinationen beider Plosive mit einem anderen alveolaren Laut, dem Frikativ /s/, untersucht. Abgesehen von Aufnahmverfahren und Material entspricht das weitere Vorgehen im wesentlichen der vorangehenden Untersuchung, allerdings wurde hier neben der prosodischen Grenze und der Wortbetonung auch der Phrasenakzent variiert.

Die obigen Ergebnisse für /kl/ und /kn/ konnten mit EMA bestätigt werden. Im hinblick auf die anderen Cluster zeigt sich, dass Plosiv+/l/ Cluster (insbesondere /kl/) stärker überlappen als Plosiv+/s/ Cluster. Dieses Muster wird sowohl von gestischer Wiederherstellbarkeit als auch von der Sonoritätshierarchie vorhergesagt, ist allerdings im Widerspruch zum DAC Modell, dass die umgekehrte Vorhersage trifft.

Phrasenakzentuierung liefert in dieser Analyse keine konsistenten, interpretierbaren Ergebnisse. Die Analyse der Einflüsse von Wortbetonung und Stärke der prosodischen Grenze zeigt allerdings Ergebnisse, die sehr gut mit Modellen wie der π-Geste und ihrer Graduierlichkeit vereinbar sind. Der erste Konsonant wird durchweg an starken prosodischen Grenzen gelängt. Im Falle von /kl/, welches am stärksten überlappt und die kürzeste Gesamtdauer hat, erreicht der Grenzeffekt auch den zweiten Konsonanten. In betonten Silben ist der zweite Konsonant durchgehend länger als in unbetonten Silben. Im Falle von /kl/ wiederum finden sich auch Hinweise auf einen Einfluss von Wortbetonung auf den ersten Konsonanten. Dieses Ergebnis zeigt, dass das Konzept, das der π-Geste für Grenzen zu Grunde liegt, in ähnlicher Form auch für Prominenz anwendbar ist. Es ist interessant, dass der Einfluss der Wortbetonung in der EMA Analyse weitaus stärker zum Ausdruck kommt als in der EPG Studie. Die Gründe für diesen Unterschied sind unklar. Ein Ansatzpunkt wäre die Tatsache, dass sich Bewegungsabläufe der Zunge aus EMA Daten sehr viel besser darstellen lassen.

Die Koordination der Cluster unter prosodischer Variation erweist sich als stabiler in /kl/ und /ks/ als in /pl/ und /ps/. Eine mögliche Schlussfolgerung ist, dass /k/+alveolar Cluster eine rigidere Koordination aufweisen als /p/+alveolar Cluster, weil durch die relative Unabhängigkeit der Artikulatoren in letzteren eine größere Variabilität möglich ist.

Variation von Stimmhaftikeit und Artikulationsort in deutschen und französischen Plosiv+/l/ Clustern

Diese Analyse beruht auf EMA und akustischen Daten von jeweils 5 französischen und deutschen Sprechern. Es wurde bereits beschrieben (Hoole, Bombien, Kühnert & Mooshammer, 2009), dass die artikulatorische Koordination in deutschen Clustern in Abhängigkeit von der Stimmhaftigkeit des ersten Konsonanten variiert. So ist die gestische Überlappung in beispielsweise /kl/ weitaus geringer als in /gl/, mutmaßlich um der Aspirationsphase zeitlich Rechnung tragen zu können. Dieser Unterschied wurde für entsprechende französische Kontraste nicht gefunden. In dieser Studie wird ein Zusammenhang zwischen diesem Unterschied zwischen den beiden Sprachen und ihren Implementierungen des Stimmhaftigkeitskontrasts hergestellt, um Rückschlüsse über die laryngeal-supralaryngeale Koordination zu ziehen.

Im Französischen gibt es für den Stimmhaftigkeitskontrast in initialen Plosiven eine gute Übereinstimmung zwischen phonologischer Beschreibung und der Phonetik, d.h. phonologisch stimmhafte Plosive sind in der Regel voll stimmhaft, während phonologisch stimmlose Plosive stimmlos realisiert werden (z.B. /b/ = [b] und /p/ = [p]). Dies ist im Deutschen weniger trivial. Phonologisch stimmhafte Plosive sind in der Regel nicht stimmhaft und phonologisch stimmlose Plosive sind stimmlos und aspiriert (/b/ = [d̥]/[t] und /t/ = [tʰ]). Gemeinsam ist beiden Sprachen, dass /b/ keine Abduktions-/Adduktionsgeste aufweist im Gegensatz zu /p/, wo dies der Fall ist. Die Sprachen unterscheiden sich darin, dass in französisch /b/ die Stimmlippen vibrieren, aber nicht in deutsch /b/. Die stimmlosen /p/ der beiden Sprachen unterscheiden sich im *timing* der oralen mit der laryngealen Geste. In der Akustik lässt sich dieser Kontrast gut mit dem Maß der *voice onset time (VOT)* (Lisker & Abramson, 1964, 1967), also dem Zeitraum zwischen Plosivlösung und Einsatz der Stimmbandschwingungen charakterisieren.

Das vorliegende Sprachmaterial enthält Zielwörter mit den simplen Silbenköpfen /b, p, g, k/ sowie den Kombinationen aus den genannten Plosiven mit /l/. Diese Wahl beruht auf der obigen Beobachtung der auffallenden Stabilität der Plosiv+/l/ Cluster. Im folgenden wird ein Cluster als stimmlos bezeichnet, wenn der Plosiv stimmlos ist (z.B. /pl/), und als stimmhaft, wenn der Plosiv stimmhaft ist (z.B. /bl/).

Nachdem festgestellt wurde, dass es sich mit den vorliegenden Daten weitestgehend so verhält, wie es aus der Literatur bezüglich Überlappung und Stimmhaftigkeitskontrast erwartet wurde (und zwar für sowohl simple als auch komplexe Silbenköpfe), können folgende neue Beobachtungen gemacht werden:

1) Die Überlappung in französischen stimmlosen wie stimmhaften Clustern entspricht eher der in den deutschen stimmlosen Clustern. Ausgehend von der Annahme das der Unterschied im Deutschen durch die Aspiration motiviert ist, ist dieses Ergebnis überraschend, da es in der Regel keine Aspiration im Französischen gibt.
2) Aufgrund der akustischen Maße hat es den Anschein, dass die laryngeale Geste im Deutschen mit dem gesamten Cluster assoziiert ist, im Französischen jedoch nur mit dem Plosiv, nicht mit dem Lateral.

Die Arbeit zeigt, dass artikulatorische Koordination sprachspezifisch ist. Grundsätzlich erlauben die Daten zwei inkompatible Rückschlüsse:

1. Im Französichen ist die glottale Geste ein Merkmal des Plosivs, während sie im Deutschen eher mit dem gesamten Silbenkopf koordiniert ist (siehe auch Kehrein & Golston, 2004; Hoole, 2006).
2. Laryngeale Merkmale sind im allgemeinen mit Silbenköpfen assoziiert. Im Französischen bilden die analysierten Cluster aber keine komplexen Silbenköpfe, sondern sind heterosyllabisch aufzufassen (Gafos, Hoole, Roon & Zeroual, 2010).
3. Stimmhaftigkeit in Plosiven und Überlappung sind aus aerodynamischen Gründen inkompatibel. In diesem Punkt ähneln sich beispielsweise Französisch und

Marokkanisches Arabisch (Zeroual & Hoole, 2010).

Eine Analyse des C-Center Effekts ist notwendig, um diese Interpretationen zu stützen. Desweiteren sollten die Rückschlüsse auf die glottale Geste mit Transilluminationsdaten der Glottis belegt werden.

Schlussbemerkungen

Die Variation des segmentellen Aufbaus von Clustern ergab, dass /k/+Alveolar Cluster eine engere und weniger variable Koordination aufweisen als /p/+Alveolar Cluster. Bezüglich der Identität des zweiten Konsonanten, erwiesen sich Plosiv+/l/ Cluster als enger und stabiler koordiniert als Plosiv+/n/ oder Plosiv+/s/. Dementsprechend zeigte /kl/ auch die geringste Gesamtverschlussdauer. Die Ergebnisse sind mit den diskutierten Modellen im Wesentlichen kompatibel, wobei der Ansatz der Wiederherstelbarkeit zum Teil der fruchtbarste war.

Die Ergebnisse der prosodischen Variation unterstützen graduierliche Modelle wie die π-Geste. Effekte die durch prosodische Grenzen und Wortbetonung entstehen sind lokal und nehmen mit Entfernung von ihrem Zentrum ab. Auf Grund der engen Koordination von /kl/ zeigen sich in diesem Cluster auch Effekte auf den distalen Konsonanten. Die Koordination selber wird in /p/ Clustern stärker beeinflusst als in /k/ Clustern, sodass an starken prosodischen Grenzen weniger Überlappung auftritt.

In velar+/l/ und labial+/l/ Clustern wirkt sich die Stimmhaftigkeit nicht auf die Variabilität in der Koordination aus. Wohl aber ist der eingangs genannte Effekt deutlich vorhanden, nach dem stimmhafte Cluster im Deutschen stärker überlappen als stimmlose Cluster, aber nicht im Französischen. Dementsprechend weisen von allen hier untersuchten Clustern /gl/ und /bl/ die meiste Überlappung auf. Die weitere Evidenz spricht dafür, dass Kombinationen von velarem Verschluss und alveolarem Lateral die größte Stabilität aufweisen. Dies deckt sich mit der Beobachtung, dass /kl/ wie eingangs berichtet in den Sprachen der Welt häufiger vorkommt als viele andere initiale Cluster.

Ein wichtiger Punkt, der bei der sprachübergreifenden Untersuchung zum Vorschein kam, ist der, dass intergesturale Koordination gerade auch über verschiedene artikulatorische Ebenen hinweg sich in verschiedenen Sprachen sehr unterschiedlich manifestieren kann. Desweiteren lassen sich mehrere Folgeuntersuchungen aus diesen Ergebnissen begründen, z.B. eine Ausweitung der sprachübergreifenden Studie durch prosodische Variation, um die Stabilität unterschiedlich-stimmhafter Cluster miteinander zu vergleichen, am besten unter Zuhilfenahme der Transilluminationstechnik, um besseren Einblick in die laryngeale Aktivität zu gewinnnen. Andere Möglichkeiten beinhalten die Miteinbeziehung anderer Cluster, um weitere Aufschlüsse über motorische Einschränkungen bzw. Freiheitsgrade zu treffen.

Literatur

Browman, C. & Goldstein, L. (1986). Towards an Articulatory Phonology. *Phonology Yearbook*, *3*, 219–252.

Browman, C. & Goldstein, L. (2000). Competing constraints on intergestural coordination and selforganization of phonological structure. *Les Cahiers de l'ICP*, *5*, 25–34.

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

Chitoran, I., Goldstein, L. & Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian. In *Laboratory phonology 7* (S. 419–448). Berlin/New York: Mouton de Gruyter.

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

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

Gafos, A., Hoole, P., Roon, K. & Zeroual, C. (2010). Variation in timing and phonological grammar in Moroccan Arabic clusters. In C. Fougeron, B. Kühnert, M. d'Imperio & N. Vallé (Hrsg.), *Laboratory phonology 10: variation, detail and representation* (S. 657–698). Berlin/New York: Mouton de Gruyter.

Hoole, P., Bombien, L., Kühnert, B. & Mooshammer, C. (2009). Intrinsic and prosodic effects on articulatory coordination in initial consonant clusters. In G. Fant, H. Fujisaki & J. Shen (Hrsg.), *Frontiers in phonetics and speech science* (S. 275–286). Beijing: Commercial Press.

Hoole, P. (2006). *Experimental studies of laryngeal articulation. Part 2: Larygeal-Oral coordiration in consonant sequences.* (Unpublished habilitation thesis, Ludwig-Maximilians-Universität, Munich). **retrieved** 26 März 2010, **from** `http://www.phonetik.uni-muenchen.de/~hoole/pdf/habilpgg_chap_all.pdf`

Kehrein, W. & Golston, C. (2004). A prosodic theory of contrast. *Phonology*, *21*, 1–33.

Kühnert, B., Hoole, P. & Mooshammer, C. (2006). Gestural overlap and c-center in selected French consonant clusters. In H. Yehia, D. Demolin & R. Laboissière (Hrsg.), *Proceedings of the 7th international seminar on speech production* (S. 40–48). UFMG. Belo Horizonte.

Lisker, L. & Abramson, A. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word*, *20*, 384–422.

Lisker, L. & Abramson, A. (1967). Some effects of context on voice onset time in English stops. *Lang Speech*, *10* (1), 1–28.

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

Selkirk, E. O. (1984). On major class features and syllable theory. In M. Aronoff & R. Oehrle (Hrsg.), *Language sound and structure* (S. 107–136). Cambridge: MIT Press.
Sievers, E. (1901). *Grundzüge der Phonetik*. Leibzig: Breitkopf und Härtel.
Zeroual, C. & Hoole, P. (2010). Mixed voicing and temporal overlap: EMA and aerodynamic observations from Moroccan Arabic. Poster presented at 12th Conference on Laboratory Phonology, Albuquerque.
Öhman, S. E. (1967). Numerical model of coarticulation. *Journal of the Acoustical Society of America*, *41* (2), 310–320.

Der Grund war nicht die Ursache sondern der Auslöser.

Franz Beckenbauer

Chapter 1

Introduction

This dissertation is an experimental investigation of how prosodic and segmental factors influence the production of word-initial German consonant clusters. Of fundamental concern here is the fact that some clusters appear to be more frequent (and maybe successful) than others in the language of the world that allow consonant clusters. For example the sequence /kl/ is readily attested in many languages, e.g. Latin 'clavis' "key": French 'clé/clef' [kleː/klɛf], Catalan 'clau' , Spanish (music) 'clavo' [klavo], Polish 'klucz' [klutʂ], Russian 'ключ' [klʉt͡ɕ], but Italian 'chiave' [kjave], Spanish 'llave' [ʎaβe]; English "claw" [klɔː]: German 'Klaue' [klaʊə], Swedish 'klöv' [kløv]. /kn/ appears to be less common, e.g. English "knee" [niː] but German 'Knie' [kniː], Swedish 'knä' [knæ], Icelandic 'hné' [n̥jeː], French 'genou' [ʒenu], Latin 'genu' ; English "knot" [nɒt] but German 'Knoten' [knoːtən], Swedish 'knut' [knʉːt], Icelandic 'hnútur' [n̥uːtʏr̥], Latin 'nodus' , French 'nœud' [nø]. The question is, now, whether it is possible to identify measurable properties in clusters that determine if a cluster will be able to remain stable diachronically. In other words: What makes a cluster a *good* cluster? In order to get to the bottom of this question Section 1.1 will review some background on the structure of consonant clusters. The results presented later will frequently touch concepts of the framework of Articulatory Phonology. Therefore a brief introduction to Articulatory Phonology with some emphasis on consonant clusters will be given in Section 1.2. Section 1.3 will give an overview of recent research on consonant clusters that is especially relevant to the theme of this work. The research aims will be summarized in 1.4.

1.1 Basic observations in consonant clusters

In general, groups of consonants without an intervening vowels or speech pauses are called consonant clusters[1]. A much observed and reported matter in phonological literature is that consonant clusters appear to follow so-called sonority hierarchies or, reversely, scales of consonantal strength in the way they are made up. Vennemann (1988) presents a scale of increasing consonantal strength as depicted in Figure 1.1.

↑ increasing consonantal strength

- voiceless plosives
- voiced plosives
- voiceless fricatives
- voiced fricatives
- nasals
- lateral liquids (*l*-sounds)
- central liquids (*r*-sounds)
- high vowels
- mid vowels
- low vowels

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

Similar scales are widely discussed in the literature (e.g. Sievers, 1901; Selkirk, 1984; Ohala & Kawasaki-Fukumori, 1997; Wright, 2004). Consonant clusters tend to follow such scales in that consonantal strength falls–or sonority rises–from the onset to the nucleus of a syllable. Reversely, consonantal strength rises from the nucleus to the coda of a syllable (cf. Head Law and Coda Law in Vennemann, 1988). Accordingly, clusters such as /#kl/ are more commonly found in the languages of the world than the reverse /#lk/ (where # denotes a syllable boundary). Sonority sequencing has been criticized by e.g./ Ohala and Kawasaki-Fukumori (1997) as being at best descriptive and not explanatory since it stands and falls with the position of the syllable boundary. Circularity arises as soon as such a boundary position has to be derived from sonority scales. Ohala and Kawasaki-Fukumori (1997) therefore propose to discard the concept of sonority (or strength) and replace it by a measure that descibes the degree of modulation in a number of acoustic parameters (amplitude, periodicity, spectral shape, F0).

[1] It should be noted from the beginning that the analyses presented in this thesis only deal with word-initial consonant clusters.

This measure should then be proportional to the survivability of a given segmental sequence.

Coming from a different direction than plain phonological categorization, Mattingly (1981) arrives at a very similar scale. Underlying his reasoning is the assumption that by coarticulation or overlap between speech gestures the transmission of information is aided by the simultaneous availablity of information on multiple sounds. This parallel transmission can be maximal in CV structures since Öhman (1967) found that the articulatory movements for the consonant and the vowel are initiated simultaneously. This is possible since in the vowel the vocal tract is not critically strictured and nothing gets in the way of the emergence of acoustic correlates. The situation is different for sequences of consonants. In the worst case two consecutive stops are likely to block each other; therefore, overlap must be less than in CV sequences. Mattingly links the extent to which sounds block each other to the closeness of articulation, i.e. to the constriction strength which easily translates to manner of articulation.

1.2 Consonant clusters in Articulatory Phonology

Articulatory Phonology is a theory developed by Browman and Goldstein (1986, 1988, 1989, 1990, 1992a, 1992b, 1995, 2000) over the last 25 years. Its core concept bears on the notion that the units of phonological contrast are so-called gestures. A gesture is understood as both a unit of phonological contrast as well as an *action unit* meaning that a gesture is at the same time representational and motor-executive. Gestures are specified by tract-variables which are associated to the articulators whose movements determine the value of the variable. Gestures are organized on tiers each of which is represented by a set of tract-variables. There are three oral tiers: 1. the tongue body has two variables, one for the constriction location, the other for constriction degree. The associated articulators are the tongue body and the jaw. 2. the tongue tip is specified almost identically with the crucial difference that the associated articulators of the tract variables are tongue body, jaw and tongue tip. 3. the lips are specified by lip aperture and lip protrusion. Upper and lower lips as well as the jaw are the relevant articulators. The velic tier has one tract-variable, velic aperture, which is associated to the velum, The glottal tier also has one tract-variable, glottal aperture, which is associated with the glottis

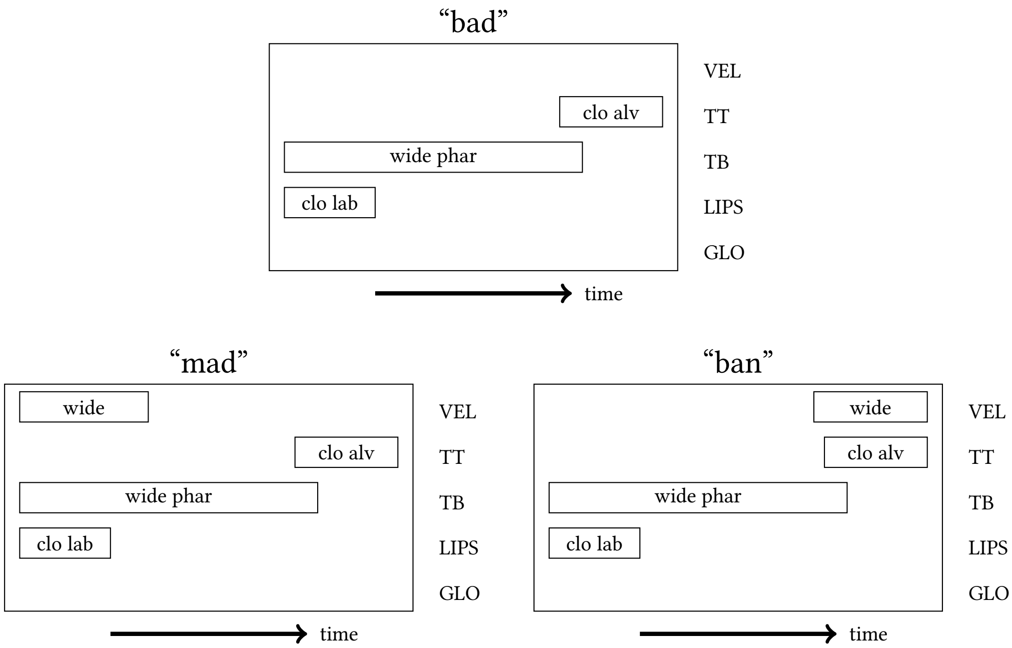

Figure 1.2: Gestural scores for the English words "mad" (upper left panel), "ban" upper right panel) and "bad" (lower panel).

Gestures are arranged in *gestural scores* to form utterances which specify when gestures are active in time. Figure 1.2 shows gestural scores of the words "bad", "ban" and "mad". They serve as an example for the use of gestures as units of phonological contrast. In traditional approaches, both "bad" and "ban" as well as "bad" and "mad" would be considered minimal pairs that only differ in one feature. The first pair differs in the manner of articulation of the coda consonant (alveolar stop vs. nasal), the second in the manner of the onset consonant (bilabial stop vs. nasal). "mad" and "ban" however would be considered to differ in both the onset as well as the coda consonant and are therefor not a minimal pair. Articulatory Phonology would state the difference between the first two pairs as the presence vs. absence of a velic gesture. Furthermore, it would consider the third pair minimal as well, since the difference can be stated by the temporal location of the velic gesture. Phonological contrast is therefore directly linked to physical events, which explains the attractivity of this gestural approach. Articulatory Phonology has also been used to explain various phenomena of running speech by gestural overlap, e.g. assimilation and deletion processes.

In contrast to other theories, Articulatory Phonology incorporates time. Gestures as action units are modeled as critically dampened mass-spring systems. According to the Task Dynamics framework (Saltzman & Kelso, 1987), the gestures' tract variables are equipped with a set of linear second order dynamics such that the damping ratio, the stiffness and the target position are entailed in the activation of tract variables (Saltzman & Munhall, 1989).

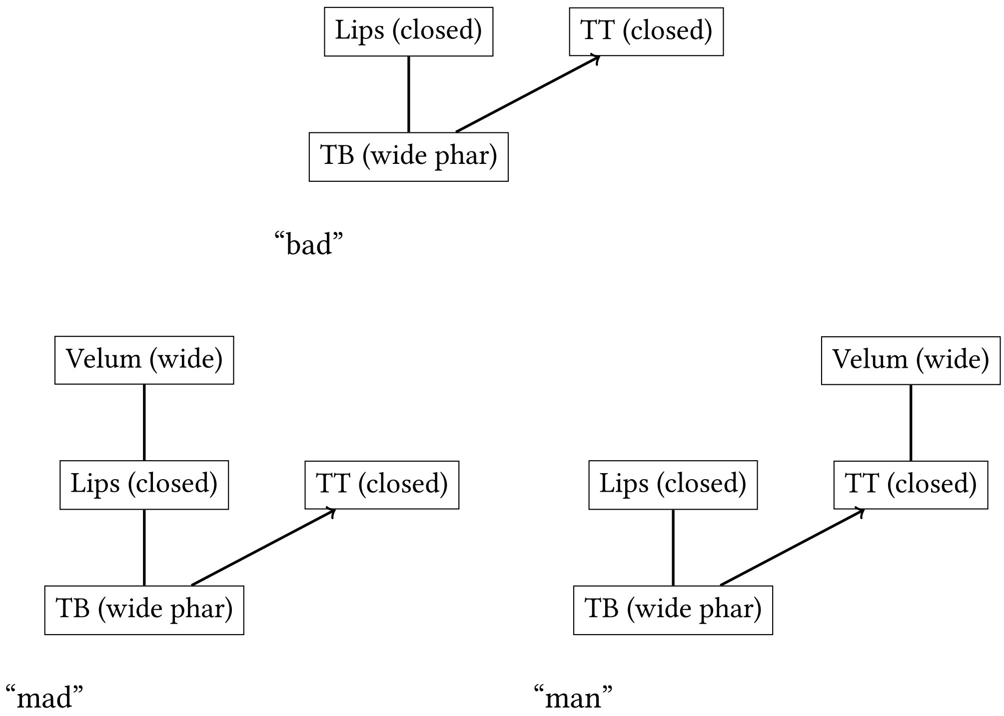

Figure 1.3: Coupling graphs for the words "bad", "ban" and "mad". —: in-phase coupling; ⟶: anti-phase coupling.

Sequentiality of gestures is in recent contributions to the gestural framework modeled in terms of coupled oscillators.[2] Gestures are associated with a planning oscillator that is coupled to adjoining gestures. Two coupling modes have emerged as being most natural: the most stable mode is the 0° / in-phase mode which applies for synchronous coordination. Simple CV structures have been observed to be synchronously coordinated (Öhman, 1967). This can only work in CV sequences since vowel gestures have lower stiffness than consonant gestures. VC and CC sequences are consequently cou-

[2]The theory of coupled oscillators can be traced back to Huygens, a Dutch physicist who in 1673 observed that two pendulum clocks mounted in close vicinity to the same wall oscillated synchronously. In fact, they resumed synchronicity a short while after being manually desynchronized.

pled anti-phase (180°) to ensure their Sequentiality. Figure 1.3 shows so-called coupling graphs for the same utterances as in Figure 1.2. Coupling graphs represent a planning system. The system attempts to settle such that all requirements are met in the best possible way. The output is a gestural score. The cases presented in Figure 1.3 appear trivial in this regard. However, the model assumes that all onset consonants are in-phase with the nucleus.

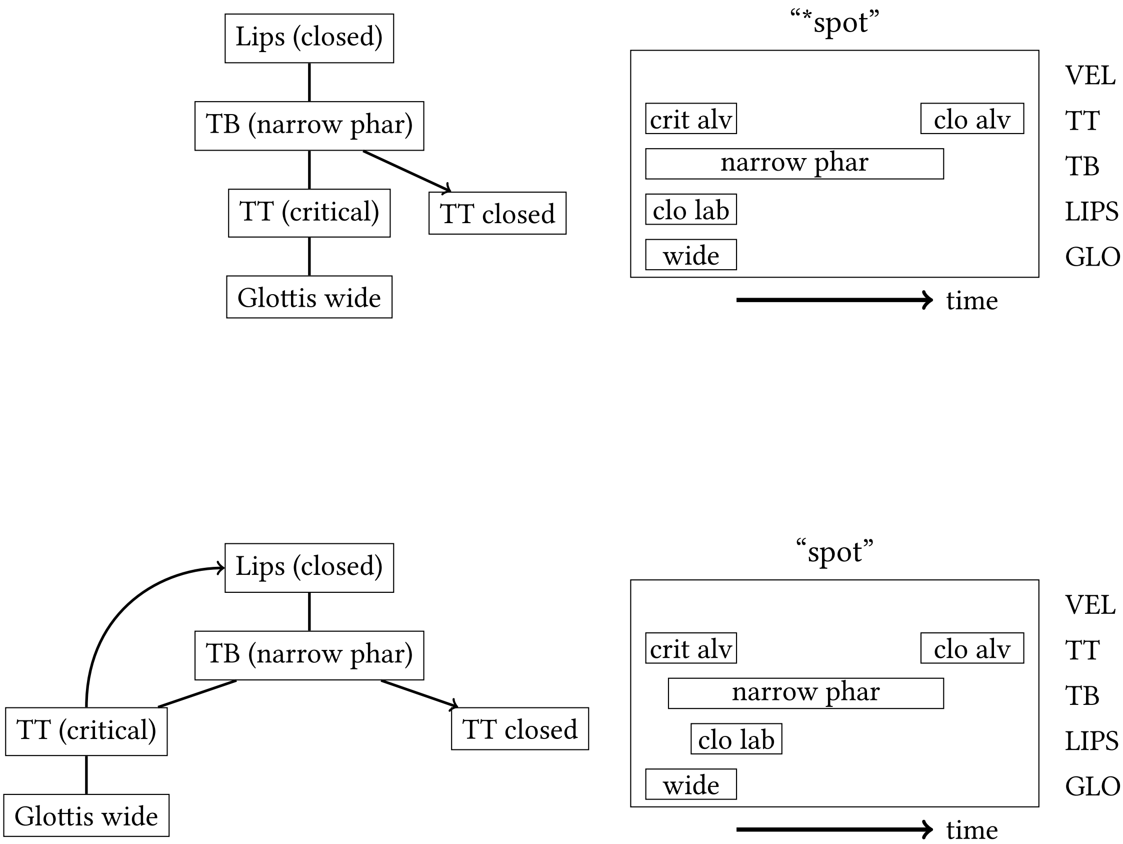

Figure 1.4: Coupling graphs and corresponding gestural scores of the utterance "spot". Upper panel: in-phase coupling of all onset consonants with the vowel; lower panel: as above but with additional anti-phase coupling between onset consonants.

Figure 1.4 illustrates coupling relations in onset clusters. The upper panel shows in-phase coupling between all onset consonants and the nucleus. The result is irregular since the complete overlap of the onset consonants inhibits sequentiality and renders the cluster irrecoverable. This problem is, however, circumvented by anti-phase couplings in consonant sequences as mentioned above. This is displayed in the lower panel of Figure 1.4. Clusters present the less trivial case for the system of planning oscillators to settle since there are conflicting constraints (Browman & Goldstein, 2000) that have to be resolved. In a real-life oscillatory system, a set of anti-phase coupled oscillators cannot simultaneously be coupled *in phase* to another. In the present case, the two gestures compete with each other for an in-phase relation to the nucleus on the planning level. Their anti-phase coupling with each other, however, results in a leftward shift of of the first (away from the vowel gesture) and a rightward shift of the second consonant (into the vowel gesture). Crucially, the center point of the gestures remains the same as if the shifts had not taken place. The model thus accounts for the C-center effect (Byrd, 1995; Honorof & Browman, 1995; Marin & Pouplier, 2010). Clusters in syllable codas do not have competing timing relations. Therefore onset clusters are considered to be timed more rigidly, allowing for less timing variability and being less overlapped than heterosyllabic or coda clusters (Byrd, 1996b; Hardcastle & Roach, 1979). For a more thorough overview of Articulatory Phonology and related models see Pouplier (2011).

1.3 Recent research on consonant clusters

1.3.1 Gestural coordination and segmental identity

It is not the case that gestural coordination in consonant clusters is the same regardless of the segmental make-up. One example is the well established place-order-effect which denotes the observation of greater overlap in front-back clusters as compared to back-front clusters (Hardcastle & Roach, 1979; Byrd & Tan, 1996, English stop-stop sequences; Chitoran, Goldstein, & Byrd, 2002, Georgian in word-initial and word-medial positions; Zsiga, 2000; Kochetov, Pouplier, & Son, 2007, Russian and Korean across word and higher boundaries). Simple manner-based approaches as outlined in 1.1 cannot account for such patterns. One concept for explaining this behavior is that of gestural recoverability (Chitoran, 2002; Silverman, 1997) which entails the assumption that the overlap between gestures must not inhibit the correct perception/recoverability of the gestures. In the case of the place-order-effect it is easily appreciable that in a sequence of e.g. /kt/ the recoverability of /k/ is only warranted up to a certain degree of overlap. If apical closure occurs before velar release the acoustical correlates for correctly identifying /k/ can not emerge. In the reverse case, there is no obstacle in the vocal tract to block /t/ release cues. However, the place-order-effect does not appear in all positions where it is conceivable. Gafos, Hoole, Roon, and Zeroual (2010) frame the relativized place order hypothesis according to which an overlap difference due to the place-order-effect can only be expected if there is enough overlap to begin with. Recently, an (extensive) extension to the coupled oscillator model has been proposed (Nam, 2007; Goldstein, Chitoran, & Selkirk, 2007; Goldstein, Nam, Saltzman, & Chitoran, 2009). Split gesture dynamics assign two planning oscillators to every gesture: one for its onset and one for its release movement. This facilitates a more fine grained specification of timing relations: in a CC sequence, anti-phase coupling from C_1 release to C_2 onset enforces less overlap than a release-to-release coupling (or a plain gesture-to-gesture) coupling would. Goldstein (2009) use this approach to account for place-order related coordination patterns in Georgian clusters and propose a similar procedure for German /kl/ vs. /kn/ clusters (Hoole, Bombien, Kühnert, & Mooshammer, 2009).

Another attempt to capture effects of segmental make-up on temporal organization in terms of coarticulation was undertaken by Recasens and Pallarès (2001) utilizing the DAC (degree of articulatory constraint) model of lingual coarticulation (Recasens, Pallarès, & Fontdevila, 1997, *et seq.*). This model classifies speech sounds by the extent of coarticulation they exert on neighboring sounds or, inversely, by the extent to which they are sensitive to coarticulation they are exposed to by neighboring sounds. Classification is accomplished by assigning DAC values to speech sounds as a function of dorso-palatal involvement (high values indicate strong, low values indicate weak resistance to coarticulation). Coarticulation is assumed to be larger the higher the difference of the DAC values of neighboring sounds. This approach is discussed at length in Chapters 2 and 3.

1.3.2 Gestural coordination and prosody

Quite a number of studies have shown that prosodic structure is reflected temporally and spatially in articulation (Pierrehumbert & Talkin, 1992; Beckman, Edwards, & Fletcher, 1992; Dilley, Shattuck-Hufnagel, & Ostendorf, 1996; Fougeron & Keating, 1997; Turk & White, 1999; Byrd & Saltzman, 2003; Cho & McQueen, 2005; Byrd, Lee, Riggs, & Adams, 2005; Lee, Byrd, & Krivokapić, 2006; Keating, 2006; Byrd, Krivokapić, & Lee, 2006; Krivokapić, 2007; Kuzla, Cho, & Ernestus, 2007; Cho & Keating, 2009; Byrd

& Choi, 2010). Pierrehumbert and Talkin (1992) analyzed effects of prosodic position and prominence on the articulation of /h/ and /ʔ/. They found that gestural magnitude is affected by both, and associate accent with a shift in a vocalic direction and boundaries with a shift in a consonantal duration. Crucially, they show that the effects of phrase boundaries are not limited to the preboundary scope, e.g. final lengthening (Edwards, Beckman, & Fletcher, 1991), but also appear in postboundary position. Postboundary effects have then been studied extensively by e.g. Fougeron and Keating (1997), Fougeron (2001), Keating, Cho, Fougeron, and Hsu (2003). Generally, consonants in postboundary position tend to have longer and stronger constrictions. Within the framework of Articulatory Phonology, boundary induced effects on articulation are accounted for in terms of the π-gesture (Byrd & Saltzman, 2003). This approach assumes an additional, prosodic tier in the gestural frame work on which prosodic gestures (π-gestures) can be activated. The activation of a π-gesture slows down the execution of other active gestures in the gestural score (much like a fermata in musical scores). The activation level of of a π-gesture first waxes towards the peak activation (located at the position of the boundary) and than wanes again which entails that boundary effects are of a graded nature. Byrd (2006) show indeed that boundary effects exist beyond the immediately adjacent segments but are of much lower magnitude.

Importantly, (Byrd & Choi, 2010) have also taken consonant clusters into consideration. They report effects of boundaries on intra-gestural and inter-gestural parameters in preboudary and postboundary tautosyllabic clusters as well as in heterosyllabic clusters spanning the boundary. Their results confirm the models prediction for (intra-gestural) strengthening effects also with regard to the gradedness of the boundary effect. Concerning inter-gestural parameters, they find that timing in heterosyllabic clusters is very much subject to prosodic variation (less overlap at strong boundaries). To a lesser degree, this effect was also found in coda clusters. Timing in onset clusters, however, is least affected by variation of boundary strength.

Apart from boundaries, prominence has also been shown to affect articulation. Turk and White (1999, among others) have shown that the domain of accentual lengthening is not limited to the stress-bearing nucleus. Instead, accentual lengthening can spread rightward (but not leftward) within the prosodic word. In an EPG and acoustic study, Cho and Keating (2009) showed that prosodic position and prominence affect articulation mostly on different scales. However, they found a cumulative effect on the contact duration of word initial consonants of position and stress and also very weakly of accent.

1.3.3 Gestural coordination beyond the oral tiers

In Articulatory Phonology the timing of glottal gestures in consonant clusters was dealt with right at the beginning. Browman and Goldstein (1986, p. 228) state two rules for the coordination of laryngeal and supra-laryngeal events:

1. If a fricative gesture is present, coordinate the peak glottal opening with the midpoint of the fricative.
2. Otherwise, coordinate the peak glottal opening with the release of the stop gesture.

Both rules are stated under the assumption that there may not be more than one glottal gesture in a syllable onset. So far, this has only been formalized in the English model of coupled oscillators in terms of a 90° delay of the glottal gesture for voiceless stops as compared to fricatives (0°). Furthermore, there are exception rules which cause the deletion of glottal gestures a) of stops in fricative-stop onset clusters (/sp, st, sk) and b) of /f/ in /sf/ onset clusters (see Nam, Goldstein, & Proctor, 2007). While the acoustic output of this model may be all right, it is not guaranteed that the model indeed mirrors the laryngeal-oral coordination of human speech. At least for simple onsets, the split gesture account might be more promising here: glottal gestures might be coupled to the constriction gestures in fricatives but to the release gesture in stops.

Hoole (2006) investigated laryngeal-oral coordination in a large inventory of syllable onsets including a number of consonant clusters by means of photoelectroglottography/transillumination. The results support the assumption that onsets do not have more than one glottal gesture. However, the support for the rules for oral-laryngeal timing find only limited support. Hoole therefore argues for an approach that captures coordination relations "in terms of the fulfillment of a set of constraints given by the aerodynamic and functional demands of each specific syllable onset" (Hoole, 2006, p. 145). For clusters with /l/, one stable finding emerged beside a variety of movement patterns that were due to inter- and intra-speaker variability, as well as the segmental make-up. By example, /pl/ was always found to be produced with more VOT than simple /p/. It is argued that this pattern emerges as a consequence of speakers exploiting physical forces that arise from aerodynamics. Crucially, this propagation of voicelessness into the sonorant is not in conflict with sonority modulation that typically underlies syllable structure.

The material analyzed by Hoole does not allow for an analysis of how the presence or absence of a glottal gesture affects inter-gestural timing of oral gestures. However, data presented in an EMA study (electromagnetic articulography) by Hoole (2009) show that there is a significant timing difference in German stop+/l/ clusters as a function of whether the stop is underlyingly voiced or voiceless (less overlap in the voiceless case). Interestingly this difference does not exist in French clusters. This topic will be revisited in detail in Chapter 4.

1.3.4 Gestural coordination and grammar

It has recently been shown that grammar significantly impacts the temporal organization. Gafos (2010) convincingly shows within the framework of Articulatory Phonology, that gestural coordination in Moroccan Arabic consonant clusters is determined by whether a cluster emerges as a result of templatic or affixal morphology. This research was enabled by the temporal nature of Articulatory Phonology.

In a related study that also uses Moroccan Arabic data, Shaw, Gafos, Hoole, and Zeroual (2009) show that the coordination relations stipulated by Articulatory Phonology can by utilized to determine the syllable affiliation of the components of initial consonant clusters. It has been much debated whether consonant cluster in Moroccan Arabic are to be considered as complex or as a sequence of simplex onsets. Shaw approach this problem assuming that complex onsets should exhibit the C-center effect. C-center alignment of consonants with respect to the following vowel is displayed in the right panel of Figure 1.5. Regardless of the number of onset consonants, their center point is stably aligned with a predefined, fixed anchor point (see also the introduction to the C-center effect above). A sequence of simplex onsets on the other hand should be aligned differently. Here, the most stable timing relation should be the interval from the right edge of the cluster (right edge of the rightmost consonant) to the same predefined fixed anchor point. This is displayed as *right edge alignment* in the left panel of

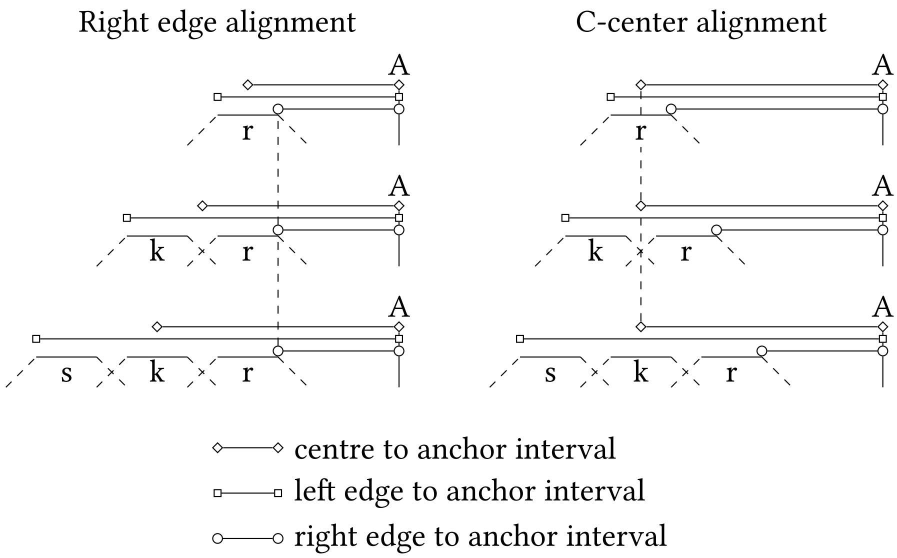

Figure 1.5: Schematic display of "right edge" and "C-center" alignment in initial consonant clusters (from Shaw, Gafos, Hoole, & Zeroual, 2009)

Figure 1.5. For right edge alignment, additional consonants that are prepended to the cluster do not affect the timing of the clusters closer to the vowel. For C-center alignment, adding a consonant shifts the other consonants such that the C-center relation is maintained. The analysis of experimental and simulated data provides evidence in support of the simplex onset hypothesis, i.e. clusters exhibit the right edge alignment rather than the C-center alignment.

Similarly, Hermes, Grice, Mücke, and Niemann (2008) present articulatory evidence that shows how the morphology of Italian *impure /s/* is reflected in gestural timing. While clusters in utterances like "la prima" do seem to exhibit the C-center alignment, an added impure /s/ as in "la sprima" does not lead to a shift of the following consonantal gestures as would be appropriate for C-center alignment. Instead impure /s/ appears to be prepended to the cluster as an additional simplex onset.

1.4 Research aims

In this work, the attempt is made to probe clusters in several ways in order to find out which coordination patterns emerge as stable. By trying to identify the properties of the cluster that can be made responsible for the stability it maybe that some information is obtained as to what makes a cluster a *good* cluster. Three ways of probing gestural coordination will be tested and discussed. 1. varying the segmental make-up of the cluster, 2. varying the prosodic environment in which a cluster appears and 3. varying glottal activity in clusters.

In Chapter 2, the clusters /sk/, /ks/, /kl/ and /kn/ will be analyzed for timing difference induced by segmental make-up and combinations of boundary strength and lexical stress using EPG (electropalatography). Chapter 3 differs from Chapter 2 in that the set of clusters is changed to include /kn/, /kl/, /ks/, /pl/ and /ps/. This change is motivated by some results of Chapter 2 and facilitated by the use of EMA (electromagnetic articulography) instead of EPG. Chapter 4 takes an entirely different approach and investigates the effect of mixed-voicing vs. full voicing on the coordination of consonant clusters cross-linguistically in French and in German stop+/l/ clusters. Chapter 5 will briefly summarize and discuss the results with respect to the research aims stated here.

Wir können so was nicht trainieren, nur üben.

MICHAEL BALLACK

Chapter 2

Segmental and prosodic effects on the articulatory coordination of word-initial consonant clusters[1]

2.1 Introduction

Speech is produced by a highly intricate interplay between articulatory actions whose underlying principles are far from being fully understood. For example, the gestural coordination of a sequence of two consonants C_1 and C_2 has been found to vary between the two extremes of total synchronicity and a very long delay. Depending on the gestures involved, the first extremum (i.e. total synchronicity / overlap) may have the following results: Assimilation and the perceptual loss of one of the consonants, diachronic metathesis of the consonants (Blevins & Garrett, 2004), or a complex doubly articulated segment (Maddieson, 1993). The opposite extremum of unconstrained delay might lead to the perception of intrusive vowels (Hall, 2003; Davidson, 2005; Davidson & Roon, 2008) for voiced consonant sequences.

This paper discusses two factors affecting the internal coordination of clusters: cluster type and prosodic variation. Our study aims at investigating production and perception related aspects contributing to the internal structure of clusters by means of

[1]A version of this chapter is available in the Journal of Phonetics (Bombien, Mooshammer, Hoole, & Kühnert, 2010). Further information in appendix B.

the temporal analysis of physiological tongue-palate contact measurements during the word-initial clusters /kl/, /kn/, /sk/ and /ks/ of 7 speakers of German. The stability of the observed patterns is furthermore tested by using prosodic variation as a probe. Boundary strength and lexical stress (confounded with accent) are varied orthogonally in order to achieve this. For several reasons, consonant coordination patterns are discussed here with regard to word-initial clusters only: Clusters in other positions have been reported to show different coordination patterns (see e.g. Browman and Goldstein (2000), Marin and Pouplier (2010)). In this current study, however, we focus on segmental composition and prosodic variation. Furthermore it has been found that final and heterosyllabic clusters are more variable in general. Since we expect only subtle prosodic effects we preferred to analyze the more stable word-initial position. The third rationale for using initial clusters is that German does not show place assimilation in this position whereas place and manner assimilations are frequently found at morpheme boundaries (Bergmann, 2008) and in word final position (Kühnert & Hoole, 2004). Clusters in word medial and word final position therefore do not play a role in this study.

2.1.1 Cluster Type

Three principles that seem to underlie and govern the temporal organization of speech gestures will be discussed in this paper[2]: manner-based ranking of overlap, recoverability of segmental content, and biomechanical/anatomical constraints. These three approaches give different reasons for observed differences in timing. They do not necessarily differ with respect to their predictions. The first principle is based on Mattingly's (1981) assumption that coarticulation, or more specifically overlap between gestures, assists the transmission of information in that information about multiple gestures is available simultaneously. This parallel transmission is supposed to increase the speed of transmission because by overlapping gestures more sounds can be uttered within an allotted time frame. Furthermore, it facilitates the recognition and recovery of gestures because it results in a robust encoding of information in the signal (Wright, 2004). For CV sequences, parallel transmission can be maximal, i.e. gestures for both

[2]We are aware of the fact that other factors also affect the timing of gestures, such as language-specific constraints, grammar (see e.g. Gafos, 2002), word frequency and phonotactic probability (Vitevitch, Armbruster, & Chu, 2004).

the C and V elements were found to be initiated simultaneously. In his seminal work, Öhman (1967) showed for VCV sequences that the gesture for the second vowel is even initiated before the consonant's gesture.

For consonant sequences, however, overlap must not prevent the emergence of acoustic correlates of any of the involved constrictions. Mattingly (1981) suggested that this restriction corresponds to the degree to which one segment allows encoding of information on the overlapping segment. In his view this lower bound of overlap follows quite neatly from the constriction degree. Accordingly, the segments with the closest constriction, the obstruents, allow the least amount of overlap. Nasals, liquids, glides and vowels permit increasingly more overlap in this order. This manner-based ranking of consonant classes also resembles sonority hierarchies as proposed by e.g. Sievers (1901) and Selkirk (1984).[3]

Further evidence for a constriction based ordering of overlap was found very recently by Kühnert, Hoole, and Mooshammer (2006): French stop + nasal clusters were produced with less overlap than stop + lateral clusters, which would also be predicted by Mattingly (1981). Violations of the sonority hierarchy within a syllable result in a more constrained phasing with a longer delay, as was for example found for word-initial stop-stop sequences in Georgian by Chitoran, Goldstein, and Byrd (2002).

There are, however, certain regularities found across languages, which cannot be explained by a manner-based hierarchy such as the very consistent place order effect. This term describes the finding that less overlap is permitted in clusters if the first segment is articulated at a place posterior to the following consonant (e.g. /kt/ or /kp/ clusters, henceforth called back-to-front) as compared to the opposite order (e.g. /tk/ or /pk/), everything else being equal. This regularity and its consequences for a universal preference of front-to-back clusters has been explained by perceptual recoverability. Since in a back-to-front sequence (e.g. /kt/) the first segment /k/ is produced posterior to the second segment (i.e. /t/), overlap can easily cause the complete deletion of the audible release of the first segment (i.e. /k/) by the following segment /t/. Hence, the recoverability of the first segment is obscured by the ongoing production of the following more anterior consonant. This situation is much less likely to occur

[3]Mattingly doesn't distinguish between stops and fricatives and combines them to the more general class of obstruents as does Sievers (1901). Selkirk (1984), on the other hand, attributes more sonority to fricatives than to stops.

in front-to-back clusters.The place order effect is extremely consistent across different languages and word positions (word-initial, word-medial, word-final, across boundaries), e.g. English stop-stop sequences across word boundaries (Hardcastle & Roach, 1979; Byrd & Tan, 1996, /d#g/, /s#g/ vs. /g#d/, /g#s/); Georgian in word-initial and word-medial positions (Chitoran, 2002, /dg/ vs. /gd/, /bg/ vs. /gb/, /$p^h t^h$/ vs. /t^hb/), Russian and Korean across word and higher boundaries (Zsiga, 2000; Kochetov, Pouplier, & Son, 2007, /pt/ vs. /kp/, /kt/). However, there seems to be a ceiling effect, meaning that only speakers who produce clusters with an overlap exceeding a lower threshold show a place order effect (see EMA results for Moroccan Arabic by Gafos, Hoole, Roon, & Zeroual, 2010). Since there are no stop-stop sequences in German in word initial position, the place-order effect cannot be tested with our data. However, this effect exemplifies that strictly manner-based approaches cannot account for all patterns. Evidence for more extensive overlap for word-medial stop-stop clusters as compared to fricative-stop and stop-fricative clusters was presented by Byrd and Tan (1996). The reason for the longer delay, if a fricative is a member of the cluster, could be that fricatives require a longer minimal stationary phase with friction noise in order to be correctly identified. According to Jongman (1989) an /s/ must have at least a duration of 50 ms in order to be identifiable. Similar findings have been presented by Meynadier, Pitermann, and Marchal (1998). This argument therefore again points in the direction of perceptual recoverability rather than a sonority-based account. Finally, Kühnert (2006, see above) do not attribute their findings to a manner-based ranking of overlap. The authors' account for the effect is that the place of the stop articulation might not be recoverable in stop + nasal clusters if the naso-pharyngeal port is opened before the stop is audibly released. In this case the only potential place cue in utterance-initial position would be distorted by nasal release because only insufficient air pressure can be built up for the production of a salient burst. This is not the case for a following lateral.

A third factor possibly governing the inter-gestural organization within clusters could be biomechanical linkages between articulators and their anatomical properties. For instance, for the difference in /tk/ vs. /kt/, Hardcastle and Roach (1979) suggest that for the tongue movement from /t/ to /k/ in /tk/ only the contraction of the *longitudinalis inferior* may be necessary, while higher complexity and extrinsic muscles are involved in what they call tongue repositioning for /kt/. If this is the case (there is hardly any data on the *longitudinalis inferior*) this could account for less co-production in the latter cluster. While the tongue repositioning account is also applicable to other oral clusters (e.g. /sʃ/ vs. /ʃs/) it does not cover observations of the place-order effect on clusters involving labials. This assumption of biomechanical linkages between articulators and their anatomical properties has been formalized within the DAC (degree of articulatory constraint) Model by Daniel Recasens with substantial evidence from EPG data mainly on Catalan (e.g. Recasens, Pallarès, & Fontdevila, 1997). The DAC model predicts that sounds produced with a high degree of articulatory involvement in the achievement of a constriction resist coarticulation from neighboring segments and at the same time exert coarticulation on these segments. This means that the coarticulatory resistance and exertion are inversely related to each other. For example, at the one end, sounds produced with active predorsal involvement, such as /s, ʃ/ and trilled /r/ or postdorsal retraction, as in dark /ɫ/, have high DAC values because they affect the neighboring segments to high degree but are only minimally influenced by them. At the other end, sounds like bilabials are specified with a low DAC value because during a labial the tongue is free to anticipate the position of the adjacent segments. According to Recasens and Pallarès (1999), dentals and alveolars, such as /t, d, n/ and clear /l/, exhibit an intermediate DAC level with the lateral showing a somewhat higher value than the others (Recasens, 2007) due to laterality requirements. With regard to the clusters analyzed in the current study, the DAC index would predict that /ks/ is produced with more overlap than /kn/ because /n/ has a lower DAC value than /s/ which exerts more coarticulation on /k/ (see Recasens & Pallarès, 2001). Clear /l/ should exert slightly more coarticulation on /k/ than /n/ due to laterality requirements as mentioned above. There have, however, been indications that German /l/ might be more resistant to coarticulation than clear /l/ in French or Spanish (Recasens, Fontdevila, & Pallarès, 1995). Accordingly, /l/'s DAC value may have to be adjusted upwards. Another view presented by Kühnert (2006) as an alternative to the perceptual recoverability account relates to the fact that, as opposed to the lateral /l/, the nasal /n/ is composed of two gestures, oral closure and velic opening. In terms of inter-articulator coupling, this added articulatory complexity might account for the observed timing differences. Unlike English /l/ (e.g. Sproat & Fujimura, 1993), German /l/ does not have a dorsal gesture.

It is a central concern of the DAC model to account for coarticulatory directionality. In the case of /sk/ and /ks/ – based on findings for the relative salience of the anticipatory and carryover effects from /s/ on /a/ in (Recasens, 1997) – it can be expected that

/s/ will exert stronger coarticulation on /k/ than *vice versa* in both cases. With regard to co-production, the DAC model makes use of another factor (Recasens, 1999, 2004; Recasens & Pallarès, 2001): Tongue repositioning, as outlined above, is needed in /ks/ as opposed to the production of /sk/. Therefore /sk/ is expected to show more overlap than /ks/. In summary, the predictions based on the DAC account yield a decrease of overlap in the following order: /sk/>/ks/>/kl/>/kn/.

2.1.2 Prosody

The second topic to be considered here is prosodic variation. It has been found in many studies (e.g. Pierrehumbert & Talkin, 1992; Fougeron & Keating, 1997; Bombien, Mooshammer, Hoole, Kühnert, & Schneeberg, 2006; Kuzla, Cho, & Ernestus, 2007; Cho, McQueen, & Cox, 2007) and for a number of languages that prosody affects the phonetic realization of segments depending on the type of prosodic variation and the segments involved. For example, prosodic phrasing generally induces a change in the temporal and spatial characteristics of the segments adjacent to the boundary, but not all segments are affected in the same way and to the same degree. For example, Fougeron and Keating (1997) and Keating, Cho, Fougeron, and Hsu (2003) found in an EPG study that lingual stops, laterals and nasals are lengthened and produced with more contact following higher boundaries. However, the fricative /s/ in French seemed to resist strengthening because of fewer articulatory and acoustic degrees of freedom. Similar interactions have been found for accent and stress: whereas tense vowels lengthen considerably in German when stressed and accented, for lax vowels only the quality but not the quantity is affected (Mooshammer & Fuchs, 2002; Hoole & Mooshammer, 2002). Applying these examples of segmental resistance to prosodic changes in the current study of consonant clusters, the question arises whether clusters are affected as a whole, i.e. the onset of the syllable as a phonological constituent, or as two independent components, i.e. sequence of consonants.

In this study, we investigate the influence of prosodic variation on initial clusters. Regarding the prosodic factors here, prosodic boundary strength and stress, two different theoretical approaches will be tested. Based on an acoustic study on realizations of /ʔ/ and /h/ in American English, Pierrehumbert and Talkin (1992) proposed that CV syllables become more consonant-like at phrasal junctures, i.e. the syllable onset lengthens and exhibits more consonant-like characteristics such as more frequent and longer glottalization. This view can also account for findings such as lower nasal air-stream for /n/ adjacent to higher boundary levels in French (Fougeron, 2001) making the nasal more obstruent-like. Accent in the Pierrehumbert and Talkin (1992) study shifts the syllable in a vocalic direction with longer durations and larger gestures. Further evidence for the differential mechanisms for signaling accent and boundary strength have been presented by e.g. Beckman, Edwards, and Fletcher (1992) and Cho and McQueen (2005). The latter, however, also provided counter-evidence to the observed strengthening effects from stop aspiration in Dutch with shorter VOTs at higher levels of prominence and prosodic boundaries. Within Pierrehumbert and Talkin's model prosodic effects vary according to the constituents of the syllable they enhance, i.e. the syllable onset is affected by prosodic boundaries and the nucleus by accent. However, no particular prediction concerning initial consonant clusters can be derived from this account.

Concerning boundaries a different view has been taken by Byrd and colleagues (e.g. Byrd, Kaun, Narayanan, & Saltzman, 2000; Byrd & Saltzman, 2003). They proposed that most of the phenomena related to phrase marking can be modeled by trans-gestural perturbations of clock rate due to a so-called π-gesture. This is an abstract non-tract prosodic boundary gesture that in earlier versions affected the stiffness of the trans-boundary gestures approximately proportionally to the boundary strength. Byrd and Saltzman (2003) replaced the stiffness approach with local clock slowing, generating temporal lengthening by lengthening the activation intervals of tract-variable gestures and the spatial strengthening by a lesser degree of overlap or truncation (see Harrington, Fletcher, & Roberts, 1995). However, it is not clear how shortening of VOT in Dutch (Cho & McQueen, 2005) and lesser velum lowering in French (Fougeron, 2001) at higher boundaries could be explained by π-gestures. An important feature of the π-gesture is that the activation strength varies smoothly, i.e. it waxes continuously towards the π-gesture's peak activation and then it wanes in a similar manner (Byrd & Saltzman, 2003; Byrd, Krivokapić, & Lee, 2006). Therefore, the prosodic effect on the constriction gestures - such as lengthening and strengthening - is strongest at the activation peak and diminishes with the distance from the peak. Generally, it has been found that temporal lengthening effects are more consistent than articulatory strengthening effects, especially when measured with EMA rather than EPG (see Keating, 2006, for an overview).[4] With respect to the current investigation the π-gesture approach would predict that the initial consonant of the cluster which is directly preceded by the boundary is affected to a greater extent than the second consonant, which is further removed from the boundary. Gestural overlap is expected to be affected in that the constriction gestures move farther apart from each other at high prosodic boundaries. Indeed, Byrd and Choi (2010) found in an EMA study of three speakers of American English that all speakers consistently lengthened the first consonant of /sp, sk, kl/ clusters for higher boundary levels. The effect on duration of the second element of these clusters was smaller and also less consistent but significant for two speakers. In an EPG study of French /kl/ clusters in two speakers, Fougeron (1998) found that effects were limited to the first consonant while the second consonant was only inconsistently influenced. Regarding the overlap between the consonants, in both studies initial clusters were relatively insensitive to prosodic changes. This gives room to the interpretation that consonants in initial clusters are more cohesive since stronger and more consistent timing effects attributed to prosodic variation were found in heterosyllabic and in coda clusters.

To our knowledge the π-gesture model has only been used for modeling the effects of prosodic boundaries. However, Saltzman, Goldstein, Holt, Kluzik, and Nam (2007) have already presented a proof of concept for the application of the π-gesture on the syllable level. Furthermore, given evidence from the literature that stress and accent are generally found to affect vowels to a greater degree than consonants (see e.g. Pierrehumbert & Talkin, 1992; Cho & Keating, 2007), the peak activation of the π-gesture for stress can be assumed to be positioned around the middle of the vowel with decreasing strength towards the onset and the coda of the stressed syllable. For accent the peak activation is probably again situated in the middle of the stressed syllable but - as was found by Turk and White (1999) - the effect spreads to the preceding and the following syllables in the same phonological word with more consistent lengthening effects on the following than on the preceding syllables. In our data, stress and accent are confounded, i.e. the initial and stressed syllable in *Claudia* also carries a pitch accent and the initial unstressed syllable in *Klausur* /klaʊ.ˈzuɐ/ precedes the accented syllable. If

[4] In yet a newer version by Saltzman, Nam, Krivokapić, and Goldstein (2008), the π-gesture is replaced by the more general μ (modulation) gestures which modulate two aspects of the vocal-tract gestures: μ_T-gestures modulate the temporal course of vocal-tract gestures such as the above described slowing down of the clock, and μ_S-gestures serve to model articulatory strengthening effects.

the π-gesture model can be applied to stress confounded with accent in the current data then the second consonant is affected by stress to a greater degree than the first one because it is closer to the peak activation of the π-gesture. We want to point out here that it is not the aim of the current study to test the π-gesture model in all its details or to implement the prosodic level stress in this model. Rather, the aim here is to provide and discuss a theoretical background for the extent and domain of prosodic effects on word-initial consonant clusters as a probe for the stability of internal structure of clusters.

2.1.3 Predictions

This section gives an overview of our predictions. Items a) - c) summarize the outcome of the three principles concerning segmental make-up as discussed in the introduction. Items d) and e) deal with prosodic variation.

a) Manner-based ordering would predict more overlap for /kl/ vs. /kn/ clusters. The same amount of overlap for /sk/ as for /ks/ can be expected (under the assumption that /s/ and /k/ have the same degree of sonority) but, as both violate the sonority sequencing constraint, less overlap can be expected than for /kl/ and /kn/ clusters.[5]

b) Similar predictions follow from perceptual recoverability, but for different reasons. Here a longer delay would be expected for /kn/ than for /kl/ in order to avoid reduction of the perceptual salience of /k/ by nasal leakage. Predictions following perceptual recoverability are restricted to the clusters /kl/ and /kn/ because /kl/ and /kn/ both consist of a velar stop and a coronal sonorant. Differences in /ks/ vs. /sk/ could be as likely due to different C1 articulators as they could be due to different C1 place of articulation.

c) Based on the assumptions of the Degree of Articulatory Constraint (DAC) model, more overlap would be predicted in /ks/ than in /kn/ and /kl/ as /n/ and /l/ have lower DAC values and thus exert less coarticulation on /k/. /kl/ and /kn/ should

[5] This is in accordance with the sonority hierarchy as proposed by Selkirk (1984). If, following e.g. Sievers (1901), stops are considered less sonorous than fricatives, /ks/ does not violate the hierarchy and less overlap should be expected here than for /sk/. This study's focus is not on corroborating either scale.

display a tendency of more overlap in /kl/ than in /kn/. /ks/ is expected to be less overlapped than /sk/ due to tongue repositioning in the former.

d) Regarding the internal coordination within clusters, the theoretical framework of the π-gesture predicts a decrease in overlap between the two consonants for higher levels of prosodic boundaries and for clusters in stressed syllables. However, the extent to which this effect takes place depends on the position of the cluster in the syllable. The timing of clusters in word-initial position is very stable (Byrd & Tan, 1996) and the interval during which the two consonants might show overlap is at some distance from the center of the prosodic effect (i.e. the prosodic boundary). Therefore we assume that changes in overlap might be very subtle. No changes in overlap duration could indicate that the overlap is specified by cluster type and therefore its variation due to prosody is highly constrained.

e) The durations of the consonants are supposed to be more susceptible to prosodic variation as compared to the overlap. If boundary strength affects the adjacent segments as predicted by the π-gesture model, then the first consonant in the cluster should lengthen to a greater degree than the consonant further away from the boundary. Palatal contact for the first consonant should also increase for higher levels of prosodic boundaries, whereas the second consonant might be less or not at all affected. The vowel duration will remain the same. For stress confounded with accent the vowel is hypothesized to be the center of the π-gesture. Since the second consonant is closer to this center it should be lengthened and possibly strengthened spatially fortemo higher levels of stress. The initial consonant should not be influenced by stress or only very slightly. This is largely in line with the account of Pierrehumbert and Talkin (1992) with the exception that this account only predicts boundary conditioned strengthening of the entire onset without being specific with regard to complex onsets.

2.2 Experiment

2.2.1 Speakers and speech material

7 speakers (5 female, 2 male) between the ages of 25 and 42 were recorded by means of EPG (Reading EPG3; 62 contacts in eight rows: 6 contacts in the front row, 8 in the remaining). All of the subjects had experience participating in EPG experiments and were equipped with custom-made pseudo palates; none of them reported any speech or hearing disorders. All speakers originate from the North or the East of Germany with long-term residence either around Kiel or Berlin without any particular dialect coloring. The target words consisted of 3 pairs, where each pair shared the initial consonant cluster but differed in lexical stress in that it was either on the first (henceforth stressed) or the second (henceforth unstressed) syllable: *Claudia* (name) /ˈklaʊ.dia/ - *Klausur* 'written exam' /klaʊ.ˈzuɐ/; *Kneipe* 'pub' /ˈknaɪ.pə/ – *Kneipier* 'pub owner' /ˈk-naɪ.ˈpjeː/; *Scarlett* (name) /ˈskaː.lət/ – *Skandal* 'scandal' /skan.ˈdaːl/. Additionally, the word *Xaver* (name) /ˈksaː.vɐ/ was included, even though no real-word could be found beginning with /ks/ stressed on the second syllable except for scientific terms rarely used by none-specialists, e.g. *Xanthan, Xylose* (orthographic *x* is canonically realized as /ks/ in German). In German, initial /ks/ is quite rare. However, the speakers are accustomed to these clusters from e.g. the name *Xaver* or *Xylophon* in the musical education of most schools. As a later addition it was only recorded for 5 of the 7 speakers, one of whom realized the initial cluster as [ts] instead of [ks]. Hence, results for /ks/ can only be presented for 4 out of seven speakers. The word preceding the test item always ended in /ɐ/ or unstressed /a/.

In order to elicit different prosodic boundaries preceding the target words, they were embedded in 4 syntactically similar contexts each: In the utterance-initial condition (U), the target word came at the beginning of the second of two sentences; in the phrase-initial condition (P), it was the first word of a sub-clause; in the list condition (L), it appeared as the third item of a list; the word-initial condition (W) had only a Prosodic Word boundary preceding the target word. All utterances were carefully designed to avoid nuclear accent on the target words. Tables A.1- A.4 show the complete speech material. The speakers were presented all utterances in randomized order in 10 repetitions yielding a total of 300 trials per speaker.

2.2.2 Measurements

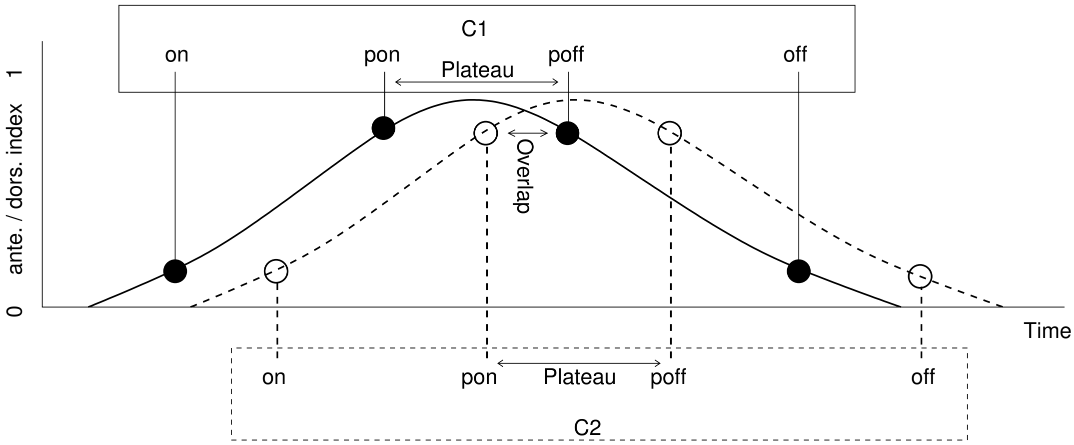

Figure 2.1: Articulatory landmarks and definition of temporal parameters

For acoustical labeling, the Munich Automatic Segmentation System (MAUS, Schiel, 1999) was applied. The output was converted and imported into the EMU (Bombien, Cassidy, Harrington, John, & Palethorpe, 2006) Speech Database System in order to facilitate hierarchical annotations. Following Cho and McQueen (2005), all utterances were assigned to one of three prosodic groups, each group defined by the prosodic boundary preceding the target word. The mapping from syntactical to prosodic boundaries is displayed in table 2.1 for all speakers and across all speakers. Obviously, the realizations of the syntactical categories may scatter across different prosodic categories and are speaker dependent. Prosodic groups were defined as follows:

1. Big Boundary (BG): a boundary tone and a pause
2. Small Boundary (SM): a boundary tone and no pause
3. Prosodic Word (WD): no boundary tone and no pause

Prosodic labeling was done by two skilled transcribers, one of them deciding the unclear cases. A pause was constituted not only by the presence of acoustical silence but also by the perception of a pause, which in turn might be evoked by final lengthening, another major cue for boundaries. Determining pauses before stops is obviously

Table 2.1: Cross-category table for mapping from syntactical to prosodic categories (for abbreviations see text).

Syntactic categories	Prosodic groups per speaker											
	f01			f02			f03			f04		
	BG	SM	WD	BG	SM	WD	BG	SM	WD	BG	SM	WD
U	39	21	0	7	40	0	45	0	0	30	5	1
P	4	51	0	0	32	15	14	30	0	4	36	2
L	2	9	49	1	4	41	1	5	39	0	3	38
W	0	0	66	0	1	54	0	0	45	0	0	44
	f05			m01			m02			all		
	BG	SM	WD	BG	SM	WD	BG	SM	WD	BG	SM	WD
U	53	2	0	56	0	0	33	11	0	263	79	1
P	30	29	1	37	20	0	6	38	0	95	236	18
L	0	1	59	26	25	11	0	1	42	30	48	279
W	0	0	59	0	0	59	0	0	53	0	1	380

problematic. Details on this problem are given below in the list of temporal parameters. Boundary tones were identified by inspecting f0 contours displayed in Emu and generated by the accompanying f0 tracking tool (tkassp/f0ana).

Articulatory landmarks in the EPG data were labeled using two indices: The anteriority index indicates the relative amount of (un-weighted) linguo-palatal contact in the anterior region (rows 1 to 5) of the pseudo-palate (number of active contacts in rows 1 to 5 divided by total number of contacts in rows 1 to 5 (e.g. Fontdevila, Pallarès, & Recasens, 1994)[6]). Here it was applied for C_2 in /kl/, /kn/ and /ks/ and for C_1 in /sk/ for which linguo-palatal contact only occurs in the anterior region. The dorsality index does the same for the posterior region (rows 6 to 8) of the pseudo palate. In order to take speaker-specific differences in dorsal stop articulation into account we applied the method by Byrd, Flemming, Mueller, and Tan (1995) and established a set of contacts unique for velar articulations for each speaker and limited the calculation of the index to this set. This profiling was not necessary for the anterior region as tongue

[6] Fontdevila (1994) provide formulas also for weighted indices. A weighted anteriority index provides a measure of how far back or front an articulation in the anterior region is. We used the unweighted versions here as we were only interested in the *amount* of contact in a specific area not the exact position of the contact.

tip articulations were always easily separable from contextual segments which were controlled for (either open vowel or velar stop). The dorsal region for speaker f03 had to be restricted to only two contacts in the last row. This restriction arose as the result of the order in which the data were analyzed: In a first step, only /kl/ clusters were examined (Bombien, Mooshammer, Hoole, Rathcke, & Kühnert, 2007), then /kn/ and then /ks/. While for the clusters /kl/ and /kn/ some contacts in the next to last row of the pseudo-palate were involved in /k/ closure formation, in /ks/ these contacts only produced noise, which had to be filtered out by further restricting the dorsal region for /ks/. The use of this procedure was necessary for one speaker only but underlines the difficulties in the analysis of velars with EPG as pointed out by Fougeron, Meynadier, and Demolin (2000).

The following articulatory landmarks were labeled (see also Fig.2.1[7]).

1. Onset and offset of constriction plateau (70% threshold) ($pon, poff$)
2. Maximum constriction at the center of the plateau

All thresholds are relative to the local maximum constriction and the local minimum constriction before/after the movement as measured in the time-course of the anteriority index for consonants with tongue front contact or the dorsality index for consonants with tongue dorsum contact. The 70% threshold criterion was defined operationally by looking at the contact patterns of all speakers. This value yielded timepoints which were most closely related with the acoustic landmarks like the offset of the preceding vowel and the burst. For analysis, the following temporal parameters were derived:

- Acoustical duration of the syllable nucleus following the cluster.
- Articulatory plateau duration of both consonants as the difference between the respective plateau offset and onset
- Plateau overlap as the time difference between plateau onset of C_2 and plateau offset of C_1, i.e. positive values indicate overlap, negative values indicate lag

[7]In Figure 2.1, the additional landmarks onset and offset of articulatory movement (20% threshold) (*on, off*) are also displayed. They are of no relevance here.

- Where applicable (see below) these parameters were also examined normalized by the interval from plateau onset of C_1 to plateau offset of C_2, to compensate for possible effects of speech rate. The standard deviation of speech rate varied from 15ms for speaker f02 to 65ms for speaker f05. To normalize a given value for C_1/C_2 plateau duration or plateau overlap it was divided by the interval from plateau onset of C_1 to plateau offset of C_2.

$$x_{norm} = \frac{x}{C2_{poff} - C1_{pon}}; \quad x \in \{C1Plateau, C2Plateau, PlateauOverlap\}$$

- The parameter pause (P) aims to serve as a means of validating the results for C_1 plateau duration. It was observed that when pauses preceded the cluster, velar contact was established at the beginning of the pause and maintained until the release of C_1 even through the longest pauses. Thus the validity of C_1 plateau duration can be questionable in co-occurrence with pauses. Pause (P) is the sum of the duration of the acoustical pause (p) preceding the target word (if present) and the difference of the acoustical duration of C_1 (C_1) and its per-speaker mean ($\bar{C}_{1s}$) (if positive):

$$P = \begin{cases} p + (C_1 - \bar{C}_{1s}) & \text{if } C_1 > \bar{C}_{1s} \\ p & \text{else} \end{cases}$$

This procedure yields a positive value for each C_1 longer than $\bar{C}_{1s}$ even where the acoustical pause p equals 0 s. Thus there are occurrences of non-zero pause values even in tokens of the conditions SM and WD where a true pause cannot be present by definition. These occurrences are not to be confused with acoustical pauses and are negligible in magnitude, see 2.3.2.1. It has to be noted that the acoustical onset of C_1=/k/ was often indeterminable when preceded by a pause. In these cases it had to be set arbitrarily just to mark the existence of a pause. C_1 durations of these cases were excluded from per-speaker mean C_{1s} calculation.

2.2.3 Statistics

Analyses of variance (ANOVA) were calculated for individual speakers and pooled over all speakers using R (R Development Core Team, 2006). For the individual speakers all valid data were included. Main effects and interactions were computed. Independent variables were prosodic group "PG" and stress level "S". In order to evaluate speaker-independent strategies, additionally ANOVAs pooled over all speakers were calculated based on the data averaged over up to 10 repetitions so that each speaker contributed only one experimental score per condition see e.g. Max and Onghena, 1999. This data reduction is necessary in order to avoid artificially inflating the error terms and degrees of freedom. Whether prosodic group and stress level affected temporal data was evaluated by calculating repeated-measures ANOVAs with the within-subject factors PG and S. Degrees of freedom were corrected by calculating the Greenhouse- Geisser epsilon in order to avoid violation of the sphericity assumption. Therefore, fractional degrees of freedom are often given in the tables. Pairwise t-tests with Bonferroni adjustments for multiple comparisons were carried out for individual statistics and for the repeated-measure ANOVAs in order to assess significant differences between the three-level-factor PG. Significance codes as given in the tables follow R's standard notation: "0 "***" 0.001 "**" 0.01 "*" 0.05" meaning that a probability between 0.05 and 0.01 ($p < 0.05$) is marked by one star, a probability between 0.01 and 0.001 ($p < 0.01$) by two stars and a probability between 0.001 and 0 ($p < 0.001$) by three stars.

2.3 Results

The results section is organized in two parts: The first part addresses the question of how sequence type affects the temporal organization of clusters. Therefore, the potential influence of prosody was ruled out by restricting the analysis to stressed /kl/, /kn/, /ks/ and /sk/ in the word-initial condition (W) as defined in section 2.2.1, i.e. not preceded by a phrase boundary. /ks/, as mentioned in section 2.2.1, is not available for all speakers. In the second part of the results section the prosodic conditions boundary and word stress will be investigated in greater detail in order to find out which characteristics of a particular cluster are stable across different prosodic conditions.

Figures 2.2, 2.3 and 2.5 show overlap patterns of the clusters under analysis in this study as bar plots. They all follow the same scheme: In the cases where C_1 and C_2 do not overlap, white space is drawn between the respective bars. Where C_1 and C_2 do overlap, this is indicated by a different gray shade. This area is to be considered part of both consonants. Standard errors are indicated at the inner edge of the respective consonant's bar which includes the overlap area, if present.

2.3.1 Cluster type

Table 2.2 shows statistical results of the comparison of the clusters. To compensate for effects of speech rate, for plateau overlap both absolute and time-normalized values were analyzed. Figure 2.2 illustrates the normalized timing patterns of all four clusters for each speaker in order to visualize inter-individual differences in overlap patterns. Figure 2.3 shows normalized (left) and non-normalized (right) timing patterns across all speakers.

The duration of the C_1 plateau is not significantly affected by the manner of the first consonant, i.e. fricative vs. stop. This is reflected by the very inconsistent results for the individual speakers. A similar picture emerges for the C_2 plateau duration.

However, we see that plateau overlap varies clearly across the four clusters. While there is always overlap in /kl/, never overlap – rather lag – in /kn/, it may be one or the other for /ks/ and /sk/. This is apparent in Figures 2.2 and 2.3 where there is always a void in-between the bars representing C_1 and C_2 for /kn/ while these bars always overlap for /kl/. Also, the standard errors for /kl/ and /kn/ do not overlap while those for /ks/ and /sk/ do, indicating that the latter clusters allow for more variability in their temporal organization.

The repeated measures ANOVA shows less overlap for /kn/ than for /sk/ and /kl/. For the single speakers, /kn/ also exhibits the least overlap, while overlap in /sk/ and /ks/ may be shorter or equal to /kl/. While in Figure 2.2 it seems that in clusters with /s/ overlap can be greater than in /kl/, this is not statistically significant. Only speaker f03 does not distinguish significantly between overlap in /kl/ and /kn/. All of this holds for both absolute and normalized data.

Overall, the most stable findings appear to be that /kn/ and /kl/ show a reversed pattern of temporal coordination: /kl/ is produced with considerable overlap between the two plateaus whereas for /kn/ a lag between the two plateaus seems to be oblig-

Table 2.2: Consonant plateau durations and overlap in /kl/, /kn/, /ks/ and /sk/ for each speaker (rows 1-7) and across all speakers (row 8; cluster /ks/ excluded, missing for f04, f05, m01) Word boundary (WD) condition only.

Spk	Measure											
	C_1 plateau duration			C_2 plateau duration			plateau overlap			norm. overlap		
	df	F	p		F	p		F	p		F	p
f01	3	/ks/ > /kn/, /kl/, /sk/						/kn/ < /kl/, /sk/, /ks/			/kn/ < /kl/, /sk/, /ks/	
	32	10.6	***		2.72	n.s.		15.0	***		14.1	***
f02	3				/sk/ > /kl/			/kn/ < /ks/ < /kl/, /sk/			/sk/ > /ks/	
	27	2.5	n.s.		3.79	*		23.0	***		19.9	***
f03	3	/kn/ > /ks/, /sk/			/sk/ > /kn/, /kl/, /ks/; /ks/ > /kn/			/kn/ < /ks/			/kn/ < /ks/	
	24	5.9	**		16.95	***		3.4	*		3.8	*
f04	2				/sk/ > /kn/, /kl/			/kn/ < /kl/, /sk/			/kn/ < /kl/, /sk/	
	19	0.8	n.s.		11.1	***		12.2	***		12.1	***
f05	2				/kl/, /sk/ > /kn/			/kn/ < /kl/, /sk/			/kn/ < /kl/, /sk/	
	26	0.2	n.s.		8.6	**		67.9	***		55.5	***
m01	2	/kn/ > /sk/						/kn/, /sk/ < /kl/			//kn/, /sk/ < /kl/	
	28	4.5	*		1.2	n.s.		52.2	***		67.0	***
m02	3	/kl/, /kn/ > /sk/ > /ks/			/ks/, /kn/, /kl/ > /sk/			/kn/, /ks/, /sk/ < /kl/			/kn/, /ks/, /sk/ < /kl/	
	27	32.7	***		17.9	***		37.7	***		45.1	***
All	2			2			1.7	/kn/ < /sk/, /kl/		1.4	/kn/ < /sk/, /kl/	
	12	2.5	n.s.	12	1.0	n.s.	10.5	16.2	**	8.6	16.3	**

Significance codes: 0 “***” 0.001 “**” 0.01 “*” 0.05. Example: For speaker f01, C_1 plateau duration is larger in /ks/ than in /kn/, /kl/ and /sk/ (highly significant).

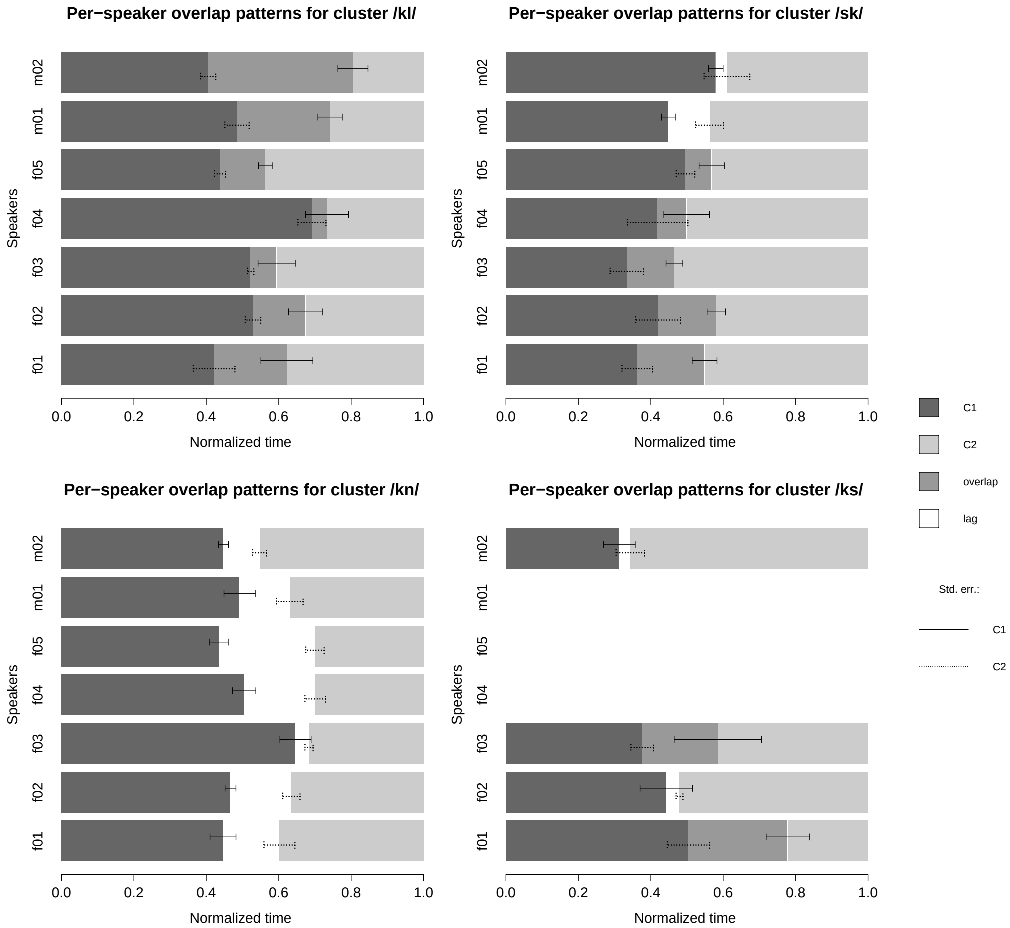

Figure 2.2: Time-normalized overlap patterns of the mean C_1 and C_2 contact plateau durations for the clusters /kl/, /kn/, /ks/ and /sk/ for all speakers. Standard errors are drawn at the inner border of the respective bar, which includes the overlap if any. Standard error bars for C1 (solid lines) are drawn slightly above those for C2 (dotted lines).

atory. Figure 2.4 illustrates this behavior. The data for these palatograms were taken from speaker f05. Both are tokens from the syntactical word-initial class with stress on the first syllable.

In the next section, the stability of the observed patterns will be tested across varying prosodic conditions. This analysis will be restricted to the clusters exhibiting the most stable patterns. Accordingly, /sk/ and /ks/ will be excluded. Further reasons for the exclusion are the asymmetrical material for /ks/ (stress variation missing, only

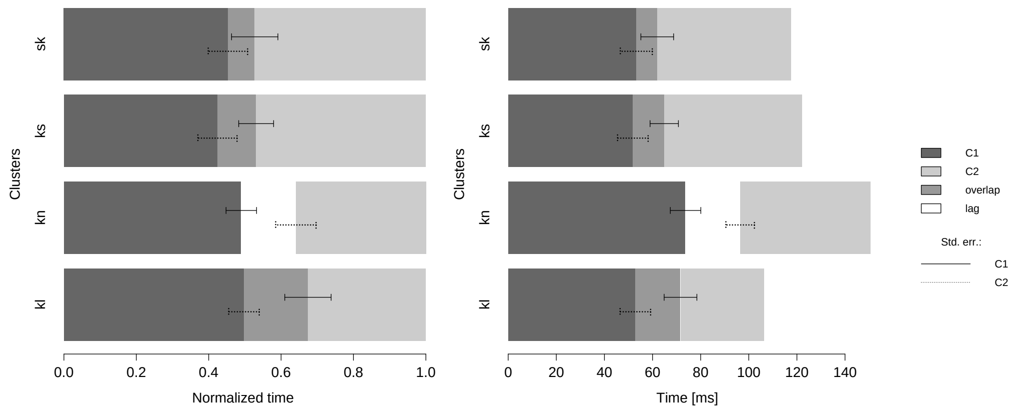

Figure 2.3: Normalized (left) and absolute (right) overlap pattern's for clusters /kl/, /kn/, /ks/ and /sk/ across all speakers. Standard errors are drawn at the inner border of the respective bar, which includes the overlap if any. Standard error bars for C1 (solid lines) are drawn slightly above those for C2 (dotted lines).

Xaver) and the problematic cross-cluster comparability: dealing with the intrinsic differences between stop-sonorant and stop-fricative clusters would be beyond the scope of this section. Furthermore, the vowels in target syllables lacked comparability to those of clusters /kl/ and /kn/ under prosodic variation. This does not affect the results of the cluster type analysis.

2.3.2 Prosody

Effects of prosodic variation are described in two parts. First, the temporal parameters C_1 plateau duration, C_2 plateau duration, plateau overlap and pause duration are considered. Then we will discuss effects in the spatial domain.

2.3.2.1 Temporal effects

Normalization of durational and overlap measures as carried out in section 2.3.1 is not applicable here since prosodic variations can be expected to influence all durational measures in a non-uniform way. As C_1 and C_2 durations are hypothesized to lengthen at strong boundaries or under lexical stress respectively, using any of these two measures for normalization could conceivably either enhance or conceal possible effects.

Table 2.3: Statistical results for for temporal parameters in cluster /kl/ under prosodic variation for each speaker (rows 1-7) and across all speakers (row 8; Prosodic groups (PG): BG, SM, WD: stress levels: S, U). Interactions (Inter.) are included if present. The degrees of freedom for the factors are fixed (PG: 2, Stress: 1).

Spk	Effect		Measure							
			C1		C2		overlap		Pause	
		df	F	p	F	p	F	p	F	p
f01	PG	70							BG > SM > WD	
				n.s.		n.s.		n.s.	151.8	***
	Stress									
				n.s.		n.s.		n.s.		n.s.
f02	PG	58	BG > SM, WD		SM > WD				BG > SM, WD	
			10.2	***	4.7	*	3.5	*	81.9	***
	Stress						S < U			
				n.s.		n.s.	4.3	*		n.s.
	Inter.								U: SM > WD	
				n.s.		n.s.		n.s.	17.3	***
f03	PG	54	BG > SM, WD						BG > SM, WD	
			20.2	***		n.s.	9.5	***	169.7	***
	Stress						S < U			
				n.s.		n.s.	30.7	***		n.s.
	Inter.						S: BG, WD > SM			
							U: BG, SM < WD			
				n.s.		n.s.	8.7	***		n.s.
f04	PG	42	BG > SM, WD				SM < WD		BG > SM, WD	
			21.8	***		n.s.	8.2	***	331.0	***
	Stress						S < U			
				n.s.		n.s.	14.3	***		n.s.
			U: SM < WD							
			10.1	***		n.s.		n.s.		n.s.
f05	PG	68	BG > SM, WD				BG < WD		BG > SM, WD	
			18.9	***		n.s.	4.1	*	86.8	***
	Stress				S > U					
				n.s.	4.0	*		n.s.		n.s.
	Inter.				SM, WD: S > U					
				n.s.	3.2	*		n.s.		n.s.
m01	PG	65	BG > SM, WD						BG > SM, WD	
			12.5	***		n.s.		n.s.	23.6	***
	Stress						S < U			
				n.s.		n.s.	6.5	*		n.s.
	Inter.			n.s.		n.s.	3.7	*		n.s.
m02	PG	52	BG > SM > WD		WD > SM				BG > SM > WD	
			12.4	***	3.7	*		n.s.	70.2	***
	Stress									
				n.s.		n.s.		n.s.		n.s.
All	PG	6.8	BG > SM, WD;						6.4 BG > SM, WD	
		1.1	14.0	*		n.s.		n.s.	1.1 43.2	***
	Stress	6								
				n.s.		n.s.		n.s.		n.s.

Significance codes: 0 “***” 0.001 “**” 0.01 “*” 0.05. Example: For speaker f05 in SM and WD condition, C_2 plateau duration is longer in stressed than in unstressed tokens.

Table 2.4: Statistical results for temporal parameters in cluster /kn/ under prosodic variation for each speaker (rows 1-7) and across all speakers (row 8; Prosodic groups (PG): BG, SM, WD: stress levels: S, U). Interactions (Inter.) are included if present. The degrees of freedom for the factors are fixed (PG: 2, Stress: 1).

Spk	Effect		Measure							
			C1		C2		overlap		Pause	
		df	F	p	F	p	F	p	F	p
f01	PG	74	BG > SM, WD				BG < WD		BG > SM, WD	
			16.7	***		n.s.	4.9	**	99.4	***
	Stress									
				n.s.		n.s.		n.s.		n.s.
f02	PG	56	BG > SM, WD						BG > SM, WD	
			21.9	***		n.s.		n.s.	81.9	***
	Stress						S < U			
				n.s.		n.s.	8.1	**		n.s.
f03	PG	48	BG > SM > WD						BG > SM, WD	
			34.8	***		n.s.		n.s.	95.3	***
	Stress		S > U				S < U			
			7.2	**		n.s.	8.6	**		n.s.
	Inter.						U: BG < SM, WD			
				n.s.		n.s.	8.7	*		n.s.
f04	PG	56	BG > SM, WD		BG < SM				BG > SM, WD	
			58.6	***	3.5	*		n.s.	33.1	***
	Stress						S < U			
				n.s.		n.s.	11.9	**		n.s.
f05	PG	74	BG > SM, WD				BG < SM, WD		BG > SM, WD	
			20.9	***		n.s.	6.9	**	226.3	***
	Stress									
				n.s.		n.s.		n.s.		n.s.
m01	PG	77	BG > SM > WD				BG < SM, WD		BG > SM, WD	
			42.1	***		n.s.	19.9	***	107.8	***
	Stress								S > U	
				n.s.		n.s.		n.s.	4.4	*
m02	PG	57	BG > SM, WD		BG < WD				BG > SM, WD	
			22.3	***	4.0	*		n.s.	241.5	***
	Stress						S < U			
				n.s.		n.s.	8.7	**		n.s.
All	PG	8.4	BG > SM, WD				11.3 BG < SM, WD		6.1 BG > SM, WD	
		1.4	53.1	***		n.s.	1.9 10.3	**	1.0 50.7	***
	Stress	6					S < U			
		1		n.s.		n.s.	6.7	*		n.s.

Significance codes: 0 "***" 0.001 "**" 0.01 "*" 0.05. Example: For speaker f03 in unstressed tokens, overlap is smaller in BG than in SM and WD.

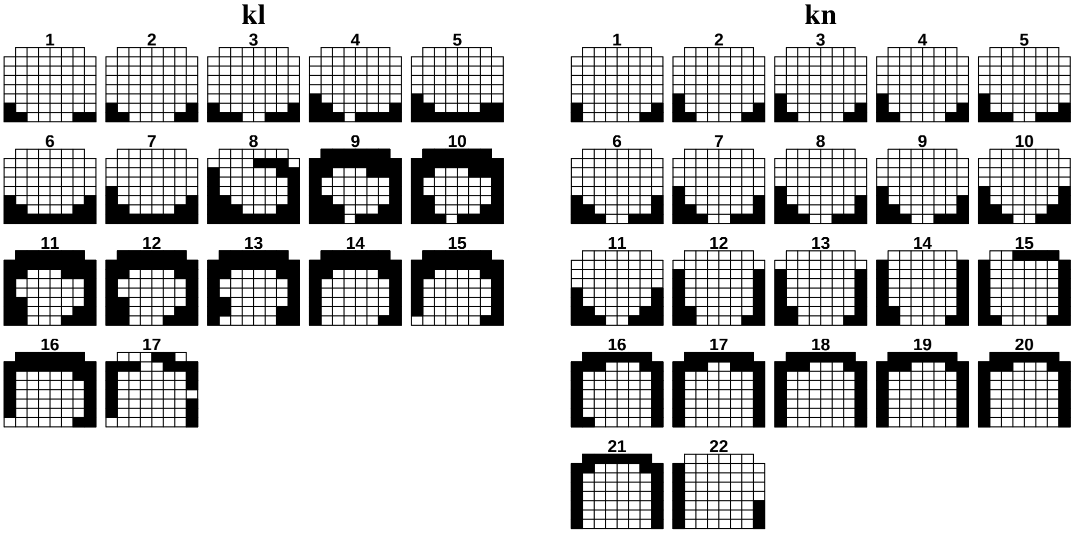

Figure 2.4: On the left: Palatograms for the cluster /kl/. Apical closure for /l/ (upper rows) is initiated distinctly before /k/ closure release (lower rows). C_1 plateau ranges from frames 4 to 10, C_2 from 9 to 16. On the right: Palatograms for the cluster /kn/. Apical closure for /n/ (upper rows) is initiated distinctly after /k/ closure release (lower rows). C_1 plateau ranges from frames 7 to 11, C_2 from 16 to 21. Frames were sampled at a rate of 100Hz. Begin and end of the data displayed corresponds to the onset of C_1 and the offset of C_2. See Figure 2.1 for the definition of these landmarks.

A separate analysis of boundary strength and stress is not feasible here since the two are varied orthogonally in our material. Tables 2.3 and 2.4 show the results of ANOVAs for all individual speakers as well as repeated measures ANOVAs across all speakers with the factors "Prosodic Group" (levels: Big Boundary (BG), Small Boundary (SM) and Word (WD)) as defined above and "Stress" (levels: stressed (S) and unstressed (U)). In Figure 2.5 the durations of the articulatorily defined consonants (dark gray and light gray) and the overlap (mid gray) or lag (white) are shown. Additionally the acoustically measured vowel duration (black) is given. This is to give evidence concerning our prediction (e) above where we assume that stress has the strongest effect on the nucleus. Intervals of syllables starting with /kl/ are presented on the left side and with /kn/ on the right side.

First, results on pause durations are presented because the boundary categories were distinguished by the presence or absence of a pause. Therefore, quite unsurprisingly, pause duration significantly distinguishes BG boundaries from SM and WD.

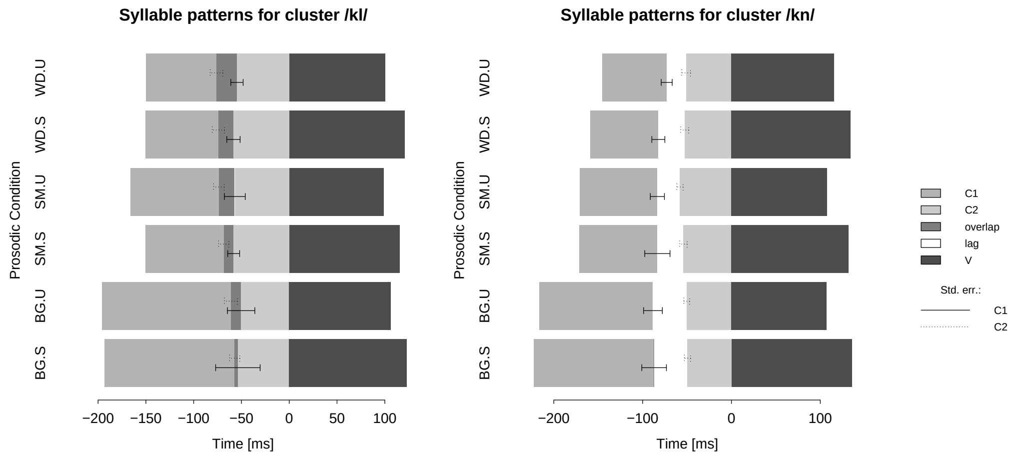

Figure 2.5: Syllable patterns of the mean C1 and C2 contact plateau durations for the clusters /kl/ and /kn/ as well as the acoustical nucleus duration across all speakers for each prosodic condition (see text). Standard errors are drawn at the inner border of the respective bar. Standard error bars for C1 (solid lines) are drawn slightly below those for C2 (dotted lines). Zero-point is aligned with the beginning of the nucleus. The patterns are drawn in a pairwise fashion (stressed S and unstressed U), one pair for each prosodic group (BG, SM, WD).

For SM and WD the durations differ only in very few cases (kl: f01, f02 and m02). This parameter mainly serves the purpose of validating the results for C_1 plateau duration. As mentioned above, full dorsal contact for /k/ was often established within and maintained throughout the pauses. In these cases it was not clear whether the constriction was intended for speech articulation or an artifact introduced by the EPG pseudo-palate, e.g. swallowing and so forth.

The duration of C_1 is clearly affected by boundary strength for both /kl/ and /kn/. Generally, we find longer plateau durations in the BG condition as compared to the weaker boundaries. Only in three cases (/kl/: m02, /kn/: f03 and m01) do plateau durations for C_1 differ significantly between the SM and WD boundary levels as well. The main difference was between BG boundary on the one hand and SM and WD on the other hand.

Effects on overlap are less consistent than those on C_1 duration. In some speakers – not necessarily the same ones – both /kl/ and /kn/ exhibit less overlap at strong boundaries than at weak boundaries. This is significant in three speakers (f03, f04, f05) for /kl/ and in four speakers (f01, f03, f05, m01) and across speakers for /kn/. Mainly, the BG boundary is distinguished from the two other levels.

Boundary strength does not, however, appear to play a role in the duration of C_2. Significant differences are very rare and directionally inconsistent (/kl/: m02 SM < WD, f02 SM > WD; /kn/: f04 BG < SM, m02 BG < WD; see tables 2.3 and 2.4). The overall insensitivity of C_2 duration to boundary strength is furthermore demonstrated across speakers in the repeated measures ANOVA. As was expected the nucleus duration was not affected by boundary strength (/kl/: $F = 2.4$, /kn/: $F = 1.7$).

No effects of stress could be observed on C_1 and C_2 plateau durations or on the duration of a pause (only speaker m01 appears to lengthen pauses before stressed /kn/). However, speakers f02, f03 and f04 produce both /kl/ and /kn/ with less overlap in stressed syllables as do speakers m01 for /kl/ and m02 for /kn/. Across speakers, no significant effect of lexical stress could be found for any of the parameters except for less overlap in stressed syllables in /kn/. The data follow our prediction with respect to the syllable nucleus being the center of the effect of stress. Nucleus durations are longer in stressed than in unstressed syllables (/kl/: $F = 4.5$ (only marginally significant), /kn/ $F = 57.0$, $p < 0.001$), as can be seen in Figure 2.5.

2.3.2.2 Spatial effects

Articulatory strengthening is often equated with an increase of palatal contact. Table 2.5 and Figure 2.6 show the effects of prosodic variation on /kl/ and /kn/ in the spatial domain, i.e. maximum contact percentage for the first and the second consonants. Boundary strength affects the contact patterns of /k/ only in /kn/ not in /kl/. For /kn/ this effect – with stronger boundaries inducing more palatal contact – is very consistent for 6 speakers and over all speakers in the repeated measures ANOVA. The strength of the boundary effect diminishes with distance from the boundary but is still significant in C_2 for three speakers and across speakers only for /kn/. Stress strengthens both consonants in some cases for the cluster /kl/ but not for /kn/: Three speakers increase the amount of palatal contact in /k/ and two in /l/. As can be seen in Figure 2.6, the spatial stress effect on /l/ in /kl/ tends to increase at lower levels of boundary strength.

Table 2.5: Statistical results for spatial parameters in clusters /kl/ and /kn/ under prosodic variation for each speaker (rows 1-7) and across all speakers (row 8; Prosodic groups (PG): BG, SM, WD; stress levels: S, U). Interactions (Inter.) are included if present. The degrees of freedom for the factors are fixed (PG: 2, Stress: 1).

Spk	Effect		/kl/					/kn/			
			C1 Max		C2 Max			C1 Max		C2 Max	
		df	F	p	F	p	df	F	p	F	p
f01	PG	70		n.s.		n.s.	74	BG, SM > WD 4.5	*	BG, SM > WD 7.2	**
	Stress			n.s.		n.s.			n.s.		n.s.
f02	PG	58		n.s.		n.s.	56		n.s.		n.s.
	Stress		S > U 15.8	***	S > U 36.1	***			n.s.		n.s.
	Inter.			n.s.		n.s		U: BG > WD SM: S < U 3.6	*		n.s.
f03	PG	54		n.s.		n.s.	48	BG > WD 6.6	**		n.s.
	Stress			n.s.		n.s.			n.s.		n.s.
f04	PG	42		n.s.		n.s.	56		n.s.	SM > WD 5.9	**
	Stress		S > U 4.6	*		n.s.			n.s.		n.s.
f05	PG	68		n.s.		n.s.	74	BG > SM, WD 35.3	***		n.s.
	Stress			n.s.		n.s.			n.s.		n.s.
	Inter.		S: SM < BG, WD 8.5	***		n.s.			n.s.		n.s.
m01	PG	65		n.s.		n.s.	77	BG > SM > WD 49.6	***		n.s.
	Stress				S > U 28.6	***			n.s.		n.s.
m02	PG	52		n.s.	5.0	*	57	BG, SM > WD 8.3	*	BG > WD 11.2	***
	Stress		S > U 6.5	*		n.s.			n.s.	U > S 13.3	***
All	PG	12		n.s.		n.s.	8.2 1.4	BG > WD 11.3	*	11.9 2.0 11.5	**
	Stress	6		n.s		n.s.	6		n.s.		n.s.

Significance codes: 0 "***" 0.001 "**" 0.01 "*" 0.05. Example: For speaker m01 in custer /kn/ maximal C_1 contact is larger in BG than in SM than in WD.

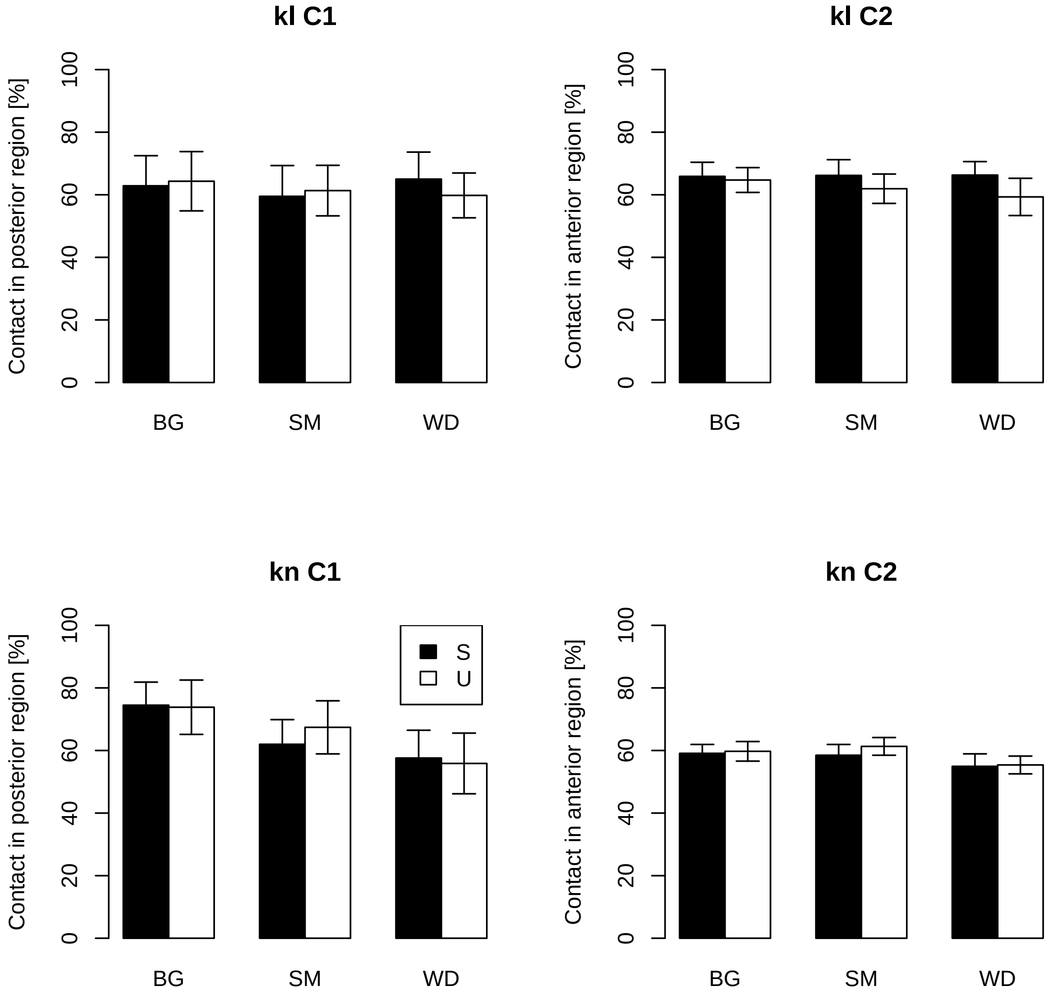

Figure 2.6: Contact percentage in in the respective region of the pseudo-palate for C_1 and C_2 ind /kl/ and /kn/, calculated across all speakers, separately for each boundary and stress level.

2.4 Summary and discussion

In this section we will summarize and discuss the results of this study according to the predictions stated in section 2.1.3. Concerning cluster type, the most obvious finding in this study is that overlap for /kl/ appears to be mandatory while the timing in /ks/ and /sk/ is less rigidly specified. /kn/ does not appear to allow for overlap, as measured here, at all. The difference between /kl/ and /kn/ was predicted correctly by manner-based ordering (a). /ks/ and /sk/, however, were assumed to overlap less than the stop+sonorant clusters, which is not confirmed. (b) The difference between /kl/ and /kn/ is accounted for by perceptual recoverability, albeit for different reasons than manner-based ordering: A lag between /k/ and /n/ would presumably prevent the stop burst being obscured by early velic opening. (c) According to the predictions of the DAC model, ordering the four clusters by the amount of overlap should yield a sequence of /sk/ > /ks/ > /kl/ > /kn/. As for /sk/ and /ks/, the predicted higher overlap in the former was not found. Rather, the two clusters behave quite similarly with large variability in the emergence of overlap. The prediction of the DAC model for /kl/ to show more overlap than /kn/ is confirmed, although this effect might actually be stronger than would be expected, if the DAC value of /l/ is assumed to be only slightly higher than that of /n/. Contrary to the predictions, overlap in /kl/ is even larger than in /ks/ and /sk/. As mentioned in 2.1.1 this might be accounted for by assigning /l/ a higher DAC value following the findings of Recasens (1995) that German /l/ appears to be less clear than clear /l/ in other languages. Indeed, if /l/ was assigned a dorsal target, as e.g. dark /l/ in Catalan, the tongue pre-dorsum would be lowered for both /k/ and /l/ and the /kl/ transition could therefore proceed without tongue repositioning as opposed to /kn/ and /ks/. Following this line of thought, Catalan and German should show a tendency for more overlap in /kl/ clusters than other languages. Data collected by Gibbon, Hardcastle, and Nicolaidis (1993), however, indicate that /kl/ clusters in Catalan show significantly more overlap than in German and other languages. Moreover, EMA data (Geumann, Kroos, & Hoole, 1999) indicate that spatial variability of the tongue dorsum in German /l/ is very high and at least certainly not less than for /n/.

Several alternative reasons might explain the consistently longer lag in /kn/ clusters: In the first place, aero-dynamic reasons, as already mentioned in section 2.1.1, might constrain the timing between the two consonants in order to avoid a velopharyngeal leakage before oral release of C_1 occurs.[8]

[8]Should this be the case, the same pattern might be predicted in initial fricative + nasal clusters, since fricatives also require a tight velo-pharyngeal closure, in order to maintain a sufficient air flow for turbulence. However, data collected by Kühnert (2006) do not support this prediction, Fricatives, with their continuous acoustic information, are presumably in less danger of becoming difficult to recover. Acoustic information on the plosives on the other hand is concentrated at the burst.

Secondly, bio-mechanical linkage could prevent early velic lowering because the tongue dorsum presses against the soft palate during velar stops. However, Kühnert (2006) show that in /pn/ and /pl/ clusters, where bio-mechanical linkage can be neglected, the timing of lips and tongue tip differs in the same way as for tongue dorsum and tip in /kn/ or /kl/ clusters.

Thirdly, it could be argued that, in terms of inter-articulator coordination, /n/ is more complex in German than /l/ since apart from the tongue tip gesture the nasal requires an additional velic opening gesture. Therefore, a larger gap, i.e. less overlap, might be induced between the consonants in /kn/ than in /kl/. As several studies showed (Krakow, 1993; Kollia, Gracco, & Harris, 1995; Byrd, Tobin, Bresch, & Narayanan, 2009), in syllable initial position the velum and the primary articulator in nasals reach their targets simultaneously. To our knowledge, however, these studies do not address how onsets of velum and oral gestures are temporally coordinated. It could be speculated that in simple nasal onsets the velic opening gesture starts earlier than the oral constriction gesture. In complex /Cn/ onsets, then, velic opening onset, and not the oral constriction gesture of the nasal, might be timed with the constriction of the preceding consonant and therefore does not start until after the release of the preceding stop's closure. Thus, given that the velic opening onset is likely to occur after the release of the preceding oral consonant, and given that the targets of the velic opening gesture and its associated oral constriction gesture are likely to be attained simultaneously, there is likely to be a substantial gap between the oral constriction gesture associated with the nasal consonant and the preceding oral constriction gesture.

Prosodic boundary strength and lexical stress were varied in the current study as a probe in order to test which of the observed patterns for clusters remain stable. For reasons of comparability (see predictions in section 2.1.3) only /kl/ and /kn/ were considered in this part. While overlap to some extent showed sensitivity to prosodic variation (less overlap at high boundaries and in stressed syllables), the range of variation was limited so that the categorical difference found between /kl/ and /kn/ remained unaffected. Therefore the assumption that the temporal coordination in /kl/ and /kn/ is highly specified and constrained by the segmental make-up of the cluster receives considerable support.

There is more evidence for changes due to prosodic variations in temporal coordination in /kn/ than in /kl/. The extent to which variation is allowed in overlap depends therefore on the segmental make-up. As was explained in Section 2.1.1, the upper limit of overlap is probably constrained by perceptual recoverability demands. The violation of the lower limit of overlap – or rather the upper limit of lag – might yield the production of a transitional vowel. Evidence for transitional vowels has been found by Davidson (2005) for illegal clusters in American English and by Gafos (2002) in Moroccan Arabic. It would be interesting to see if a lag in /kl/ would induce the perception of such a transitional vowel. If so, it would explain why speakers avoid the drifting apart of the consonant gestures in /kl/. Accordingly, for /kn/ the upper limit of lag before perception of a transitional vowel would be higher.

Apart from the internal structure, the consonants themselves are affected by prosodic variation in both the temporal and the spatial domain. The strength of the boundary affects mainly the duration of C_1's plateau in both /kl/ and /kn/, i.e. /k/ was lengthened at higher boundaries. Articulatory strengthening was restricted to C_1 in /kn/ and at higher boundaries only. The second consonant is not sensitive to boundary strength, i.e. we could not replicate the findings of Byrd and Choi (2010) who found lengthening of C_2 in onset clusters in two out of three speakers. Additionally, the pause duration was the most consistently affected measure in this study. As was pointed out in Section 2.2.2, during the pause at big boundaries speakers varied in their timing of C_1 constriction: frequently constriction was achieved at the beginning of the pause, resulting in overlong plateau durations. In these cases, the lengthening of C_1 is reducible to the occurrence of a pause. Indeed, C_1 plateau duration mainly distinguished big boundaries from lower boundary levels. However, the occurrence of a pause cannot be made responsible for the effects observed on overlap. In summary we found stronger effects of boundary strength on duration and palatal contact of C_1. The overlap was affected less consistently and the second consonant only spatially in /kn/. Stress, on the other hand, only influenced the duration of the nucleus (longer in stressed syllables) and the overlap (less in stressed syllables).

In the introductory section, we proposed two models explaining how segments are affected by prosodic variation. We will discuss effects of boundary strength first. Pierrehumbert and Talkin suggest (for CV and VC syllables) that initial strengthening shifts the articulatory magnitude in a more consonantal direction. No specific predictions concerning consonant clusters can be derived from this account, but it is confirmed to the extent that articulatory strengthening takes place. The π-gesture approach more specifically predicts a decrease of the effect with distance from the boundary. In fact, our data corroborate this prediction with regard to the diminishing effects going from C_1 to overlap and C_2.

However, there is no simple way of modeling the differential behavior of /kl/ and /kn/ induced by prosodic variation within the framework of π-gestures. /kn/ is more susceptible to effects of this kind than /kl/ in the temporal and the spatial domain. We assume that this is strongly related to the internal structure of /kl/ vs. /kn/: The former shows overlapping consonant plateaus during which the tongue is highly constrained by multiple affordances. Apart from the central alveolar contact, lateral aperture is required to produce an /l/. In /kl/ clusters the tongue is further constrained by a simultaneous dorsal closure. In /kn/ on the other hand, contact patterns are less constrained because the dorsal constriction for /k/ and the apical constriction for /n/ are produced sequentially, i.e. there is a lag between the consonant plateaus. In so being less constrained, the components of the cluster have more degrees of freedom for adjustments to prosodic variation. /kl/ behaves in this respect similarly to what Fougeron and Keating (1997) found for /s/, namely that this consonant is less susceptible to articulatory strengthening at higher prosodic boundaries.

For stress, the π-gesture approach would predict the largest impact on the syllable's nucleus and a continuous decrease from the nucleus to the onset. In our data, we find a *dis*continuity: While C_2 as the closest segment to the nucleus remains largely unaffected, there is a significant decrease in overlap in stressed syllables. Evidence for the discontinuity of prosodic effects was also found by Turk and Shattuck-Hufnagel (2007) on the syllable level: In their data, final lengthening affected the main-stress-syllable and the rime of domain final words but skipped the phonological material in between these two syllables. Final lengthening is thus unevenly or discontinuously distributed.

2.5 Conclusions

The major finding of this study is that the gestural coordination for /kl/ is categorically different as compared to /kn/ with an obligatory lag between the consonant plateaus for /kn/ and overlap for /kl/. This is accounted for by all three principles introduced in section 2.1.1: perceptual recoverability, manner-based ordering and the DAC model.

However, while the recoverability based account does not make prediction for the clusters involving /s/, neither the manner-based ordering nor the DAC model can account for the internal structure of these clusters. Prosodic variation can influence the differential coordination between the clusters' consonants only within certain limits determined by the segmental make-up of the clusters. Our results give evidence that effects due to prosodic variation are rather subordinate to segmental setup and specifically that stop+nasal sequences play a special role. This might be of particular interest to research in sound change as well, e.g. loss of /k/ in English *knee* due to unmet parallel transmission requirements in terms of insufficient overlap.

Prosodic variation was successfully applied as a probe to investigate the stability of the internal organization within clusters in finding the limits in timing variation that prosodic conditioning induced. Furthermore, we found that not only do different segments display different susceptibility to prosodic variation but also groups of segments, such as clusters. In agreement with Articulatory Phonology (Browman & Goldstein, 1992a) and especially the notion of C-Center coordination (Browman & Goldstein, 1988, 2000; Byrd, 1995), we therefore assume that the temporal coordination (here in terms of overlap) is part of the phonological specification. Prosodic variation in clusters on the other hand appears to have limits determined by segmental setup. This is in accordance with limits of prosodic variation on singleton consonants such as the highly constrained /s/ (e.g. Shadle & Scully, 1995).

Jede Seite hat zwei Medaillen.

Mario Basler

Chapter 3

A durational EMA investigation of segmental and prosodic effects on word-initial German stop+alveolar clusters

3.1 Introduction

This chapter investigates effects of segmental make-up and prosodic variation on the production of word-initial stop+alveolar clusters in German by means of EMA (Electromagnetic Articulography). In the EPG (Electro Palatography) study in Chapter 2, /kl/ was found to be produced with a particularly high degree of overlap. Prosodic variation did not influence the inter-consonantal coordination. One can therefore speculate that the tight gestural coordination in /kl/ clusters is phonologically specified. On the one hand, the large amount of overlap is surprising since both consonants are lingually articulated. On the other hand, the tongue tip is to a certain extent independent of the tongue back. Tongue tip raising can therefore occur relatively early during the velar occlusion. At first glance /kn/ clusters can be expected to behave likewise since both clusters involve the same lingual articulators. However, the data presented in the previous chapter and also in Hoole, Bombien, Kühnert, and Mooshammer, 2009 showed that /kn/ exhibits substantially less overlap than /kl/ which enforces the consideration of other factors as well. Quite obviously, the second consonant (C_2) in /kn/ features a total oral occlusion and velar lowering as opposed to C_2 in /kl/ with a central occlusion and no velar lowering. C_1 burst in /kn/ must occur before oral occlusion and velic lowering in order to prevent nasal venting during closure and burst and thus the loss of the burst characteristics which are the major cue for stop place recoverability. Overlap is therefore disprefered in this cluster. If, on the other hand, /k/ is released into the lateral /l/, both the burst characteristics and the laterality are preserved. In Chapter 2, /kn/ overlap was also found to be significantly affected by prosodic variation in terms of boundary strength in contrast to /kl/. This can be interpreted as a less rigid specification in the mental lexicon for /kn/. One aim of this study is to revisit this issue with EMA data.

Further attention will be given to another line of thought. /kl/ was produced with considerable overlap in spite of both articulators being articulated lingually. It should be interesting to see how strong overlap is in a cluster whose articulators are more independent from each other. The most suitable candidate for comparison with /kl/ is /pl/ since it is the only other stop+lateral cluster in German. This choice also allows for an attempt to answer the question whether /l/ is particularly suited for overlap due to its articulatory properties. To this end two additional clusters are selected with C_2 replaced by another alveolar: /ks/ and /ps/.

Apart from the segmental considerations outlined above, this study also deals with the influence prosodic variation has on the observed patterns. As mentioned above, the internal coordination of /kl/ in terms of gestural overlap was not found to be influenced by prosodic variation in contrast to that of /kn/. A comparison of /l/ and /s/ as C_2 can also shed light on possibly different mechanisms applied to express prosodic strengthening as a function of manner of C_2. This study systematically varies three factors which are likely to influence the production of word-initial consonant clusters: Position, accent and stress. Throughout this text, accent and stress denote properties of two different prosodic constituents: (lexical) stress refers to the prosodic word while accent is associated with a phrase.

3.1.1 EMA vs. EPG

EMA (Electromagnetic Articulography) and EPG (Electropalatography) are very different systems which yield very different information. The following will point out that they are complementary rather than competing methods. EMA allows for position and motion tracking of sensor coils in 3D space whereas EPG measures the contact of a pseudo palate fitted to the hard palate and the tongue. One EPG data frame is in essence nothing more than a fixed number of binary states each indicating whether a corresponding electrode on the artificial palate is in contact with the tongue at a given point in time. Information about actual articulatory movements can only sparsely be deduced from the dynamics of contact strength and location. The distribution of contacts on the artificial palate and their number (62 in the Reading EPG3 System Hardcastle, Gibbon, & Jones, 1991) allows for a detailed analysis of linguopalatal contact in both the anterior-posterior as well as the lateral dimension of the hard palate. It should be noted that EPG measures the contact location/area on the pseudo palate but not the part of the tongue that is engaged in this contact.

EMA data is not limited to lingual articulation and most decidedly not about contact but about position. EMA point coordinates are obtained by solving complex nonlinear equations whose input is the demodulated electrical signal induced into the sensor coil by six magnetic field transmitters (see Hoole & Zierdt, 2010). Barring computing errors, EMA allows for position tracking of sensors at any given time. The system used in this investigation (Carstens AG 500) uses a maximum of 12 sensors which can be positioned anywhere within the recording space and–in the case of speech recordings–attached to any accessible part of the head. EMA sensors can interfere with each other if spaced too closely. In the case of lingual articulations it is therefore necessary to decide how to place sensors on the tongue. Data density for the tongue will accordingly always be lower in EMA recordings than in EPG. Apart from a custom fitted pseudo palate for each speaker EPG requires only little effort in both the recording as well as the analysis. EPG is therefore a good choice for investigations which focus on the amount and location of linguopalatal contact. For other articulations, e.g labial articulations, EPG is insufficient and EMA is more promising in spite of a much larger effort during recording and data processing. EMA is also appropriate for investigations focusing on precisely those articulations where linguopalatal contact is not involved, e.g. tongue lowering kinematics in open vowels. But from EMA alone it cannot be detected whether contact with the palate was made. It also does not provide information about the distance from the palate to a sensor.

3.1.2 Segmental make-up

Segmental make-up and its consequences are discussed with respect to three principles which are partially contradictory and partially congruent in their predictions on how consonant cluster components are coordinated.

1. the manner based ranking with implications from Mattingly (1981), Ohala (1992) and Wright (2004)

2. perceptual recoverability as outlined in Byrd (1996b, 1996a), Chitoran, Goldstein, and Byrd (2002), Silverman and Jun (1994) and Silverman (1997).

3. Biomechanical constraints as formalized in the DAC model (e.g. Recasens, Pallarès, & Fontdevila, 1997)

Intergestural coordination within clusters is investigated here by means of temporal overlap. Several measures have been proposed, e.g the interval of time in which two successive gestures or parts of them are active simultaneously or the temporal distance between certain landmarks of these gestures. Details will be given in Section 3.2.5.

3.1.2.1 Manner based ordering

Manner based ordering arises from Mattingly's 1981 assumption that an underlying principle of speech is the parallel transmission of information. This view regards the simultaneous presence of information about multiple gestures, in other words gestural overlap, as beneficial for speech in that it increases the speed in which information is transmitted. True parallel transmission can only be achieved in simple CV structures, since the consonant and the vowel are activated simultaneously, as shown by Öhman (1967). Sound sequences can, however, be optimized in terms of parallel transmission by ordering them according to the closeness of the sounds' constrictions which corresponds to a manner based ranking (obstruents, nasals, liquids, glides, vowels). To maximize the parallel transmission, speech sounds should be arranged to cycle from maximal constriction to minimal constriction in rank order and then back to constriction in reverse order. This is in fact a commonly found syllable pattern and according to Wright (2004) can optimize speech output in terms of perceptual cue robustness. Manner based ranking resembles sonority hierarchies as proposed by Sievers (1901) or Selkirk (1984). It is well worth mentioning here, that sonority hierarchies do not always treat obstruents as one single class. Wright (2004), for example, presents the following order as a common version of the sonority hierarchy: stops, fricatives, nasals, liquids, glides, vowels. Gestural overlap of two sounds is then assumed to correspond to the rank order, being maximal when they represent the opposite ends of the rank order and less when the distance on the rank order decreases. The least overlap is to be expected for sonority reversals, e.g. /nt lt/. The following grouping arises from the application of this principle to the clusters subject to this investigation:

1) /kl pl/ Obstruent+Liquid clusters are the most preferred clusters in the given range which is why most overlap is expected here. The rank order does not make any prediction concerning place of articulation. No differences can therefore be inferred for the C_1 stops.

2) /kn/ Since the consonants in this cluster are neighbors in the rank order, minimal overlap or at any rate less overlap than in 1).

3) /ks ps/ These clusters present the case where both consonants belong to the same class, the obstruents. In terms of parallel transmission, this appears to be the worst case and no overlap is expected or less overlap than in 1) and in 2).

3.1.2.2 Perceptual recoverability

Consonant sequences following the perceptual recoverability principle are coordinated in such a way that essential perceptual cues of one consonant will not be hidden by another consonant. Chitoran (2002), for instance, show that Georgian front-to-back clusters (e.g. /#tk/) are more overlapped than back-to-front clusters (e.g. /#kt/). This cross-linguistically consistent so-called *place order effect* (Hardcastle & Roach, 1979; Byrd & Tan, 1996; Zsiga, 2000; Kochetov, Pouplier, & Son, 2007) and the overall preference of front-to-back clusters are attributed to perceptual recoverability. In back-to-front stop-stop sequences, the second consonant is likely to delete the audible release of the first consonant, a major cue for recoverability, should overlap exceed some limit. This is not the case for front-to-back sequences. Purely manner based systems can not explain the place order effect and therefore, while no stop-stop clusters occur in German domain-initially, perceptual recoverability is used here to expand on the predictions made by the manner based account whenever possibly.

For the difference between /kn/ and /kl/ it was argued in Chapter 2 that early velic lowering in /kn/ is likely to impair the recoverability of stop burst characteristics and is therefore avoided. On this basis it was and is assumed that /kn/ is produced with less overlap than /kl/. While the manner based model predicts the same, the present argument provides stronger and physically graspable grounds for this assumption. Similarly, perceptual recoverability predicts less overlap in stop-/s/ clusters than in stop-/l/ clusters not because of sonority classes but because /s/ requires a certain amount of stationary frication (50 ms, Jongman, 1989) to be reliably recovered. Furthermore, the articulation of /s/ is heavily constrained due to precise formation constraints (Stone, Faber, Raphael, & Shawker, 1992). As for the two places of C_1 stop articulation, no predictions are made here and therefore no difference in overlap is assumed. Only the place-order-effect might indicate that clusters with /p/ as C_1 should be more overlapped than clusters with C_1=/k/. The effect has, however, been demonstrated to be most consistent in stop-stop clusters. In stop-continuant clusters, on the other hand, complete masking of C_1 by C_2 is rather unlikely.

To summarize, for the clusters under analysis here, perceptual recoverability predicts yalmost the same order of clusters sorted by overlap as the manner based ranking does: /kn/ /ks, ps/ < /kl, pl/. The exception is the fact that no difference for stop+/n/ vs. stop+/s/ clusters is directly derivable.

3.1.2.3 DAC Model

The DAC (degree of articulatory constraint) model of lingual coarticulation (Recasens, 2007; Recasens & Espinosa, 2006; Recasens & Pallarès, 2001; Recasens, 1999; Recasens & Pallarès, 1999; Recasens, 1997; Recasens, Fontdevila, & Pallarès, 1995) formalizes the anatomical properties of articulators and the assumption that they are biomechanically linked. Speech sounds are associated with DAC values that indicate the extent to which a sound is sensitive to coarticulatory effects of neighboring segments and to which extent a sound exerts coarticulation on neighboring segments. [p], for instance, is minimally constrained since it does not have an articulatory target for the tongue body and therefore assigned a low DAC value (DAC=1). The same holds for the vowel [ə]. Sounds with considerable tongue-dorsum involvement as [ʃ, n, i] and dark [ɫ], on the other hand, are highly constrained and specified for the maximal DAC value (DAC=3). Notably, dorsovelar sounds as [k] are largely subject to or outcomes of *blending* (see Recasens & Espinosa, 2006). They are therefore assigned lower DAC values than dorsopalatal sounds such as [c]. Sounds which have only indirect dorsal involvement such as alveolars are assigned intermediate DAC values (DAC=2). Laterality requirements (see Recasens, 2007) on the tongue dorsum however cause a slight raise of the DAC value for clear [l]. Added to that, Recasens (1995) find that German /l/ is less susceptible to coarticulation than clear [l] in French or Spanish. The DAC value of the fricative /s/ is raised to 3 due to the requirement of a precise formation of a medial groove (see Stone, 1992, as mentioned above).

For consonant clusters, it is assumed that two sounds with equal DAC values will repel each other as the degree of articulatory constraint of a sound is inversely related to the degree to which it exerts coarticulation on neighbors. In this case no gestural overlap is to be expected. The opposite is the case for clusters whose components have different DAC values. Let DAC of C_1 be 3 and DAC of C_2 be 1 then C_1 will strongly influence C_2 because the latter is only minimally constrained. C_2, in contrast, hardly exerts coarticulation on C_1 which is highly constrained. This setting does allow for gestural overlap.

While directionality of coarticulation is a major concern of the DAC model, it is of nearly no relevance to this study. Whether a sound is more sensitive to or influential in anticipatory than carryover effects does not make a statement concerning the amount of articulatory overlap between two adjacent segments. The model does, however, incorporate a means of predicting coarticulation differences resulting from the order of consonant sequences (e.g. /kt/ vs. /tk/) as suggested by Hardcastle and Roach (1979). In the given example complex *tongue repositioning* would be required for the transition from /k/ to /t/ in /kt/ while for /tk/ simply the contraction of the *longitudinalis inferior* would be sufficient. This account needs to be substantiated with data on the aforesaid muscle which unfortunately is very hard to come by.

The following is an attempt to rank the clusters of this study according to the amount of overlap they display as predicted by the DAC model.

/kl/ /l/ with a slightly raised intermediate DAC exerts coarticulation on /k/. Overlap is expected.

/pl/ Labials and velars should allow for a similar amount of coproduction. Therefore overlap for /pl/ should be as for /kl/

/kn/ /n/ with an intermediate DAC value exerts coarticulation on /k/ but not as much as /l/. Less overlap is expected than for /kl/

/ks/ /s/ with a maximal DAC value should exert even more coarticulation than /l/. Accordingly, more overlap is expected in this cluster.

/ps/ /ps/ should behave to /ks/ as /pl/ to /kl/, i.e. a similar amount of overlap is expected as in /ks/.

The following order of overlap will tentatively be assumed: /kn < kl+pl < ks+ps/

3.1.3 Prosodic variation

Two influential accounts of prosodically conditioned phonetic detail have been presented in the previous chapter. They are therefore only briefly summarized here. a) Pierrehumbert and Talkin (1992) observed that for simple CV syllables a preceding phrase boundary shifts the entire syllable in a consonantal direction in terms of gestural magnitude. Prominence–here in terms of phrasal accent–, on the other hand, shifts the entire syllable in a vocalic direction. In other words, boundaries strengthen the syllable onset whereas prominence affects the nucleus. b) The π gesture approach Byrd, Kaun, Narayanan, and Saltzman (2000), Byrd and Saltzman (2003), Byrd, Krivokapić, and Lee (2006)–associated with the framework of Articulatory Phonology (e.g. Browman & Goldstein, 1992a)–formalizes effects of boundary strength in terms of a so called *π-gesture* which acts as a local decelerator on gestural scores. Articulatory gestures in the vicinity of such a π-gesture are slower and at the same time stronger. The π-gesture, the *prosodic* gesture, itself is associated with prosodic boundaries and its size correlates with the strength of the prosodic boundaries. Articulatory gestures at utterance boundaries should therefore be under the influence of a large π-gesture and should display considerable lengthening and strengthening as opposed to articulatory gestures remote from any prosodic boundary.

Unlike Chapter 2, this chapter also considers phrasal accent as an additional contributor to prominence besides lexical stress. Quite a number of studies have investigate the influence of phrasal accent (Turk & Shattuck-Hufnagel, 2007; Cho & McQueen, 2005; Mooshammer, 2010; de Jong, 2004; Fougeron, 2001; Cambier-Langeveld, 2000; Cambier-Langeveld & Turk, 1999; Turk & White, 1999; Meynadier, Pitermann, & Marchal, 1998; Turk & Sawusch, 1997; Dilley, Shattuck-Hufnagel, & Ostendorf, 1996; Harrington, Fletcher, & Roberts, 1995; Eefting, 1991) on articulation. These works primarily focus on the domain of so-called accentual lengthening of segments associated with phrasal accent. The impact of accent on a word is manifested most strongly on the primary stressed syllable in terms of prosodic features (e.g. pitch, duration, amplitude). Effects of accent are, however, also found to go beyond the lexically stressed syllable–predominantly in terms of rightward spreading but there also appears to be a reliable albeit small leftward propagation (see esp. Turk & White, 1999). Recent contributions by Cho and McQueen (2005) and Cho and Keating (2009) are of particular interest here because they investigate articulatory and acoustic variation as a function of accent, stress and position. Previous works never considered more then two of the prosodic factors. In the case of the latter (Cho & Keating, 2009), articulatory data are discussed which offers some opportunities for comparison with the data in the current investigation. In an EPG study of the constructed names *Nebaben* /nɛbəbɛn/ and *Tebabet* /tɛbəbɛt/, they varied position (phrase initial vs. phrase medial), lexical stress (first vs. last syllable) and phrasal accent (narrow focus on test word vs. other word) in order to test by means of articulatory and acoustic analysis of the first CV syllable a) whether position and prominence affect articulation along the same physical dimensions, b) whether positional effects are local to the boundary-adjacent segment, c) for the domain of accentual effects and d) whether stress and accent cumulatively attribute to prominence. Focusing on the measures relevant to this study, they found that spatial measures of the consonant were only affected by position, while most measures of the vowel and nasal duration were affected by prominence rather than by position. Some measures (EPG seal duration, /t/ VOT, and vowel amplitude) appeared to be influenced by both prominence and position. Positional effects were therefore found not to be local to the boundary adjacent segment. The data rather support a graded effect of boundary strength decreasing with distance from the boundary which is in line with the π-gesture approach. The data also confirm that accentual effects are not local to the primary stressed syllable. Tokens with primary stress on the final syllable were just as much affected by accent in terms of e.g. seal duration as tokens with primary stress on the first syllable. On the other hand, some vowel measures emerged to be subject only to effects of stress. Solely measures of energy (nasal, /t/ burst, vowel amplitude) turned out to be affected in a cumulative fashion. The measures of greatest interest for this study are those concerning the consonantal onset of the first syllable of the target word. More precisely–since this study deals with durational parameters derived from EMA data–the only transferable measure is EPG seal duration, which is fortunately one of the few to show effects of all three prosodic parameters. In the case of the present study, there are two consonants in the onset of the initial syllable. This provides grounds for testing the graded nature of positional effects. For a graded effect, the boundary's impact on a word-initial CC cluster should be more potent on C_1 than on C_2. It should be interesting to see whether the reverse picture emerges for prominence, i.e. whether prominence related lengthening effects can be regarded as graded, too.

3.1.4 Predictions

Based on the models presented in the introduction the following predictions are made. In shorthand the C_1 stops /p/ and /k/ will be represented as /S/ (e.g. /Ss/ vs. /Sl/ clusters) and the alveolar C_2s will be represented as /C/ (e.g. /pC/ vs. /kC/ clusters).

Segmental make-up

S1) For the difference between /kl/ and /kn/ it is assumed in accordance with all models and the previous chapter that /kn/ should be less tightly timed than /kl/.

S2) For the four-way contrast of /kl/ – /pl/ – /ps/ – /ks/, the models vary. According to the DAC model, /Ss/ clusters should be more overlapped than /Sl/ clusters. Conversely, manner based ranking and perceptual recoverability both predict more overlap in /Sl/ than in /Ss/ clusters. None of the models give rise to the expectation that C_1 place of articulation should determine the amount of overlap. The articulators in /pC/ clusters are anatomically less dependent than in /kC/ clusters. It is therefore hypothesized that /pC/ clusters can overlap to a larger extent.

Prosodic variation

P1) Prosodic boundaries that may or may not precede the cluster will have the strongest impact at their centers, i.e. the peak of the π-gesture. A strong effect of boundary strength / prosodic position on C_1 is therefore assumed. C_2, being further removed from the center of the effect, is less affected if at all since the effect of boundaries is considered to be of a graded nature. Prosodically induced variation in articulatory overlap might reveal information on lexical specification of consonantal cohesion, i.e. should strong changes of overlap induced by prosodic variation emerge one might conclude that intergestural timing of the cluster in question is not specified by its segmental make-up. Byrd and Choi (2010) find less overlap (higher latency) in clusters adjacent to a strong prosodic boundary (i.e. phrase initial clusters).

P2) Lexical stress is assumed to be centered on the stressed syllable's nucleus which renders C_2 closer to the point of impact than C_1. Consequently, a stronger effect is expected on C_2 than on C_1. Effects of stress on overlap are interpreted in the same way as effects of boundary strength. This prediction is made with some reserve, since no consistent effect of stress was found in Chapter 2. However, effects of stress on the syllable onset have been observed in the literature (e.g. Bombien, Mooshammer, Hoole, Rathcke, & Kühnert, 2007; Cho & McQueen, 2005, for a subset of the data presented in Chapter 2).

P3) Effects of accent are assumed to be found independently of lexical stress, i.e. accentual lengthening of C_2 is expected in stressed and unstressed tokens. Since Cho and Keating (2009) did not find a cumulative effect of stress and accent, accentual lengthening should be just as strong in stressed as in unstressed tokens. According to Turk and White (1999), however, accentual lengthening should be weaker in syllables preceding the stressed syllable since leftward spreading of accent was only rarely encountered. As accent concurs with lexical stress, it's effect should be stronger on C_2 than on C_1 if accentual lengthening is of a graded nature.

Segmental make-up and prosodic variation At this point, there is no empirical basis or literature to support any assumption about different behavior of the different clusters under prosodic variation. But, of course, it will be discussed.

3.2 Method

3.2.1 Speech material

Unlike the previous chapter which analyzed the clusters (/kl kn sk ks/), this study deals with the sequences /kn kl pl ks ps/[1]. For each cluster, one or two disyllable German target words beginning with the cluster were selected: one with lexical stress on the first syllable (henceforth: the stressed condition; e.g. 'Psalmen' [ˈpsal.mən] (psalms)) the other with lexical stress on the second syllable (henceforth: the unstressed condition; e.g. 'Psalmist' [psal.ˈmɪst] (psalmist)). All target words were embedded into four carrier phrases each to elicit different levels of prosodic boundaries before the target word. In each case, the target word – and therefore the cluster – was preceded by an unstressed open vowel ([a, ɐ]). In the *utterance initial* condition (U), the target word is the first word in the second of two consecutive sentences. The *phrase initial* condition (P) has the target word as the first word in a sub-clause. In the *list* (L) condition the target word appears as the third of four items. Finally, in the *word* condition (W), the target word is preceded by a simple word boundary. Additionally, the target word was embedded into a deaccentuation context (D) to be compared with class W for effects of accentuation. The complete speech material can be found in appendix A.2.[2]

3.2.2 Recordings

Four speakers (3 female, 1 male; age: 20 - 25) were recorded by means of EMA. The female subjects are of urban Bavarian (Munich) origin with a standard-like German pronunciation. The male speaker originates from Ingolstadt with quite some dialectal

[1]The inclusion of /pn/ would be highly favourable in this analysis. However, to keep the size of the corpus within certain limits a number of clusters had to be left out including /pn/ due to its very low functional load in German.

[2]While similar, the material for the clusters /kl kn sk ks/ in this chapter is not entirely identical with the material in chapter 2. In order to enforce prenuclear accents in all conditions some utterances had to be changed.

coloring. No speech or hearing disorders were reported. The speech material was presented to the speakers for reading five times in randomized order. The prompting system used for the recordings also triggered the simultaneous acquisition of EMA and audio data. Raw EMA amplitudes were stored on hard disk for subsequent processing. Audio data were recorded on a multichannel DAT device (Sony PC208Ax) together with a synchronization impulse for later segmentation.

3.2.3 Prosodic grouping

In order to analyze the impact of varying the prosodic boundary preceding the target clusters, all utterances were assigned to one of three prosodic groups: 1) BiG boundary (BG), 2) SMall boundary (SM), 3) prosodic WorD boundary (WD). To this end, the utterances were divided into three types of phrases on the level of *intermediate phrases* (see Beckman & Pierrehumbert, 1986): 1) preceding phrases, 2) the target phrase and 3) following phrases.

Only one or two phrases per utterance are of immediate relevance in this study: One of them is quite obviously the target phrase as it contains the cluster. The cluster appears at the very beginning of the phrase with one exception: When the cluster is not preceded by a phrase boundary but merely a prosodic word boundary, the target phrase constitutes the entire utterance with the target word somewhere medial. In all other cases, the strength of the boundary was determined by properties of the transition between the target phrase and the immediately preceding phrase as described in Peters (2006). On the part of the preceding phrase, these properties are the presence or absence of a pause and of final lengthening as well as the quality of the boundary tone (low or high). On the part of the target phrase, the parameter step was labeled as up, down or equal depending on the shift of the f_0 onset relative to the f_0 offset of the preceding phrase.

As alluded to above, all utterances whose target phrase was not preceded by another phrase were assigned to the WD group (prosodic word). In accordance with Peters (2006), utterances with boundaries involving either a pause or the combination of a low boundary tone plus final lengthening plus an up-step of f_0 across the boundary were classed as BG (big boundary). All utterances with a boundary that did not meet the requirements for the BG group, i.e. all remaining utterances, were assigned to the

Table 3.1: Mapping of utterance type to prosodic groups (separate for both stress categories). For abbreviations see text.

Prosodic grouping	U	P	L	W	D
S					
Pi	92	65	23	1	0
Pm	0	29	74	97	0
U					
Pi	78	55	26	1	0
Pm	0	20	50	79	0

Table 3.2: Contingency table for accentuation analysis. For abbreviations see text.

	S	U
A	98	80
D	94	72

SM group (small boundary).

In the end, however, very few tokens were categorized as big boundaries. Consequently, a binary opposition was established between phrase medial (Pm) consisting of all items categorized as WD and phrase initial (Pi) consisting of both the SM and BG categories. Table 3.1 gives an overview of how the positional categories map the syntactically defined utterance types as listed in appendix A.2.

It has to be noted that for the analysis of accentuation no prosodic grouping was performed. Instead, all utterances of class W (prosodic word) were classified as accented (and for that purpose labeled A) and all utterances of class D (deaccented) were classified as deaccented. A contingency table is given in Table 3.2. In the case of the accentuation analysis, both prosodic factors (accent and stress) were fully determined by the speech material's design. In the analysis of the impact of prosodic position and stress the latter is also fully determined by the design, unlike prosodic position categories which were assigned *post hoc*.

3.2.4 Data processing

The physiological data were analyzed in Matlab and Emu. Semi-automatic algorithms computed the time points of articulatory landmarks within a given interval using EMA-coil trajectories (or signals derived from such trajectories, see below) and their velocity signals, i.e. their first derivatives. Different types of velocity signals were used depending on the articulator in question. The following outlines which articulators, EMA coils, trajectories and velocities were involved in the analysis of the sounds in the focus of this work.

TB The articulatory trajectory resulting from the EMA coil glued to a position on the tongue dorsum is referred to as TB (tongue back). Its vertical (up-down) component was used for the analysis of the velar stops /k g/ because they are produced by lifting the tongue dorsum to the soft palate. Accordingly, the velocity signal was computed as the first derivative of the vertical component only.

TT The TT trajectory captures the movement of the tongue tip. It was utilized for the analysis of the coronal consonants /l n s/. Coronal constrictions can involve both tongue tip lifting and tongue tip fronting. Therefore, the tangential velocity was used for landmark detection. The tangential velocity (v_t) is defined as the square root of the sum of the squared first derivatives of the trajectory's vertical and anterior-posterior dimensions (v_x, v_y): $v_t = \sqrt{v_x^2 + v_y^2}$. The lateral component can be excluded here since it does not substantially contribute to articulations of the tongue back.

LA The bilabial stops /p b/ are special in that two active articulators, the lower and the upper lip, are involved in two dimensions, vertical and anterior-posterior. This can best be captured by computing the Euclidean distance between the respective EMA coils as a measure of lip aperture (LA). The Euclidean distance d between two points p, q in a three-dimensional space is calculated as $d(p,q) = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2 + (p_z - q_z)^2}$. In this analysis, the Euclidean distance was calculated in a two-dimensional space since only motion in the vertical and the anterior-posterior dimension but not in the lateral dimension were of interest. Landmarks were detected using the first derivative of the resulting signal.

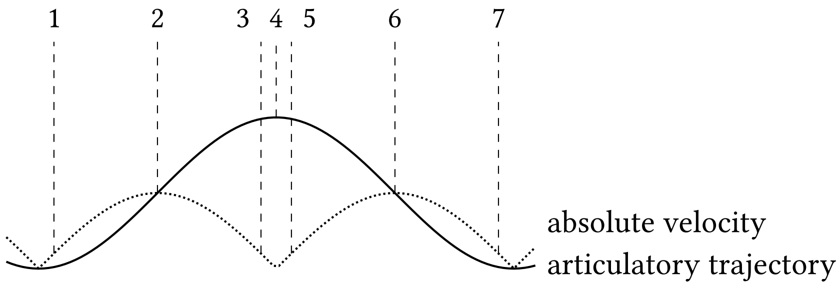

Figure 3.1: Schematic display of landmark positioning. 1: onset of gesture; 2: maximum velocity in onset; 3: begin of constriction plateau; 4: maximum constriction; 5: end of constriction plateau; 6: maximum velocity in offset; 7: offset of gesture. 1, 3, 5, 7 positioned using 20% threshold, see text.

Figure 3.1 shows the positioning of articulatory landmarks by reference to a trajectory and its absolute velocity. Maximum onset velocity (2), maximum constriction(4) and maximum offset velocity (6) are are easily detectable from the respective signal. The other landmarks, onset and offset of the gesture (1, 7) and begin and end of the constriction plateau (3, 5), are interpolated values and represent the 20% threshold of the difference between two adjacent extrema in the velocity signal, e.g. begin of constriction plateau (3) is position at the 20 % threshold between maximum onset velocity and maximum constriction (where velocity is zero). The 20% threshold method has been found to yield the most stable results when compared to other static and dynamic thresholds. Using zero crossings (or local minima for tangential velocities) does not yield reliable landmarks because sometimes more than one zero crossing can occur during and after the the target phase (see Kroos, 1996). Figure 3.2 is a real example of the word "Claudia" uttered by speaker *f01*.

3.2.5 Measurements of articulatory timing

Effects of segmental setup on the temporal coordination of consonant clusters are analyzed here in terms of gestural overlap. A large number of studies have applied a nearly equally large number of measures for articulatory timing or gestural overlap. Oliveira, Yanagawa, Goldstein, and Chitoran (2004) have in return made a comparison of various measures of articulatory timing in consonant clusters across a number of corpora. The most stable measures, i.e. the ones with the lowest variation coefficient, turned out to be measures of latency between the two consonants' gestures, specifically between their targets (begin of constriction plateau in Figure 3.1) and usually normalized by C_1 formation duration, i.e. the interval from gestural onset to target attainment. Here, target latency will be defined as the difference between T3 and B3 in Figure 3.2. A similar measure has been applied in EMA studies by e.g. Byrd and Choi (2010). The measure used in Chapter 2, plateau overlap, was defined as the interval between C_1 constriction plateau release (B5 in Figure 3.2) and C_2 constriction plateau onset (T3 in Figure 3.2), i.e. the overlapping section of both plateaus (positive) or the lag between them (negative). Both measures, plateau overlap and target latency, will be employed in this study for consistency with respect to Chapter 2 and for better comparability with other studies. It has to be noted, however, that the notion of a plateau in the EMA data is different from that in EPG data. In the latter, the plateau boundaries were placed with respect to the time-course of the contact percentage in the anterior or the posterior region of the EPG palate. In the EMA data, plateaus were defined with respect to the trajectories of coils attached to the articulators. In consequence, EMA plateaus in the case of lingual articulators can be expected to be shorter because linguo-palatal contact may be present before EMA plateau attainment and may also extend until after EMA plateau release. Assuming that EMA plateaus and EPG plateaus are still centered around the same target, EMA plateaus of two subsequent consonants are less likely to overlap since they are shorter than EPG plateaus.

3.2.6 Statistics

The R environment (R Development Core Team, 2009) was used for statistical computing and the preparation of statistical graphics (see Sharpsteen & Bracken, 2009, for LaTeX-ready output). Tests across speakers were conducted using an R extension for linear mixed-effects modeling (LME) by Bates and Maechler (2009) (see Baayen, 2008, for a detailed description) specifying subject as a random factor. Mixed model-ling was preferred over repeated measures ANOVAs mainly because of two major disadvantages in the latter: a) Repeated measures ANOVAs are calculated on cell means, which requires data manipulation and entails loss of data. b) Repeated measures ANOVAs rely on a balanced design of the data. Both these points can be disregarded with linear mixed-effects model-ling.

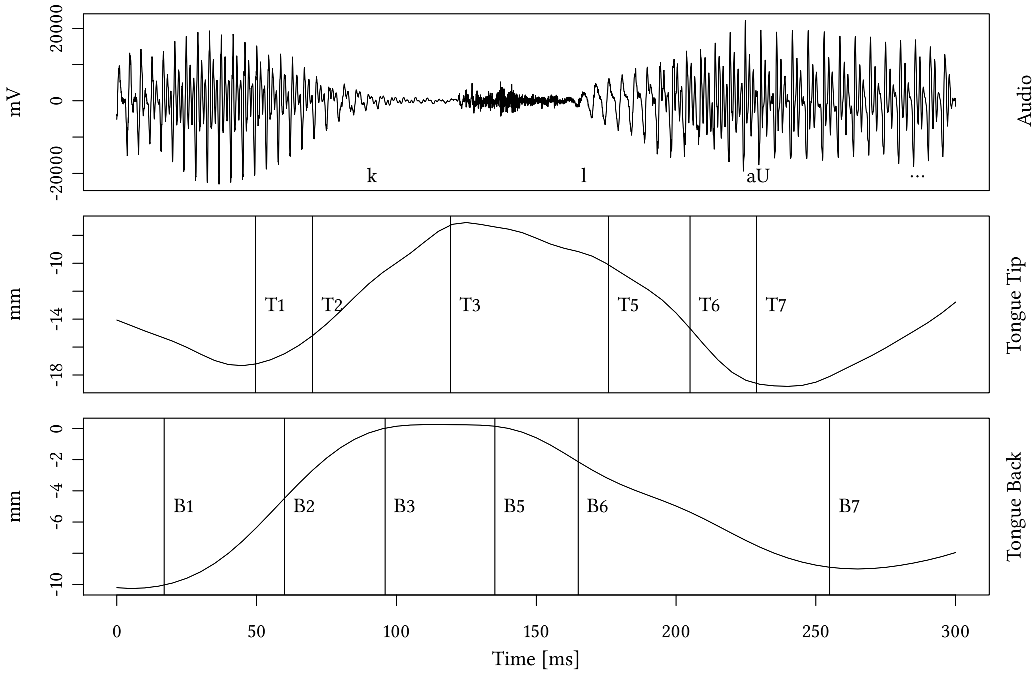

Figure 3.2: Extraction of temporal parameters in a case of the word "Claudia". Vertical lines correspond to articulatory landmarks labeled according to articulator (T = tongue tip, B = tongue back) and landmark number as defined in Figure 3.1 (max. constriction (4) omitted for clarity).

In order to avoid collinearity between the predictors, they were coded and centered. All interactions between the predictors were included in the models if not explicitly stated otherwise. Log-likelihood tests for goodness of fit were applied in order to assess whether models improved by allowing the slopes of the fixed effects to vary with the random factor subject. In case of model improvement (indicated by increased log-likelihood) by inclusion of another random factor additional analysis on the speaker level was conducted. In the present case, inclusion of random factors was never necessary or justified. The actual set of necessary predictors was determined by the individual problems and will be presented and explained in the subsections of Section 3.3. Statistical results of the fixed effects are presented by the estimates of the regression coefficients of the model β (for the intercept, this is the grand mean in case of centered factors) and the standard error of β.

A drawback in LME modelling is that degrees of freedom in the denominator of the resulting F statistics cannot be estimated with sufficient reliability. Baayen (2008) proposes an anticonservative approach which amounts to subtracting the number of levels in the relevant factors from the total number of observations. In the more complex cases in this study this number can easily reach values above 1500. Following Reubold, Harrington, and Kleber (2010), the degrees of freedom in the denominator were rather arbitrarily set to 60 to avoid obtaining significance for only small changes in the F value.

Table 3.3: Dependent variables and predictors of the statistical models and their descriptions. See Figures 3.1 and 3.2 for reference points relevant for measurement calculation.

Dep. Var.	Description	Note
PLATC1	$= B5 - B3$: C_1 plateau duration (constriction duration)	
PLATC2	$= T5 - T3$: C_2 plateau duration (constriction duration)	
TLAT	$= T3 - T5$: target latency: interval from C_1 plateau onset to C_2 plateau onset	larger values mean longer latencies
TLATN	= TLAT $/B3 - B1$: target latency normalized by interval from C_1 gestural onset to C_1 constriction onset (formation)	
POVER	$= B5 - T3$: plateau overlap: overlapping interval of C_1 plateau and C_2 plateau or the lag between them	positive for overlap, negative for lag
POVERN	= POVER $/T5 - B3$: plateau overlap normalized by the interval from C_1 plateau onset to C_2 plateau offset	
Predictor	**Description**	**Note**
C1P	C_1 place of articulation, velar /k/ or labial /p/	
C2M	C_2 manner of articulation, lateral /l/ or nasal /n/	
POS	prosodic position, phrase-initial (Pi) or phrase-medial (Pm), i.e. adjacent to a phrase boundary or not	
STRESS	lexical stress, on the target syllable (stressed, S) or on not on the target syllable (unstressed, U)	
ACC	phrasal accent, deaccentuated (D) or accented (A)	

3.3 Results

The results will be presented in two parts. Section 3.3.1 will deal with the clusters /kl/ and /kn/ exclusively. Section 3.3.2 will then turn to the clusters /kl pl ks ps/ where C_1 place and C_2 manner of articulation are varied systematically. The reasons for this separation are twofold: a) a separate analysis of /kl/ and /kn/ allows for a better comparison to the results in Chapter 2 and b) dropping /kn/ in the second set of clusters enables a clean 2×2 design of the data set. Both parts will follow the same structure. In a first step, the temporal properties that are due to segmental make-up will be presented based on prosodically unmarked tokens, i.e. only tokens of the syntactical class W (i.e. simple word boundary) with stress on the first syllable of the target word are included. In the second step, the emphasis will lie on how these properties are affected by prosodic variation. The following temporal parameters will be analyzed: 1. the duration of C_1 plateau 2. the duration of C_2 plateau 3. the plateau overlap (normalized where applicable) 4. the target latency (normalized where applicable). The results will be visualized as vertical bar charts with the durations of C_1 plateau, C_2 plateau and plateau overlap juxtaposed. The duration of target latency is indirectly derivable as the interval from C_1 plateau onset (left edge) to plateau overlap offset. Each part of the results section will be headed by a table compressing the relevant statistics. These include interactions which substantially add to the understanding of the data. Detailed statistics tables can be consulted in 3.A

3.3.1 /kl/ vs. /kn/

In this section, the results for the comparison of /kl/ and /kn/ clusters as a function of segmental make-up (C_2 manner) and prosodic variation (position, accent and stress) are presented.

3.3.1.1 Segmental make-up

As evident from the overview given in Table 3.4, articulatory timing varies considerably as a function of C_2 manner. Both the latency measure and the overlap measure show significant responses to C_2 manner variation, with higher significance in the absolute measures as compared to the normalized data. For plateau overlap, /kl/ overlaps 21±4

Measures	C_2 manner (l/n)
POVERN	l > n**
POVER	l > n***
TLATN	n > l**
TLAT	n > l***
PLATC1	l < n*
PLATC2	*n.s.*
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05	

Table 3.4: Overview of segmental effects on articulatory measures in /kl/ vs. /kn/ clusters.

ms more than /kn/ ($F[1,60] = 28.47, p < 0.001$) or 15±4 % in the relative measure ($F[1,60] = 11.00, p < 0.01$). Analogously, /kn/ has longer latency than /kl/: 26±5 ms ($F[1,60] = 26.16, p < 0.001$) or 21±3 % in the normalized measure ($F[1,60] = 9.13, p < 0.01$). It is noteworthy that C_1 plateau duration is 5±3 ms longer in /kn/ than in /kl/ ($F[1,60] = 4.24, p < 0.05$). No such effect is found for C_2 plateau duration. Overall, this amounts to a shorter total duration in /kl/ than in /kn/ primarily due to less overlap in the latter. These results are visualized in Figure 3.3. They are consistent with the results presented in Chapter 2 as well as with the prediction made in P1.

3.3.1.2 Prosodic variation

Table 3.5: Overview of segmental effects as well as position and stress on articulatory measures in /kl/ vs. /kn/ clusters.

Measure	Position (Pi/Pm)	Stress (S/U)	C_2 manner (l/n)
PLATC1	Pi > Pm ***	*n.s.*	*n.s.*
PLATC2	Pi > Pm (/kl/)	S > U **	*n.s.*
	POS × C2M *		
POVER	*n.s.*	U > S ***	l > n ***
TLAT	Pi > Pm ***	S > U *	n > l ***

Position and stress Table 3.5 gives an overview of the effects of varying position and stress on the articulation of the consonant clusters /kl/ and /kn/. The prevalent coordination pattern of more plateau overlap (POVER) and shorter target latency (TLAT)

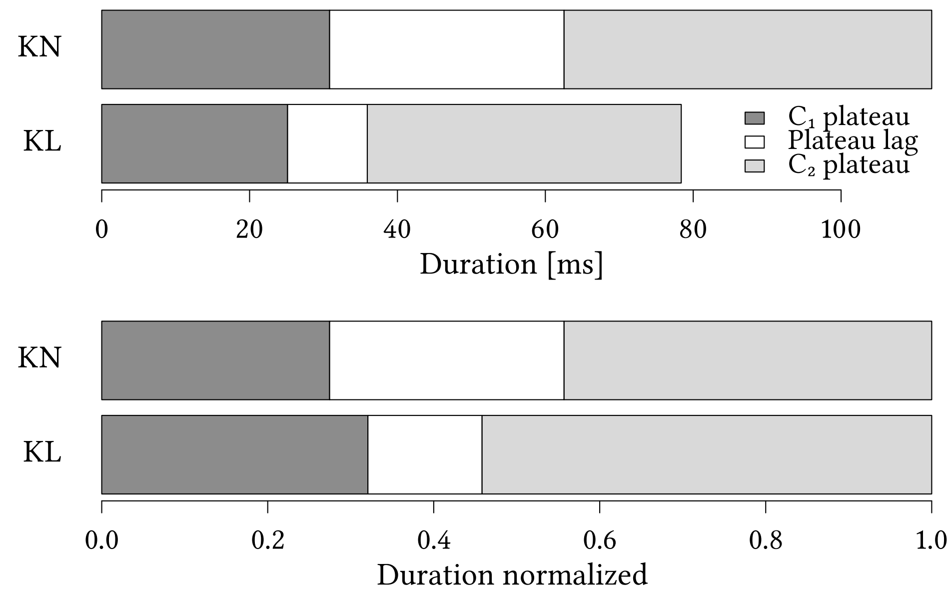

Figure 3.3: Overlap pattern of /kl/ and /kn/ clusters. C_1 plateau, plateau lag and C_2 plateau durations.

observed in /kl/ clusters in Section 3.3.1.1 is resistant to prosodic variation. /kn/ is by average 31±2 ms less overlapped than /kl/ ($F[1, 60] = 163.63, p < 0.001$). While no effect of position is detectable for plateau overlap, stress significantly shortens overlap by an average of 8±2 ms ($F[1, 60] = 13.45, p < 0.001$). Similar to the results for overlap, target latency is 29±3 ms longer in /kn/ than in /kl/ ($F[1, 60] = 81.36, p < 0.001$). There is also an effect of stress albeit not as clear as for plateau overlap: stressed items have 8±3 ms longer target latency than unstressed items. Unlike plateau overlap, however, target latency is quite sensitive to position: phrase-initially, the latency is 18±3 ms longer than phrase-medially ($F[1, 60] = 30.95, p < 0.001$).

In Chapter 2, effects of prosodic position on plateau overlap were also found to be more pronounced in /kn/ than /kl/ with less overlap at higher boundaries. Here, there is no such effect on plateau overlap. With regard to prediction P1 there appears to be more evidence for an approach favoring the specification of intergestural timing by segmental make-up. However, the results for target latency do not seem to fit into such an approach. The stress sensitivity of coordination measures as observed here is stronger in comparison to the findings in Chapter 2 where a consistent effect of stress

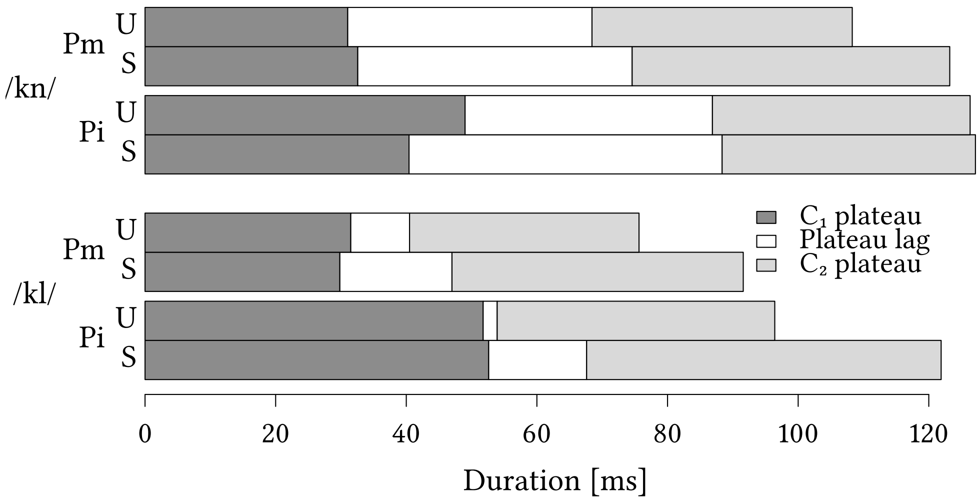

Figure 3.4: Overlap pattern of /kl/ and /kn/ clusters as a function of prosodic position (phrase initial (Pi) or phrase medial (Pm)) and lexical stress (stressed (S) and unstressed (U)).

on plateau overlap was only found in /kn/. This confirms previous findings (Bombien, 2007) rather than a specification by segmental make-up, see prediction P2.

There is very clear evidence for articulatory strengthening induced by prosodic position as predicted in P1. C_1 plateau duration is by average 17±2 ms longer phrase-initially than phrase-medially ($F\,[1,60] = 85.57, p < 0.001$). For /kl/ but not for /kn/, the strengthening effect even spreads onto C_2 plateau duration, as evident from the interaction of position and C_2 manner ($F\,[1,60] = 7.11, p < 0.01$). Evidence for a graded nature of boundary effects is even stronger here based on EMA data than for the EPG data in Chapter 2. As for prediction P2, C_2 plateau durations in stressed items are indeed longer than in unstressed items (7±2 ms, $F\,[1,60] = 9.35, p < 0.01$). Effects of stress do not extend to C_1 plateau duration. Stress-induced lengthening of platC2 could not be observed in the EPG data in Chapter 2. Effects of position, stress and C_2 manner are displayed in Figure 3.4. Detailed statistics are located in Table 3.11.

Table 3.6: Overview of segmental effects as well as accent and stress on articulatory measures in /kl/ vs. /kn/ clusters.

Measure	Accent (A/D)	Stress (S/U)	C_2 manner (l/n)
PLATC1	D < A (/kn/)**	*n.s.*	/kl/ > /kn/ *
	Accent × C_2 manner		
PLATC2	*n.s.*	S > U ***	*n.s.*
POVER	A < D *	*n.s.*	/kl/ > /kn/ ***
TLAT	A > D **	*n.s.*	/kl/ < /kn/ ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

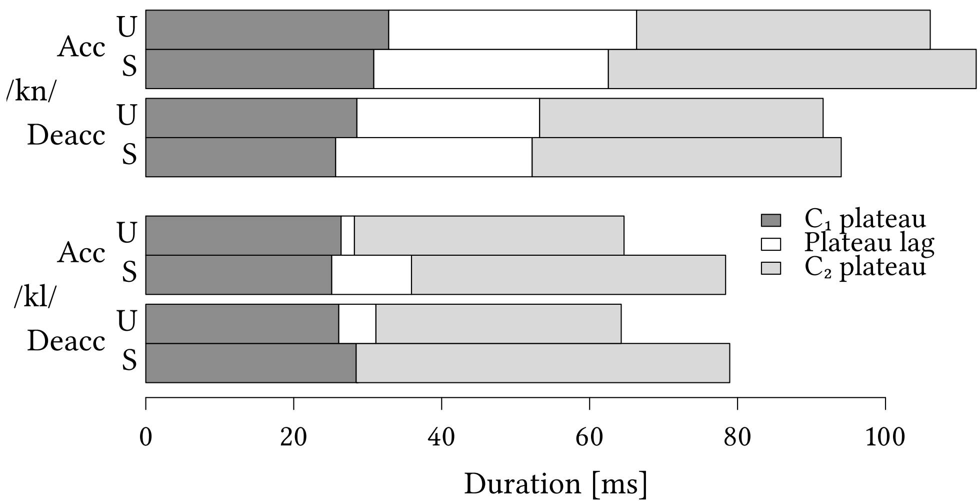

Figure 3.5: Overlap pattern of /kl/ and /kn/ clusters as a function of accentuation (accented (A) and deaccented (D)) and lexical stress (stressed (S) and unstressed (U)).

Accent and stress The overview in Table 3.6 clearly shows that the reoccuring observation of the difference in timing between /kl/ and /kn/ is not overridden by varying the prosodic factors accent and stress. /kn/ overlaps 25±2 ms less than /kl/ ($F[1, 60] = 145.11, p < 0.001$) and its target latency is 28±2 ms longer than that of /kl/ ($F[1, 60] = 135.42, p < 0.001$). This is so inspite of the effect that varying accent has on both measures: plateau overlap is 5±2 ms longer in deaccented items than in accented items

($F\,[1,60] = 7.89, p < 0.01$). The target latency is accordingly 7±2 ms longer in accented items (A) than in deaccented items ($F\,[1,60] = 9.78, p < 0.01$). Since there are no interactions of C_2 manner and accentuation the effect of C_2 manner or–in other words–cluster type on both measures can be regarded as very stable.

Strengthening effects due to prosodic variation are most prominent for stress on C_2 plateau duration which is on average 9±3 ms longer in stressed items than in unstressed items ($F\,[1,60] = 13.75, p < 0.001$) as predicted in P2. Accentual lengthening, however, was only scarcely encountered. Merely in /kn/ was C_1 plateau duration in accented items found to be 6±2 ms longer than in deaccented items. Longer C_1 plateau durations in /kl/ than in /kn/ by 3±1 ms support the corresponding finding in Section 3.3.1 but given the size of the effect the support must be regarded as rather weak. Prediction P3 is met with respect to the independence of accent and stress: accentual effects on C_1 plateau and on plateau overlap were found regardless of the location of lexical stress. It is somewhat surprising, however, that accentual lengthening skips C_2 entirely.

3.3.2 Stop + alveolar clusters

Similar to Section 3.3.1, this section describes the effects of segmental setup and prosodic variation on consonant clusters. Here, the set of clusters consists of /kl/, /ks/, /pl/ and /ps/. Since not only C_2 manner but also C_1 place is varied, both factors will be analyzed in terms of segmental make-up.

3.3.2.1 Segmental make-up

Table 3.7 summarizes the effects of segmental make-up. In both the relative and the absolute measure, plateau overlap is affected by C_2 manner. It is less in stop+/s/ clusters (/Ss/) than in stop+/l/ clusters (/Sl/). The difference amounts to 13±5 ms ($F\,[1,60] = 7.67, p < 0.01$). The relative difference is 10±5 % ($F\,[1,60] = 4.92, p < 0.05$). Accordingly, target latencies are 20±3 ms longer in /Ss/ clusters than in /Sl/ clusters ($F\,[1,60] = 33.98, p < 0.001$) or 26±7 % in the normalized measure ($F\,[1,60] = 15.56, p < 0.001$). The latter also appears to be influenced by C_1 place of articulation: The latency is 35±7 % longer in /pC/ than in /kC/ clusters ($F\,[1,60] = 14.73, p < 0.001$). This effect is not present for the absolute latency measure. Regarding prediction P2, these findings are in agreement with the manner based approach but not with the DAC

Table 3.7: Overview of segmental effects on articulatory measures in /kl/, /pl/, /ks/ and /ps/ clusters.

Measure	C_1 place (p/k)	C_2 manner (l/s)
POVER	*n.s.*	/l/> /s/ **
POVERN	*n.s.*	/l/> /s/ (/k/) *
		C_1 place × C_2 manner
TLAT	*n.s.*	/s/> /l/ ***
TLATN	/p/ > /k/ ***	/s/> /l/ ***
PLATC1	/p/ > /k/ ***	/s/ > /l/ (/p/) **
		C_1 place × C_2 manner
PLATC2	/p/ > /k/ *	/l/ > /s/ (/p/) *
		C_1 place × C_2 manner

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

model. Figure 3.6 illustrates these results. Detailed statistics are presented in Table 3.13.

The plateau durations of both consonants highly depend on the segmental make-up. PLATC1 is on average 8±2 ms longer for /pC/ than for /kC/ clusters ($F[1,60] = 13.38, p < 0.001$). PLATC1 is 8±2 ms longer when C_2 is the sibilant rather than the lateral ($F[1,60] = 10.35, p < 0.01$) but according to the interaction of C_1 place and C_2 manner of articulation, this is only true in /pC/ clusters ($F[1,60] = 5.63, p < 0.05$). The finding of longer plateau durations in /p/ than in /k/ is in accord with previous findings as reported by e.g. Byrd (1993) and is considered a universal according to (Maddieson, 1997). PLATC2 in /pC/ clusters exceeds that of /kC/ clusters by 7±3 ms ($F[1,60] = 5.15, p < 0.05$). C_2 manner likewise effects PLATC2 towards 7±3 ms shorter durations in the fricative than in the lateral ($F[1,60] = 6.07, p < 0.05$). As per the interaction of C_1 place and C_2 manner of articulation, the manner effect is only present in /pC/ clusters ($F[1,60] = 6.89, p < 0.05$). This behavior might be considered an aerodynamically conditioned effect: the /l/ plateau is shorter after /k/ than after /p/ in onset clusters because the velar's stop burst is filtered by the articulatory setup to contain laterality cues. This is not the case for /pl/ clusters since the lateral constriction precedes the source of the stop burst viz. the lips. Given that the timing in /pl/ and /kl/ is very similar, /l/ is lengthened after /p/ to ensure a sufficient amount of laterality cues in the signal. Details are presented in Figure 3.6 and Table 3.13.

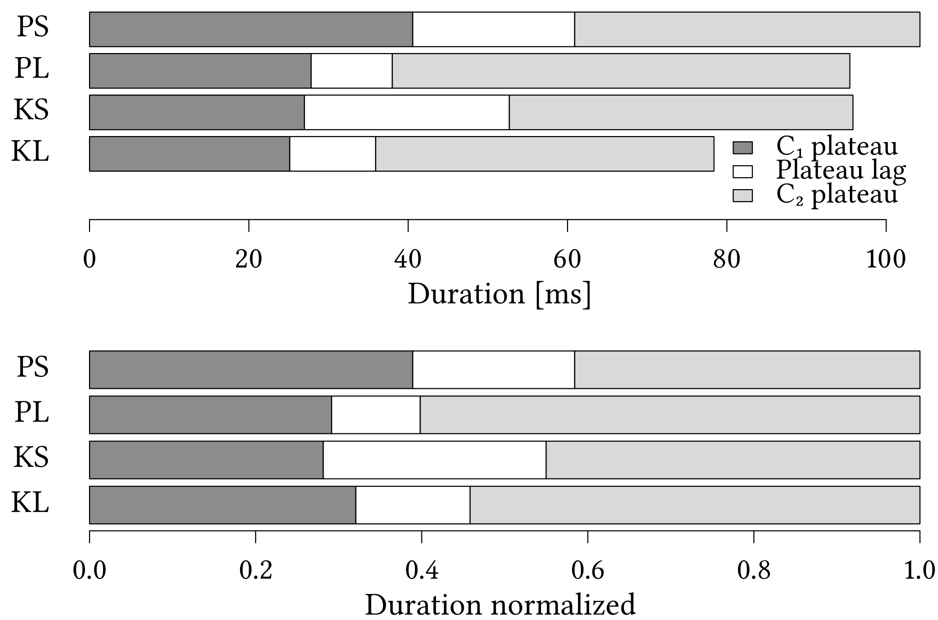

Figure 3.6: Overlap pattern of /kl/ and /kn/ clusters. C_1 plateau, plateau lag and C_2 plateau durations.

3.3.2.2 Prosodic variation

Table 3.8: Overview of segmental effects as well as the effects of position and stress on articulatory measures in /kl/, /pl/, /ks/ and /ps/ clusters.

Measure	Position (Pi/Pm)	Stress (S/U)	C_1 place (k/p)	C_2 manner (l/s)
PLATC1	Pi > Pm ***	S > U (Pi) ** POS × STRESS	p > k (Pm) *** POS × C1P	s > l (p) ** C1P × C2M
PLATC2	*n.s.*	S > U ***	p > k (Pi) * POS × C1P	s > l (U) * STRESS × C2M
POVER	Pi < Pm (p) *** POS × C1P	S < U (k)/(s) * C1P × STRESS C2M × STRESS	*n.s.*	l < s ***
TLAT	Pi > Pm ***	S > U (Pi) POS × STRESS	p > k *	s > l ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

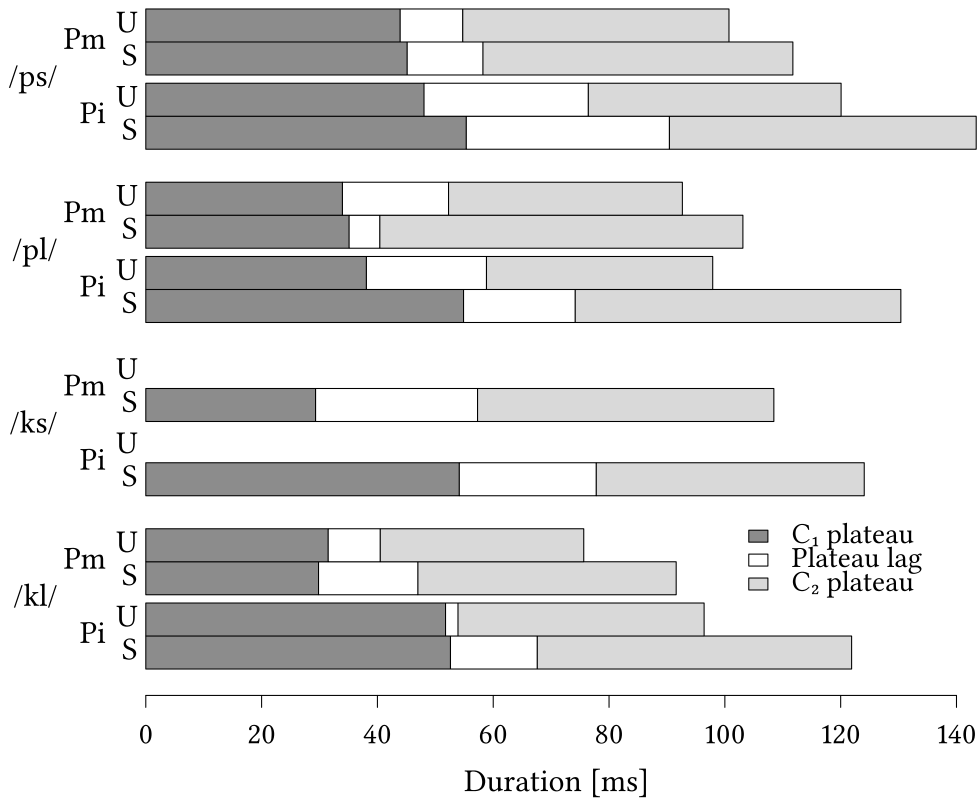

Figure 3.7: Overlap pattern of /kl/, /ks/, /pl/ and /ps/ clusters as a function of prosodic position (phrase initial (Pi) or phrase medial (Pm)) and lexical stress (stressed (S) and unstressed (U)).

Position and stress Details on the effects of segmental make-up in combination with variation of position and stress are presented in Figure 3.7 and Table 3.14. The overview in Table 3.8 clearly indicates that the measures of articulatory coordination are affected differently by segmental and prosodic factors. Concerning the segmental make-up, the common effect is that of C_2 manner: /s/ shows 6±2 ms less plateau overlap with the preceding stop than /l/ ($F\,[1,60] = 21.79, p < 0.001$). The target latency in /s/ is accordingly 11±3 ms longer for the fricative than for the lateral ($F\,[1,60] = 38.63, p < 0.001$). Target latency is also influenced by C_1 place of articulation: It is 6±3 ms longer for bilabials than for velars ($F\,[1,60] = 6.59, p < 0.05$). This shows that the segmentally conditioned coordination patterns as observed in Section 3.3.2.1 are not affected by prosodic variation.

There is no prosodic simple main effect in plateau overlap. In /pC/ clusters–but not in /kC/–overlap is shorter phrase-initially than phrase-medially ($F\,[1,60] = 22.18, p < 0.001$). Two two-way interactions emerge for the effect of stress on overlap: 1) stress interacts with C_1 place of articulation such that stressed /kC/ clusters are 18±5 ms less overlapped than their unstressed counterparts ($F\,[1,60] = 6.53, p < 0.05$). This obviously only applies to /kl/ clusters since the unstressed condition is missing for /ks/ clusters. 2) stress also interacts with C_2 manner ($F\,[1,60] = 5.07, p < 0.05$). Stressed stop+/s/ clusters are produced with 12±5 ms less overlap than unstressed stop+/s/ clusters. Again, since the unstressed condition is missing for /ks/ clusters this can only apply to /ps/ clusters.

Prosodic effects on target latency are different since they do not interact with segmental factors. A simple main effect of position gives evidence that latencies are by average 21±3 ms longer phrase-initially than phrase-medially ($F\,[1,60] = 70.43, p < 0.001$). Stressed items have 8±3 ms longer latencies than unstressed items ($F\,[1,60] = 7.90, p < 0.01$). Due to an interaction with position, however, this effect is restricted to phrase-initial occurrences (pos:strss: $F\,[1,60] = 8.23, p < 0.01$).

The variations of the plateau durations induced by segmental make-up presented in 3.3.2.1 are partially preserved and partially altered in prosodic variation. PLATC1 is longer before /s/ than before /l/ ($F\,[1,60] = 8.95, p < 0.01$), but only if C_1=/p/ according to the C1P × C2M interaction ($F\,[1,60] = 7.77, p < 0.01$). The above observation of longer plateau durations in /p/ than in /k/ is now restricted to the phrase-medial position ($F\,[1,60] = 18.44, p < 0.001$). Phrase-initially the difference is not maintained.

As in Section 3.3.1.2, the data show very clear effects of prosodically conditioned lengthening. Phrase-initial clusters have by average 18±2 ms longer C_1 plateau durations than phrase-medial clusters ($F\,[1,60] = 111.64, p < 0.001$). There is also an effect of stress on the C_1 plateau duration but only in terms of an interaction with position ($F\,[1,60] = 6.97, p < 0.05$): Phrase-initially, stressed clusters have longer C_1 plateau durations than phrase-medial. This finding is incongruent with previous findings presented in this work. Neither Chapter 2 nor Section 3.3.1 of this Segmental and prosodic effects (EMA) reports an influence of stress on PLATC1. This interaction can be interpreted as a cumulative effect of position and stress in the sense of Cho and

Keating (2009).

PLATC2 is more variable when viewed within prosodic variation as compared to Section 3.3.2.1: After /p/, C_2 plateaus are longer than after /k/ (4±2 ms, $F[1,60] = 4.74, p < 0.05$) but the C1P × POS interaction restricts this to phrase-initial occurrences ($F[1,60] = 5.23, p < 0.05$). The lateral's plateau duration is only shorter than the fricative's in unstressed clusters ($F[1,60] = 5.20, p < 0.05$). The considerable influence of stress on PLATC2 found in Section 3.3.1.2 for the clusters /kl/ and /kn/ is confirmed here. C_2 plateau durations are on average 11±2 ms longer in stressed than in unstressed clusters ($F[1,60] = 45.68, p < 0.001$) this is in agreement with the findings in 3.3.1.2 and prediction P2.

Table 3.9: Overview of segmental effects as well as the effects of accent and stress on articulatory measures in /kl/, /pl/, /ks/ and /ps/ clusters. Three-way interactions are described in the text.

Measure	Accent (A/D)	Stress (S/U)	C_1 place (k/p)	C_2 manner (l/s)
PLATC1		U> S (A); S> U (D)	p > k ***	s > l (p)***
		ACC × STRESS		C1P × C2M
PLATC2		S > U ***		
		C1P × ACC × STRESS		
		C2M × ACC × STRESS		
POVER				l > s (k) ***
		C2M × ACC × STRESS		C1P × C2M
TLAT			p > k (l) **	l > s ***
		C2M × ACC × STRESS	C1P × C2M	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

Accent and stress Table 3.9 is a synopsis of the combined effects of segmental make-up and the variation of phrasal accent and lexical stress. Segmental effects on the plateau overlap are generally in accord with those found in Section 3.3.2.1. /l/ allows for more overlap with the preceding stop than /s/ ($F[1,60] = 14.75, p < 0.001$). But here this is only true in /kC/ clusters as the interaction of C1P and C2M indicates ($F[1,60] = 16.54, p < 0.001$). There is one (three-way) interaction of C2M, ACC and

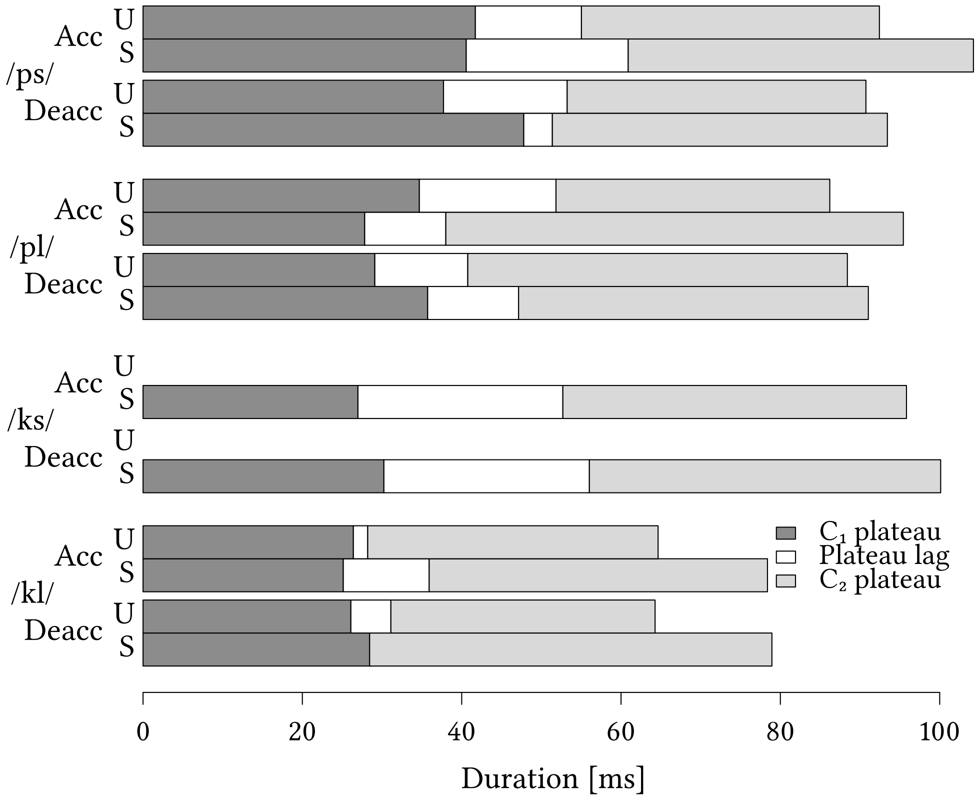

Figure 3.8: Overlap pattern of /kl/, /ks/, /pl/ and /ps/ clusters as a function of accentuation (accented (A) and deaccented (D)) and lexical stress (stressed (S) and unstressed (U)).

STRESS which is rather obscure and points towards less overlap in stressed stop+/s/ clusters under accent. Almost the same applies to target latency. Stop+/s/ clusters have longer latencies than stop+/l/ clusters ($F[1,60] = 54.10, p < 0.001$). There is one additional effect here regarding C_1 place. Target latency is longer in /pC/ than in /kC/ clusters ($F[1,60] = 30.09, p < 0.001$). The interaction of C1P and C2M, however, restricts this effects to clusters with C_2 being the lateral ($F[1,60] = 7.52, p < 0.01$). And then there is another of those obscure three-way interactions of C2M, ACC and STRESS which the author completely fails to make sense of ($F[1,60] = 6.27, p < 0.05$).

Segmentally conditioned effects on C_1 plateau duration are basically the same as observed in Section 3.3.2.1. Longer plateaus are found in /p/ than in /k/ ($F[1,60] = 64.22, p < 0.001$). PLATC1 is also longer when C_2 is the fricative rather than the lateral ($F[1,60] = 29.57, p < 0.001$), but as the C1P × C2M interaction indicates this is only so in /pC/ clusters ($F[1,60] = 9.75, p < 0.01$). There is only one effect related to prosodic variation which is the interaction of ACC and STRESS ($F[1,60] = 0.03, p > 0.05$). Under accent, stressed items have shorter C_1 plateaus than unstressed while, reversely, stressed items have *longer* C_1 plateaus when *deaccented.*

The segmental effects on PLATC2 are not maintained under variation of accent and stress. There is a main effect for stress confirming results of Section 3.3.1.2: C_2 plateau duration is 9±2 ms longer in stressed than in unstressed items. However, according to the three-way interaction of C1P, C2M and STRESS ($F[1,60] = 8.02, p < 0.01$), PLATC2 is only longer in the stressed conditions when /kC/ clusters occur under accented and when /pC/ clusters occur deaccented and also in accented stop+/l/ clusters (C2M × ACC × STRESS: $F[1,60] = 6.96, p < 0.05$).

3.4 Summary and discussion

The results are summarized and discussed with respect to the predictions given in Section 3.1.4.

3.4.1 Segmental setup

For the difference between /kl/ and /kn/ it is assumed—in accordance with all models and the previous chapter—that /kn/ should be less tightly timed then /kl/ in terms of overlap (but not precision). This is unconditionally confirmed in the present study. The finding is further corroborated by the analysis of target latency: Consecutive articulatory targets are achieved with shorter delay in /kl/ than in /kn/. Normalized target latency also proves to be a measure which can be successfully predicted by the cluster's components in the second set of clusters. It is longer after bilabial than after velar stops and shorter before the lateral than before the fricative. The following order arises from sorting the four clusters by the observed normalized target latency: /ps/ > /pl/, /ks/ > /kl/. Plateau overlap on the other hand is very sensitive to C_2 manner but not to C_1 place. According to the DAC model, /Cs/ clusters should be more overlapped than /Cl/

clusters due to stronger exertion of coarticulation. This assumption finds no support in the present data. Contrarily, manner based ranking and perceptual recoverability both predict more overlap in /Cl/ than in /Cs/ clusters. This is indeed the observed pattern. It is interesting to see that overlap in stop+/s/ clusters patterns with the repeatedly observed overlap in /kn/ rather than /kl/. It was a conclusion of Chapter 2 that /kn/ displaying so much less overlap and more prosodically induced variation in overlap than /kl/ might play a special role. The current data allow for the reverse: stop+/l/ and especially /kl/ appear exceptionally stable in overlapping to a rather high extent. The C_1 conditioned difference in normalized target latency and also in target latency in the analysis presented in Section 3.3.2.1 is not present in plateau overlap. This calls for a discussion of measures of intergestural coordination which is given below in Section 3.4.3. While the manner based approach accounts for the C_2 manner related differences, it fails to predict the latency differences which are due to C_1 place. Longer latencies in front-to-back clusters (as /ps pl/) than in back-to-front clusters (as /kl kn ks/) also don't pattern with the place-order-effect which should have the reverse outcome in order to prevent that C_1 release cues not be masked by C_2. In sum, none of the principles introduced in Section 3.1.2 succeeds in predicting the patterns observed here.

3.4.2 Prosodic variation

The results concerning prosodic boundaries, lexical stress and phrasal accent will be summarized one by one in this section. It was assumed that strong **prosodic boundaries** exert lengthening on the adjoining segments in a graded manner, i.e. strongest on the immediately adjacent segments and less strong on more remote segments. In accord with this prediction the present data show that C_1 is significantly longer phrase-initially than phrase medially. This finding is robust across all clusters under analysis in this study. As for the graded nature of positional effects, boundary conditioned lengthening of C_2 is only found in /kl/ clusters. In all other clusters, C_2 appears to be insensitive to boundary strength. This will be further discussed below with regard to overlap differences between the clusters. Prosodically induced changes in intergestural timing were speculated to be related to compositional specifications of the cluster. More precisely, the specification of a cluster which is resistant to influences of prosodic variation was hypothesized to be more rigid than the specification of clusters more prone to prosodically induced changes. /kC/ clusters are obviously more rigidly specified than /pC/ clusters since the latter allow for a much more pronounced effect of position on plateau overlap. This is illustrated by the interactions of cluster type and position in Table 3.14 and the corresponding Figure 3.7. Corresponding effects on target latency appear to be more general here since they also apply to /kC/ clusters. This will be discussed comparatively to plateau overlap in Section 3.4.3.

Lexical stress was assumed to be centered on the stressed syllable's nucleus which renders C_2 closer to the point of impact than C_1. Consequently, a stronger effect was expected on C_2 than on C_1. The data support this assumption across the board. PLATC2 is significantly longer in stressed than in unstressed position. There is also some evidence that effects of stress are to some extent graded. PLATC1 especially in /pC/ clusters exhibits some stress-induced lengthening in phrase-initial position. Stress shortens plateau overlap in /kl/ clusters and in /ps/ clusters. In all clusters considered in 3.3.2, however, there is a clear effect on target latency which is longer in stressed position phrase-initially. This is another point to be discussed in Section 3.4.3.

Effects of **accentual variation** are contradicting and unclear. The main assumption was that the largest impact would fall on the nucleus-adjacent consonant, i.e. C_2. But this is not the case. Instead, PLATC2 remains largely unaffected by varying accent apart from an obscure interaction with stress, C_1 place and C_2 manner. Generally, it was assumed that deaccentuation removes articulatory strength. Only in a small subset of the data (/kn/ but not /kl/ in Section 3.4.3) was this expectation confirmed as PLATC1 was shorter in deaccented than in accented position. With regard to temporal coordination, only the results of the comparison of /kl/ and /kn/ are in accord with the predictions. Accented items show less plateau overlap and longer target latencies than deaccented tokens which is in line with the assumption that prominence pulls gestures apart. This was not confirmed for the second set of cluster (/kl ks pl ps/).

3.4.3 Temporal coordination

Especially in the analysis of prosodically driven variation, plateau overlap and target latency yield very different results. Often, TLAT varies according to the predictions–i.e. longer latencies in phrase initial position or under prominence–where no such change is visible in POVER. But a similar difference is also observed in the analysis concerning the segmental setup in Section 3.3.2.1. As Figure 3.6 illustrates, target latency is significantly longer in /pC/ than in /kC/ clusters. Furthermore, target latency is longer for C_2 = /s/ than for C_2 = /l/. Plateau overlap on the other hand, turns out to vary as a function of C_2 only in /kC/ clusters with more overlap in /kl/ than in /ks/ (and also /kn/, see Figure 3.3). There is virtually no difference between overlap in /pl/ and /ps/ clusters.

Considering C_1 plateau durations at this point sheds light on the differences between the measures for temporal coordination because of course TLAT includes the entire duration of PLATC1[3]. TLAT is the relation of C_1 plateau onset to C_2 plateau onset while POVER is the relation of C_1 offset to C_2 onset. Figure 3.6 shows that PLATC1 clearly depends on the place of articulation in that plateau durations for /p/ are longer than those for /k/. This is in agreement with previous findings: Maddieson (1997) reports–based on observations by e.g. Byrd (1993)–that stop closures universally appear to be longer in bilabials than in velars. PLATC1 is also affected by the manner of articulation of C_2. For /kC/ clusters, it is longer when C_2 is the alveolar nasal than when it is the lateral or the fricative. For /pC/ clusters, PLATC1 is considerably longer when C_2 is the fricative than when it is the lateral.

As discussed above, C_1 is consistently lengthened in phrase-initial position. Again, the duration of PLATC1 plays a crucial role in TLAT in all five clusters as TLAT is lengthened as a function of prosodic position in parallel to PLATC1. This parallel variation is not found for POVER. In this regard, only /pC/ clusters emerge as sensitive to positional effects but never /kC/ clusters. Stress induced effects on TLAT are consistent such that clusters in stressed syllables tend to have longer latencies than clusters in unstressed syllables. At a closer look, however, it turns out that there are parallel effects of stress in /pC/ clusters on both TLAT and PLATC1. In /kC/ clusters, on the other hand, reduced POVER in stressed clusters leads to an increase of TLAT while PLATC1 remains unaffected by stress.

Plateau overlap seems to reflect segmental properties which are governed by certain principles which in turn stipulate how closely two oral articulations may follow each other. Target latency on the other hand takes these segmental properties into account as well as prosodic and C_1 specific factors. Effects on TLAT can be viewed as the summed

[3]TLAT could conceivably contain only portions of PLATC1 in case of positive occurrences of POVER. There were, however, no such occurrences in the present data.

effects on PLATC1 and POVER. In itself or on its own, TLAT is not a reliable descriptor of the segmental properties of a cluster. Maybe it can be regarded as parameter reflecting speech planning e.g. in terms of *phase windows* (Byrd, 1996a; Saltzman & Byrd, 2000) although a measure based on gestural onsets and not of target attainment might be more appropriate here, see below.

These considerations have to be taken into account when interpreting the results of TLAT and POVER. In doing so, the focus now returns to the question whether (and if so, how) temporal coordination of a cluster is specified by the segments it consists of. The comparison of /pC/ and /kC/ clusters indicates that articulator independence plays a crucial role in inter-gestural timing. Contrary to the expectation that articulator independence as in /pC/ clusters allows for more plateau overlap the results show the greatest value of plateau overlap in /kl/ clusters in spite of the relative dependency of tongue back and tongue tip. While plateau overlap in /ks/ and /kn/ clusters may be less than in /pC/ clusters, overlap remains unaffected by positional variation in all /kC/ while in /pC/ it varies as a function of position (lesser in initial position). As it is, independence of articulators seems to allow for more coordinatory variation while strong dependence appears to constrain the coordination possibilities to narrow windows. This is compatible with *phase windows* as described below.

In Articulatory Phonology, consecutive gestures are considered to be phased with each other in stipulated relations. Viewing gestures as full cycles (360°) of a critically dampened oscillator, phasing relations are expressed as the angle of the first gesture at which the second gesture is initiated. A phasing relation of 240° would therefore mean that the second gesture is initiated at target achievement of the first gesture (240° of 360°). According to Byrd (1996a), such relations are primarily determined by the type of gestures involved (/VC/, /CC/, /CV/) and by syllable position (e.g. /C#C/ vs. /#CC/). These factors constrain the totally available range of phasing relations (0° — 360°) to more limited *phase windows* which can be further narrowed by complex interactions of e.g. prosodic factors and segment identity. Should it be permissible to regard successive gestures' target achievements as the gestural phase angle which is subject to phasing relations,[4] then the present data indeed demonstrate how complex

[4]Ideally, a measure based on the gestural onsets, not the targets, should be used in such an investigation. This, however, requires normalization of intra-gestural timing: a first look at the kinematics of the complete gestures for velar and bilabial stops indicates that not only are gestural durations in the velar

these interactions can be. In tongue-back–tongue-tip clusters plateau overlap appears to be constrained by biological, mechano-inertial constraints to such an extent that the prosodic factor position is not effective. Successive gestures executed by independent articulators are less constrained in this regard and therefore position can influence the size of the phase window to a greater extent. Also, the fact that /pl/ and /ps/ overlap to the same extent while there is much more overlap in /kl/ and much less in /ks/ seems to indicate that /p/ sets an upper limit for overlap which offers a sufficient delay for both /s/ or /l/ formation. In /ks/ on the other hand, groove formation requirements prolong this delay until the stop gesture, i.e. the tongue back gesture is sufficiently released. This is not necessary in /pC/ clusters as the labial constriction does not constrain the tongue tip. /kl/ is the cluster in which the maximal observed overlap occured. Possibly, /l/ is extraordinarily well suited for coproduction with /k/. One might speculate that there is a minimal requirement for a central constriction in the alveolar region which merely involves the muscles for tongue tip raising. Laterality, or a lateral airstream may than be obtained passively by the impetus of the stop burst thus providing sufficient information for recoverability. This would render /l/ a rather unconstrained sound at least in this setting, meaning that /l/ has a lower DAC value after /k/ than in other positions.

A system that produces phasing relations as observed here in terms of target latency needs to know about biological constraints and inter-gestural cohesion. One such system is Nam, Goldstein, and Proctor, 2007title which is being developed at Haskins Labs (Nam, 2007) and also builds upon Articulatory Phonology. Underlying TaDA's gestural phasing is a planning model that employs coupled oscillators. Based on the observation that articulatory movements for syllable onset and nucleus are initiated at the same time (Öhman, 1967), onset consonants and the following vowel are considered to be coupled *in-phase* (simultaneous oscillation). In onsets that consist of more than one consonant in-phase coupling is assumed between the vowel and each of the consonants in the onset. This, of course, would yield total overlap of the onsets constituents and therefore additional *anti-phase* couplings are specified between consecutive consonants to enforce sequentiality. Competing constraints (Browman & Goldstein, 2000;

longer than in the bilabial but also target attainment and release occur later within the gestural cycle. This is probably due to the inertial properties of the articulators i.e. the tongue-dorsum is slower than the lower lip. Normalization for these effects is beyond the scope of this study.

Nam & Saltzman, 2003) emerge because physically two (or more) oscillators cannot be coupled to each other anti-phase and at the same time in-phase to a third. The gestural model reconciles this competition. In the case of a cluster with two consonants C_1 is shifted leftwards while C_2 is shifted rightwards (into the vowel). The temporal midpoint of both consonants, however, retains the same timing to the vowel as the midpoint of a simplex onset would. The output of the model therefore reflects the frequently reported *c-center effect* (Browman & Goldstein, 1988; Honorof & Browman, 1995; Byrd, 1995; Kühnert, Hoole, & Mooshammer, 2006). To accomplish different phasing relations for different clusters as observed in this study, the system needs to incorporate coupling strength in order to make the coordination more tight (as in e.g. /kl/) or more loose (as in e.g. /kn/). In this case, coupling strength can not be an intrinsic parameter of the model but rather an external property of segmental identity and cannot exceed certain limits to ensure that overlap does not inhibit gestural recoverability. Crucially, the framework of Articulatory Phonology (and hence also TaDA) does not contain a concept of segmental identity. The results presented here indicate that maybe it should. Again, using a split-gesture model might resolve some of the conflicts: Coupling C_2 to the release gesture of the stop in the case of /kn/ should cause a right-ward shift of /n/ and reduce overlap substantially. It is, however, not entirely clear how to motivate this drastically different coordination plan from a articulatory perspective.

Another line of research (Gafos, 2002; Shaw, Gafos, Hoole, & Zeroual, 2009; Gafos, Hoole, Roon, & Zeroual, 2010) bears on the fact that different languages can have different coordinations in consonant sequences. Coordinations follow a language-specific grammar and are based on the cluster's position (pre-vocalic, post-vocalic, inter-vocalic). In this approach, coordination is specified by reference to landmarks identical to those used in this study. The landmarks in Figure 3.1 are referred to as follows 1: onset, 3: target, 5: release, 7: r offset. In a cluster with large overlap *release* of C_1 would be coordinated with *target* of C_2. In a cluster with little overlap it would be a coordination of *r offest* and *onset* of C_1 and C_2 respectively. This is in stark contrast to the approach favored in the phase windows framework where onset of C_2 is coordinated with onset of C_1 by a certain phase window (i.e. a phase angle out of a window of possible values). The approach is being successfully applied to analyze complexity of syllable structure (so far limited to Moroccan Colloquial Arabic). An important difference between this line of research and the present study lies in that very fact. While the influence of segmental identity on coordination is acknowledged it is phonological structure which is of main concern.

3.4.4 Common effects of the factors of prosodic variation

Cho and Keating (2009) found only one parameter that was affected by position as well as both phrasal accent and lexical stress. In their analysis, the EPG seal duration of the initial consonant in two /CVCVCVC/ words was jointly (but not cumulatively) affected by the three prosodic factors in terms of lengthening. It was one aim of the present study to see if and how this finding is projected from simple to complex syllable onsets. Unlike in Cho and Keating (2009), the material here did not follow a fully crossed three-by-three design, i.e. especially effects of boundary strength and accent were not simultaneously observed. Therefore, this section is split into the discussions of position and stress on the one hand and accent and stress on the other.

3.4.4.1 Position and stress

Fougeron (2001) found that effects of boundary strength were local to the segment immediately adjacent to the boundary. For the most part, this finding is supported by the present data as the strongest impact of prosodic position was observed on PLATC1. This is also in line with the findings presented in the EPG study in Chapter 2. The case of stress presents the mirror image in a sense. Here, too, the effect was local in being strongest on the segment closest to the stressed nucleus. This is somewhat in contrast with the findings of the previous chapter as no significant effect of stress on PLATC2 was found there. There is no immediately obvious reason for this difference. A starting point might be that articulatory trajectories are immediately available from EMA but not from EPG data. The temporal resolution of the EMA data can therefore be regarded as more exact than the derived indeces used in the EPG study. The results do, however, corroborate the findings of Cho and Keating (2009) and Bombien (2007) where stress-induced lengthening of prenuclear consonants was observed. The π-gesture approach predicts that effects of boundary strength should be graded, i.e. waxing towards the boundary position and waning afterwards. This is confirmed by Cho and Keating (2009) in that not only parameters of the consonant but also of the vowel underwent some sort of strengthening or lengthening that could be attributed to a stronger preceding boundary. The present data does not include information about vowel production but at least for the cluster /kl/ there is evidence for the graded nature of boundary effects as not only platC1 but also platC2 were lengthened in phrase-initial position. This has interesting implications for the post-boundary temporal scope of the π-gesture as discussed by Byrd (2006). In post-boundary /$C_1VC_2VC_3V$/ sequences, Byrd find prosodic lengthening only for the first consonant (while C_2 and C_3 undergo compensatory shortening). The strongest effect is found for C_1 closing movements followed by a less consistent effect on the opening movement of C_2. The authors conclude that boundary effects are strongest at the juncture and wane afterwards. It is not immediately clear from their study whether the π-gesture's influence spreads in terms of articulatory units (as in closing or open movements) or in terms of time (as in ms). The current data support a spreading in terms of time rather than articulatory units: Boundary effects that stretch beyond C_1 are only found in /kl/ clusters which also happen to have the shortest overall duration and overlaps to a very high degree. In the other clusters, the C_2 plateau outreaches the scope of the π gesture and no lengthening effect is detectable.

Were the π-gesture approach applicable to lexical stress as well, as to some extent assumed in this study, then effects of stress should be graded in the same way as effects of position. It is well established that the nucleus of a stressed syllable bears the strongest effect of stress since the nucleus is its primary domain. In finding such a consistent effect in platC2, the assumption of gradedness then finds considerable support, since it is definitely not only the nucleus which is affected. But there is some evidence that the domain of stress effects is even wider spread since phrase-initially there is some lengthening of /platC1/ in stressed clusters. The exception here is /kn/ which also happend to be the longest cluster. This gives further evidence to the gradedness of the π-gesture. In this light there is less reason to assume a discontinuity of the spreading of stress effects mentioned in the end of Chapter 2. As in the previous chapter, there is a shortening effect of stress on the plateau overlap of /kC/ clusters. But since there is some evidence that stress effects can also affect platC1 a rather continuous pattern emerges now. Also given that in the clusters /kl ks pl ps/ stress lengthened C_1 plateaus phrase-initially one might argue that position and stress do in fact affect articulation along the same scale.

3.4.4.2 Accent and stress

As mentioned above, the effects of accent are conflicting and unclear. A common effect such as lengthening of seal duration as in Cho and Keating (2009) cannot be found. Some reservations have to be made with respect to the elicitation of the accentuation contrast. In the present speech material, target words containing broad focus were compared to target words with accent deliberately moved to the preceding word by deaccentuation. For one thing, the syntactical structure of the deaccentuation test sentences was rather complex and the speakers may not always have succeeded in producing them as intended. Further, in the speech material used by Cho and Keating (2009), narrow focus was either on the target word or removed as far as possible (utterances consisted of 5 syllables). Here, prenuclear accented items were compared to deaccented items where a very narrow focus preceded the target word immediately but no target word ever bore narrow focus. It is not unprobable that this difference yields different results.

3.4.4.3 Summary

A common object of positional effects and effects of prominence as presented by Cho and Keating (2009) in terms of lengthening of seal duration is not found in the same way here. Accentual effects proved to be uninterpretable in the current speech material, and even in the well-designed material of Cho and Keating (2009) the effect of accent on EPG seal duration of the initial consonant is rather weak compared to the effects of position and stress. However, if one were to project the findings of this study on words with complex onsets (#C_1C_2V) to words with simple onsets (#CV), positional effects on C_1 would be mapped to the simple onset as well as stress-induced effects on C_2. Furthermore, the strongest lengthening effect on C_1 was observed under the *combined* effects of initial position and stress.

3.4.5 The measurement of stop plateaus

Figure 3.9 displays gestural segmentation of a /pl/ cluster constructed analogously to the depiction of /kl/ in Figure 3.2 on page 64. As mentioned in Section 3.2.5, lip aperture was used as the measure for gestural segmentation of bilabial articulation. Since it is

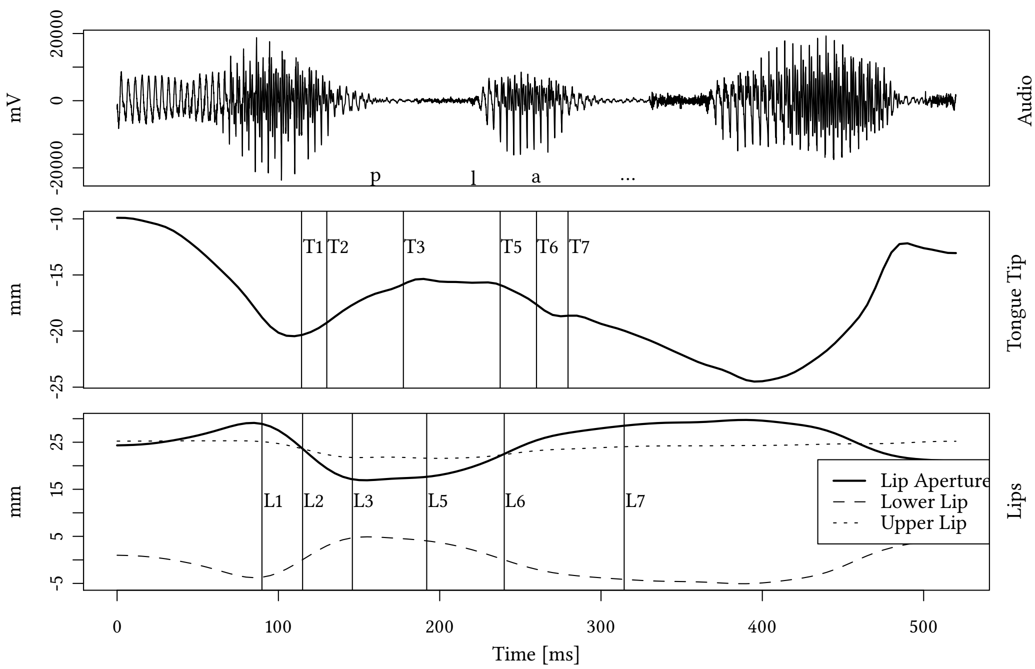

Figure 3.9: Extraction of temporal parameters in a case of the word ”Plakat”. Vertical lines correspond to articulatory landmarks labeled according to articulator (T = tongue tip, L = lip aperture) and landmark number as defined in Figure 3.1 (max. constriction (4) omitted for clarity). The lower panel additionally displays the vertical positions of the sensors attached to the upper and lower lip (shifted 30 mm up for better display).

not a directly recorded signal but the Euclidean distance between the sensors placed on the lower lip and on the upper lip, the vertical components of the latter two are included in the figure for illustrative purposes. Comparing Figure 3.2 to Figure 3.9 it is interesting to see that /k/ burst in the waveform occurs in the middle of constriction plateau while /p/ burst is well aligned with the plateau offset. A natural explanation lies in the looping trajectory of /k/ in 3.10, a movement pattern well addressed by Mooshammer, Hoole, and Kühnert (1995). The early burst in /k/ may well be due to forward movement of the tongue back, pobably away from the palate, during the phase labeled constriction plateau, i.e. the closure is already released even though the tongue dorsum is still in a high position. Movement in the anterior-posterior dimension is not captured by the analysis procedure for this sensor which is based on vertical movement only. But even if forward movement were registerd by the procedure (e.g. by use of the tangential velocity in the landmarking algorithm), it would not necessarily improve the measure

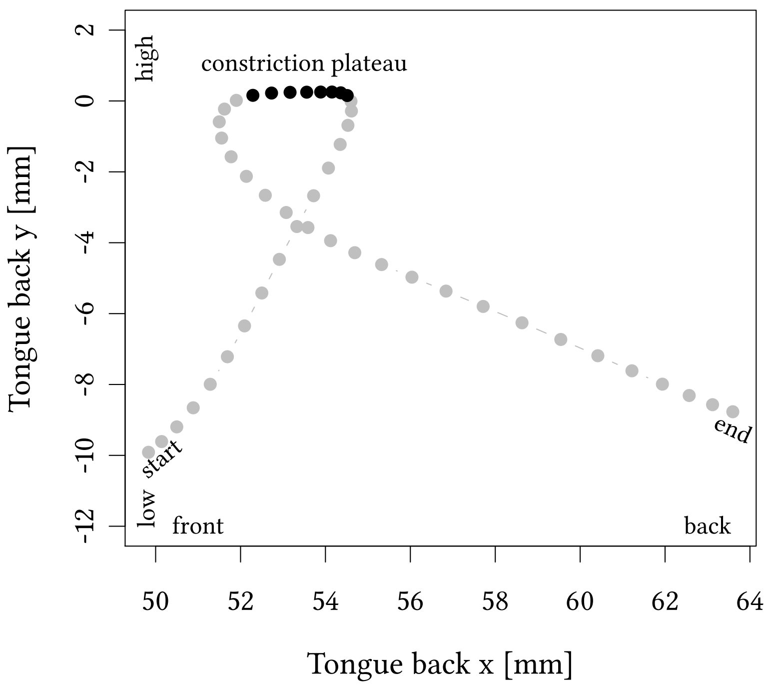

Figure 3.10: xy trajectory of the tongue dorsum sensor during /k/ in /kl/ (same utterance as in Figure 3.2. One point every 5 ms (EMA sampling frequency 200 Hz), low density meaning fast movement, high density meaning slow movement. Constriction plateau as defined in Figure 3.1 marked by black points.

since the quasi stationary phase between two velocity troughs might shorten unduly. More ideal would be a measure based on the least distance from the tongue back sensor to the palate as obtained e.g. by palate traces.

There is one important implication: Should constriction plateau offset in /k/ occur earlier then overlap in /kC/ clusters would decrease, too, so that /kC/ clusters could in fact turn out to be less overlapped than /pC/ clusters. This would support the initial hypothesis, of more overlap in /pC/ clusters due to articulator independence. The conclusions drawn with regard to the exceptionally high overlap in /kl/ clusters might then not be justified. However, /kl/ remains the shortest of all clusters presented in this study and positional effects stretching across C_1 and overlap to influence even C_2

are still encountered only in /kl/.

3.5 Conclusion

The main conclusions of this study can be drawn from the way prosodic variation affects intra-gestural properties on the one hand and intergestural properties on the other hand. Intergestural properties such as plateau overlap and target latency analyzed under prosodic variation allow for the conclusion that the gestural coordination in initial consonant clusters is stipulated nonuniformly by the segmental make-up. It is obvious that neither an account solely based on manner of articulation (sonority hierarchies) nor one largely disregarding manner (DAC) can account for the patterns observed here. Segmental make-up also appears to determine the amount of coordination variation due to prosodic variation. /kC/ clusters are less variable in this regard than /pC/ clusters. It is argued that higher variability in /pC/ clusters is due to higher interarticulatory independence

Concerning intragestural properties, C_1 plateaus are consistently lengthened phrase-initially while C_2 plateaus are longer in stressed syllables than in unstressed syllables. The impact of both position and stress is therefore strongest on the immediately adjacent consonant. The results provide sound evidence for the graded nature of both prosodic factors provided the property in question (C_1 or C_2 plateau) be in the range of the prosodic effect. This is especially the case for /kl/ whose summed plateaus and overlap yield the shortest total duration observed among all clusters investigated in this study. Its tight coordination may account for the fact that in /kl/ both consonant plateaus are affected by both position and stress. It would be interesting to see whether these results can be replicated for other similarly tightly coordinated clusters. This study has put a strong focus on timing in initial consonant clusters and therefore only durational measures were presented. An analysis of spatial effects induced by prosodic and segmental variation would be a welcome addition. Such a study has the potential of shedding light on the differences that are due to manner of articulation to a much larger extent than a purely durational study can.

3.A Statistic tables

3.A.1 kn kl

Table 3.10: Segmental make-up: C_1-duration, C_2-duration, plateau overlap and target latency in /kl/ and /kn/ clusters are analyzed as a function of manner of C_2 (lateral axproximant or nasal).

platC1	β	SE(β)	t-value	p	sig. diff.
(Intercept)	28.03	1.71	16.39		
c2m	−5.74	2.79	−2.06	*	/kn/ > /kl/
platC2	β	SE(β)	t-value	p	sig. diff.
(Intercept)	46.29	6.03	7.68		
c2m	−7.05	4.45	−1.59		
povern	β	SE(β)	t-value	p	sig. diff.
(Intercept)	−0.21	0.02	−9.53		
c2m	0.15	0.04	3.32	**	/kl/ > /kn/
pover	β	SE(β)	t-value	p	sig. diff.
(Intercept)	−21.47	2.55	−8.41		
c2m	21.01	3.94	5.34	***	/kl/ > /kn/
tlatn	β	SE(β)	t-value	p	sig. diff.
(Intercept)	0.68	0.08	8.81		
c2m	−0.21	0.07	−3.02	**	/kn/ > /kl/
tlat	β	SE(β)	t-value	p	sig. diff.
(Intercept)	49.57	2.60	19.06		
c2m	−26.62	5.20	−5.11	***	/kn/ > /kl/

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

Table 3.11: Position and stress: C_1-duration, C_2-duration, plateau overlap and target latency in /kl/ and /kn/ clusters are analyzed as a function of manner of prosodic position and lexical stress in addition to manner of C_2 (lateral axproximant or nasal).

platC1	β	SE(β)	t-value	p	sig. diff.
(Intercept)	40.39	3.23	12.51		
pos	16.78	1.81	9.28	***	Pi >Pm
strss	−0.97	1.76	−0.55		
c2m	2.15	1.76	1.23		
pos:strss	−1.40	3.51	−0.40		
pos:c2m	6.67	3.51	1.90		
strss:c2m	1.08	3.51	0.31		
pos:strss:c2m	6.11	7.03	0.87		

platC2	β	SE(β)	t-value	p	sig. diff.
(Intercept)	43.13	3.39	12.71		
pos	1.86	2.56	0.73		
strss	7.35	2.49	2.95	**	S >U
c2m	1.99	2.49	0.80		
pos:strss	−3.54	4.99	−0.71		
pos:c2m	13.52	4.99	2.71	*	/kl/: Pi >Pm
strss:c2m	6.53	4.99	1.31		
pos:strss:c2m	12.93	9.99	1.30		

pover	β	SE(β)	t-value	p	sig. diff.
(Intercept)	−25.26	7.22	−3.50		
pos	−1.21	2.54	−0.48		
strss	−8.49	2.46	−3.44	***	U > S
c2m	31.35	2.46	12.72	***	/kl/ > /kn/
pos:strss	−3.27	4.93	−0.66		
pos:c2m	7.13	4.93	1.45		
strss:c2m	−1.38	4.93	−0.28		
pos:strss:c2m	0.35	9.87	0.04		

tlat	β	SE(β)	t-value	p	sig. diff.
(Intercept)	65.65	6.49	10.12		
pos	17.86	3.27	5.47	***	Pi > Pm
strss	7.51	3.17	2.37	*	S > U
c2m	−29.19	3.17	−9.20	***	/kn/ >/kl/
pos:strss	1.89	6.35	0.30		
pos:c2m	−0.52	6.35	−0.08		
strss:c2m	2.45	6.34	0.39		
pos:strss:c2m	5.56	12.71	0.44		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

Table 3.12: Accent and stress: C_1-duration, C_2-duration, plateau overlap and target latency in /kl/ and /kn/ clusters are analyzed as a function of accentuation (accented or deaccented) and lexical stress (stressed or unstresssed) in addition to manner of C_2 (lateral axproximant or nasal).

platC1	β	SE(β)	t-value	p	sig. diff.
(Intercept)	27.93	1.49	18.74		
acc	1.57	1.14	1.37		
strss	−0.92	1.14	−0.80		
c2m	−2.96	1.14	−2.59	*	/kl/ < /kn/
acc:strss	−1.69	2.28	−0.74		
acc:c2m	−6.40	2.29	−2.80	**	/kn/: D < W
strss:c2m	2.87	2.29	1.26		
acc:strss:c2m	−4.80	4.57	−1.05		

platC2	β	SE(β)	t-value	p	sig. diff.
(Intercept)	42.09	3.77	11.16		
acc	0.22	2.56	0.09		
strss	8.81	2.55	3.45	***	S > U
c2m	−2.18	2.55	−0.85		
acc:strss	−1.55	5.11	−0.30		
acc:c2m	−5.94	5.12	−1.16		
strss:c2m	5.90	5.11	1.15		
acc:strss:c2m	−20.08	10.23	−1.96		

pover	β	SE(β)	t-value	p	sig. diff.
(Intercept)	−16.39	5.25	−3.12		
acc	−5.18	2.09	−2.49	*	W < D
strss	−1.21	2.08	−0.58		
c2m	24.86	2.08	11.93	***	/kl/ > /kn/
acc:strss	−5.41	4.17	−1.30		
acc:c2m	3.43	4.17	0.82		
strss:c2m	−2.85	4.17	−0.68		
acc:strss:c2m	−16.63	8.35	−1.99		

tlat	β	SE(β)	t-value	p	sig. diff.
(Intercept)	44.31	4.67	9.49		
acc	6.79	2.41	2.81	**	W > D
strss	0.30	2.41	0.13		
c2m	−27.81	2.41	−11.52	***	/kn/ > /kl/
acc:strss	3.68	4.83	0.76		
acc:c2m	−9.86	4.83	−2.04		
strss:c2m	5.65	4.83	1.17		
acc:strss:c2m	11.95	9.67	1.24		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

3.A.2 Stop+alveolar Clusters

Table 3.13: Segmental make-up: C_1-duration, C_2-duration, plateau overlap and target latency in /kl/, /ks/, /pl/ and /ps/ clusters are analyzed as a function of place of C_1 (velar or labial) and manner of C_2 (lateral axproximant or nasal).

platC1	β	SE(β)	t-value	p	sig. diff.
(Intercept)	30.01	3.22	9.31		
c1p	−8.39	2.34	−3.58	***	/p/ > /k/
c2m	−7.54	2.34	−3.22	**	/s/ > /l/
c1p:c2m	11.12	4.68	2.37	*	/p/: /s/ > /l/

platC2	β	SE(β)	t-value	p	sig. diff.
(Intercept)	46.11	8.09	5.70		
c1p	−6.93	2.98	−2.32	*	/p/ >/k/
c2m	7.36	2.98	2.47	*	/l/ > /s/
c1p:c2m	−15.67	5.97	−2.63	*	/p/: /l/ > /s/

povern	β	SE(β)	t-value	p	sig. diff.
(Intercept)	−0.18	0.07	−2.73		
c1p	0.06	0.05	0.14		
c2m	0.10	0.05	2.22	*	/l/ >/s/
c1p:c2m	0.06	0.09	0.68		

pover	β	SE(β)	t-value	p	sig. diff.
(Intercept)	−17.61	6.58	−2.67		
c1p	−3.06	4.51	−0.68		
c2m	12.47	4.51	2.77	**	/l/ > /s/
c1p:c2m	4.81	9.01	0.53		

tlatn	β	SE(β)	t-value	p	sig. diff.
(Intercept)	0.83	0.10	8.02		
c1p	−0.25	0.07	−3.74	***	/p/ > /k/
c2m	−0.26	0.07	−3.95	***	/s/ > /l/
c1p:c2m	0.05	0.13	0.41		

tlat	β	SE(β)	t-value	p	sig. diff.
(Intercept)	47.58	4.04	11.78		
c1p	−5.32	3.43	−1.55		
c2m	−20.01	3.43	−5.83	***	/s/ > /l/
c1p:c2m	6.29	6.86	0.92		

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05

Table 3.14: Prosodic variation (boundary and stress): C_1-duration, C_2-duration, plateau overlap and target latency in /kl/, /ks/, /pl/ and /ps/ clusters are analyzed as a function of place of C_1 (velar or labial), manner of C_2 (lateral axproximant or nasal), prosodic position (phrase initial or phrase medial) and lexical stress (stressed or unstressed).

platC1	β	SE(β)	t-value	p	sig. diff.
(Intercept)	43.49	4.30	10.12		
c1p	−1.80	1.79	−1.01		
c2m	−5.27	1.80	−2.93	**	/s/ > /l/
pos	18.27	1.78	10.29	***	Pi > Pm
strss	2.92	1.78	1.64		
c1p:c2m	4.03	4.07	0.99	**	/p/: /s/ > /l/
c1p:pos	13.61	3.58	3.80	***	Pm: /p/ > /k/
c2m:pos	0.58	3.60	0.16		
c1p:strss	−9.34	4.07	−2.30		
c2m:strss	5.19	4.07	1.28		
pos:strss	5.79	3.57	1.62	*	Pi: S > U
c1p:c2m:pos	−12.37	8.14	−1.52		
c1p:pos:strss	−13.18	8.14	−1.62		
c2m:pos:strss	9.06	8.14	1.11		

platC2	β	SE(β)	t-value	p	sig. diff.
(Intercept)	48.73	6.29	7.75		
c1p	−4.10	2.13	−1.93	*	/p/ > /k/
c2m	−0.50	2.14	−0.23		
pos	2.99	2.11	1.42		
strss	11.09	2.12	5.23	***	S > U
c1p:c2m	−6.35	4.83	−1.31		
c1p:pos	6.05	4.26	1.42	*	Pi: /p/ > /k/
c2m:pos	4.93	4.28	1.15		
c1p:strss	−8.22	4.84	−1.70		
c2m:strss	10.81	4.84	2.23	*	U: /s/ > /l/
pos:strss	1.14	4.24	0.27		
c1p:c2m:pos	18.16	9.67	1.88		
c1p:pos:strss	7.24	9.68	0.75		
c2m:pos:strss	−4.40	9.68	−0.45		

pover	β	SE(β)	t-value	p	sig. diff.
(Intercept)	−16.42	5.33	−3.08		
c1p	4.69	2.37	1.98		
c2m	5.71	2.38	2.40	***	/l/ >/s/
pos	−2.90	2.35	−1.23		
strss	−5.38	2.36	−2.28		
c1p:c2m	−2.94	5.38	−0.55		
c1p:pos	17.39	4.74	3.67	***	/p/: Pi < Pm

continued on next page

Table 3.14 *continued from previous page*

C2M:POS	7.07	4.76	1.49		
C1P:STRSS	−18.31	5.38	−3.40	*	/k/: S < U
C2M:STRSS	12.37	5.38	2.30	*	/s/ S < U
POS:STRSS	−5.33	4.72	−1.13		
C1P:C2M:POS	−9.12	10.76	−0.85		
C1P:POS:STRSS	8.38	10.77	0.78		
C2M:POS:STRSS	−9.26	10.77	−0.86		

TLAT	β	SE(β)	t-value	p	sig. diff.
(INTERCEPT)	59.90	3.26	18.37		
C1P	−6.49	2.64	−2.46	*	/p/ > /k/
C2M	−10.99	2.65	−4.14	***	/s/ > /l/
POS	21.33	2.61	8.16	***	Pi > Pm
STRSS	8.30	2.63	3.16	**	S > U
C1P:C2M	6.95	6.00	1.16		
C1P:POS	−3.79	5.28	−0.72		
C2M:POS	−6.47	5.31	−1.22		
C1P:STRSS	8.91	6.00	1.48		
C2M:STRSS	−7.16	6.00	−1.19		
POS:STRSS	11.16	5.26	2.12	**	Pi: S > U
C1P:C2M:POS	−3.07	12.00	−0.26		
C1P:POS:STRSS	−21.42	12.01	−1.78		
C2M:POS:STRSS	18.17	12.01	1.51		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

Table 3.15: Prosodic variation (accent and stress): C_1-duration, C_2-duration, plateau overlap and target latency in /kl/, /ks/, /pl/ and /ps/ clusters are analyzed as a function of place of C_1 (velar or labial), manner of C_2 (lateral axproximant or nasal), accent accented or deaccented) and lexical stress (stressed or unstressed).

PLATC1	β	SE(β)	t-value	p	sig. diff.
(INTERCEPT)	31.63	3.76	8.40		
C1P	−9.93	1.39	−7.13	***	/p/ > /k/
C2M	−5.63	1.40	−4.02	***	/s/ > /l/
ACC	−2.19	1.36	−1.61		
STRSS	2.03	1.38	1.47		
C1P:C2M	10.55	3.18	3.32	**	p: /s/ > /l/
C1P:ACC	−0.17	2.78	−0.06		
C2M:ACC	0.06	2.80	0.02		
C1P:STRSS	1.69	3.15	0.54		
C2M:STRSS	−5.31	3.18	−1.67		
ACC:STRSS	−8.59	2.77	−3.10	***	A: U > S; D: S>U
C1P:C2M:ACC	0.98	6.36	0.15		

continued on next page

Table 3.15 *continued from previous page*

c1p:acc:strss	9.35	6.29	1.49		
c2m:acc:strss	−1.08	6.36	−0.17		

platC2	β	SE(β)	t-value	p	sig. diff.
(Intercept)	42.48	4.65	9.14		
c1p	−3.09	2.20	−1.40		
c2m	3.97	2.21	1.79		
acc	2.65	2.15	1.24		
strss	8.98	2.19	4.11	***	S > U
c1p:c2m	−4.78	5.02	−0.95		
c1p:acc	3.12	4.40	0.71		
c2m:acc	−7.29	4.43	−1.65		
c1p:strss	2.11	4.97	0.42		
c2m:strss	4.77	5.02	0.95		
acc:strss	−0.96	4.38	−0.22		
c1p:c2m:acc	−21.62	10.05	−2.15		
c1p:acc:strss	−37.50	9.94	−3.77	**	/k/: A: S>U; /p/: D: S>U
c2m:acc:strss	26.51	10.05	2.64	*	/l/: A: S>U

pover	β	SE(β)	t-value	p	sig. diff.
(Intercept)	−13.07	4.00	−3.27		
c1p	−0.31	2.70	−0.11		
c2m	9.07	2.72	3.34	***	/l/ > /s/
acc	−0.67	2.64	−0.25		
strss	0.53	2.69	0.20		
c1p:c2m	18.66	6.17	3.02	***	/k/: /l/ > /s/
c1p:acc	8.76	5.41	1.62		
c2m:acc	−5.99	5.44	−1.10		
c1p:strss	−5.05	6.11	−0.83		
c2m:strss	0.78	6.17	0.13		
acc:strss	−13.18	5.38	−2.45		
c1p:c2m:acc	−28.02	12.35	−2.27		
c1p:acc:strss	−21.35	12.21	−1.75		
c2m:acc:strss	26.65	12.34	2.16	*	/s/: A: S > U

tlat	β	SE(β)	t-value	p	sig. diff.
(Intercept)	44.74	1.19	37.55		
c1p	−9.59	2.44	−3.93	***	/p/ > /k/
c2m	−14.73	2.46	−6.00	***	/s/ > /l/
acc	−1.50	2.38	−0.63		
strss	1.44	2.43	0.59		
c1p:c2m	−8.23	5.58	−1.48	**	/l/: /p/ > /k/
c1p:acc	−8.96	4.88	−1.83		
c2m:acc	6.10	4.91	1.24		
c1p:strss	6.65	5.52	1.20		
c2m:strss	−6.10	5.57	−1.09		

continued on next page

Table 3.15 *continued from previous page*

ACC:STRSS	4.72	4.86	0.97		
C1P:C2M:ACC	29.46	11.15	2.64		
C1P:ACC:STRSS	30.59	11.03	2.77		
C2M:ACC:STRSS	−27.93	11.15	−2.50	*	sumpm

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

Das nächste Spiel ist immer das nächste.

Matthias Sammer

Chapter 4

Some implications of C_1-voicing for the timing of word-initial French and German consonant clusters

4.1 Introduction

Hoole, Bombien, Kühnert, and Mooshammer (2009) found that voicing—in a phonological sense—conditions gestural overlap in word-initial consonant clusters in German. More specifically, in sequences such as /gl/ and /bl/ where C_1 underlyingly is voiced, overlap of the two consonants' constriction plateaus is higher than in /kl/ and /pl/ where C_1 is voiceless. The motivation for this study emerges from the comparison to French word-initial consonant clusters of the same segmental make-up: Gestural overlap in French /gl/ vs. /kl/ and /bl/ vs. /pl/ pairs does not seem to differ. It stands to reason to link this cross-linguistic difference to another better known fact: French and German are considered to differ in the means of implementing voicing contrasts in initial stop consonants. French accomplishes the contrast by the use of (true) voicing whereas German employs aspiration: phonologically voiced stops are usually not voiced while phonologically voiceless stops are in fact voiceless but also post-aspirated. Traditionally, aspiration does not play a role in French. Voice-onset time (VOT), the duration from stop release to onset of phonation, has commonly been employed to characterize this difference; see details below. This study aims at using synchronous articulatory (EMA) and acoustic data to make inferences on the coordination of laryngeal and supra-laryngeal articulations.

4.1.1 Voice onset time and the voicing contrast in French and German word-initial stops

Lisker and Abramson (1964) state that characterizing voicing contrasts in stops can most fruitfully be accomplished using Voice Onset Time (VOT). Other measures or acoustic properties on their own fail to account for the various different mechanisms that the world's languages employ to create the voicing contrast. In a condensed view of English, true physiological voicing, i.e. "'the presence of a glottal buzz"' (Lisker & Abramson, 1964, p. 384) or its absence, reliably separates word-medial and final /b d g/ from /p t k/ but it fails word-initially since there both groups are generally produced without vocal fold vibration. Aspiration, on the other hand, distinguishes /p t k/ from /b d g/ in word-initial and medial position but is less successful word-finally since there is oftentimes no aspiration in /p t k/ and not even an audible release in /b d g/. Lisker and Abramson (1964) therefore conclude for English that neither voicing nor aspiration alone can account for the phonological voicing contrast. VOT, then, is the temporal distance from the release of the stop in question to the onset of voicing. This distance can be i) positive (long and short lags), for example in voiceless aspirated stops where voicing starts after the release of the stop, ii) negative, for example in voiced/prevoiced stops, where voice onset is prior to the release of the stop and iii) zero in voiceless unaspirated stops. According to Maddieson (2009, among others) languages with bimodal voicing contrasts typically pattern into having either a prevoiced–short lag opposition as for example French (also Caramazza & Yeni-Komshian, 1974) and Spanish or a short lag –long lag opposition as for example English and German.

Few sources are available for gestural coordination in (voiceless) stop sounds for French and German. In their fiberscopic analysis, Benguerel, Hirose, Sawashima, and Ushijima (1978) state that in French the glottal devoicing gesture is timed with the occlusion such that glottal aperture starts at the same time as the oral occlusion and ends at or short after the release. In a very recent work, Hoole (2006) using fiberscopic transillumination analyzes laryngeal-oral coordination in a large set of consonants both as singletons and in clusters in word initial position. The set includes both /p/ as well as /pl/ (velars are usually considered unsuitable for fiberscopic recordings since movements of the tongue root can interfere with the fiberscope and perturb the resulting transillumination signals). In his data, glottal abduction starts after the onset of oral occlusion with peak glottal opening (PGO) in close vicinity to the release of the oral occlusion. This means that a considerable portion of the glottal aperture remains after the occlusion which is quite contrary to the patterning described for French data.

The study of Klatt (1975) is of immediate importance since unlike most other studies it also deals with VOT in English consonant clusters. Among others, it covers the clusters under analysis in this study: /kl gl/ and /pl bl/ as well as the corresponding singleton stops. The most general finding is, of course, that voiceless stops have considerably longer VOT than voiced stops thus confirming the typical Germanic dichotomy of short lag vs. long lag. Furthermore, three observations are relevant to the study at hand. The first is the universal (Maddieson, 1997) fact that place of articulation has an effect on VOT with longer values for labial than for lingual stops. This has been demonstrated in quite a number of studies for a range of languages before Klatt and afterwards, e.g. Lisker and Abramson (1964, 1967), Weismer (1979), Crystal and House (1988), Docherty (1992), Nearey and Rochet (1994), Cho and Ladefoged (1999), Hoole (2006). Some of these works, but not all, also find that velars have longer VOT than apicals. Secondly, VOT has cross-linguistically been found to vary as a function of the following vowel's height (Fischer-Jørgensen, 1972, e.g.). After high vowels, VOT is generally longer than after low vowels. Both phenomena have been attributed to the fact that VOT lengthens as a function of constriction degree of the vocal tract after stop release. In other words, a slow depletion of supraglottal air pressure due to a slow opening of the vocal tract impedes the onset of voicing. In the case of velar stops, the inertia of the tongue causes the vocal tract's constriction to decrease much slower after the release than for labials. Similarly, high vowels, e.g. /i/, must in themselves be regarded as constrictions unlike low vowels for which the tongue and jaw lowering cause the vocal tract to assume its most unconstricted state.

4.1.2 Models of glottal timing

On the phonetic surface, the phonological concept of voicing in German stops is a matter of aspiration rather than of voicing (/t/ = [t^h], /d/ = [t]). An alternative but widely accepted view by Kohler (1995) uses the terms *fortis* and *lenis* for this distinction, since apart from aspiration closure duration and burst intensity are essential as well (/t/ = [t^h], /d/ = [d̥]). In phonology, a number of features have been proposed for the distinction of such pairs: Keating (1984) proposed ±voice as purely abstract features to account for different implementations of VOT differences in stop pairs in different languages. This is a departure from the notion that phonological features are physically based.

Articulatory Phonology (AP) suggests that the voicing contrast be modeled by the absence or presence of glottal opening-and-closing gestures (Browman & Goldstein, 1986; Goldstein & Browman, 1986) to provide a closer mapping between phonological and physical categories. AP assumes that different voicing contrasts in stops (e.g. English vs. French) are implemented by language specific phasing relations of the oral gesture with the glottal gesture. A recent review on how glottal gestures are incorpoarted into AP is given in Best and Hallé (2010). This work also outlines some of the difficulties for phonology that arise from the different means languages apply to implement the voicing contrast and generally confirms that oral-laryngeal coordination has not sufficiently been dealt with in phonological theory.

4.1.3 Research questions

The following summarizes the research questions (RQ) addressed in this study. The first questions basically aim at establishing well known patterns for French and German as they have previously been reported. More interesting, however, is whether these patterns also pertain in clusters. Based on the literature the following results are expected for VOT and the occlusion duration:

1. German stop voicing contrast is realized by a short-lag/long-lag opposition in VOT whereas in French the opposition is one of prevoiced/short-lag.
   (a) This also applies to clusters.
2. /k/ has shorter occlusion than /p/
   (a) This also applies to clusters.

3. /p/ has shorter VOT than /k/
   (a) This also applies to clusters.
4. Stop occlusion duration is shorter in stop+/l/ clusters than in simple onsets.
5. VOT is longer in stop+/l/ clusters than in simple onsets.

Based on the previous findings,

6. overlap is expected to be sensitive to stop voicing in German. Larger overlap is expected with voiced stops than with voiceless. Overlap in French clusters is insensitive to voicing. There might be less overlap in labial+/l/ than in velar+/l/ clusters according to some corresponding evidence in Chapter 3.

Two possibilities for cluster timing in French emerge:

7. Overlap in French stop+/l/ clusters is more like overlap in *voiced* clusters in German
8. Overlap in French stop+/l/ clusters is more like overlap in *voiceless* clusters in German

Considering the case of German, a reasonable assumption would be that plateau overlap is less in the context of a voiceless stop in order to temporally accommodate the glottal gesture or more precisely the aspiration which is due to glottal timing. A perceptual motivation behind this might be the more gradual sonority modulation (Ohala, 1992; Ohala & Kawasaki-Fukumori, 1997): As it is, i.e. large lag between the stop and /l/, a sonority profile of the following order is likely to emerge: voiceless stop – aspiration – voiceless lateral (fricative?) – voiced lateral – vowel. This sequence could be expressed as a series of uniform rises in sonority. Increasing overlap might lead to a fully devoiced lateral and, importantly, a voiceless transition from the lateral into the vowel which presents a rather stark rise in sonority compared to the previous modulations.In this case one could assume that the greater lag in the voiceless case is the result of a rightward shift of the lateral. This argumentation, of course, bears on the notion that the lateral is considered to be underlyingly voiced. In the case of voiceless stop+fricative clusters, there should be no requirement of a voiced C_2-vowel transition.[1]

Since in French the glottal gesture must be timed differently in order to account for the fact that less or no aspiration occurs, there is no need to shift the lateral to the right

[1]Considering that voiceless stop+fricative clusters are rather exceptional in German (apart from affricates and (mostly Greek) loanwords) one might argue that the lack of continuity in the sonority modulation is the reason for the dispreference of such.

since it is at no risk of undergoing total devoicing. It would therefore seem appropriate to assume that French clusters are timed more like voiced German clusters, i.e. with more overlap than voiceless German clusters (RQ 7).

4.2 Method

4.2.1 Speakers and speech material

Five speakers each of French and German were recorded by means of EMA. The test corpora (French and German) are part of a larger project and were designed to contain all possible word onsets of the respective language. For each word onset, two words were selected one with a low back vowel the other with a high front vowel following the onset (e.g. 'Bad' [ba:t] (bath) and 'Biest' [bi:st] (biest).[2] This study uses only subsets of these corpora containing the simple onsets /b/, /g/, /p/ and /k/ as well as the same consonants forming complex onsets with a following /l/. These subsets are presented in appendix A.3. The choice of material was based on previous findings that /kl/ clusters exhibit the most stable coordination patterns of the clusters analyzed and because it has a fully-voiced counterpart /gl/. /pl/ and /bl/ were chosen because they present the only other pair of clusters with this voicing contrast in German that does not involve velic activity. The target words were embedded in carrier sentences which had three slots for the target words:

German Ich sage wieder *«word#1»* oder *«word#2»* oder *«word#3»*.
I say again «word#1» *or* «word#2» *or* «word#3».

French Je vois *«word#1»* ou *«word#2»* ou *«word#3»*.
I see «word#1» *or* «word#2» *or* «word#3».

The randomization routine ensured that this study's target words were distributed equally between the first and the second position (not the third). The relevant EMA sensors were placed on the upper and lower lip, below the lower incisors, on the tongue tip, tongue mid and tongue back.

[2]Write a note about the change in corpus design.

4.2.2 Extraction of temporal parameters

Articulatory landmarks were labeled as described in Section 3.2.4. Normalization of plateau overlap, however, was carried out by dividing the absolute data by the constriction plateau duration of C_2 rather then by the interval from C_1 constriction plateau onset to C_2 constriction plateau offset. The rationale behind this decision is tied to the expectation that C_2 constriction duration is the least variable possible parameter. While C_2 is always the voiced alveolar lateral /l/ C_1 varies in terms of both place of articulation and voicing. Both factors are likely to affect C_1 constriction duration. a) A large body of work has found closure duration in bilabial stops to be universally longer than in velar stops of the same voicing type (e.g. Byrd, 1993; Maddieson, 1997). b) It has been shown that stop durations can vary as a function of phonological voicing (Fuchs, 2005). In order to validate whether C_2 is indeed more stable and therefore in this context a more reliable normalization operand, the extent of sensitivity of C_2 constriction plateau duration to variation of C_1 place and voicing will be reported in the beginning of the results section.

The acoustical measures, C_1 occlusion duration and voice onset time, were defined as follows: Occlusion duration starts at the beginning of the occlusion as determined from waveform and spectrogram. It ends at the stop's release. VOT is here defined to start at occlusion offset. It ends at the onset of periodicity following the stop burst. This interval maybe zero but not negative. Phonologically voiced stops in German vs. French differ in that French stops are fully voiced whereas German stops usually are not. A measure of voice lead/voicing during closure/negative VOT would be appropriate to capture this difference. However, a glottal abduction-adduction gesture is involved in neither French nor German voiceless stops and consequently no inferences can be made about the coordination of laryngeal and oral articulation's. A measure of voice lead was therefore not included in this study.

4.2.3 Statistics

Although the material allows for the analysis of effects on VOT due to vowel quality as described above this aspect is not pursued here. Instead the data are pooled in this regard. Should any statistical bluring result from this, it should only enhance the power of significances found in other regards. R (R Development Core Team, 2009) was used

Table 4.1: Predictors and dependent variables used in this chapter's statistics.

Predictor	Description
LANG	Language: French or German (FR/DE)
PLAC	Place of articulation: (bi-)labial or velar (L/V)
VOX	Voicing: voiced or voiceless (phonologically) (+V/-V)
COMP	Complexity: complex or simple onsets (C/S)
SPK	Speakers
Variable	**Description**
VOT	Voice Onset Time in ms
OCC	Occlusion of the stop in ms (as measured in the acoustics)
C1P	C_1 plateau duration in ms
C2P	C_2 plateau duration in ms
POVER	Plateau overlap in ms
TLAT	Target latency in ms

Table 4.2: Percentage of presence of VOT in voiced French simple stop onsets.

	Speakers				
	ff01	ff02	fm01	fm02	fm03
VOT PRESENT					
Labial	0%	20%	0%	10.5%	0%
Velar	5%	80%	0%	0%	60%

to fit linear mixed effect models to the data as described in Section 3.2.6. Table 4.1

4.3 Results

In this section, the language specific patterns of plateau overlap in stop+/l/ clusters and of VOT in simple and complex onsets are established.

4.3.1 Voice onset time and occlusion in singletons

VOT In the German data, VOT was present after all stop bursts regardless of the stops' voicing. As anticipated based on the literature, the situation is different in the French

Measure	Predictor Place (Vel/Lab)	Voicing (+V/-V)	Language (DE/FR)
VOT	Vel > Lab ***	-V > +V ***	DE > FR ***
Occlusion	Lab > Vel *** DE (plac × lang) ***	*n.s.*	FR > DE **
C_1 plateau	Vel > Lab ***	-V >+V **	*n.s.*

Table 4.3: Summary of main effects on simplex onsets. Interactions are only presented when they contribute crucially to the understanding of the data.

data. Only the voiceless stops are consistently followed by an interval of voicelessness. There is some variability for the voiced stops but the general assumption is supported here that voicing is present when the stop is released ($VOT = 0ms$). Furthermore, voicing was present throughout the occlusion phase of the voiced French stops except for the cases were VOT was positive, too. A contingency table of the presence of VOT after voiced stops in French is given in Table 4.2. For two speakers (*ff02* and *fm03*) voiced velar stops show a tendency for being produced with following VOT. All other occurrences are rather exceptional.

A mixed model was fitted to VOT duration with SPK as a random factor. The detailed output of the model is given in Table 4.5.

Generally, the stops produced by German speakers have 37±3 ms longer VOT than the stops produced by French speakers ($F[1,60] = 95.5, p < 0.001$). The place of articulation further determines the amount of VOT ($F[1,60] = 327.9, p < 0.001$) in that labial stops have 19±1 ms shorter VOT than velar stops. The strongest effect is–quite naturally–that of voicing ($F[1,60] = 1620.6, p < 0.001$). Voiceless stops have significantly longer VOT (43±1 ms) than voiced stops. Unraveling the interaction of language and voicing ($F[1,60] = 216.1, p < 0.001$) reveals that the voicing effect is much stronger in German (58±2 ms; $F[1,60] = 1492.4, p < 0.001$) than in French (27±2 ms; $F[1,60] = 321.2, p < 0.001$). Furthermore, the language difference (more VOT in German) is much more pronounced in voiceless stops (52±6 ms; $F[1,60] = 72.4, p < 0.001$) than in voiced stops (20±2 ms; $F[1,60] = 99.7, p < 0.001$), although there seems to be less variability in the latter.[3] Another prominent interac-

[3] A more consistent language effect in voiced stops would probably emerge for a measure like voicing during closure/voice lead because /b/ in French is fully voiced while it is voiceless in German. However, the results do not contribute to the understanding of laryngeal-oral coordination, see also Section 4.2.2

tion is that of voice and place ($F[1,60] = 51.3, p < 0.001$). In velars, voicing accounts for 51±2 ms VOT difference ($F[1,60] = 873.8, p < 0.001$) but only for 35±1 ms ($F[1,60] = 880.4, p < 0.001$) in labials which is still considerable but also considerably less. As for the effect of place (more VOT in velars than in labials), the difference is larger in voiceless (26±2 ms; $F[1,60] = 216.0, p < 0.001$) than in voiced stops (10±1 ms; $F[1,60] = 146.8, p < 0.001$). Finally, there is a weak interaction of language and place ($F[1,60] = 4.4, p < 0.05$) which untangled adumbrates a slightly stronger language effect on velars than on labials and a slightly stronger place effect in German than in French. In both cases, a variation of only about 4–5 ms is explained. By and large, the interactions point towards cumulative effects of the factors. Stops with the least VOT have the following properties: voiced, bilabial, French. Changing any of these properties (within the range analyzed here) will add to the amount of VOT additively such that the stops with the largest VOT are voiceless, velar and German. Based on these results it can be established that in terms of VOT the present data adhere to the commonly found patterns regarding the language specific voicing contrasts in singleton stop onsets. An overview of VOT durations as a function of language, voicing and place is given in the upper panel of Figure 4.1 along with the durations for acoustical occlusion to which the focus turns now.

Occlusion As above for VOT, a mixed model was fitted to the occlusion duration which is presented in detail in Table 4.5. Labials have longer occlusion durations than velars (16±1 ms; $F[1,60] = 198.3, p < 0.001$) which is not surprising since it is in line with corresponding universal findings. Voicing by itself does not influence the occlusion duration. Language, on the other hand, accounts for slightly longer occlusion durations in French than in German (28±10 ms; $F[1,60] = 7.4, p < 0.01$). However, the effect of place interacts with both language and voicing. The place×language interaction ($F[1,60] = 72.1, p < 0.001$) indicates that the place induced difference is much greater in German (26±1 ms; $F[1,60] = 308.9, p < 0.001$) than in French (6±2 ms; $F[1,60] = 10.3, p < 0.01$). Furthermore, language related differences (longer occlusions in French) are only significant in velars (37±9 ms; $F[1,60] = 16.2, p < 0.001$) but not in labials (17±11 ms; $F[1,60] = 2.2, p > 0.05$). The interaction of place and voice ($F[1,60] = 17.1, p < 0.001$) is due to a greater place-related difference in voiceless stops (21±2 ms; $F[1,60] = 148.5, p < 0.001$) than in voiced stops (10±2

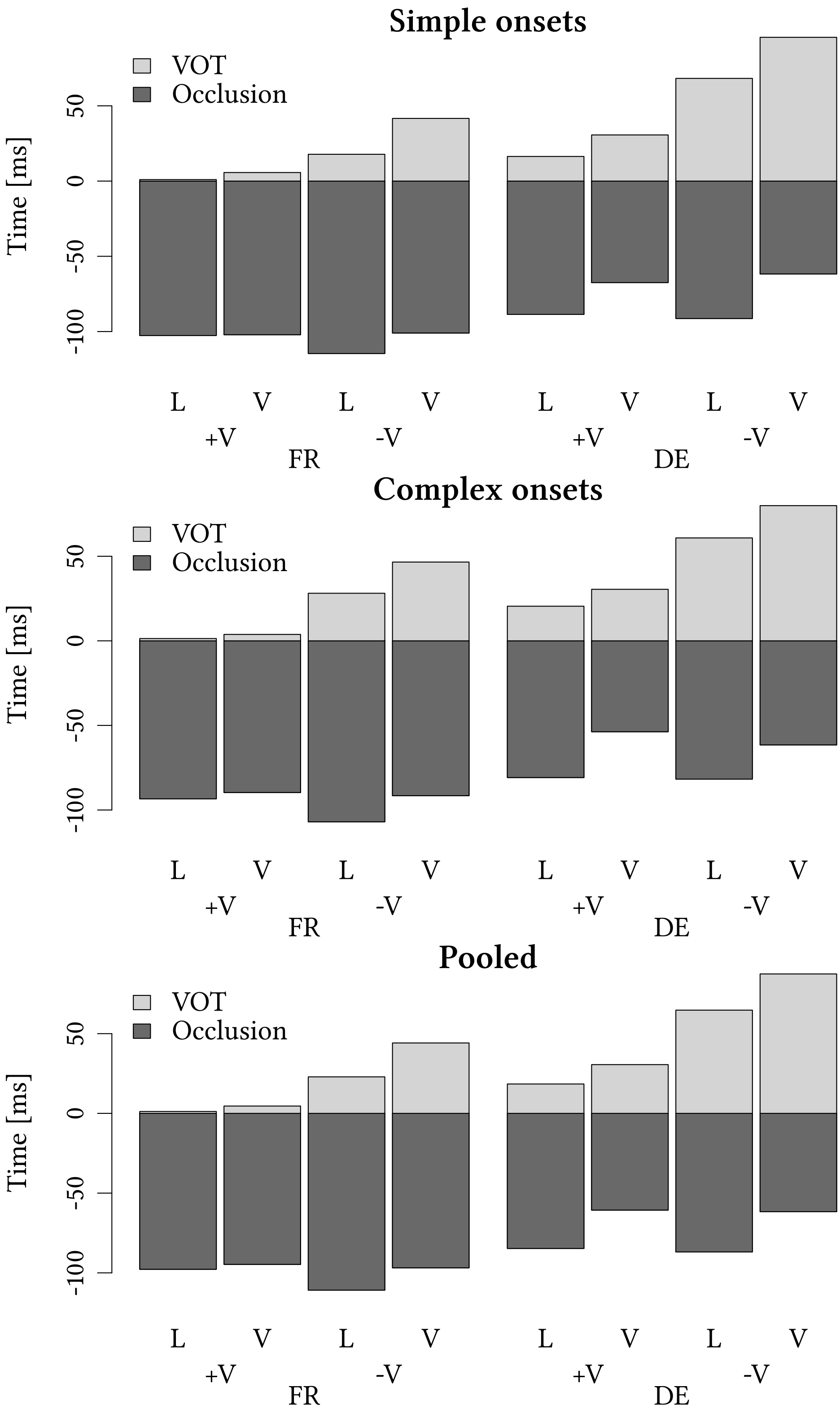

Figure 4.1: Mean durations of acoustical occlusion and VOT in simple (upper panel) and complex (middle panel) onsets as a function of language (French (FR) vs German (DE)), voicing (+V vs. -V) and place (labial (L) vs. velar (V)) of articulation. Lower panel displays the pooled data. 0 alignment at stop release.

ms; $F[1,60] = 62.1, p < 0.001$). Additionally, there is a voicing effect in labials (8±1 ms; $F[1,60] = 30.5, p < 0.001$) but not in velars. The interaction of voicing and language ($F[1,60] = 8.9, p < 0.01$), finally, can be broken down to a tiny voicing effect in French (5±2 ms longer in voiceless stops; $F[1,60] = 8.5, p < 0.01$) but not in German stops. In sum, as evident from Figure 4.1, German stops have slightly shorter occlusion durations than French stops. Furthermore, German stops are subject to a systematic place-induced variation whereas French stops remain rather stable.

C_1 plateau duration As a measure expected to behave in parallel to the occlusion duration, the stops' constriction plateaus will now be at the center of attention. A mixed model was fitted to the stops' plateau durations (also referred to as C_1 plateau especially in the context of clusters) in analogy to the inspection of VOT and occlusion duration. The details are presented in Table 4.5. There are in total two simple main effects to report. Place of articulation effects the plateau duration such that velars have significantly larger durations than labials (14±2 ms; $F[1,60] = 61.8, p < 0.001$). This is quite contrary to the findings of occlusion duration above where the reverse pattern was found. The reason probably lies in the technique of measuring the constriction plateau of velars which only involves vertical movement. It is more than conceivable that tongue-palate contact is released by a fronting movement of the tongue which is not captured by the measure applied here and which occurs before the lowering of the tongue dorsum (see Section 3.4.5). The second effect is that of voicing. Voiced stops have shorter durations than voiceless stops (5±2 ms; $F[1,60] = 7.9, p < 0.01$). No language-specific differences were encountered.

Summary This paragraph summarizes the results presented in this section with regard to the research questions posed in Section 4.1.3. The voicing contrast is reflected differently by means of VOT in German and French as expected (RQ 1): German has a short-lag—long-lag opposition while French has a voiced—long-lag opposition. Another expectation (RQ 2) is met by /k/ having shorter occlusion than /p/. Finally, RQ 3 is confirmed since /p/ indeed has shorter VOT than /k/. The importance of these findings is that the present data agree with the occlusion and VOT patterns commonly found for German and French simple onsets.

The results for C_1 plateau duration were expected to closely match those of occlusion duration which they do not. Regarding the effect of place of articulation, this is attributed here to the measurement technique applied for EMA data of velar constrictions. It is surprising, however, that the plateau duration is sensitive to voicing–in a manner predictable from the literature–but occlusion duration is not.

An unexpected result is that language has an effect on the occlusion duration as well and not only on VOT. It is shorter in German than in French. In the light of the reverse effect on VOT (more VOT in German than in French) this indicates some kind of a trade-off effect.

4.3.2 Voice onset time and occlusion in clusters

	Predictor			
Measure	Place (Vel/Lab)	Voicing (+V/-V)	Language (DE/FR)	Complexity (C/S)
VOT	Vel > Lab ***	-V > +V ***	DE > FR ***	*n.s.*
Occlusion	Lab > Vel *** DE > FR	-V > +V *** FR, Lab, C	FR > DE ** Vel	S > C ***
C_1 plateau	Lab < Vel ***	-V > +V *** FR, Lab		S >C ***

Table 4.4: Summary of effects on both complex and simplex onsets. Interactions are only presented when they contribute crucially to the understanding of the data.

After establishing standard patterns in the simple onsets above in Section 4.3.1 the data set for the analysis now broadens to include not only simple but also complex onsets. This allows for a comparative analysis of different onset complexities in the same statistical model by adding the factor complexity.

VOT Following the above order, VOT will be considered first. Table 4.6 lists the details of the statistical model. First of all, there is no main effect of complexity, but all remaining factors significantly affect VOT. Velars have generally 16±1 ms longer VOT than labials ($F[1, 60] = 470.1, p < 0.001$), voiced stops have 42±1 ms shorter VOT than voiceless stops ($F[1, 60] = 3208.8, p < 0.001$) and in French stops VOT is 33±4 ms shorter than in German stops ($F[1, 60] = 72.2, p < 0.001$). All of this is in line with the results presented above in Section 4.3.1 as are the two way interactions not involving complexity so that in general the same picture emerges: Bilabial voiced stops in French have the least VOT which increases if any of these properties changes. (plac:vox: $F[1,60] = 83.2, p < 0.001$; plac:lang: $F[1,60] = 9.0, p < 0.01$; vox:lang: $F[1,60] = 199.2, p < 0.001$). The three-way interaction of place, voicing and language ($F[1,60] = 5.3, p < 0.05$) indicates that in voiced onsets the effect of language (more VOT in German) is stronger in velars than in labials and the effect of place (more VOT in velars) is stronger in German than in French.

Several interactions also involve complexity which is of major interest in this section. Complexity interacts with place ($F[1,60] = 14.5, p < 0.001$) such that the effect of place is weaker in complex onsets (13±2 ms; $F[1,60] = 48.8, p < 0.001$) than in simple onsets (20±2 ms; $F[1,60] = 90.0, p < 0.001$). Additionally, in velars VOT is 6±2 ms shorter when they are part of a cluster than when they form simple onsets ($F[1,60] = 7.0, p < 0.05$). This is not the case for labial stops. Complexity also interacts with language ($F[1,60] = 31.9, p < 0.001$). While the language effect obtains to the same extent in both complex and simple onsets, there is a tendency for 4±2 ms less VOT in complex onsets but only in German ($F[1,60] = 3.5, p = 0.068$). The interaction of place, voicing and complexity ($F[1,60] = 1.3, p > 0.05$) indicates that the place effect is stronger in voiced simple than in voiced complex onsets. Reversely, the place effect is stronger in voiceless simple than in voiceless complex onsets. A complexity effect (longer VOT in complex onsets) is found in voiceless velars and very small complexity effect in voiced labials.

According to the interaction of voicing, language and complexity ($F[1,60] = 55.0, p < 0.001$) there is a complexity effect for voiceless stops in French but the reverse in German (VOT shorter in simple onsets). The language effect is stronger in simple voiceless cases than voiced but reversely it is stronger in complex voiceless cases than voiced. Complexity definitely adds to the complexity of the data but overall there is little support for the predictions made from the literature, that complex onsets have a longer VOT than simple onsets.

Occlusion Turning to occlusion again–now for both simplex and complex onsets–a mixed model was fitted to occlusion as a function of place, voicing, language and complexity. Place of articulation accounts for an average of 16±1 ms shorter occlusions in velars than in labials ($F[1,60] = 437.9, p < 0.001$). Occlusion duration is 27±9 ms longer in French than in German onsets ($F[1,60] = 8.4, p < 0.01$). The interaction of place and language ($F[1,60] = 110.7, p < 0.001$) indicate that German onsets show a stronger place effect (25±1 ms; place: $F[1,60] = 534.4, p < 0.001$) than French onsets (8±1 ms; $F[1,60] = 34.3, p < 0.001$) while on the other hand the effect of language is significant only in velars (35±8 ms; $F[1,60] = 17.1, p < 0.001$) but not in labials (18±10 ms; $F[1,60] = 2.7, p > 0.05$). This is a slight departure from the results presented in 4.3.2 where occlusion duration in French was less susceptible to place variation. Another departure is that here there is actually an effect of voicing resulting in 4±1 ms longer occlusions in voiceless stops than in voiced ($F[1,60] = 31.1, p < 0.001$). In spite of its low variability the effect must be considered rather low. In fact, as the interaction of voice and place ($F[1,60] = 12.9, p < 0.001$) suggests, voicing does not have an effect at all on velars but only on labials (8±1 ms; $F[1,60] = 52.9, p < 0.001$). The interaction of voice and language ($F[1,60] = 9.6, p < 0.01$) furthermore shows that voicing has no effect on occlusion duration in German but only in French (7±1 ms; $F[1,60] = 28.6, p < 0.001$). This leads to the three-way interaction of place, voicing and language ($F[1,60] = 13.4, p < 0.001$) which is due to the fact, that voicing is only effective on occlusion duration in French labial stops (14±2 ms; $F[1,60] = 79.2, p < 0.001$). There is so far no substantial difference concerning occlusion duration between the full data set analyzed here and the set of simple onsets analyzed above.

The question how complexity influences occlusion duration is addressed now. Complexity has a main effect causing 9±1 ms longer occlusion in simple than in complex onsets ($F[1,60] = 121.4, p < 0.001$). The interaction of voicing and complexity ($F[1,60] = 6.9, p < 0.05$) shows that the voicing effect in French labial stops is further restricted to complex onsets (6±1 ms; $F[1,60] = 21.8, p < 0.001$) which explains why no voicing effect was found for simple onsets above. This is further corroborated by the three-way interaction of place, voicing and complexity ($F[1,60] = 7.8, p < 0.01$). The four-way interaction which would round up the picture fails to reach significance by an inch ($F[1,60] = 4.0, p = 0.05$).

C_1 plateau duration As in the analysis of C_1 plateau duration in simple onsets above, there are main effects of place ($F[1,60] = 88.0, p < 0.001$) and voicing ($F[1,60] = 15.6, p < 0.001$). The plateau in labials is by average 11±1 ms shorter than in velars and 4±1 ms shorter in voiced than in voiceless stops. Both results confirm the above findings but they are doubtlessly weaker. This might be connected to the additional main effect of complexity ($F[1,60] = 52.9, p < 0.001$) which shortens the plateau duration by about 8±1 ms. Indeed the place effect is stronger in the simple onsets (15±2 ms; $F[1,60] = 61.1, p < 0.001$) than in the complex onsets (7±1 ms; $F[1,60] = 28.8, p < 0.001$) as the interaction of place and complexity suggests ($F[1,60] = 12.3, p < 0.001$) but there are no other interactions involving complexity. Instead there is an interaction of voicing and language ($F[1,60] = 19.1, p < 0.001$). Language itself does not have a main effect, neither here nor in the analysis of simple onsets only above. However, the voicing effect is restricted to French (10±1 ms; $F[1,60] = 47.5, p < 0.001$) and not significant in German. The weak interaction of place and voice ($F[1,60] = 6.1, p < 0.05$) points towards a voicing effect in labials only (7±1 ms; $F[1,60] = 82.1, p < 0.001$) but not in velars. Finally, an interaction of place and language ($F[1,60] = 5.1, p < 0.05$) indicates that the place effect is slightly stronger in French than in German.

Summary This section summarizes the results presented in this section on both simple and complex onsets. The findings meet the expectations in that the VOT contrast with regard to voicing and language applies to simple onsets as well as to complex onsets (RQ 1a). Furthermore, the place of articulation effect for VOT (more VOT after velar than labial stops) and occlusion duration (longer for labials than for velars) is present in complex onsets as well (RQ 2a and 3a). It can therefore be established that the patterns summarized above for simple onsets also pertain in clusters.

Complexity itself also affects the parameters in question. Occlusion duration is indeed shorter in complex than in simple onsets (RQ 4). VOT, on the other hand, was expected to be longer in complex onsets (RQ 5). This can not be confirmed with the present data.

C_1 plateau duration yields similar results than for simple onsets only above, i.e. the reverse effect of place of articulation (due to measurement technique) and longer duration in voiceless stops. Complexity affects C_1 plateau duration in the same way it affects occlusion duration: longer simple than in complex onsets.

4.3.3 Plateau overlap and C_2 plateau duration

C_2 plateau As mentioned above, C_2 plateau duration is taken into consideration as a possible candidate for normalization of plateau overlap. The reason behind this is the idea that C_2 being the only segmental constant in the consonant clusters considered here might turn out to be insensitive to variation of C_1 place and voicing as well as the language. They do not. Voicing has a highly significant effect in that /l/ has 5±1 ms longer plateaus after voiceless stops than after voiced ($F[1, 60] = 14.5, p < 0.001$). Similarly, /l/ plateaus are 4±1 ms longer after velar than after labial stops ($F[1, 60] = 10.0, p < 0.01$). In spite of there consistency, both effects are obviously rather small. A source of much higher variation is language as seen in Table 4.7. While the difference between C_2 plateau durations in German and French is not significant, the languages differ in the strength of variation. The grand means across all speakers of the respective language group and the corresponding standard error are 41±3 ms for French and 57±8 ms for German. The employment of C_2 plateau duration for overlap normalization is therefore questionable.

Plateau overlap A mixed model was designed to calculate the effects of language, voicing and place of articulation on plateau overlap. Language by itself does not have a significant effect ($F\,[1, 60] = 2.4, p > 0.05$). There is a strong main effect of voicing ($F\,[1, 60] = 87.2, p < 0.001$) suggesting that there is generally 13±1 ms more overlap in clusters with voiced than with unvoiced stops. The interaction with language, however ($F\,[1, 60] = 37.6, p < 0.001$), calls for a closer inspection for each language. For the German data, there is indeed a very significant effect of voicing ($F\,[1, 60] = 100.7, p < 0.001$) which accounts for about 21±2 ms more overlap in voiced clusters. The corresponding effect for French is rather marginal ($F\,[1, 60] = 5.9, p < 0.05$), the overlap difference between voiced and unvoiced clusters being only 4±2 ms. Examining the data for language specific differences per voicing category shows that clusters with voiceless stops overlap to a similar extent. In clusters with voiced stops, on the other hand, there is a marginally significant difference ($F\,[1, 60] = 6.4, p < 0.05$) pointing towards 22±9 ms more overlap in German than in French. Overall, the effect of voicing on overlap is present in German but not in French which confirms the expectations. The interactions of voice and language further provides good evidence, that, if anything, overlap in French clusters is more similar to German voiceless clusters but different from German voiced clusters. Then there is a main effect of place of articulation ($F[1,60] = 34.4, p < 0.001$) causing 8±1 ms longer overlap in velar+/l/ clusters. This is a striking result because the effect is much stronger than the corresponding finding in Section 3.3.2.1. The interaction of place and voice ($F[1,60] = 7.4, p < 0.01$) might offer an explanation for these different findings. While it obtains across both voicing conditions, in clusters with voiced stops the effect is much stronger (12±2 ms: $F[1,60] = 33.2, p < 0.001$) than in clusters with voiceless stops (4±2 ms: $F[1,60] = 5.6, p < 0.05$). In Chapter 3, however, only voiceless C_1 were considered i.e. where the place effect is rather weak.

4.3.4 The voiceless phase

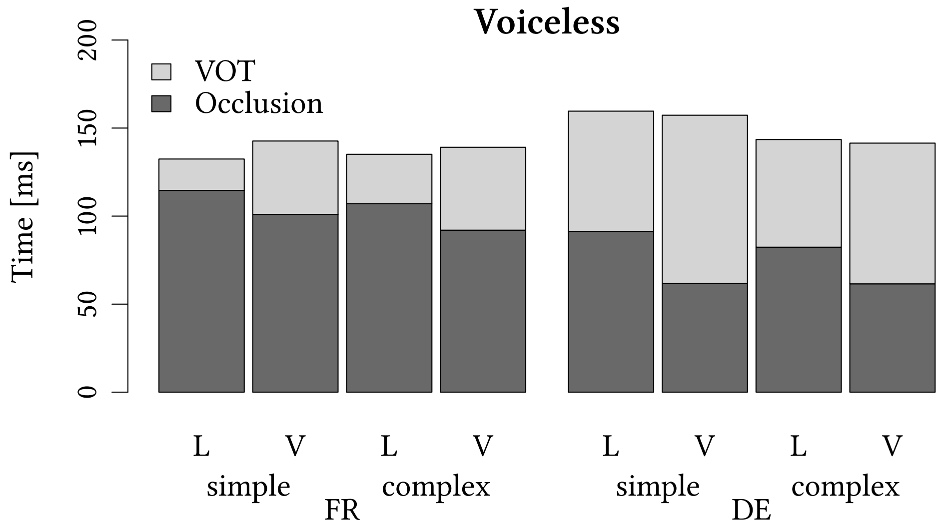

Figure 4.2: Mean durations of acoustical occlusion and VOT in simple and complex onsets as a function of language (French (FR) vs German (DE)), complexity (complex vs. simple) and place (labial (L) vs. velar (V)) of articulation. Alignment at occlusion onset.

Considering the above results, a combination of effects attracts attention especially. It seems in Figure 4.1 that the combined durations of occlusion and VOT in the case of voiceless stops is rather stable. Figure 4.2 is a condensed version of Figure 4.1 with all voiced tokens removed and an opposition of complex vs. simple onsets. Importantly, the onset of the occlusion is used as line-up point in order to better illustrate the relative stability of the voiceless phase. Mainly the timing of the stop's burst within this interval varies as a function of place of articulation and language. For place, of course, this is not a new observation (Weismer, 1980; Cho & Ladefoged, 1999) and it has been argued that underlyingly the glottal devoicing gesture is the same in all cases. Based on findings that show longer VOT in stop+/l/ clusters as compared to simple stop onsets (Hoole, 2006) discusses several possibilities. The most "radical" possibility proposes lengthening of the glottal gesture due to the addition of the sonorant. To test this here, a mixed model is fitted to a subset of the data including only voiceless stops with place, language and complexity as predictors and the sum of occlusion and VOT duration as the dependent variable. Place of articulation affects the total voiceless duration such that it is 3±1 ms longer in velar contexts than in labial contexts ($F\,[1,60] = 5.7, p < 0.05$), a weak effect that barely scrapes significance. Language on its own does not affect the duration of the voiceless phase but complexity does ($F\,[1,60] = 51.9, p < 0.001$): Complex onsets have on average 10±1 ms shorter voiceless durations than simple onsets.

The interactions bring language into play. Place and language interact ($F\,[1,60] = 17.1, p < 0.001$) such that the place effect described above is only significant in the French speakers ($F\,[1,60] = 23.4, p < 0.001$) where the voiceless phase is 9±2 ms longer in velar than in labial context. The interaction of language and complexity ($F\,[1,60] = 25.7, p < 0.001$) is due to the fact that the complexity effect above is only significant in the German data where complex onsets have 16±2 ms shorter voiceless phases than simple onsets ($F\,[1,60] = 71.2, p < 0.001$). Detailed statistics are presented in Table 4.8.

4.4 Summary and discussion

Most literature-based expectations concerning VOT and occlusion duration were met with only one exception (see below). The voicing contrast for each language was realized as usual (RQ 1): short-lag/long-lag opposition in German vs. a voiced/short-lag opposition in French. RQ 1a asked whether this patterning also obtained in clusters which the data confirm. It is worth noting that the very short VOT lag after voiced French stops results from the labeling convention applied here: Even in the voiced cases, a VOT interval was labeled should voicing cease during the stop's release. Additionally, one speaker (*ff02*) regularly produced both /g/ as well as /gl/ with a short aspiration phase.

Furthermore, occlusion durations are indeed less in velars than in labials both in single (RQ 2) as well as in complex onsets (RQ 2a). Reversely, VOT is longer in velars than in labials as posited in RQ 3 for singletons and RQ 3a for clusters.

The effect of complexity on occlusion duration follows RQ 4 in that stops in clusters have shorter occlusions than singleton stops.

The exception to the literature-based expectations is the extent of the influence complexity has on VOT. Based on previous works it was assumed in RQ 5 that VOT should lengthen when a sonorant is added as compared to singleton stops. This assumption finds no support in the data, neither in French nor in German.

It was assumed in RQ 6 that plateau overlap should follow the pattern observed previously. Indeed, plateau overlap in German stop+/l/ clusters varies as a function of stop voicing (more overlap/shorter lag in voiced clusters) while it remains stable across both voicing conditions in French. Furthermore, the question was raised whether overlap in French (both voiced and voiceless) clusters should turn out to be rather like in voiced (RQ 7) or voiceless (RQ 8) clusters. There is clear evidence in support of the German voiceless pattern, i.e. there is always a considerable lag in French clusters, c.f. RQ 8. This is contrary to the argumentation presented in the introduction which was in favor of RQ 7, i.e. overlap in French clusters should pattern as in voiced clusters in German since there is no need to accommodate a glottal gesture/aspiration phase. This surprising result will be further discussed below. It is worth noting here, however, that there does not seem to be a difference of variability as a function of voicing, i.e. neither voiced nor voiceless clusters exhibit greater stability than the other.

Some more results need to be reviewed that were not explicitly covered by the research questions. Occlusion durations (and along with them the stops' plateau durations) tend to be longer in French than in German. While this was not directly predicted, it is well compatible with the results obtained for VOT and the total phase of voicelessness (in the case of voiceless stops). There are within-language differences between German and French concerning the total voiceless phase (place effect in French, complexity effect in German) but between each other, they do not differ substantially. Concerning occlusion duration and VOT on the other hand the languages differ con-

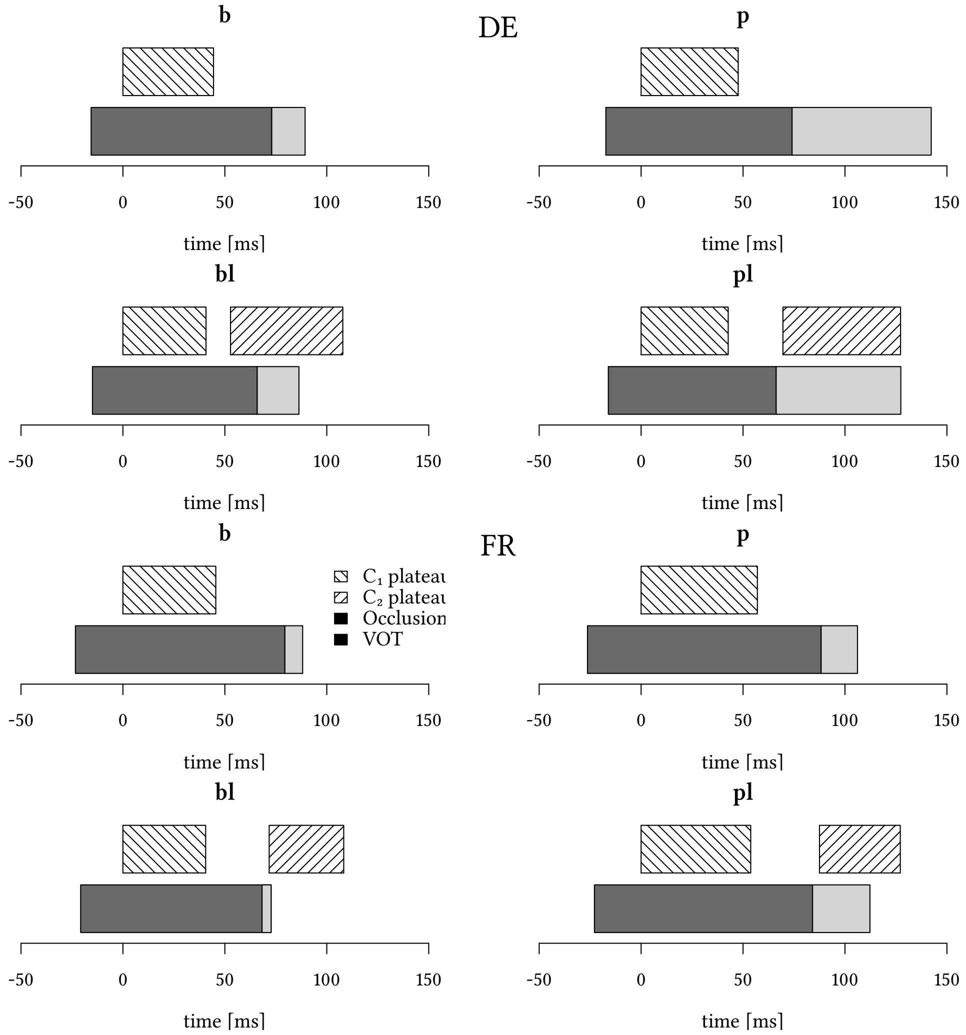

Figure 4.3: Occlusion and VOT aligned with articulatory plateaus of singleton bilabial stops and bilabial+/l/ clusters for German (upper four panels and French (lower four panels). Zero alignment at plateau onset of the stop.

siderable in such a way that higher VOT and lower occlusion duration in German vs. lower VOT and higher occlusion duration in French add up to more or less the same total voiceless duration. In essence this supports previous statements that the timing of the stop release relative to the voiceless phase is fundamentally different between German and French: In French, stop release occurs much later during the voiceless phase than in German. What is new here is that underlyingly French and German stops might have a quantitatively very similar glottal gesture. The results are strongly

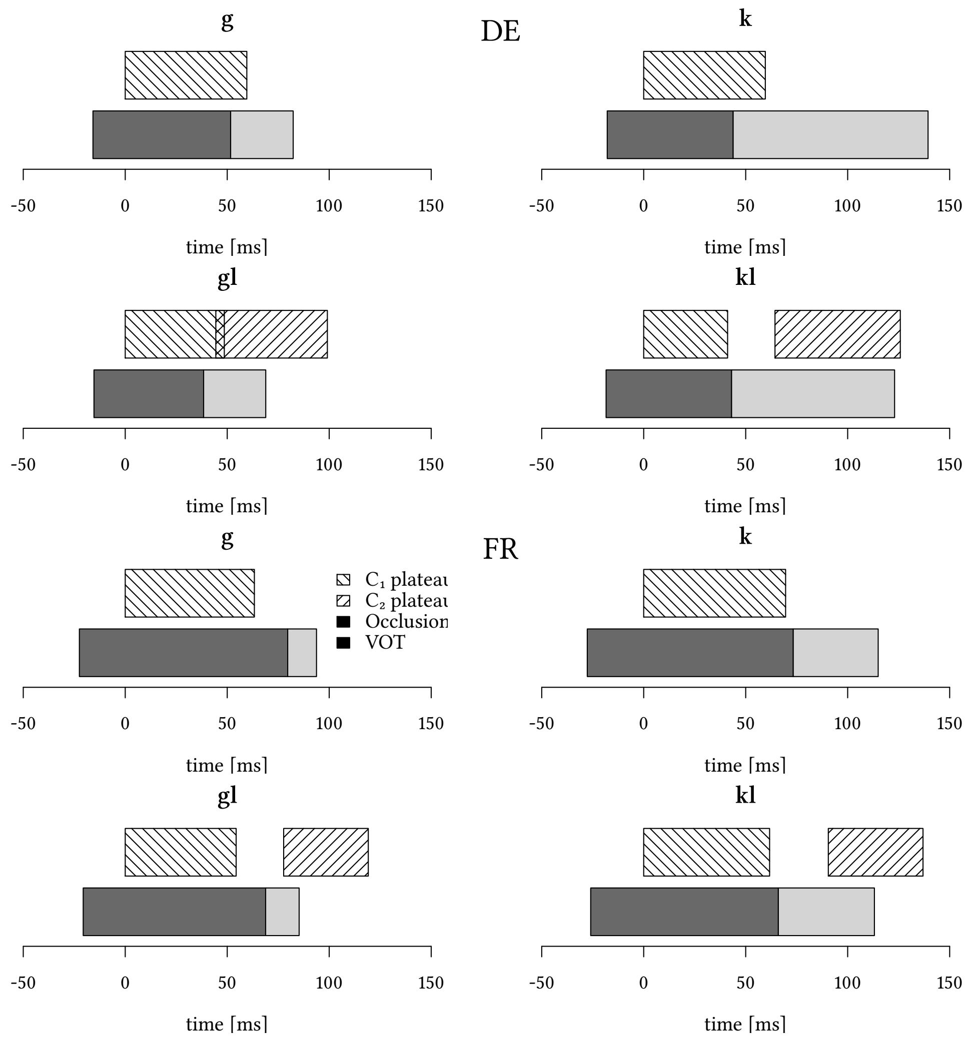

Figure 4.4: Occlusion and VOT aligned with articulatory plateaus of singleton velar stops and velar+/l/ clusters for German (upper four panels and French (lower four panels). Zero alignment at plateau onset of the stop.

reminiscent of place-related effects discussed by Hoole (2006) where stop burst occurs earlier in velars than in bilabials within the glottal gesture.

Finally, there is a tendency for /l/ plateaus to be shorter in French clusters than in German. The difference is not significant in the mixed model but the languages strongly differ with regard to the extent of variability in /l/ plateau production.

Figures 4.3 and 4.4 put the picture together. The figures show the alignment of acoustical (lower bars) and articulatory (upper bars) events separated by place, voicing complexity and language. The point of departure in this study is clearly visible in terms of overlap relations in the second row of panels for German, where overlap varies as a function of C_1 voicing, and in the fourth row for French, where overlap is little regardless of voicing. The patterning of acoustic occlusion and VOT in relation to the articulatory landmarks indicates that glottal timing in French clusters is plain different from the timing in German clusters. This is particularly obvious from the timing of the second consonant. In the introduction it was argued that C_2 may undergo rightward shift in order to accommodate the glottal gesture. This may or may not be true for German but is evidently not for French where Figures 4.3 and 4.4 and the statistics convey the impression that C_2 shifts rightward regardless of the voicing in C_1. In fact the amount of the interconsonantal plateau lag is as large in all French clusters as in the German voiceless clusters in spite of consistently less VOT in French than in German. Furthermore, it appears that, in French clusters, C_2 is not as much under the influence of the glottal gesture as it is in German. This brings up the question to what domain laryngeal properties belong: segments or syllable constituents. In the discussion of German data, Hoole (2006) cites Kehrein and Golston (2004) who conclude their analysis of laryngeal contrast in a large variety of languages with the statement that laryngeal features are properties of the syllable constituents rather than of segments. The German data presented agree with this concept but not the French data where C_2 seems removed from both the stop as well as the devoicing gesture.

As a final measure which might shed some light on these timing differences the distance between voice onset and C_2 plateau offset was computed as a percentage of C_2 plateau duration, i.e. the portion of the C_2 plateau that is not devoiced. The data for this measure are restricted to contain complex onsets only to account for the circumstance that simple onsets do not have a C_2. Values between 0% and 100% indicate the point of voice onset within the constriction plateau of /l/. Values above 100% arise when voicing sets in before C_2 target attainment, negative values when voicelessness outlasts C_2 plateau offset. Table 4.9 displays the statistics output of a mixed model fitted to this measure as a function of place, voicing and language. A corresponding illustration is given in Figure 4.5. Significant simple main effects emerge for all three predictors, no interactions are encountered. The voiced portion of the C_2 plateau is on average 26±6% longer after labial than after velar stops ($F[1, 60] = 18.1, p < 0.001$). This is in line with longer VOT after velars than after labials. 74±6% of variation are, quite naturally, due to

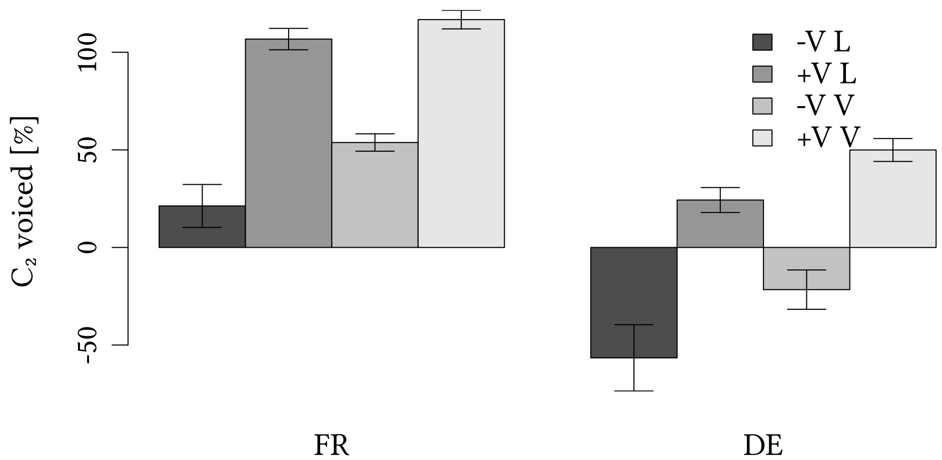

Figure 4.5: Voiced portion of the C_2 plateau in complex onsets as a function of language (FR vs. DE), place of articulation (L vs. V) and voicing (+V vs. -V).

voicing. Since voice onset is earlier for voiced stops, the voiced portion of C_2 plateaus is also larger after voiced stops ($F[1, 60] = 145.0, p < 0.001$). Most importantly here, however, is the effect language has on this measure. The C_2 plateau has 76±18% more voicing in French clusters than in German ($F[1, 60] = 18.4, p < 0.001$). Since in French VOT is comparably small and the lag between the consonantal plateaus generally high, this should come to no surprise. But the voiced portion is bigger in French than in German in spite of the tendency for C_2 duration being larger in German than in French. This result is a further indication that glottal timing in onset clusters considerably depends on language specific grammar: In German, the glottal gesture could be regarded a property of the entire onset (Hoole, 2006) with only marginal voicing at the right edge of the underlyingly voiced sonorant C_2. In the French clusters analyzed here, on the other hand, the glottal gesture appeared to be already receding before C_2 or in other words: C_2 hardly undergoes any devoicing. Interestingly, first results in an ongoing study indicate that this is not the case for /Cr/ clusters.

Another line of thought emerges from adhering to Kehrein and Golston's (2004) idea of associating the glottal gesture with the entire syllable onset rather than with one of its constituents. Should this idea be proved true then the two consonants in the French clusters are not part of a complex onset since otherwise the glottal gesture would span both consonants. Rather they should be parsed heterosyllabically (c.f. Shaw, Gafos, Hoole, & Zeroual, 2009) or C_1 could be regarded as extrasyllabic (cf. Rialland, 1994). Extrasyllabicity can be ruled out here since according to Rialland stop+liquid clusters form onsets (unlike e.g. stop+nasal where the stop would be considered extrasyllabic).

Heterosyllabicity, as presented for Moroccan Arabic (Gafos, Hoole, Roon, & Zeroual, 2010; Shaw, 2009; Gafos, 2002) is of greater interest here since the account is based on physiological data rather than on phonological rules. In short these works promote the idea that the coordination in word-initial clusters informs about syllable structure: Cluster that exhibit the C-center effect are considered complex syllable onsets. Clusters that do not exhibit this effect are parsed as a series of heterosyllabic onsets. The C-center effect has not been shown for French clusters (nor have there been, to the author's awareness, any published attempts to do that) and the present data do not allow for the required analysis. However, the timing of VOT and the cluster constrictions in French indicates that the clusters do not form complex onsets. Accordingly they should not exhibit the C-center effect. Interestingly, this would put the isochrony of French as a syllable timed language (Pike, 1945) at stake since additional onsets should add to the syllable duration. However, the issue of isochrony will not be discussed further following Liberman (2008) who warns against "the whole idea of *stress-timed* vs. *syllable-timed* languages, which is a gigantic tangled intellectual thicket that's easy to get into and hard to get out of" (my emphasis). It would be interesting to test for the C-center effect with appropriate data. This will be done in the near future since the date is available for some of the French speakers. Furthermore, an analysis of initial French clusters under prosodic variation in comparison to the results obtained for German in 23 might be rewarding. Additional evidence for a heterosyllabic parse in Frech clusters would be obtaind should timing in French clusters be more susceptible to prosodic variation than in German clusters.

There might even be a connection between syllable structure and the implementation of voicing contrasts in a language: French and Moroccan Arabic are similar here since both have fully voiced +V stops and disprefer overlap in mixed-voicing clusters (c.f. Zeroual & Hoole, 2010). More generally, one might say that true voicing and overlap are in some way incompatible: /bl/ in German exhibits high overlap but the stop is phonetically not voiced. /bl/ in French is fully voiced but there is very low overlap. In the case of French, the low overlap would assist the maintenance of voicing since intra-oral pressure can drop inbetween the two consonants. In the case of German this is simply not necessary.

Kehrein and Golston (2004, p. 26) do not rule out the possibility that within a syllable constituent a laryngeal feature has a stronger association to one segment than to another. In other words, in order to be property of a syllable onset a laryngeal feature does not necessarily have to spread equally across all segments involved in the onset. Furthermore, the inferences made concerning the timing of the glottal gesture are based on measurements of acoustical occlusion and VOT. Data obtained by laryngeal transillumination should be much better suited to shed light on the issues discussed here.

4.A Statistics tables

Table 4.5: Effects of place, voicing and language on acoustic and articulatory measures in simple onsets.

VOT	β	SE(β)	t-value	p	sig. diff.
(Intercept)	37.17	1.74	21.41		
lang	37.26	3.47	10.75	***	DE > FR
plac	18.91	1.09	17.35	***	Vel > Lab
vox	−43.39	1.08	−40.20	***	-V > +V
lang:plac	4.79	2.19	2.19	*	plac(DE) > plac(FR) lang(Vel) > lang(Lab)
lang:vox	−31.82	2.16	−14.71	***	vox(DE) > vox(FR) lang(-V) > lang(+V)
plac:vox	−15.53	2.16	−7.18	***	plac(-V) > plac(+V) vox(Vel) > vox(Lab)
lang:plac:vox	5.31	4.33	1.22		

Occlusion	β	SE(β)	t-value	p	sig. diff.
(Intercept)	90.44	5.12	17.65		
plac	−16.21	1.17	−13.81	***	Lab > Vel
vox	−1.77	1.16	−1.52		
lang	−28.33	10.23	−2.77	**	FR > DE
plac:vox	11.17	2.33	4.80	***	plac(-V) > plac(+V) Lab: -V > +V

continued on next page

Table 4.5					*continued from previous page*
PLAC:LANG	−19.84	2.35	−8.42	***	plac(DE) > plac(FR) Vel: FR > DE
VOX:LANG	6.92	2.33	2.97	**	FR: -V > +V
PLAC:VOX:LANG	−5.79	4.66	−1.24		

C_1 plateau	β	SE(β)	t-value	p	sig. diff.
(INTERCEPT)	55.79	7.10	7.86		
PLAC	14.42	1.87	7.73	***	Vel > Lab
VOX	−5.17	1.85	−2.80	**	-V > +V
LANG	−6.53	14.18	−0.46		
PLAC:VOX	3.89	3.70	1.05		
PLAC:LANG	−1.60	3.74	−0.43		
VOX:LANG	7.22	3.70	1.95		
PLAC:VOX:LANG	−1.99	7.42	−0.27		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

Table 4.6: Effects of place, voicing, language and complexity on acoustic and articulatory measures in simple and complex onsets.

VOT	β	SE(β)	t-value	p-value	sig. diff.
(Intercept)	35.79	1.90	18.80		
plac	15.59	0.74	21.09	***	Vel > Lab
vox	−41.72	0.74	−56.61	***	-V > +V
lang	32.54	3.80	8.55	***	DE> FR
comp	−1.05	0.74	−1.42		
plac:vox	−13.86	1.48	−9.40	***	plac(-V) > plac(+V) vox(Vel) > vox(Lab)
plac:lang	4.20	1.48	2.84	**	plac(DE) > plac(FR) lang(Vel) > lang(Lab)
vox:lang	−20.62	1.48	−13.98	***	vox(DE) > vox(FR) lang(-V) > lang(+V)
plac:comp	−5.54	1.48	−3.75	***	plac(C) < plac(S) Vel: C < S
vox:comp	3.10	1.47	2.10		
lang:comp	−8.37	1.48	−5.67	***	DE: S > C
plac:vox:lang	6.64	2.95	2.25	*	lang(+V, Vel) > lang(+V, Lab)
plac:vox:comp	3.61	2.95	1.22		
plac:lang:comp	−2.16	2.96	−0.73		
vox:lang:comp	21.90	2.95	7.42	***	plac(+V, S) > plac(+V, C)
plac:vox:lang:comp	2.08	5.90	0.35		

Occlusion	β	SE(β)	t-value	p-value	sig. diff.
(Intercept)	86.02	4.51	19.07		
plac	−16.41	0.80	−20.56	***	Lab > Vel
vox	−3.98	0.80	−5.00	***	-V > +V
lang	−26.98	9.01	−2.99	**	FR > DE
comp	−8.81	0.80	−11.06	***	S > C
plac:vox	6.46	1.59	4.06	***	Lab: -V > +V
plac:lang	−17.00	1.60	−10.63	***	plac(DE) > plac(FR) Vel: FR > DE
vox:lang	4.72	1.59	2.96	**	FR: -V > +V
plac:comp	−0.91	1.59	−0.57		
vox:comp	−4.31	1.59	−2.71	*	C: -V > +V
lang:comp	2.67	1.59	1.67		
plac:vox:lang	−11.96	3.19	−3.75	***	FR Lab: -V > +V
plac:vox:comp	−9.03	3.18	−2.84	**	C Lab: -V > +V
plac:lang:comp	5.60	3.19	1.75		
vox:lang:comp	−4.41	3.19	−1.38		
plac:vox:lang:comp	−12.74	6.37	−2.00		

continued on next page

Table 4.6 *continued from previous page*

C_1 plateau	β	SE(β)	t-value	p-value	sig. diff.
(Intercept)	51.68	5.58	9.26		
plac	10.79	1.15	9.41	***	Vel > Lab
vox	−4.23	1.14	−3.71	***	-V > +V
lang	−8.01	11.15	−0.72		
comp	−8.18	1.14	−7.16	***	S > C
plac:vox	5.92	2.29	2.59	*	Lab: -V > +V
plac:lang	−5.07	2.30	−2.21	*	plac(FR) > plac(DE)
vox:lang	9.88	2.29	4.32	***	FR: -V > +V
plac:comp	−8.02	2.29	−3.51	***	plac(S) > plac(C)
vox:comp	1.85	2.28	0.81		
lang:comp	−3.02	2.29	−1.32		
plac:vox:lang	0.09	4.58	0.02		
plac:vox:comp	3.32	4.57	0.73		
plac:lang:comp	−4.94	4.58	−1.08		
vox:lang:comp	5.06	4.57	1.11		
plac:vox:lang:comp	4.98	9.15	0.54		

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05

Table 4.7: Effects of place, voicing and language on C_2 plateau duration and plateau overlap in simple and complex onsets.

C_2 plateau	β	SE(β)	t-value	p-value	sig. diff.
(Intercept)	49.00	4.16	11.77		
lang	16.15	8.32	1.94		
vox	−4.83	1.24	−3.88	***	-V > +V
plac	3.84	1.24	3.09	**	Vel > Lab
lang:vox	−2.40	2.49	−0.97		
lang:plac	−4.57	2.49	−1.84		
vox:plac	−3.12	2.49	−1.26		
lang:vox:plac	−4.29	4.98	−0.86		

Plateau overlap	β	SE(β)	t-value	p-value	sig. diff.
(Intercept)	−21.85	4.53	−4.83		
lang	14.20	9.05	1.57		
vox	12.84	1.36	9.41	***	V+ > -V
plac	8.10	1.36	5.94	***	Vel > Lab
lang:vox	16.73	2.73	6.13	***	vox(DE) > vox(FR) +V: DE > FR
lang:plac	3.65	2.73	1.34		
vox:plac	7.50	2.73	2.75	**	plac(+V) > plac(-V)
lang:vox:plac	8.40	5.46	1.54		

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05

Table 4.8: Effect of place, language and complexity on the combined duration of occlusion and VOT in voiceless clusters.

Occlusion + VOT	β	SE(β)	t-value	p-value	sig. diff.
(Intercept)	143.99	6.24	23.06		
plac	3.09	1.33	2.33	*	Vel > Lab
lang	13.50	12.48	1.08		
comp	−9.51	1.33	−7.16	***	S > C
plac:lang	−10.74	2.66	−4.04	***	FR: Vel > Lab
plac:comp	−4.33	2.66	−1.63		
lang:comp	−13.51	2.66	−5.08	***	DE: S > C
plac:lang:comp	10.02	5.32	1.88		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

Table 4.9: Effects of place, language and voicing on the relative position of voice onset within the C_2 plateau duration.

C_2 voiced	β	SE(β)	t-value	p-value	sig. diff.
(Intercept)	34.19	8.85	3.86		
plac	26.04	6.19	4.21	***	Vel < Lab
vox	74.33	6.19	12.01	***	-V < +V
lang	−76.31	17.69	−4.31	***	DE < FR
plac:vox	−13.97	12.38	−1.13		
plac:lang	9.02	12.37	0.73		
vox:lang	0.53	12.37	0.04		
plac:vox:lang	17.24	24.75	0.70		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

> **Ich glaube nicht, dass wir das Spiel verloren hätten, wenn es 1:1 ausgegangen wäre.**
>
> BAYERN-MANAGER ULI HOENEß

Chapter 5

Conclusion

The final chapter will briefly review the results of the three previous chapters with respect to the three means of variation outlined in the introductory Section 1.4 that were used to probe the stability of intra-gestural and inter-gestural coordination in word-initial consonant clusters.

5.1 Variation of the segmental make-up

The clusters analyzed in Chapter 2 consisted of lingual clusters only (/kl/, /kn/, /ks/, /sk/), a choice resting upon the recording method: EPG can only register linguo-palatal contact which rules out all speech sounds in whose articulation linguo-palatal contact plays a subordinate rule, if any at all. It turned out that the clusters involving the sibilant both showed high degrees of variability in their inter-gestural coordination which may be related to special requirements for sibilants. These requirements are intrinsic to the sibilants and might not therefore have an impact on intergestural timing. Consequently, these requirements do not necessarily result in coordinative invariance. However, it is also not the case that these requirements entail the high variability which was found in the present data so that this issue remains unresolved here and should receive further attention in future research, see below. Stable coordination patterns emerged for /kl/ and for /kn/, the former with strong overlap, the latter with a considerable lag between the constriction plateaus. While the direction of this difference was correctly predicted by the models discussed in the introduction, only gestural recoverability was able to account for the striking extent of this difference. It is conceivable that in strongly overlapped productions of /kl/ the burst cues can be transmitted sufficiently through the lateral channels. For /kn/ on the other hand, these cues are at stake should velic lowering and/or the apical stop occur to early.

This finding for /kl/ and /kn/ was reproduced in the EMA study presented in Chapter 3. This study additionally compared /kl/ with /ks/, /pl/ and /ps/ under the assumption that the lower interdependence between the lips and the tongue tip (as compared to the tongue back and the tongue tip) would allow for more overlap in /p/+alveolar than in /k/+alveolar clusters. This assumption, however, found no support in the data. Instead, /p/ clusters appeared to be less overlapped and slightly more variable than /k/ clusters. This was even more so in the voiced clusters analyzed in Chapter 4. It was also assumed that greater coordinatory stability should appear in stop +/l/ clusters. Indeed, for both stops, the combination with /l/ emerged as more overlapped and less variable than the combination with /s/ which is in line with the results for /ks/ and /sk/ in the EMA study. It is however also true, that /k/ clusters have tighter and less variable coordination than /p/ clusters. The overall duration of /kl/ clusters is shorter than that of the other clusters.

5.2 Variation of prosody

Chapters 2 and 3 successfully applied prosodic variation in terms of boundary strength and lexical strength to word-initial consonant clusters. In Chapter 3 phrasal accent was varied as well albeit without success, and it will therefore be excluded from discussion here. In all clusters, boundary strength consistently affected the duration of the C_1 plateau which was longer at strong boundaries than at weak boundaries. Only in the case of /kl/, however, did this effect carry over to C_2 (EMA data only). This can be interpreted as a graded lengthening effect of the boundary which is stronger on C_1 than on C_2. There is also a significant effect of whether lexical stress falls on the first syllable (whose onset is the cluster) or not. The C_2 plateau is considerably longer in stressed syllables than in unstressed syllables. It is again /kl/ where this effect extends further than in the other cluster, that is to say: in /kl/ clusters the C_1 plateau lengthens as a function of lexical stress, too. Effects of stress are therefore also apparently of graded nature. This is especially manifest since the graded propagation–both for stress and for boundary related effects–appears in the cluster with the shortest total constriction interval, /kl/. In the other clusters, especially the stop+/s/ clusters, the consonant distal to the effect, i.e. C_2 for boundary effects and C_1 for stress effects, is simply too remote to be lengthened. The results therefore are in favor of the π-gesture approach (Byrd & Saltzman, 2003) and with regard to the gradedness confirm and extend previous findings (Byrd, Krivokapić, & Lee, 2006; Krivokapić, 2007).

Another factor contributing to the impression of gradedness is the influence on overlap. Under prosodic variation, the coordination in /kl/ is the least variable of all clusters, closely followed by the other tongue-back–tongue-tip cluster /ks/. The same has been found in the analysis of segmental make-up summarized above, but not quite as distinct. It can therefore be stated that prosodic variation is a helpful tool for probing stability in consonant clusters.

5.3 Variation in the larynx

Chapter 4 attempted to shed light on oral-laryngeal coordination in French and German by combining the analysis of VOT and EMA data. It is quite a departure from the points discussed hitherto in that it expands from purely supra-laryngeal properties of clusters in one language to oral-laryngeal coordination cross-linguistically. This summary will have to reach back a little before it gets to the point of discussing the goodness of clusters. The mixed-voicing cluster /kl/ was chosen because in the analysis so far it has emerged as the most stable and because it has a fully voiced counterpart /gl/. The pair /pl/ and /bl/ was chosen because it presents a good case for testing against. Crucially, of course, both clusters are common in both German and French. The point of departure was the overlap difference between mixed-voicing clusters and fully voiced clusters in German clusters (more overlap in /bl/ than in /pl/) which is not present in French clusters. This has been linked to the voicing difference between the languages. For German it has been assumed that mixed-voicing clusters overlap to a lesser extent in order to prevent the aspiration from fully devoicing the lateral. In other words, aspiration which is due to the oral-glottal timing has to be accommodated. Consequently the oral gestures move apart. Since in French timing is different and aspiration does not usually play a role, a shifting of the oral gestures appears to be superfluous and should not occur. Oral coordination in French clusters, fully voiced and with mixed voicing, should resemble the coordination in German fully voiced clusters. This expectation is not borne out. On the contrary, the oral gestures in French clusters are coordinated as in the mixed-voicing clusters in German. As expected, the acoustics show that sonorant devoicing is much stronger in German than in French. Taken together, it appears as if the glottal gesture in French is associated with the stop only, while it seems to be a property of the entire onset in German (see Kehrein & Golston, 2004; and also Hoole, 2006). These results can be linked to work on Moroccan Arabic (e.g. Shaw, Gafos, Hoole, & Zeroual, 2009) where similar timing differences are considered to be crucial for syllable affiliation. To be more clear, it might be possible to argue for a hetero-syllabic parse in the French clusters. The crucial point here is, however, that there are considerable timing differences between German and French clusters. One cannot rule out that in French clusters criteria for cluster goodness might be entirely different than in German. However, Kühnert, Hoole, Mooshammer, and Bombien (2008) show that one important result of this work is true for both German and French although less consistent in the latter: /kl/ clusters are more overlapped than /kn/ clusters. It maybe that the implementation of the voicing contrast conceals the fact that some segmentally conditioned coordination patterns can nonetheless be found cross-linguistically. This conclusion needs to be substantiated with data on more languages. As for the coordination stability, there is no indication that fully voiced clusters should be more stable than mixed-voicing clusters or *vice versa*.

5.4 Conclusion and outlook

Is /kl/ a *good* cluster? It appears so. It is a cluster that is frequently encountered in the languages of the World and it is readily compatible with the models of sonority modulation, does not stand in the way of recoverability and it exhibits considerable gestural overlap. It may well be that it is the combination of these properties–and very likely others as well–that makes a cluster successful or diachronically stable cross-linguistically or *good*. The fact that the cluster /kn/ has been lost in English could be regarded as evidence for an approach to sound change in which these properties play a crucial rule. It is, however, not clear how such properties would have to be weighted. Consider for example the role of stability. /ks/ does not lag far behind /kl/ in terms of stability, but it is much less overlapped and less preferable in terms of sonority modulation. Also, it plays a very marginal role in German. But then again, the role of /pl/ is not so marginal in German in spite of the presence of /pfl/ due to the High German consonant shift. Nevertheless, it is more variable and less overlapped than /kl/ even in spite of having less articulatory interdependence. A comparison with English where the functional load of /pl/ clusters is higher than in German might shed light on this issue.

The results presented in Chapters 2 to 4 are based on kinematic patterns as they are actually produced by speakers. The analysis segmental and prosodic variation allows for the interpretation that some kinematic patterns are more prefered than others. This interpretation, however, cannot be proven solely on the basis of the articulographic data. Instead, perception tests should be carried out in order to assess whether kinematic patterns which have *not* been observed and which are assumed to be disprefered by the speaker are also disprefered by the listener. Articulatory synthesis systems (e.g. TaDA: Nam, Goldstein, & Proctor, 2007; or VocalTractLab: Birkholz, Jackèl, & Kröger, 2007; Birkholz & Kröger, 2006; Birkholz, 2007) can be used to systematically vary kinematic patterns within a continuum including both observed and unobserved coordination relations and to create acoustic stimuli. Using such stimuli in perception experiments should show whether speakers avoid the unobserved patterns because they are acoustically and perceptorily unfavorable. Articulatory synthesis is superior to other synthesis systems for such a task because its acoustic output is based on articulatory trajectories which can be manipulated in time and space. The emergence of epenthetic or transitional vowels in consonant clusters (/CC/ → /C^{e}C/) for example is easy to simulate in articulatory synthesis by adding to the phase angle of the coordination relations, i.e. by pulling apart the consonants' trajectories.

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

- /gl/ appears to be less frequent in German than /kl/. The present investigations do not have an explanation for this or the parallel bilabial case, i.e. /bl/ vs. /pl/. One might speculate that the already weak release cues of the voiced stops do not benefit from the following lateral such that the entire combination is less salient. It should be interesting to review this patterning under prosodic variation possibly including /gn/ vs. /kn/ although the addition of velic activity would complicate matters.
- Spatial properties have been entirely disregarded in the analysis of EMA data

so far. It is not unlikely that the differences between the coordination patterns of clusters involving /s/ and other clusters can be traced back to requirements that are better reflected in space than in time. It might play a role, for example, that German /s/ is typically articulated laminally while /l/ is articulated apically. Furthermore, /s/ has been found to be produced with a very clear somatosensory goal (see Perkell, 2004; Ghosh, 2010) – more so perhaps than other sounds. On the other hand, high variability in /s/ clusters may arise because in order to meet these special requirements (Stone, Faber, Raphael, & Shawker, 1992) the tongue might have to rid itself from various contextual factors before target achievement. It should be interesting to devise and conduct experiments that specifically control for such factors.

- Coordination patterns in French and German appear to differ considerably. There are some links to research on Moroccan Arabic here, that definitely justify future research in this area. The association of the glottal gesture within the onset can be interpreted to suggest hetero-syllabic affiliation of the clusters' consonants in French. Investigations of the C-center or more generally on the alignment in both French and German onsets would shed light on this issue.
- Data in French and Moroccan Arabic indicate that true voicing in the stop in sequences such as /bl/ is incompatible with overlap. Aerodynamically, this could be accounted for by the requirement of an intermediate release of supra-glottal air pressure by means of inter-gestural lag in order to facilitate the maintenance of glottal vibration. A very interesting test case would be Dutch, which in spite of being a Western Germanic language (presumably with Germanic overlap patterns as opposed to those encountered in French) has a voicing contrast of prevoiced — short-lag as in French rather than a short-lag — long-lag as in German.
- It was argued concerning the difference between /kl/ and /pl/ that rather than producing /pl/ with more overlap because the articulators are independent of each other, the strong interdependence of the tongue-tip and the tongue-dorsum might impose constraints that force the production of /kl/ into a very narrow window of overlap and variability. This is not easily reconciled with current, yet unpublished research by Marianne Pouplier and Štefan Beňuš on syllabic /l/ in Slovak. Their data show little overlap in initial /Cl/ clusters regardless of whether /l/ constitutes the syllable nucleus or is part of the onset. On the other

hand, Slovak has a prevoicing — short-lag contrast as French does. It may well be that voicing contrast and degree of overlap are typologically linked.

As for the *goodness* of clusters, it might turn out that what presents itself as good in one language might appear in a less favorable form in another language due to constraints which do not exist in the former (e.g. voicing contrast implementation). Future research should involve the identification of such constraints. This point emphasizes the importance of research which promotes the incorporation of language specific grammar in the application of physically based models of phonology (Gafos, 2002; Shaw, 2009; Gafos, Hoole, Roon, & Zeroual, 2010).

Appendix A

Speech Material

A.1 Complete speech material (EPG)

Table A.1: Utterances for cluster /kl/

	stress on first syllable
Utterance initial	*Thomas studiert in Fulda.* ***Claudia*** *geht noch zur Schule.*
	'Thomas goes to college in Fulda. Claudia ist still in school.'
Phrase initial	*Olga sagt immer,* ***Claudia*** *sei noch zu jung.*
	'Olga always says that Claudia is still too young.'
List	*Thomas, Peter,* ***Claudia*** *und Elke fahren in den Süden.*
	'Thomas, Peter, Claudia and Elke are driving south.'
Word initial	*Gestern war* ***Claudia*** *noch gesund.*
	'Yesterday, Claudia was still OK.'
	stress on second syllable
Utterance initial	*Die Arbeit war super.* ***Klausur*** *und mündliche Prüfung waren nicht so toll.*
	'The thesis was great. Written and oral exams were not as good.'
Phrase initial	*Tine sagt immer,* ***Klausur*** *schreiben macht Spaß.*
	'Tine always says it's fun to write exams.'
List	*Hausarbeit, Wetter,* ***Klausur*** *und Erkältung machen schlechte Laune.*
	'Housework, weather, written exams and a cold cause sulkiness.'
Word initial	*Morgen muss sie wieder* ***Klausur*** *schreiben.*
	'Tomorrow she has to write a test again.'

Table A.2: Utterances for cluster /kn/

	stress on first syllable
Utterance initial	*Peter ist Fussballtrainer.* ***Kneipe*** *und Stadion sind sein Leben.*
	'Peter is a football coach. Pub and stadium are his life.'
Phrase initial	*Thomas sagt immer,* ***Kneipe*** *oder Café machen zu viel Arbeit.*
	'Thomas always says a pub or a coffee shop are too much work.'
List	*Restaurant, Bar,* ***Kneipe*** *und Disco wollen sie heute noch besuchen.*
	'The plan to visit a restaurant, a bar, a pub and a disco today.'
Word initial	*Sie arbeitet in einer* ***Kneipe*** *als Kellnerin.*
	'She works in a pub as a waitress.'
	stress on second syllable
Utterance initial	*Walter trinkt gerne Vodka.* ***Kneipier*** *ist sein Traumberuf.*
	'Walter likes Vodka. He dreams of being a pub owner.'
Phrase initial	*Peter sagt immer,* ***Kneipier*** *ist ein schöner Beruf.*
	'Peter always says that pub owner is a nice job.'
List	*Koch, Kellner,* ***Kneipier*** *oder Barkeeper würde er gern werden.*
	'He would like to be cook , waiter, pub owner or barkeeper.'
Word initial	*Er wollte immer* ***Kneipier*** *werden.*
	'He always wanted to be a pub owner.'

Table A.3: Utterances for cluster /sk/

	stress on first syllable
Utterance initial	*Olga studiert in Jena.* ***Scarlett*** *geht noch zur Schule.*
	'Olga goes to college in Jena. Scarlett is still in school.'
Phrase initial	*Walter sagt immer,* ***Scarlett*** *sei zu jung.*
	'Walter always says that Scarlett is still too young.'
List	*Peter, Walter,* ***Scarlett*** *und Olga fahren in den Süden.*
	'Peter, Walter, Scarlett and Olga are driving south.'
Word initial	*Gestern war* ***Scarlett*** *noch gesund..*
	'Yasterday Scarlett still was well.'
	stress on second syllable
Utterance initial	*Walter hört immer Schlager. "****Skandal*** *um Rosi" mag er besonders gern.*
	'Walter likes Schlager music. "Skandal um Rosi" is his favourite.'
Phrase initial	*Peter sagt immer, "****Skandal*** *um Rosi" geht ihm auf die Nerven..*
	'Peter always says, "**Skandal** um Rosi" gets on his nerves.'
List	*Affäre, Schickeria,* ***Skandal*** *und Betrug gehören in Thomas Kolumne..*
	'Affairs, jet set, scandals and deceit are part of Thomes' column.'
Word initial	*Das war der größte* ***Skandal*** *im letzten Jahr.*
	'It was last year's greatest scandal.'

Table A.4: Utterances for cluster /ks/

stress on first syllable

Utterance initial	*Volker studiert in Jena.* ***Xaver*** *geht noch zur Schule.*
	'Volker goes to college in Jena. Xaver is still in school'
Phrase initial	*Walter sagt immer,* ***Xaver*** *sei zu jung..*
	'Walter always says that Xaver is still too young.'
List	*Inge, Walter,* ***Xaver*** *und Elke fahren in den Süden.*
	'Inge, Walter, Xaver and Elke are driving south.'
Word initial	*Am Montag war* ***Xaver*** *noch gesund.*
	'On monday Xaver was still well.'

A.2 Complete speech material (EMA)

Table A.5: Utterances for cluster /kl/. Target words highlighted in bold, contrastive accent (where applicable) in asteriscs.

Stress on first syllable	
Utterance initial	*Thomas studiert in Fulda.* ***Claudia*** *geht noch zur Schule.*
	'Thomas goes to college in Fulda. Claudia ist still in school.'
Phrase initial	*Olga sagt immer,* ***Claudia*** *sei zu jung.*
	'Olga always says that Claudia is too young.'
List	*Thomas, Peter,* ***Claudia*** *und Elke fahren in den Süden.*
	'Thomas, Peter, Claudia and Elke are driving south.'
Word initial	*Gestern war* ***Claudia*** *noch frisch und gesund.*
	'Yesterday, Claudia was still OK.'
Deaccented	*Das Buch hat nicht *Dieter** ***Claudia*** *gegeben, sondern Peter.*
	'It is not Dieter who gave the book to Claudia but Peter.'
Stress on second syllable	
Utterance initial	*Die Arbeit war super.* ***Klausur*** *und mündliche Prüfung waren nicht so toll.*
	'The thesis was great. Written and oral exams were not as good.'
Phrase initial	*Tine sagt immer,* ***Klausur*** *schreiben macht Spaß.*
	'Tine always says it's fun to write exams.'
List	*Hausarbeit, Wetter,* ***Klausur*** *und Erkältung machen schlechte Laune.*
	'Housework, weather, written exams and a cold cause sulkiness.'
Word initial	*Morgen muss sie weder* ***Klausur*** *noch Examen schreiben.*
	'Tomorrow she has to write a test again.'
Deaccented	*Morgen wird nicht *Walter** ***Klausur*** *schreiben, sondern Volker.*
	'It is not Walter who will write a test tomorrow but Volker.'

Table A.6: Utterances for cluster /kn/. Target words highlighted in bold, contrastive accent (where applicable) in asteriscs.

Stress on first syllable	
Utterance initial	*Peter ist Fussballtrainer.* ***Kneipe*** *und Stadion sind sein Leben.* 'Peter is a football coach. Pub and stadium are his life.'
Phrase initial	*Thomas sagt immer,* ***Kneipe*** *oder Café machen zu viel Arbeit.* 'Thomas always says a pub or a coffee shop are too much work.'
List	*Restaurant, Bar,* ***Kneipe*** *und Disco wollen sie heute noch besuchen.* 'The plan to visit a restaurant, a bar, a pub and a disco today.'
Word initial	*Sie arbeitet in einer* ***Kneipe*** *als Kellnerin.* 'She works in a pub as a waitress.'
Deaccented	*Bier schmeckt nicht in *seiner** ***Kneipe*** *am besten, sondern in meiner.* 'Beer is not best in his pub but in mine.'
Stress on second syllable	
Utterance initial	*Walter trinkt gerne Vodka.* ***Kneipier*** *ist sein Traumberuf.* 'Walter likes Vodka. He dreams of being a pub owner.'
Phrase initial	*Peter sagt immer,* ***Kneipier*** *ist ein schöner Beruf.* 'Peter always says that pub owner is a nice job.'
List	*Koch, Kellner,* ***Kneipier*** *oder Barkeeper würde er gern werden.* 'He would like to be cook , waiter, pub owner or barkeeper.'
Word initial	*Er wollte immer* ***Kneipier*** *werden.* 'He always wanted to be a pub owner or a bar keeper.'
Deaccented	*Früher ist nicht *Peter** ***Kneipier*** *gewesen, sondern Volker.* 'It is not Peter who formerly was a pub owner but Volker.'

Table A.7: Utterances for cluster /ks/. Target words highlighted in bold, contrastive accent (where applicable) in asteriscs.

Stress on first syllable	
Utterance initial	*Volker studiert in Jena.* ***Xaver*** *geht noch zur Schule.* 'Volker goes to college in Jena. Xaver is still in school'
Phrase initial	*Walter sagt immer,* ***Xaver*** *sei zu jung..* 'Walter always says that Xaver is still too young.'
List	*Inge, Walter,* ***Xaver*** *und Elke fahren in den Süden.* 'Inge, Walter, Xaver and Elke are driving south.'
Word initial	*Am Montag war* ***Xaver*** *noch gesund.* 'On monday Xaver was still well.'
Deaccented	*Das Buch hat nicht *Eva** ***Xaver*** *gegeben, sondern Walter.* 'It is not Eva who gave the book to Xaver but Walter.'

Table A.8: Utterances for cluster /ps/. Target words highlighted in bold, contrastive accent (where applicable) in asteriscs.

Stress on first syllable	
Utterance initial	*Elke singt gerne Lieder.* ***Psalmen*** *singt sie auch.*
	'Elke likes to sing songs. She also sings psalms.'
Phrase initial	*Thomas sagt immer,* ***Psalmen*** *seinen altmodisch.*
	'Thomas always says that psalms are old-fashioned.'
List	*Sprüche, Lieder,* ***Psalmen*** *und Verse kann sie auswendig.*
	'She knows quotations, songs, psalms and verses by heart.'
Word initial	*David hat viele der* ***Psalmen*** *und Lieder verfasst.*
	'David has composed many of the psalms.'
Deaccented	*Es soll nicht *Elke** ***Psalmen*** *singen, sondern Anna.*
	'It is not Elke who will sing psalms but Anna.'
Stress on second syllable	
Utterance initial	*David war König von Juda.* ***Psalmist*** *war er auch..*
	'David was the king of Juda. He also was a psalmist.'
Phrase initial	*Peter sagt immer,* ***Psalmist*** *wäre er gern.*
	'Peter always says that he would like to be a psalmist.'
List	*Hirte, Kämpfer,* ***Psalmist*** *und König ist David gewesen.*
	'David was a shepard, a fighter, a psalmist, and a king.'
Word inital	*Im Alter ist er* ***Psalmist*** *und Sänger gewesen.*
	'In old age he was a psalmist and a singer.'
Deaccented	*Also ist nicht *Peter** ***Psalmist*** *gewesen, sondern Paul.*
	'So it was not Peter who was a psalmist but Paul.'

A.3 Speech material for the voicing study

	French			
	Voiced		Voiceless	
Onset	high vowel	low vowel	high vowel	low vowel
velar				
simplex	—	gâte	kif	cap
complex	glisse	glace	clique	claque
labial				
simplex	bique	bac	pic	pâte
complex	blini	blatte	plisse	plaque
	German			
velar				
simplex	gib	gab	kies	kahl
complex	glied	glas	klean	klag
labial				
simplex	biest	bad	piep	pack
complex	blieb	blatt	plitsch	plan

Table A.9: Material for the analysis of voicing in clusters in French and German stop+/l/ clusters

Appendix B

Previous works

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

Bombien, L. (2006). Voicing alterations in Icelandic - a photoglottographic and acoustic investigation. *AIPUK*, *37*, 63–82. **retrieved from** http://www.ipds.uni-kiel.de/pub_exx/aipuk/aipuk_37/37_6_Bombien_akfra.pdf

Bombien, L., Cassidy, S., Harrington, J., John, T., & Palethorpe, S. (2006). Recent developements in the EMU Speech Database System. In *Proceedings 11th SST conference* (pp. 313–316). Auckland.

Bombien, L., Mooshammer, C., Hoole, P., & Kuehnert, B. (2008). Prosodic effects on articulatory coordination in initial consonant clusters in German. *The Journal of the Acoustical Society of America*, *123* (5), 3331–3331

Bombien, L., Mooshammer, C., Hoole, P., Kühnert, B., & Schneeberg, J. (2006, Dec.). An EPG study of initial /kl/ clusters in varying prosodic conditions in German. In *Proc. 7th issp* (pp. 457–460). Ubatuba, Brazil.

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

Bombien, L., Mooshammer, C., Hoole, P., & Kühnert, B. (2010). Prosodic and segmental effects on EPG contact patterns of word-initial German clusters. *Journal of Phonetics*, *38* (3), 388–403.

Hoole, P., Bombien, L., Kühnert, B., & Mooshammer, C. (2009). Intrinsic and prosodic effects on articulatory coordination in initial consonant clusters. In G. Fant, H. Fujisaki & J. Shen (Eds.), *Frontiers in phonetics and speech science* (pp. 275–286). Beijing: Commercial Press.

Hoole, P., & Bombien, L. (2010). Velar and glottal activity in Icelandic. In S. Fuchs, P. Hoole, C. Mooshammer & M. Żygis (Eds.), *Between the regular and the particular in speech and language* (pp. 171–204). Frankfurt: Peter Lang.

Kühnert, B., Hoole, P., Mooshammer, C., & Bombien, L. (2008). Influences of manner and voicing on articulatory coordination in German and French initial consonant clusters. *The Journal of the Acoustical Society of America*, *123* (5), 3740–3740

Mády, K., Bombien, L., & Reichel, U. D. (2008). Is Hungarian losing the vowel quantity distinction? In *Proc. 8th international seminar on speech* (pp. 445–448). Strasbourg.

Bibliography

Baayen, R. H. (2008). *Analyzing linguistic data: a practical introduction to statistics using R* (1st ed.). Cambridge: Cambridge University Press.

Bates, D., & Maechler, M. (2009). *Lme4: Linear mixed-effects models using S4 classes.* R package version 0.999375-32. **retrieved from** `http://CRAN.R-project.org/package=lme4`

Beckman, M., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory analysis. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology: gesture, segment, prosody* (pp. 68–86). Cambridge: Cambridge University Press.

Beckman, M., & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*. Phonology Yearbook 3, *3*, 255–309.

Benguerel, A., Hirose, H., Sawashima, M., & Ushijima, T. (1978). Laryngeal control in French stop production: a fiberscopic, acoustic and electromyographic study. *Folia Phoniatrica et Logopaedica*, *30* (3), 175–198

Bergmann, P. (2008, Aug.). Assimilation within complex words in German. Poster presented at the workshop Consonant Clusters and Structural Complexity, Munich.

Best, C. T., & Hallé, P. A. (2010, Jan.). Perception of initial obstruent voicing is influenced by gestural organization. *Journal of Phonetics*, *38* (1), 109–126

Birkholz, P. (2007). Control of an articulatory speech synthesizer based on dynamic approximation of spatial articulatory targets. In *Interspeech 2007 — Eurospeech* (pp. 2865–2868). Antwerp, Belgium.

Birkholz, P., Jackèl, D., & Kröger, B. J. (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., & Kröger, B. J. (2006). Vocal tract model adaptation using magnetic resonance imaging. In *7th international seminar on speech production (issp'06)* (pp. 493–500). Ubatuba, Brazil.

Blevins, J., & Garrett, A. (2004). The evolution of metathesis. In B. Hayes, R. Kirchner & D. Steriade (Eds.), *Phonetically based phonology* (pp. 117–156). Cambridge: Cambridge University Press.

Bombien, L., Cassidy, S., Harrington, J., John, T., & Palethorpe, S. (2006). Recent develeopements in the EMU Speech Database System. In *Proceedings 11th SST conference* (pp. 313–316). Auckland.

Bombien, L., Mooshammer, C., Hoole, P., Kühnert, B., & Schneeberg, J. (2006, Dec.). An EPG study of initial /kl/ clusters in varying prosodic conditions in German. In *Proc. 7th issp* (pp. 457–460). Ubatuba, Brazil.

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

Bombien, L., Mooshammer, C., Hoole, P., & Kühnert, B. (2010). Prosodic and segmental effects on EPG contact patterns of word-initial German clusters. *Journal of Phonetics*, *38* (3), 388–403.

Browman, C., & Goldstein, L. (1986). Towards an Articulatory Phonology. *Phonology Yearbook*, *3*, 219–252.

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

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

Browman, C., & Goldstein, L. (1990). Tiers in Articulatory Phonology. In J Kingston & M. E. Beckman (Eds.), *Papers in laboratory phonology 1: between the grammar and the physics of speech* (pp. 241–376). Cambridge: Cambridge University Press.

Browman, C., & Goldstein, L. (1992a). Articulatory Phonology: an overview. *Phonetica*, *49* (3-4), 155–180.

Browman, C., & Goldstein, L. (1992b). Targetless schwa: an articulatory analysis. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology: gesture, segment, prodosy* (pp. 26–56). Cambridge: Cambridge University Press.

Browman, C., & Goldstein, L. (1995). Dynamics and Articulatory Ponology. In R Port & T van Gelder (Eds.), *Mind as motion: dynamics, behavior, and cognition* (pp. 175–193). Cambridge, MA: MIT Press.

Browman, C., & Goldstein, L. (2000). Competing constraints on intergestural coordination and selforganization of phonological structure. *Les Cahiers de l'ICP*, *5*, 25–34.

Byrd, D., & Choi, S. (2010). At the juncture of prosody, phonology, and phonetics - the interaction of phrasal and syllable structure in shaping the timing of consonant gestures. In C. Fougeron, B. Kühnert, M. d'Imperio & N. Vallé (Eds.), *Papers in Laboratory Phonology 10: variation, detail and representation* (pp. 31–60). Berlin/New York: Mouton de Gruyter.

Byrd, D. (1993, Feb.). 54,000 American stops. *UCLA Working Papers in Phonetics*, *83*, 97–116.

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

Byrd, D. (1996a). A phase window framework for articulatory timing. *Phonology*, *13* (2), 139–169. **retrieved from** `http://www.jstor.org/stable/4615480`

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

Byrd, D., Flemming, E., Mueller, C. A., & Tan, C. C. (1995). Using regions and indices in EPG data reduction. *Journal of Speech and Hearing Research*, *38* (4), 821–827.

Byrd, D., Kaun, A., Narayanan, S., & Saltzman, E. (2000). Phrasal signatures in articulation. In M. Broe & J. Pierrehumbert (Eds.), *Papers in laboratory phonology 5: acquisition and the lexicon* (pp. 70–87). Cambridge: Cambridge University Press.

Byrd, D., Krivokapić, J., & 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., Lee, S., Riggs, D., & 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., & Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, *31* (2), 149–180.

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

Byrd, D., & Tan, C. 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., & Turk, A. E. (1999). A cross-linguistic study of accentual lengthening: Dutch vs. English. *Journal of Phonetics*, *27* (3), 255 –280

Caramazza, A, & Yeni-Komshian, G. (1974). Voice onset time in two French dialects. *Journal of Phonetics*, *2*, 239–245.

Chitoran, I., Goldstein, L., & Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian. In *Laboratory phonology 7* (pp. 419–448). Berlin/New York: Mouton de Gruyter.

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

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

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, *27* (2), 207 –229

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

Cho, T., McQueen, J. M., & Cox, E. A. (2007). Prosodically driven phonetic detail in speech processing: the case of domain-initial strengthening in English. *Journal of Phonetics*, *35* (2), 210–243.

Crystal, T. H., & House, A. S. (1988). Segmental durations in connected-speech signals: current results. *The Journal of the Acoustical Society of America*, *83* (4), 1553–1573

Davidson, L. (2005). Addressing phonological questions with ultrasound. *Clinical Linguistics & Phonetics*, *19* (6-7), 619–633.

Davidson, L., & Roon, K. (2008). Durational correlates for differentiating consonant sequences in Russian. *Journal of the International Phonetic Association*, *38* (2), 137–165.

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

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

Docherty, G. J. (1992). *The timing of voicing in british English obstruents.* Berlin, New York: Walter de Gruyter.

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

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

Fischer-Jørgensen, E. (1972). 'p t k' et 'b d g' francais en position intervocalique accentuee. In A Valdman (Ed.), *Papers in linguistics and phonetics to the memory of pierre delattre* (pp. 143–200). The Hague: Mouton.

Fontdevila, J, Pallarès, M. D., & Recasens, D. (1994). The contact index method of EPG data reductions. *Journal of Phonetics*, *22*, 141–154.

Fougeron, C. (1998). *Variations articulatoires en début de constituants prosodiques de différents niveaux en francais.* (Thèse de doctorat, Université Paris III).

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

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

Fougeron, C., Meynadier, Y., & Demolin, D. (2000). 62 vs 96 electrodes: a comparative analysis of Reading and Kay Elemetrics EPG pseudo-palates. In *Proceeding of the 5th seminar on speech production : models and data* (pp. 309–312). Kloster Seeon, Bavaria.

Fuchs, S. (2005, Aug.). *Articulatory correlates of the voicing contrast in alveolar obstruent production in German.* (PhD thesis, Zentrum für allgemeine Sprachwissenschaft, Berlin).

Gafos, A. (2002). A grammar of gestural coordination. *Natural Language and Lingustic Theory*, *20* (1), 169–337.

Gafos, A., Hoole, P., Roon, K., & Zeroual, C. (2010). Variation in timing and phonological grammar in Moroccan Arabic clusters. In C. Fougeron, B. Kühnert, M. d'Imperio & N. Vallé (Eds.), *Laboratory phonology 10: variation, detail and representation* (pp. 657–698). Berlin/New York: Mouton de Gruyter.

Geumann, A., Kroos, C., & Hoole, P. (1999). Are there compensatory effects in natural speech? In *Proceedings of the 14th icphs* (pp. 399–402). San Francisco.

Ghosh, S. S., Matthies, M. L., Maas, E., Hanson, A., Tiede, M., Ménard, L., …Perkell, J. S. (2010). An investigation of the relation between sibilant production and somatosensory and auditory acuity. *The Journal of the Acoustical Society of America*, *128* (5), 3079–3087

Gibbon, F., Hardcastle, W., & Nicolaidis, K. (1993). Temporal and spatial aspects of lingual coarticulation in /kl/ sequences: a cross-linguistic investigation. *Language and Speech*, *36* (2-3), 261–277.

Goldstein, L., Chitoran, I., & Selkirk, E. (2007). Syllable structure as coupled oscillator modes: evidence from Georgian vs. Tashlhhiyt Berber. In W. Barry & J. Trouvain (Eds.), *Proceedings of the xvi. international congress of phonetic sciences.*

Goldstein, L., Nam, H., Saltzman, E., & Chitoran, I. (2009). Coupled oscillator planning model of speech timing and syllable structure. In G. Fant, H. Fujisaki & J. Shen (Eds.), *Frontiers in phonetics and speech science.* Beijing: Commercial Press.

Goldstein, L., & Browman, C. (1986). Representations of voicing contrasts using articulatory gestures. *Journal of Phonetics*, *14*, 339–342.

Hall, N. (2003). *Gestures and segments: vowel intrusion as overlap.* (PhD thesis, University of Haifa).

Hardcastle, W. J., Gibbon, F. E., & Jones, W. (1991). Visual display of tongue-palate contact: electropalatography in the assessment and remediation of speech disorders. *Br J Disord Commun*, *26* (1), 41–74.

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

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

Hermes, A., Grice, M., Mücke, D., & Niemann, H. (2008). Artuculatory indicators of syllable affiliation in word initial consonant clusters in Italian. In R. Sock, S. Fuchs &

Y. Laprie (Eds.), *Proceedings of the 8th international seminar on speech production* (pp. 429–432).

Honorof, D., & Browman, C. (1995). The center or edge: how are consonant clusters organized with respect to the vowel. In *Proceedings of the xiiith icphs* (Vol. 3, pp. 552–555). Stockholm.

Hoole, P., Bombien, L., Kühnert, B., & Mooshammer, C. (2009). Intrinsic and prosodic effects on articulatory coordination in initial consonant clusters. In G. Fant, H. Fujisaki & J. Shen (Eds.), *Frontiers in phonetics and speech science* (pp. 275–286). Beijing: Commercial Press.

Hoole, P. (2006). *Experimental studies of laryngeal articulation. Part 2: Larygeal-Oral coordiration in consonant sequences.* (Unpublished habilitation thesis, Ludwig-Maximilians-Universität, Munich). **retrieved** 26 Mar. 2010, **from** `http://www.phonetik.uni-muenchen.de/~hoole/pdf/habilpgg_chap_all.pdf`

Hoole, P., & Mooshammer, C. (2002). Articulatory analysis of the German vowel system. In P. Auer, P. Gilles & H. Spiekermann (Eds.), *Silbenschnitt und Tonakzente* (pp. 129–152). Tübingen: Niemeyer.

Hoole, P., & Zierdt, A. (2010). Five-dimensional articulography. In B. Maassen & P. van Lieshout (Eds.), *Speech motor control: new develpements in basic and applied research* (pp. 331–350). Oxford: Oxford University Press.

Huygens, C. (1673). *Horoloquium oscilatorium.* Paris.

Jongman, A. (1989). Duration of frication noise required for identification of English fricatives. *Journal of the Acoustical Society of America*, *85* (4), 1718–1725.

Keating, P., Cho, T., Fougeron, C., & Hsu, C. (2003). Domain-initial articulatory strengthening in four languages. In *Laboratory phonology 6* (pp. 143–161). Cambridge: Cambridge University Press.

Keating, P. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, *60* (2), 286–319. **retrieved from** `http://www.jstor.org/stable/413642`

Keating, P. (2006). Phonetic encoding of prosodic structure. In J. Harrington & M. Tabain (Eds.), *Speech production: models, phonetic processes and techniques* (pp. 167–186). New York: Psychology Press.

Kehrein, W., & Golston, C. (2004). A prosodic theory of contrast. *Phonology*, *21*, 1–33.

Klatt, D. H. (1975). Voice Onset Time, Frication, and Aspiration in Word-Initial Consonant Clusters. *Journal of Speech and Hearing Research*, *18* (4), 686–706. **retrieved from** `http://jslhr.asha.org/cgi/content/abstract/18/4/686`

Kochetov, A, Pouplier, M., & Son, M. (2007). Cross-language differences in overlap and assimilation patterns in Korean and Russian. In J. Trouvain & W. Barry (Eds.), *Proceedings of the 16th international congress of the phonetic sciences* (pp. 1361–1364). Saarbrücken, Germany.

Kohler, K. J. (1995). *Einführung in die Phonetik des Deutschen* (2nd ed.). Grundlagen der Germanistik. Berlin: Erich Schmidt.

Kollia, H. B., Gracco, V. L., & Harris, K. S. (1995). Articulatory organization of mandibular, labial, and velar movements during speech. *Journal of the Acoustical Society of America*, *98* (3), 1313–1324.

Krakow, R. A. (1993). Nonsegmental influences on velum movement patterns: syllables, sentences, stress, and speaking rate. In M. Huffmann & R. A. Krakow (Eds.), *Nasals, nasalization, and the velum* (Vol. 5, pp. 87–116). Phonetics and Phonology. San Diego: Academic Press, Inc.

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

Kroos, C. (1996). *Kinematische Analyse der Gespanntheitsopposition im Standarddeutschen.* (MA thesis, Institut für Phonetik und sprachliche Kommunikation der Ludwig-Maximilians-Universitat München).

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

Kühnert, B., Hoole, P., & Mooshammer, C. (2006). Gestural overlap and c-center in selected French consonant clusters. In H. Yehia, D. Demolin & R. Laboissière (Eds.), *Proceedings of the 7th international seminar on speech production* (pp. 40–48). UFMG. Belo Horizonte.

Kühnert, B., Hoole, P., Mooshammer, C., & Bombien, L. (2008). Influences of manner and voicing on articulatory coordination in German and French initial consonant clusters. *The Journal of the Acoustical Society of America*, *123* (5), 3740–3740

Kühnert, B., & Hoole, P. (2004). Speaker-specific kinematic properties of alveolar reductions in English and German. *Clinical Linguistics and Phonetics*, *18* (6-8), 559–575.

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

Liberman, M. (2008, May 5). Slicing the syllabic bologna. **retrieved from** `http://languagelog.ldc.upenn.edu/nll/?p=124`

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word*, *20*, 384–422.

Lisker, L., & Abramson, A. (1967). Some effects of context on voice onset time in English stops. *Lang Speech*, *10* (1), 1–28.

Maddieson, I. (1993). Investigating Ewe articulations with electromagnetic articulography. *Forschungsberichte - Institut für Phonetik und Sprachliche Kommunikation der Universität München*, *31*, 181–214.

Maddieson, I. (1997). Phonetic universals. In W. J. Hardcastle & J. Laver (Eds.), *Handbook of phonetic sciences* (pp. 619–639). Oxford: Blackwell.

Maddieson, I. (2009, June). *Patterns of sounds.* Cambridge: Cambridge University Press.

Marin, S., & Pouplier, M. (2010). Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control*, *14* (3), 380–407. **retrieved from** `http://journals.humankinetics.com/mc-current-issue/mc-volume-14-issue-3-july-2010/temporal-organization-of-complex-onsets-and-codas-in-american-english-testing-the-predictions-`

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

Max, L., & Onghena, P. (1999). Some issues in the statistical analysis of completely randomized and repeated measures for speech, language and hearing research. *Journal of Speech, Language and Hearing Research*, *42* (2), 261–270.

Meynadier, Y., Pitermann, M., & Marchal, A. (1998). Effects of contrastive focal accent on linguopalatal coarticulation in the French [kskl] clusters. In *Proceedings of the fifth international conference on spoken language processing* (Vol. 5, pp. 1871–1874).

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

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

Mooshammer, C., Hoole, P., & Kühnert, B. (1995). On loops. *Journal of Phonetics*, *23* (1), 3–21.

Nam, H. (2007). Syllable-level intergestural timing model: split-gesture dynamics focusing on positional asymmetry and moraic structur. In J. Cole & J. I. Haulde (Eds.), *Papers in laboratory phonology 9: phonology and phonetics.* Berlin, New York: de Gruyter.

Nam, H., Goldstein, L., & Proctor, M. (2007). TADA (TAsk Dynamics Application). available for download `http://www.haskins.yale.edu/tada_download/`.

Nam, H., & Saltzman, E. (2003). A competitive, coupled oscillator model of syllable structure. In *Proceedings of the 15th international congress of phonetic sciences.* Barcelona.

Nearey, T., & Rochet, B. (1994). Effects of place of articulation and vowel context on VOT production and perception for French and English stops. *Journal of the International Phonetic Association*, *24* (01), 1–18.

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

Ohala, J. J., & Kawasaki-Fukumori, H. (1997). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In S. Eliasson & E. H. Jahr (Eds.), *Language and its ecology: essays in memory of Einar Haugen* (100, pp. 343–365). Trends in Linguistics. Studies and Monographs. Berlin: Mouton de Gruyter.

Oliveira, L., Yanagawa, M., Goldstein, L., & Chitoran, I. (2004). Towards standard measures of articulatory timing. *The Journal of the Acoustical Society of America*, *116* (4), 2643–2644. **retrieved from** `http://link.aip.org/link/?JAS/116/2643/4`

Perkell, J. S., Matthies, M. L., Tiede, M., Lane, H., Zandipour, M., Marrone, N., ...Guenther, F. H. (2004). The distinctness of speakers' /s/–/ʃ/ contrast is related to their auditory discrimination and use of an articulatory saturation effects. *Journal of Speech, Language and Hearing Research*, *47* (6), 1259–1269

Peters, B. (2006). *Form und Funktion prosodischer Grenzen im Gespräch [Form and function of prosodic boundaries in conversation].* (PhD dissertation, Christian-Albrechts-Universität, Kiel).

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology ii: gesture, segment, prosody* (pp. 90–127). Cambridge: Cambridge University Press.

Pike, K. (1945). *The intonation of American English.* Linguistics. Ann Arbor: University of Michigan.

Pouplier, M. (2011). Atoms of phonological representation. In M. van Oostendorp, C. J. Ewen, E. Hume & K. Rice (Eds.), *The Blackwell Companion to Phonology* (Chap. 17, Vol. 1). John Wiley & Sons (in press).

R Development Core Team. (2006). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna. **retrieved from** R Foundation for Statistical Computing: `http://www.R-project.org`

R Development Core Team. (2009). *R: a language and environment for statistical computing*. ISBN 3-900051-07-0. R Foundation for Statistical Computing. Vienna, Austria. **retrieved from** R Foundation for Statistical Computing: `http://www.R-project.org`

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

Recasens, D., & Pallarès, 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., & Pallarès, M. (2001). Coarticulation, blending and assimilation in Catalan consonant clusters. *Journal of Phonetics*, *29* (02), 273–201.

Recasens, D. (1999). Lingual coarticulation. In W. J. Hardcastle & N. Hewett (Eds.), *Coarticulation: Theory, Data and Techniques* (pp. 80–104). Cambridge: Cambridge University Press.

Recasens, D. (2004). The effect of syllable position on consonant reduction (evidence from Catalan consonant clusters). *Journal of Phonetics*, *32* (3), 435–453

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

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

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

Reubold, U., Harrington, J., & Kleber, F. (2010). Vocal aging effects on f0 and the first formant: a longitudinal analysis in adult speakers. *Speech Communication*, *In Press, Uncorrected Proof*, –

Rialland, A. (1994). The phonology and phonetics of extrasyllabicity in French. In P. Keating (Ed.), *Phonological structure and phonetic form* (pp. 136–159). Cambridge: Cambridge University Press.

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

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

Saltzman, E., & Kelso, J. S. (1987). Skilled actions: a task-dynamic approach. *Psychological Review*, *94*, 84–106.

Saltzman, E., Nam, H., Krivokapić, J., & Goldstein, L. (2008). A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. In *Proceedings of fourth conference on speech production* (pp. 175–184). Campinas.

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

Schiel, F. (1999, Aug.). Automatic phonetic transcription of non-prompted speech. In *Proc. 14. icphs* (pp. 607–610). San Francisco.

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

Shadle, C. H., & Scully, C. (1995). An articulatory-acoustic-aerodynamic analysis of [s] in vcv sequences. *Journal of Phonetics*, *23*, 53–66.

Sharpsteen, C., & Bracken, C. (2009). *Tikzdevice: A device for R graphics output in PGF/TikZ format.* R package version 0.4.8. **retrieved from** `http://CRAN.R-project.org/package=tikzDevice`

Shaw, J., Gafos, A., Hoole, P., & Zeroual, C. (2009). Temporal evidence for syllabic structure in Moroccan Arabic: theory, experiment, model. *Phonology*, *26* (To appear

2009 in special issue of Phonology on "'Relations between Phonological Models and Experimental Data"'), 187 –215.

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

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

Silverman, D., & Jun, J. (1994). Aerodynamic evidence for articulatory overlap in Korean. *Phonetica*, *51* (4), 210–220.

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

Stone, M., Faber, A., Raphael, L. J., & Shawker, T. H. (1992). Cross-sectional tongue shape and linguopalatal contact patterns in [s], [ʃ] and [l]. *Journal of Phonetics*, *20* (2), 253–270.

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

Turk, A. E., & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, *35* (4), 445–472.

Turk, A. E., & White, L. (1999). Structural influences on accentual lengthening in English. *Journal of Phonetics*, *27* (2), 171–206.

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.

Vitevitch, M., Armbruster, J., & Chu, S. (2004). Sub-lexical and lexical representations in speech production: effects of phonotactic probability and onsetdensity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *30* (2), 514–529.

Weismer, G. (1979). Sensitivity of voice-onset time (VOT) measures to certain segmental features in speech production. *Journal of Phonetics*, *7*, 197–204.

Weismer, G. (1980). Control of the voicing distinction for intervocalic stops and fricatives: some data and theoretical considerations. *Journal of Phonetics*, *8*, 427–438.

Wright, R. (2004). A review of perceptual cues and cue robustness. In B. Hayes, R. Kirchner & D. Steriade (Eds.), *Phonetically based phonology* (pp. 34–57). Cambridge: Cambridge University Press.

Zeroual, C., & Hoole, P. (2010). Mixed voicing and temporal overlap: EMA and aerodynamic observations from Moroccan Arabic. Poster presented at 12th Conference on Laboratory Phonology, Albuquerque.

Zsiga, E. C. (2000). Phonetic alignment constraints: consonant overlap and palatalization in English and Russian. *Journal of Phonetics*, *28* (1), 69–102.

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

Acknowledgements

Ich sage nur ein Wort: *Vielen Dank*

Horst Hrubesch

I thank Tine Mooshammer who managed to push me much, much further into this profession than I initially wanted. Her advice not only in scientific matters and her friendship have been indispenable to me.

Phil Hoole provided advice and help when I really needed them and guided me through this work with mostly slack reins. Thanks also for letting me go to all these cool conferences and for giving me a job. I am grateful to Jonathan Harrington for encouragement and strategic advice.

Quite a number of people were at one time or the other involved in prosodic, acoustic and articulatory labeling of the data analyzed here. For that I wish to express my gratitude to Elizabeth Heller, Yuuki Era, Angelika Berwein, Jennifer Schneeberg, Tamara Rathcke, Janine Lilienthal, Nina Redlingshöfer and Adriana Shehu. Most of the EPG data were collected at ZAS Berlin, thanks to Susanne Fuchs and Jörg Dreyer.

I am also grateful to Gilbert. I am not entirely sure for what but I'm open to suggestions.

I have mostly had great luck with my office mates. Thank you very much to Tina, to Katalin, Uwe, Maria Paola and Raphael, and to Sus'l for *happy times*. To complete this list I should definitely move in with Felicitas at some time. Anyway, I think I have made my point. I'm very grateful to all participants of the combined PhD and Post Doc meetings for valuable comments and suggestions.

Ich konnte mir der Unterstützung meiner Familie im Norden und im Westen immer sicher sein und bin dafür sehr dankbar.

I do not have the right words (nor do I think this is the right place) to thank my partner René Bombien in the way I feel would be appropriate for bearing with me through all the strain and stress that came along with the preparation of this work. I couldn't have done it without you and I'm glad you are still there.

I am privileged to have a fish tank located next to my desk at home. Thanks to all the fish and shrimp for every so often removing my attention from the computer screen to a colorful and lively display of ingestion and mating. Once in a while it is good to be reminded of what life really is about.

Lebenslauf

NAME:	Lasse Bombien
MATRIKELNR.:	10100765
GEBURTSDATUM:	08.12.1977
GEBURTSORT:	Kiel
FAMILIENSTAND:	verpartnert
KONFESSION:	evangelisch
WOHNORT:	Faberstraße 10 81373 München
SCHULAUSBILDUNG:	1984 – 1989 Grundschule Dänischenhagen 1989 – 1994 Ricarda-Huch-Schule Kiel (Gymnasium) 1994 – 1995 Menntaskóli við Hamrahlið Reykjavík/Island 1995 – 1998 Ricarda-Huch-Schule Kiel
SCHULABSCHLUSS:	Abitur (Note 2,4)
ZIVILDIENST:	1998 – 1999 Ambulante Altenpflege in Kiel
STUDIUM:	1999 – 2006 an der Christian-Albrechts-Universität zu Kiel
HAUPTFACH:	Phonetik und digitale Sprachverarbeitung
NEBENFÄCHER:	Nordische Philologie Informatik für Nebenfächler
STUDIENABSCHLUSS:	Magister (Note 1,4)
BERUF:	seit 06.2006 wissenschaftlicher Mitarbeiter am Institut für Phonetik und Sprachverarbeitung der Ludwig-Maximilians-Universität München

München, 5. April 2011