

Dissertationen der LMU

Band 2

Conceição Cunha

Die Organisation von Konsonantenclustern und CVC-Sequenzen in zwei portugiesischen Varietäten

Universitätsbibliothek der
Ludwig-Maximilians-Universität

Conceição Cunha

Die Organisation von Konsonantenclustern und CVC-Sequenzen
in zwei portugiesischen Varietäten

Dissertationen der LMU München

Band 2

Die Organisation von Konsonantenclustern und CVC-Sequenzen in zwei portugiesischen Varietäten

von

Maria da Conceição Gomes da Cunha

Herausgegeben von der
Universitätsbibliothek der Ludwig-Maximilians-Universität
Geschwister-Scholl-Platz 1
80539 München

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie. Detaillierte bibliografische Daten sind im Internet unter <http://dnb.dnb.de> abrufbar.

Verlagshaus Monsenstein und Vannerdat OHG Münster
www.mv-wissenschaft.com

© Autor Conceição Cunha 2015

Open-Access-Version dieser Publikation verfügbar unter:
<http://nbn-resolving.de/urn:nbn:de:bvb:19-183346>

ISBN: 978-3-95925-003-0 (Druckausgabe)

ISBN: 978-3-95925-004-7 (elektronische Version)

**Die Organisation von
Konsonantenclustern und CVC-
Sequenzen in zwei portugiesischen
Varietäten**

Inauguraldissertation
zur Erlangung des Doktorgrades der Philosophie an der
Ludwig-Maximilians-Universität München

vorgelegt von

Maria da Conceição Gomes da Cunha

aus Vila Nova de Gaia

2012

Erstgutachter: Prof. Dr. Jonathan Harrington
Zweitgutachter: Prof. Dr. Thomas Krefeld
Datum der mündlichen Prüfung: 02. 07. 2012

Inhaltsverzeichnis

Abbildungsverzeichnis	XIII
Tabellenverzeichnis	XIIIV
Abkürzungsverzeichnis	XVII
Vorwort.....	XIX
1 Einleitung.....	1
1.1 Gegenstand der Untersuchung	3
1.2 Gliederung der Arbeit.....	5
2 Die Entwicklung von Clustern und CVC-Sequenzen.....	7
2.1 Obstruent-Liquid-Cluster.....	10
2.1.1 Stabilität und Instabilität in der Entwicklung von tautosilbischen Clustern im Portugiesischen.....	10
2.1.2 Die Entwicklung der Obstruent-Liquid-Cluster im romanischen Sprachvergleich	20
2.1.3 Synchrone Variabilität bei Liquid-Clustern im Portugiesischen.....	23
2.1.4 Variation bei Obstruent-Liquid-Clustern im romanischen Vergleich	31
2.2 Reine Obstruentencluster im Portugiesischen	33
2.2.1 Die Entwicklung reiner Plosivcluster	34
2.2.2 Synchrone Variabilität bei Obstruentenclustern.....	36

2.2.3	Die theoretische Modellierung von Plosivclustern im Portugiesischen	37
3	Die Herausbildung des portugiesischen Vokalsystems	41
3.1	Betonte orale Vokale	41
3.2	Nasale Vokale	43
3.3	Unbetonte orale Vokale.....	44
3.4	Einfügung und Auslassung von Vokalen	45
3.4.1	Vokalepenthese oder Svarabhakti Vokal im Portugiesischen	45
3.4.2	Svarabhakti Vokale im romanischen Vergleich	46
3.4.3	Vokaltilgung im EP.....	47
3.4.4	Vokaltilgung in anderen Sprachen.....	48
3.4.5	Entstehung weiterer Konsonantencluster.....	49
3.4.6	Entstehung komplexer Konsonanten- gruppen.....	50
4	Konsonantenkoordination mit Liquiden	53
4.1	Einführender Forschungsstand.....	53
4.1.1	Konsonantencluster und Sonoritäts- hierarchien	54
4.1.2	Die dynamische Grundlage der Artikulatorischen Phonologie	57
4.1.3	Definition einer Geste	60
4.1.4	Koordination von CVC-Sequenzen und Clustern.....	62

4.1.5	Anpassungen in der gesprochenen Sprache	69
4.1.6	Fragestellung und Hypothesen	72
4.1.6.1	Stimulustyp: Cluster vs. $\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$	73
4.1.6.2	Sprechervarietäten: BP und EP	74
4.1.6.3	Artikulationsart: zentrale vs. laterale Liquide	75
4.1.6.4	Artikulationsort: bilabialer vs. velarer C_1	75
4.1.6.5	Betonung: betonte vs. unbetonte Kondition	77
4.2	Methode	79
4.2.1	Elektromagnetische Artikulographie	79
4.2.2	Versuchspersonen	81
4.2.3	Sprachmaterial	81
4.2.4	Akustische und artikulatorische Stimuli	83
4.2.4.1	Weiterverarbeitung der physiologischen Rohdaten	83
4.2.4.2	Artikulatorische Segmentierung	84
4.2.5	Statistische Auswertung	86
4.3	Die zeitliche intergestische Koordinierung	88
4.3.1	Definition der Variablen	88
4.3.2	Ergebnisse	91
4.4	Räumliche intragestische Koordination (<i>Phasing</i>)	98
4.4.1	Definition der Variablen	98
4.4.2	Ergebnisse	99

4.4.3	Korrelation beider Faktoren.....	104
4.4.4	Diskussion der wichtigsten Ergebnisse.....	105
4.5	Zusammenfassung und Ausblick.....	108
5	Die Koordination von Plosiven.....	113
5.1	Einführender Forschungsstand.....	113
5.1.1	Die Übertragung von Plosiven.....	115
5.1.2	Fragestellung und theoretischer Rahmen...117	
5.1.3	Hypothesen und Vorhersagen.....	119
5.1.3.1	Sprachvarietäten: BP und EP.....	120
5.1.3.2	Stimulustyp: Cluster vs. <C ₁ et> vs. <C ₁ ut>...121	
5.1.3.3	Artikulationsort: bilabialer vs. velarer C ₁	121
5.1.3.4	Betonung: betonte vs. unbetonte Kondition.....	122
5.2	Methode.....	123
5.2.1	Sprachmaterial.....	123
5.2.2	Statistische Auswertung.....	124
5.3	Die zeitliche intergestische Koordinierung.....	125
5.3.1	Definition der Variablen.....	125
5.3.2	Ergebnisse.....	126
5.3.3	Diskussion der wichtigsten Ergebnisse.....	130

5.4	Räumliche intragestische Koordination.....	133
5.4.1	Definition der Variablen.....	133
5.4.2	Ergebnisse.....	134
5.4.2.1	Korrelation beider Faktoren.....	137
5.4.2.2	Diskussion der wichtigsten Ergebnisse.....	139
5.5	Schlussfolgerungen und Ausblick	140
6	Die Perzeption von Clustern und CVC-Sequenzen.....	145
6.1	Einführender Forschungsstand	145
6.1.1	Perzeption in der Sprachwissenschaft	146
6.1.2	Aktueller Stand der Forschung.....	148
6.1.3	Kategoriale Wahrnehmung.....	151
6.2	Fragestellung.....	152
6.2.1	Identifikation des Stimulustyps.....	152
6.2.2	Wahrnehmung von Koordinations- unterschieden.....	152
6.2.3	Hypothesen und Vorhersagen	154
6.2.3.1	Sprechervarietät: BP vs. EP	154
6.2.3.2	Stimulustyp: Cluster vs. $\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$	155
6.2.3.3	Artikulationsort: bilabialer vs. velarer C_1	156
6.2.3.4	Betonung: betonte vs. unbetonte Kondition.....	157
6.2.3.5	Hörervarietät (BP vs. EP).....	158

6.3	Methode.....	159
6.3.1	Stimuli	159
6.3.2	Perzeptionsexperiment.....	161
6.3.3	Statistische Auswertung.....	163
6.4	Ergebnisse	164
6.4.1	Identifikation des Stimulustyps	164
6.4.1.1	Plosiv-Liquid-Token	165
6.4.1.2	Plosiv-Token	167
6.4.1.3	Zwischenfazit.....	169
6.5	Diskussion.....	184
6.5.1	Sprechervarietät: BP vs. EP	184
6.5.2	Stimulustyp: Cluster vs. <C ₁ eC ₂ > vs. <C ₁ uC ₂ >	185
6.5.3	Artikulationsort: bilabialer vs. velarer C ₁	186
6.5.4	Betonung: betonte vs. unbetonte Kondition.....	188
6.5.5	Hörervarietät: BP vs. EP	188
6.6	Schlussfolgerung und Ausblick.....	189

7	Conclusio	193
7.1	Cluster vs. CVC-Sequenzen.....	194
7.2	Varietätsspezifische Unterschiede	197
7.3	Artikulationsort und Produktionsrichtungseffekt.....	199
7.4	Betonung.....	201
7.5	Schlusswort	202
8	Literaturverzeichnis	205

Abbildungsverzeichnis

Abbildung 1: Skala der Sonorität / Konsonantischen Stärke (nach Restle & Venemann, 2001:1312).....	56
Abbildung 2: Markierungspunkte des Zungendorsums (TB) für den Konsonanten /k/ im Wort <i>crata</i> von S2	85
Abbildung 3: Schematische Darstellung der Plateau-Distanz ...	89
Abbildung 4: Schematische Darstellung der Überlappung / Synchronisierung der Gesten	89
Abbildung 5: Dauer des intermediären Abstands	91
Abbildung 6: Übersicht der Synchronisierung der Gesten nach Liquid-Gruppen und Varietäten getrennt.....	95
Abbildung 7: Maximale Geschwindigkeit des vertikalen Verlaufs der Schließungsgeste der Zungenspitze	100
Abbildung 8: Räumlicher Umfang der Zungenspitzenbewegung	102
Abbildung 9: Korrelation des Umfangs und der maximalen Geschwindigkeit der Zungenspitzen-geste	104
Abbildung 10: Dauer des intermediären Abstands.....	126
Abbildung 11: Übersicht der Synchronisierung der Gesten nach Liquid-Gruppen und Varietäten getrennt.....	129
Abbildung 12: Maximale Geschwindigkeit des vertikalen Verlaufs der Schließungsgeste der Zungenspitze (TTY)	134

Abbildung 13: Räumlicher Umfang der Zungenspitzenbewegung	136
Abbildung 14: Korrelation des Umfangs und der maximalen Geschwindigkeit der Zungenspitzenbewegung nach Stimulustyp	138
Abbildung 15: Schematische Darstellung der Hypothesen für den Zusammenhang zwischen Produktion und Perzeption von Clustern und CVC-Sequenzen	154
Abbildung 16: Übereinstimmung der entsprechenden Perzeptionsantworten und intendierten Produktionsstimuli bei Plosiv-Liquid-Token	165
Abbildung 17: Anteil der entsprechenden Perzeptionsantworten und intendierten Produktionsstimuli bei reinen Plosiv-Token	168
Abbildung 18: Antworten zu den produzierten Plosiv-Liquid-Clustern (anteilig).....	172
Abbildung 19: Anteilige Antworten zu den produzierten CVC-Sequenzen mit Liquid.....	175
Abbildung 20: Anteilige Antworten zu den intendierten reinen Plosivclustern.....	177
Abbildung 21: Anteilige Antworten zu den intendierten Plosiv-CVC-Sequenzen.....	182

Tabellenverzeichnis

Tabelle 1: Obstruent-r-Cluster im Klassischen Latein und gegenwärtigen Portugiesisch	11
Tabelle 2: Obstruent-Lateral-Cluster im Klassischen Latein und Portugiesischen	16
Tabelle 3: Obstruentencluster im Portugiesischen	33
Tabelle 4: Die betonten Vokale vom Klassischen Latein bis zur Iberoromania	42
Tabelle 5: Beispiele von komplexen Konsonantengruppen	51
Tabelle 6: Liste der aufgenommenen Stimuli mit den Zielclustern und Sequenzen	82
Tabelle 7: Liste der aufgenommenen Plosiv-Token mit den Zielclustern und Sequenzen	124
Tabelle 8: Liste der portugiesischen Zielwörter mit deutscher Übersetzung	160
Tabelle 9: Liste der Stimuli für das Perzeptionsexperiment ...	161

Abkürzungsverzeichnis

AltPT	Altportugiesisch
BP	Brasilianisches Portugiesisch
C	Konsonant
EP	Europäisches Portugiesisch
Fr.	Französisch
It.	Italienisch
Kat.	Katalanisch
KL	Klassisches Latein
Okz.	Okzitanisch
PT	Portugiesisch
Rum.	Rumänisch
Sp.	Spanisch
V	Vokal
VL	Vulgärlatein

Vorwort

Bei dieser Arbeit handelt es sich um meine für die Veröffentlichung überarbeitete Dissertation, die im Sommersemester 2012 an der Fakultät für Sprach- und Literaturwissenschaften der Ludwig-Maximilians-Universität München eingereicht und am 2. Juli 2012 erfolgreich verteidigt wurde. Die Dissertation entstand im Rahmen des *Linguistischen Internationalen Promotionsprogramms (LIPP)* am Institut für Phonetik und Sprachverarbeitung und am Institut für Romanische Philologie. Zum Gelingen der Dissertation haben zahlreiche Personen und Institutionen beigetragen, bei denen ich mich an dieser Stelle herzlich bedanken möchte.

Meinen Betreuern – Prof. Dr. Jonathan Harrington, Prof. Dr. Thomas Krefeld und PD. Dr. David Restle – gebührt mein besonderer Dank für die fortwährende Betreuung und die fachliche Unterstützung. Der regelmäßige Austausch in den Doktoranden-Kolloquien und ihre kontinuierliche Ansprechbarkeit bei Fragen haben mir ein ideales Umfeld geschaffen, in dem diese Arbeit entstehen konnte. Den beteiligten Hochschullehrern und den Mitpromovenden des *Linguistischen Internationalen Promotionsprogramms* danke ich für anregende Diskussionen in den Kolloquien und die schöne gemeinsame Zeit. Der Programmkoordinatorin Dr. Caroline Trautmann danke ich für optimale Rahmenbedingungen und Ihr offenes Ohr.

Dem *Zentrum für Sprach- und Literaturwissenschaft* und dem *Graduate Center* der Ludwig-Maximilians-Universität München danke ich für die finanzielle Unterstützung sowie dem *Linguistischen Internationalen Promotionsprogramms* für die Finanzierung des Forschungsaufenthalts in Campinas, Brasilien, diverse Tagungsreisen und die Organisation von Tagungen und Workshops.

Phil Hoole danke ich für die Skripten für die Nachverarbeitung der artikulatorischen Daten, Florian Schiel für die automatische Segmentierung der akustischen Daten und Christoph Draxler für das Aufsetzen des Perceptionsexperiments. Lasse Bombien danke ich für die Unterstützung mit EMU und den inspirierenden Austausch und Susanne Walzl für Ihre Hilfe mit den artikulatorischen EMA-Aufnahmen. Weiter möchte ich Katalin Mády und Uwe Reichel für die Aufnahme in das Büro, das entspannte Arbeitsklima und die vielen wertvollen Anregungen danken. Claudia Höhn danke ich für den letzten Schliff dieses Manuskriptes.

Großer Dank gebührt auch Patricia de Crignis, Sylvia Jaki, Tabea Reiner und Ursula Stangel für das erfolgreiche Mitwirken im Promotionsprogramm sowie im Sprecheramt, bei der Organisation mehrerer Tagungen, der Herausgabe des ersten LIPP-Journals, die geschätzte Freundschaft und nicht zuletzt den unermüdlichen Einsatz beim Korrekturlesen. Weiterhin danke ich meiner Familie und den vielen Freunden für ihre ununterbrochene Unterstützung.

Remus Gergel danke ich für die langjährige und unermüdliche Unterstützung und widme ihm diese Arbeit.

1 Einleitung

Nome: *Deixim Branco*

Ausschnitt aus einer Rechnung vom 23.09.2010 von der Postfiliale in der Avenida da República, Vila Nova de Gaia

Auf den ersten Blick sieht für einen Portugiesisch-Sprecher der Name auf der Rechnung sehr außergewöhnlich aus. Denkt man an die phonetische Lautkette [ˈdeʃiˈbr̃aku], bekommt man die gleiche Lautsequenz wie bei dem Ausdruck *deixe em branco*, wörtl. ‚Lassen Sie es leer / Lassen Sie es unausgefüllt‘. Erst dann wird einem das Missverständnis zwischen Sprecher und Hörer in der konkreten Situation klar: Der Sprecher wollte keinen Namen auf der Rechnung eingetragen haben, der Hörer hat es nicht verstanden und bildete die Lautkette um zu einer zu seiner Frage passenden Sequenz. D.h. die Lausequenz wurde als Vor- und Nachname wiedergegeben, obwohl ein solcher Vorname in der Sprache nicht vorhanden ist. Der Grund für das Missverständnis ist vermutlich, dass der Sprecher in dem ersten Wort ein Segment weggelassen hat /deʃ(i)ˈ/, ohne welches der Hörer das intendierte Wort nicht rekonstruieren konnte.

Durch solche Vokal-Auslassungen – die sogenannten Vokaltilgungen – stehen die benachbarten Konsonanten direkt nebeneinander. Dadurch können in der Sprache reine Konsonantenfolgen entstehen – Konsonantencluster – die abhängig von der Ausprägung der Vokaltilgung bezüglich der Anzahl der involvierten Elemente und der Komplexität steigen können. Solche Konsonantencluster aus Konsonant-Vokal-Konsonant (CVC) Verbindungen im Europäischen Portugiesisch (EP) werden in der vorliegenden Arbeit mit lexikalischen Konsonantenclustern verglichen und den gleichen Wörtern aus Brasilianischem Portugiesisch (BP) gegenübergestellt. Ziel der Arbeit ist die Produktion von Konsonanten im Rahmen der Artikulatorischen

Phonologie auf der Grundlage gegenwärtiger empirischer Daten zu analysieren und ihre Auswirkung auf die Perzeption darzustellen. Hierbei strebt sie an, zu einem besseren Verständnis der Verwechselbarkeit von Clustern und CVC-Sequenzen aus der Interaktion zwischen Produktion und Perzeption beitragen zu können.

Das besondere Augenmerk liegt dabei auf zwei Besonderheiten dieser Sprache aus synchroner und diachroner Sicht: Zum einen findet man eine innerhalb der Romanistik außergewöhnlichen Verwechselbarkeit bei der Realisierung von Konsonantenclustern und CVC-Sequenzen portugiesischer Varietäten. Dies liegt daran, dass Konsonantencluster in beiden Varietäten mit einem intermediären vokalischen Element realisiert werden können ([pɪlãtɐ] *planta* ‘plant’ im EP (Veloso 2007a) und [kapitar] *captar* ‘fangen’ im BP (Bisol 1999, Mateus and d’Andrade 2000)), aber Vokalelisionen in den gleichen Varietäten für die Tilgung mancher unbetonten Vokale sorgen und somit für die Entstehung von phonetischen Clustern aus lexikalischen CVC-Sequenzen (z.B. [flisiˈdad] *felicidade* ‘Fröhlichkeit’ und [ˈfrido] *ferido* ‘verletzt’ im EP und z. B. [aˈbɔbra] anstatt /aˈbɔbɔra/ *abóbora* ‘Kürbis’ im BP (Bisol 2000, 2010). Vokalepenthese sind im BP verbreiteter und Vokaltilgungen treten im EP häufiger auf, so dass im BP beide Formen eher mit einem intervokalisches Element realisiert werden können und somit [apiˈtar] sowohl einen Ausschnitt aus dem Wort *captar* ‘fangen’ als auch *tapetar* ‘Teppich legen’ darstellen kann. Im EP sorgen häufige Vokaltilgungen dafür, dass Konsonantencluster in der gesprochenen Sprache in Anzahl und Komplexität massiv zunehmen (Mateus and d’Andrade, 2000:42-44) und manche Wortpaare wie *crer* ‘glauben’ und *querer* ‘wollen’ homophon zu [ˈkɾɛf] wurden (Mateus & Martins 1982). Zum anderen ging die diachronische Entwicklung der Cluster vom Klassischen Latein bis hin zum Portugiesischen bereits einen ungewöhnlichen Weg mit absolut stabilen Obstruent-r-Clustern und instabilen Clustern mit Lateralen, die mal zu Affrikaten mal zu zentralen Liquiden wurden.

1.1 Gegenstand der Untersuchung

Aus dem Literaturüberblick in den folgenden zwei Kapiteln lässt sich entnehmen, dass lexikalische Cluster als reine Konsonantensequenzen oder mit einem (epenthetischen/svarabhahatischen) intermediären Vokal beschrieben werden und lexikalische CVC-Sequenzen durch Vokaltilgungen ihren intermediären Vokal verlieren können, so dass sie mit oder ohne lexikalischen Vokal auftreten können. Aus der historischen Entwicklung geht hervor, dass Obstruent-l-Cluster und reine Plosivcluster aus dem Klassischen Latein mehreren Lautwandelprozessen unterlagen, während Obstruent-r-Cluster eine stabile Entwicklung aufwiesen. Außerdem bestand die Grundlage der aktuellen portugiesischen Cluster sowohl aus lateinischen Clustern als auch aus den entsprechenden CVC-Sequenzen nach einer Vokaltilgung im Klassischen Latein, die meistens eine parallele Entwicklung aufwiesen. So betrachtet zeigt sich die Unterscheidung zwischen lexikalischen Clustern und CVC-Sequenzen auch in ihrer historischen Entwicklung als nicht stabil.

Daraus entstehen die zentralen Fragen dieser Arbeit und so strebt die vorliegende Arbeit an, anhand von akustischen und physiologischen Daten zu untersuchen, inwiefern zwei Konsonanten in der Produktion der Sprache eine unterschiedliche Organisation bzw. Strukturierung aufweisen, je nachdem, ob sie als Abfolge von Konsonanten oder mit einem intermediären Vokal vorkommen, wie sich die Koordination in der Produktion auf die Perzeption in einem Sprachsystem auswirkt und wie sich die Interaktion zwischen Produktion und Perzeption in der Sprachentwicklung zum Ausdruck bringen lässt.

Methodisch beschäftigt sich die vorliegende Arbeit im Rahmen der Artikulatorischen Phonologie (Browman & Goldstein 1986, 1988) mit der Koordination von Konsonanten in reinen Konsonantenclustern und CVC-Sequenzen in zwei portugiesischen Varietäten und mit ihrer Auswirkung auf die

Perzeption der Sprache. In dieser Orientierung schlieÙe ich mich der perzeptiver Linguistik von Krefeld & Pustka (2010:9) an, für die Produktion und Perzeption „wie zwei Seiten derselben kommunikativen Medaille“ empirisch und theoretisch jede linguistische Analyse bereichern und längst nicht ausschließlich phonetische Bereiche sind.

Das ausgewählte Sprachsystem war aufgrund seiner Besonderheiten sowohl synchronisch als auch diachronisch betrachtet das Portugiesische. Synchron weisen portugiesische Varietäten optionale Neutralisierung sowohl von lexikalischen Konsonantenclustern zu Sequenzen mit intermediären Vokalen durch Epenthese (CC zu CeC, vor allem im BP, Bisol 1999) als auch von lexikalischen CVC-Sequenzen zu Konsonantenfolgen durch Vokaltilgung (CVC zu CC, vor allem im EP, Vigário 2003, Veloso 2007a) auf. In der diachronischen Entwicklung aus dem Lateinischen unterlag das Galicisch-Portugiesische einer Vielzahl von Lautwandelprozessen, die vor allem den Konsonantismus betrafen und dieser Varietät eine besondere Färbung innerhalb des romanischen Kontextes verliehen. Im 17. Jahrhundert fand in der europäischen Varietät eine starke rhythmische Umstrukturierung statt, die Auswirkungen auf den Stellungswechsel der Klitika (zur Enklisis im EP), auf die Erhöhung der unbetonten Vokale im EP und auf den gesamten rhythmischen Eindruck beider Varietäten hatte. Somit wurde der Grundstein für die Ausdifferenzierung beider Varietäten gelegt (Frota et al. 2008).

Konkret untersucht die Arbeit für eine Auswahl von lexikalischen Clustern (bestehend aus Plosiv und Liquid oder aus zwei Plosiven) und entsprechenden lexikalischen CVC-Sequenzen mit intermediären <e> und <u>, inwiefern sie sich in der Sprachproduktion unterscheiden und diese Unterschiede eine Auswirkung auf die Perzeption der Sprache hervorrufen. Dabei wird die Rolle von Faktoren wie Sprechervarietät (EP vs. BP), C₁-Artikulationsort (labial vs. velar), Betonung, Artikulationsart bei den Clustern mit Liquiden (lateral vs. zentral) und Hörervarietät bei der Perzeption mit einbezogen.

1.2 Gliederung der Arbeit

Die Arbeit besteht aus insgesamt sieben Kapiteln: Nach der Skizzierung der historischen Entwicklung von Konsonantenclustern im 2. Kapitel, folgt im Kapitel 3 die Herausbildung des Vokalsystems sowie vokalische Prozesse wie Epenthesen und Tilgungen mit einer groben Situierung des Portugiesischen im romanischen Kontext.

Im 4. Kapitel werden nach einer Einführung in den theoretischen Rahmen der Artikulatorische Phonologie und der ausführlichen Vorstellung der Methode die Ergebnisse der physiologischen Analysen zur Koordination von Konsonanten präsentiert und diskutiert. Es werden dabei beide Konsonantengesten (intergestische Koordination) in Obstruent-Liquid Clustern und in den entsprechenden CVC-Sequenzen analysiert sowie der Verlauf und die maximale Geschwindigkeit der apikalen Geste der Zungenspitze für die Produktion der Liquide (intra-gestische Koordination) in zwei Portugiesisch-Varietäten näher beleuchtet.

Die Analyse medialer reiner Plosivclustern und entsprechender CVC-Sequenzen sind das Untersuchungsobjekt des 5. Kapitels mit der gleichen Methode wie im Kapitel zuvor nach einer knappen Skizze der Besonderheiten dieser Gruppen in der Produktion.

Im 6. Kapitel wird eine Perzeptionsstudie mithilfe der zuvor gewonnen Daten der beiden Portugiesisch-Varietäten aus den vorherigen Kapiteln diskutiert. Dabei wird untersucht, ob lexikalische Cluster und lexikalische CVC-Sequenzen in Bezug auf möglicherweise neutralisierte Formen in der Perzeption identifiziert werden können und inwiefern dabei der

Zusammenhang zwischen intendierten¹ und wahrgenommenen Token sich durch Koordinationsunterschiede erklären lässt.

Es folgt eine abschließende *Conclusio* bestehend aus einer Zusammenfassung der Ergebnisse, ihrer Positionierung in Bezug auf die Entwicklung der Sprache und einem Ausblick auf zukünftige Anwendungs- und Erweiterungsmöglichkeiten der aktuellen Forschungsergebnisse.

¹ Der Begriff ‘intendiert’ bezieht sich in dieser Arbeit nicht auf kognitive Prozesse beim Sprechen oder darauf, was der Sprecher sagen wollte, sondern bezeichnet einfach das Token, das die Probanden laut Stimuluspräsentation produzieren sollten.

2 Die Entwicklung von Clustern und CVC-Sequenzen

Das anstehende Kapitel stellt die Entwicklungsgeschichte der lateinischen Formen bis hin zum Portugiesischen im romanischen Vergleich dar, bei der die mündliche Übertragung eine zentrale Rolle für die Herausbildung der einzelnen Sprachen und ihrer Sprachsysteme spielte. Im Anschluss darauf folgt der Überlieferungsweg der portugiesischen Cluster und unbetonten Vokale, die bei CVC-Sequenzen Variation hervorrufen, dem aktuellen Forschungsstand gegenüber. Ziel dieses Kapitels ist hierbei herauszufinden, inwiefern Variation zwischen Clustern und CVC-Sequenzen in der Entwicklung bereits vorhanden war und wie aktuelle Studien diese umreißen.

Die Überlieferung des lateinischen Wortschatzes bis hin zu den verschiedenen romanischen Sprachen fand innerhalb mehrerer Zyklen statt, weshalb sich ein lateinischer Laut nicht selten zu unterschiedlichen Lauten in der Romania entwickelte. So wurden beispielsweise intervokalische nasale und laterale Konsonanten in den meisten romanischen Sprachen beibehalten, im Galicisch-Portugiesischen² jedoch zwischen dem 8. und dem 10. Jahrhundert getilgt und das intervokalische (-l-) wurde im Rumänischen zu /r/ rhotarisiert (Müller 2011: 134). So wurde aus dem Klassisch Lateinischen (KL) *mōla* im Rumänischen *moară*, im Italienischen, Spanischen, Katalanischen und im Okzitanischen

² Galicisch-Portugiesisch oder Galicisch war die Sprache des Nordwestens der Iberischen Halbinsel. Historisch gesehen ist Portugiesisch eine Ausweitung der galicischen Sprache im Nordwesten, die nach einer ausgiebigen dichterischen Produktivität im 13. Jahrhundert auf beiden Seiten der politischen Grenze in der zweiten Hälfte des 14. Jahrhunderts begann, sich auszudifferenzieren. Die sprachlichen Gemeinsamkeiten des Nordportugiesischen und Galicischen wurden durch den Einfluss der Sprache der später eroberten Regionen im Süden etwas geschwächt (Jensen 1999: 10, Castro 1991, Williams 1961).

mola, im Französischen *meule* und im Portugiesischen *mó*. Ebenso wurde aus dem KL *bōna* im Rumänischen *bună*, im Italienischen *buona*, im Spanischen *buena*, im Katalanischen und im Okzitanischen *bona*, im Französischen *bonne* und im Portugiesischen *boa* (Jensen 1999: 203-205, Koike & Macedo 1992). Gegenwärtige intervokalische Laterale im Rumänischen, Galicischen³ und Portugiesischen und intervokalische Nasale im Galicischen und Portugiesischen wurden zu einem späteren Zeitpunkt entlehnt.

Um die Herausbildung des Wortschatzes in einem Sprachsystem zu verfolgen und seine unterschiedlichen Überlieferungsformen zu präzisieren, unterscheidet man in der historischen Übertragung zwischen Erbwörtern (*popular words*), die mündlich von einer Generation zur nächsten übertragen wurden und dabei einer Reihe von regulären phonologischen Prozessen unterlagen, und gehobenen, bildungssprachlichen Wörtern oder Latinismen (*learned words* oder *latinisms*), die keine ununterbrochene Übertragung durchliefen, sondern aus dem Klassischen Latein importiert wurden, um lexikalische Lücken im technischen oder medizinischen Bereich zu füllen. Aufgrund dieser Übertragungslücke blieben die gehobenen Wörter in den meisten romanischen Sprachen sehr nah an der ursprünglichen Form des Klassischen Lateins und sind daher nicht bedeutsam für die Herausbildung der einzelnen Sprachsysteme (Castro 1991: 90-126, Jensen 1999: 46-47, Williams 1961: 28-29). Dies ist daran erkennbar, dass die gehobenen Wörter abweichende marginale Lösungen für die lautliche Entwicklung aufweisen, die wenig über das Sprachsystem aussagen.

In jeder romanischen Sprache findet man nicht selten im aktuellen Lexikon unterschiedliche Formen für verwandte Wörter. Der Grund dafür liegt genau an dem Zeitpunkt der Überlieferung und die Anwendung damit assoziierter phonologischer Prozesse. So ist das lateinische Substantiv *ōcŭlu* ein Erbwort im Romanischen und wurde im Rumänischen zu *ochi*, im Italienischen zu *occhio*, im

³ Zur Schreibweise siehe Kabatek (1996).

Spanischen zu *ojo*, im Katalanischen zu *ull*, im Französischen zu *oeil* und im Portugiesischen zu *olho*, während das entsprechende Adjektiv *ōcūlāre* zu den gehobenen Latinismen gehört und in allen romanischen Sprachen nahe an der ursprünglichen Form blieb: im Rumänischen *ocular*, im Italienischen *oculare*, im Spanischen, Katalanischen und Portugiesischen *ocular*, im Französischen *oculaire* (Jensen 1999: 46-47). Die Alternanz zwischen Erbwörtern und gehobenen Latinismen hat dazu beigetragen, dass im aktuellen Lexikon mitunter zwei verschiedene Lexeme mit demselben lateinischen Stamm auftreten: Aus KL *hōspītāle* wurde im Französischen, Spanischen und Portugiesischen sowohl *hostel/hotel* (bzw. *hôtel*) 'Hotel' wie *hospital* (bzw. *hôpital*) 'Krankenhaus'. Das KL *pēnsāre* wurde im Spanischen Portugiesischen zu *pesar* 'wiegen' und *pensar* 'denken'.

Als Mischform zwischen Erbwörtern und gehobenen Wörtern definiert man außerdem die sogenannten halbgehobenen Wörter (*semi-learned words*). Diese wurden relativ früh zur jeweiligen Sprache hinzugefügt, so dass sie manche der phonologischen Prozesse durchliefen und gleichzeitig durch die unterbrochene Überlieferung andere wiederum nicht und somit Eigenschaften der gehobenen Wörter aufweisen. Ein Beispiel dafür ist Lat. *saccūlu* > Fr. *siècle*, das einerseits die für Erbwörter typische Diphthongierung des betonten Vokals und andererseits das gehobene Cluster /kl/ enthält (Jensen 1999: 47, Wireback 1997).

Zusammenfassend gesagt gab es in der Herausbildung der aktuellen romanischen Sprachen eine unterschiedliche Prägung des zugrunde liegenden lateinischen Wortschatzes, je nachdem ob die Wörter mündlich überliefert wurden oder die Überlieferung unterbrochen wurde. Die mündlich überlieferten Formen gehören zum Erbwortschatz und nur dieser war für die Konstituierung einer Sprache ausschlaggebend. Daher sind Erbwörter für die phonologische Beschreibung der Sprache besonders wichtig (Castro 1991, Jensen 1999, Koike & Macedo 1992, Mattos & Silva 1991, 1994). Die gegenwärtige Variation bei portugiesischen Clustern wird mit der späten Überlieferung begründet: D.h. die

Variation würde ausschließlich später etablierte Cluster (Obstruent-l- und Obstruentencluster) treffen, die im Grunde nicht als lexikalische Cluster des Portugiesischen angesehen werden (Velooso 2006, 2007b, Mateus et al. 2005, Mateus & d'Andrade 2000, 1998).

2.1 Obstruent-Liquid-Cluster

Das Hauptinteresse an den portugiesischen Obstruent-Liquid-Clustern liegt in der relativ stabilen, diachronischen Entwicklung der Cluster mit R-Lauten im Vergleich zu der besonders instabilen Clustern mit Lateralen. Im romanischen Vergleich zeigen Letztere im Portugiesischen eine einzigartige Entwicklung mit mehreren Lautwandeln (Velooso 2007a: 122, Mateus & d'Andrade, 2000: 40). Dies veranlasst Velooso (2006: 130, 2007: 127) zu der These, dass l-Cluster keine tauto-⁴, sondern heterosilbische Cluster seien. Ansonsten herrscht Einstimmigkeit, dass alle Obstruenten-Liquid-Cluster aufgrund ihrer steigenden Sonorität zum Nukleus einer einzigen Silbe gehören (Mateus & d'Andrade 2000: 40, 1998). Die Frage nach der Zugehörigkeit zu einer oder zwei Silben wird in der vorliegenden Arbeit lediglich am Rande gestellt, da ihr Schwerpunkt auf dem Zusammenhang zwischen unterschiedlich stabiler Entwicklung und der synchronen Variabilität von Clustern vs. CVC-Sequenzen liegt.

2.1.1 Stabilität und Instabilität in der Entwicklung von tautosilbischen Clustern im Portugiesischen

Im Klassischen Latein waren Cluster mit jedem beliebigen Plosiv und den labio-dentalen Frikativen /f, v/ als erstem Element und zentralen oder lateralen Liquididen als zweitem Element in wortinitialer und -medialer Position möglich. In der Übertragung bis hin zum Portugiesischen wurde bei der Obstruent-r-Gruppe – parallel zu den einfachen Plosiven (KL *lupus* > Port. *lobo*) – der

⁴ Tautosilbische Cluster werden in einer Silbe analysiert (Brot), während heterosilbische Cluster durch eine Silbengrenze in zwei Silben aufgesplittet werden (sin.gen, vgl. z.B. Selkirk (1984) *homorganic vs. heterorganic clusters*).

mediale stimmlose Plosiv stimmhaft. Abgesehen davon wurde diese Gruppe über die verschiedenen Zyklen in das portugiesische Lexikon übernommen (Velooso 2007a: 122, Mateus & Andrade 2000: 40).

Position	Cluster	KL	Port.	D-Glosse
initial	[br]	bracchium	<i>braço</i>	‘Arm’
	[dr]	drăcōnis ⁵	<i>dragão</i>	‘Drache’
	[gr]	grātuitus	<i>gratuito</i>	‘gratis’
	[fr]	frăctūra	<i>fractura</i>	‘Fraktur’
	[pr]	pretium	<i>preço</i>	‘Preis’
	[tr]	tremere	<i>tremar</i>	‘zittern’
	[kr]	crēdere	<i>crer</i>	‘glauben’
medial	[br]	fībra	<i>fibra</i>	‘Faser’
	[dr]	cyllīndrus	<i>cilindro</i>	‘Zylinder’
	[gr]	sēgregāre	<i>segregar</i>	‘ausschließen’
	[fr]	īnfrā	<i>infra-</i>	‘unten’
medial	[pr]	caprā capricōrnus	<i>cabra</i> <i>capricórnio</i>	‘Ziege’ ‘Steinbock’
	[tr]	petra astrum	<i>pedra</i> <i>astro</i>	‘Stein’ ‘Stern’
	[kr]	lacrima lucrum	<i>lágrima</i> <i>lucro</i>	‘Träne’ ‘Gewinn’

Tabelle 1: Obstruent-r-Cluster im Klassischen Latein und gegenwärtigen Portugiesisch (mit Beispielen aus Jensen 1999: 175-181, Boyd-Bowman 1954: 110-111 und eigenen Ergänzungen⁶).

⁵ Das portugiesische Wort stammt aus dem Genitiv *drăcōnis* und nicht aus dem Nominativ *drăcō*; [dr] ist ein griechisches Cluster, das allerdings im KL in vielen Lehnwörtern vorhanden war (Schmeiser 2009: 499).

So gelangten Obstruent-r-Cluster ziemlich früh in den Erbwortschatz des späteren Galicisch-Portugiesischen und wurden in den weiteren Zyklen ohne Änderungen immer zahlreicher. In der tabellarischen Übersicht werden die einzelnen Cluster anhand von Beispielen dargestellt.

Weitere Obstruent-r-Cluster entstanden aus lateinischen CVC-Sequenzen, die nach Tilgung unbetonter Vokale zu reinen Konsonantenclustern wurden. Dies betraf einige prätonische⁷ /e/ des Vulgärlatein, die ursprünglich aus dem Klassischen Latein stammenden Vokale ē, ĭ, ĕ, wie an den folgenden Beispielen zu sehen: KL *bēriillum* > Port. *brilho*, KL *quĩritāre* > Port. *gritar*, KL *sũpĕrāre* > Port. *sobrar*. Diese Vokaltilgung wird so früh datiert, dass die entstandenen Wörter dem Erbwortschatz angehören. Parallel zu den klassisch lateinischen Clustern wurden bei diesen Clustern ebenso mediale stimmlose Plosive stimmhaft (*apĕrĩre* > Port. *abrir*, KL *lĕpĕre* > Port. *lebre*, parallel zu KL *capra* > Port. *cabra*, Boyd-Bowman 1954: 122-123, Williams 1961: 56-57, 224-233).

Im Gegensatz zur relativ stabilen Entwicklung der Obstruent-r-Cluster und trotz der vielen akustischen und artikulatorischen Ähnlichkeiten zwischen zentralen und lateralen Liquiden (Proctor 2009: 5-16) unterlagen Obstruent-l-Cluster in ihrer Entwicklung bis hin zum Portugiesischen einer Reihe von Lautwandelprozessen sowohl in wortinitialer als auch in -medialer Position (Jensen 1999: 174-181, Castro 1991).

Im Erbwortschatz wurden stimmlose Obstruenten plus /l/ in reinen Clustern zu stimmlosen Affrikaten [tʃ] palatalisiert und später in der Entwicklung des Galicisch-Portugiesischen zu stimmlosen Frikativen [ʃ] vereinfacht, d.h. KL [pl, fl, kl] > [tʃ] > [ʃ]. So wurde beispielsweise KL *flamma* > Port. *chama* und

⁶ Die Beispiele wurden mithilfe der *Dicionários Porto Editora* (2007, 2008) vereinheitlicht und ergänzt. Die deutschen Übersetzungen folgen den portugiesischen und nicht den lateinischen Wörtern.

⁷ Prätonisch (pretonic): vor der hauptbetonten Silbe.

KL *plumbum* > Port. *chumbo*. Dieser Erbwortschatz wurde durch die Aufnahme halbgehobener Wörter erweitert, bei denen der laterale Liquid /l/ zu einem zentralisierten /r/-Laut rhotarisiert wurde (meistens ein Tap [r] im Portugiesischen, KL *plaga* > Port. *praga*). Durch die Aufnahme von gehobenen Wörtern mit dem lateralen Cluster (KL *plātanus* > Port. *plātano*) ab dem 14. und 15. Jahrhundert in das portugiesische Lexikon erweitern sich die resultierenden Formen in wortinitialer Position auf drei (/ʃ/, /Cr/, und /Cl/, wobei C=Obstruent), die im gegenwärtigen Lexikon koexistieren. Initiale Obstruent-l-Cluster mit dem gleichen Artikulationsort (/tl/, /dl/) waren im Klassischen Latein nicht vorhanden.

Nach Tilgung kurzer Vokale entstand zu einem frühen Zeitpunkt des Lateinischen eine Reihe von medialen Clustern: Das mediale Cluster /kl/ muss so früh entstanden sein, dass es noch dem lateinischen Primär-Wortschatz zugeordnet wird (Jensen 1999: 240, Menéndez-Pidal 1926: 152). Parallel zum initialen /kl/ unterlag mediales /kl/ ebenso einer Palatalisierung, dessen Ergebnis allerdings [ʎ] war, wie an den folgenden Beispielen ersichtlich: KL *ōcūlum* > Port. *olho* ['oʎu]⁸ 'Auge', KL *genūcūlum* > Port. *joelho* 'Knie', KL *aurīcūla* > (VL *orīcla* >) Port. *orelha* [o'reʎə] 'Ohr'.

Wurden unbetonte Vokale getilgt, entstanden überdies mediale /tl/-, /bl/- und /gl/-Cluster. /tl/ entstand früh im Vulgärlatein, wurde allerdings bereits im Laufe des Vulgärlateins durch /kl/ ersetzt, so dass es nie zu einem gängigen Cluster im Lateinischen wurde. So alternierten bereits im Vulgärlatein solche Dubletten wie *vetulus non veclus*, *vitulus non viclus* (vgl. Jensen 1999: 240-241). Bei /bl/ erfuhren mediale lexikalische Cluster aus dem Klassischen Latein und die im Vulgärlatein durch Vokaltilgung

⁸ Die Transkribierung erfolgte nach den Konventionen für die breite Transkription für EP (Mateus & d'Andrade 1998), die sich vor allem durch die Vokalqualität vom BP unterscheidet. Da die Vokale an dieser Stelle nicht relevant sind, werden die varietätsspezifischen Unterschiede vorerst ausgeklammert.

entstandenen Cluster eine parallele Entwicklung: KL *ōbligare* > Port. *obrigar*, KL *nōbīle* > Port. *nobre*, KL *parābōla* > Port. *palabra*⁹). Eine andere Entwicklungsmöglichkeit bei den durch Vokaltilgung entstandenen Clustern war die Tilgung ihres ersten Elements: KL *fabūlāre* (nach Vokaltilgung: *fablar*) > Port. *falar* (Jensen 1999: 236-237).

Die Entwicklung von /gl/ fand so langsam statt, dass sich in manchen Wörtern nie ein Cluster entwickelte. Man schreibt diesen Formen eine langsame Entwicklung zu, weil sie zu alt sind, um als gehobene Wörter eingestuft zu werden (Jensen 1999: 242-243, Koike & Macedo 1992). Es folgt eine tabellarische Übersicht der Entwicklungsformen der Cluster mit Lateralen, um die durch mehrere Lautwandel entstandene Variabilität zu veranschaulichen.

Position	Cluster	KL	Port.	D-Glosse
initial	[bl]	blandu blasphēmia	<i>brando</i> <i>blasfēmia</i>	‘schwach’ ‘Blasphemie’
	[dl]	-		
	[gl]	glans globus	<i>glande</i> <i>globo</i>	‘Eichel’ ‘Kugel’
	[fl]	flamma flaccus flūctuāre	<i>chama</i> <i>fraco</i> <i>flutuar</i>	‘Flamme’ ‘schlapp’ ‘Wellen schlagen’

⁹ In der Weiterentwicklung wurden im Portugiesischen manche /b/ zu /v/, so dass das Wort aktuell *palavra* geschrieben wird (Jensen 1999: 237).

	[pl]	plumbum plaga plātānus	<i>chumbo</i> <i>praga</i> <i>plātano</i>	‘Blei’ ‘Schlag’ ‘Platane’
	[tl]	-		-
	[kl]	aclāmāre clāvus clīnicus	<i>chamar</i> <i>cravo</i> <i>clínico</i>	‘rufen’ ‘Nagel’ ‘Arzt’
medial	[bl]	öbligare	<i>obrigar</i>	‘zwingen’
		trībūlāre	<i>‘trilbar’</i>	‘Pfad’
	[gl]	fabūlāre publicus coagūlāre ¹⁰ agglomerare	<i>falar</i> <i>público</i> <i>coalbar</i> <i>aglomerar</i>	‘sprechen’ ‘öffentlich’ ‘gerinnen’ ‘anhäufen’
	[fl]	afflāre sūfflāre inflammāre	<i>achar</i> <i>soprar</i> <i>enflamear</i>	‘finden’ ‘blasen’ ‘entflammen’

¹⁰ Alle diese CVC-Sequenzen des Klassischen Lateins, die früh Vokaltilgung unterlagen und sich parallel zu den Clustern entwickelten, sind ebenso im gegenwärtigen Portugiesisch in gelehrten Wörtern vorhanden (KL *coagūlāre* > Port. *coalbar*, *coagular*; KL *aurīcūla* > Port. *orelba*, *aurícula*, KL *ocūlo* > Port. *olho*, *óculo*).

[pl]	dūplu	<i>dobro</i>	‘doppelt’
	templum	<i>templo</i>	‘Tempel’
[tl]	vētūlus	<i>velho</i>	‘alt’
	tītūlu	<i>título</i>	‘Titel’
[kl]	aurīcūla	<i>orelha</i>	‘Ohr’
	ecclēsiā	<i>igreja</i>	‘Kirche’
	acclāmātiō	<i>aclamação</i>	‘Beifall’

Tabelle 2: Obstruent-Lateral-Cluster im Klassischen Latein und Portugiesischen (angelehnt an Schmeiser, 2009a: 499 mit ergänzenden Beispielen aus Jensen 1999: 175-181, Boyd-Bowman 1954: 110-111 und eigenen Ergänzungen¹¹).

Die große Variabilität der Obstruent-l-Cluster wurde ausführlich beschrieben, ohne dass sich eine Erklärung hierfür etablierte (Castro 1991, Mattos & Silva 1991, Schmeiser 2009a). Für den Lautwandel der Laterale zu zentralen Liquiden macht Vennemann (1993: 326-330, 1989: 17-21, auch Restle & Vennemann 2001: 1315-1316) die allgemeine Tendenz zur zunehmenden Kopfqualität der reinen Konsonantenfolgen verantwortlich. Die Kopfqualität steigt der Reihe nach von den Clustern mit Nasal zu denen mit Lateral und r-Lauten (/KN/ → /Kl/ → /Kr/). Aus diesem Grund hatte der Wandel von /l/ zu /r/ nach Obstruent eine Steigerung der Kopfqualität zur Folge und somit eine Verbesserung des Systems zu einer stabilen Zielgruppe (Obstruent-r-Cluster):

Das Ziel dieser Entwicklung war das folgende verbesserte System [...] Köpfe aus Obstruent plus r sind bisher überall stabil; sie sind ja zum Teil sogar das Ziel des Abbaus der l-Gruppen. (Restle & Vennemann 2001:1316)

Wie Schmeiser (2009a: 504) bereits nahelegte, ist die weitere Analyse anhand des distinktiven Merkmals [+/-kontinuierlich]

¹¹ Die Beispiele wurden mithilfe der Dicionários Porto Editora (2007, 2008) vereinheitlicht und ergänzt.

von Mateus & d'Andrade (2000), Mateus (1990), u.a. auch in der Lage, Rhotazismus zu erklären. Nach dieser Analyse wurde bei frühen Clustern der Lateral [-kontinuierlich] und unterlag aufgrund einer Restriktion von nicht kontinuierlichen Lauten in tautosilbischen Clustern einem lateralen Rhotazismus zu /r/. Mit einem gestischen Ansatz bietet Schmeiser (2009a: 504) eine Beschreibung von Palatalisierung, Rhotazismus und Erhaltung des Clusters innerhalb desselben Modells. Dabei wird die Palatalisierung durch die starke Überlappung beider Konsonantengesten erklärt:

[...] patrimonial transmission results in palatalization. In gestural terms, the two consonants overlap so much that neutralization of the gesture occurs [...] (Schmeiser 2009a: 506).

Es ist allerdings schwer nachzuvollziehen, wie eine starke Überlappung der Konsonantengesten allein den Wandel von einem Obstruent-l-Cluster [pl], [kl] und [fl] zu einer stimmlosen Affrikate [tʃ] erklären soll. Wie kann zeitliche Überlappung alleine die Änderung der Artikulationsart von Lateral zu dem frikativen Teil der Affrikate erklären? Der Rhotazismus des Laterals zu einem Tap in halbgehobenen Clustern soll aufgrund geringerer Überlappung beider Konsonantengesten zustande gekommen sein:

In the case of semi-learned transmissions, the cluster in G-P almost always undergoes rhotacization of the lateral. In gestural terms, the two gestures are restricted to intersecting within the imposed Phase Window (Schmeiser 2009a: 506).

Sollte sich die Koordinierung beider Konsonanten zueinander lediglich an einer weiteren Koordinierung der r-Cluster im Vergleich zu den l-Cluster unterscheiden, wäre der logischere Lautwandel derjenige hin zur entsprechenden CVC-Sequenz mit Lateral. Wie Browman & Goldstein (1990a, b) für das Wortpaar [bɹaɪ] vs. [b(ə)ɹaɪ] (bray vs. beret) zeigten, unterscheiden sich solche Wortpaare mit einem oder keinen pretonischen Vokal

durch die Synchronisierung beider Konsonantengesten zueinander. Daher tendierten Hörer dazu, die gleiche Konsonantenabfolge (/b/ + /r/) mit Überlappung als Cluster (bray) und ohne Überlappung als CVC (beret) wahrzunehmen (Browman & Goldstein 1990b). Oder, um es mit den Worten Schmeisers auszudrücken:

[...] in the case of learned transmissions, inter-gestural timing is such that the target of the Latin lateral is aligned with the release of the obstruent. In this way, the gestures are still intersecting within the Phase Window, but overlap more [...] (Schmeiser 2009a: 507).

Gehobene Wörter oder Latinismen werden in der zitierten Arbeit wiederum durch eine weitere Synchronisierung für das Erhalten des Lateral erklärt, deren Ausmaß zwischen der von Erbwörtern und halbgehobenen Wörtern einzustufen ist. Bei dieser Arbeit ist es sehr bedauerlich, dass sie keine Daten analysiert, um die aufgestellten Hypothesen zu testen. Inzwischen haben empirische Untersuchungen gezeigt, dass die Gestenkoordination von Konsonantenclustern mit zentralen Liquiden weiter auseinander liegt als mit Lateralen (Pouplier & Beňuš 2011; Hoole et al. 2013.). Dies war ebenso der Fall für das Portugiesische in Cunha (2012), allerdings zeigt diese Studie, dass der räumliche Gestenumfang und die Geschwindigkeit einen zusätzlichen Einfluss auf die Unterscheidung von /r/- und /l/-Clustern¹² ausüben.

Die drei wichtigsten Punkte des historischen Rückblicks können wie folgt zusammengefasst werden:

- (i) Obstruent-r-Cluster waren stabil, während Obstruent-l-Cluster diachronisch einer Reihe von Lautwandelprozessen unterlagen. Dies hatte zur

¹² In beiden Varietäten ist der zentrale Liquid ein alveolarer Tap /r/ (regional auch Trill /r/ sowie andere Varianten möglich, vgl. Mateus & Andrade 2000, Silva 1996), während die Laterale als klares /l/ im BP (Silva 1996, Pagan & Wertzner 2007) und als dunkles /ɫ/ im EP realisiert werden (Andrade 1999, Martins et al. 2011, Oliveira et al. 2011).

Folge, dass sich laterale Cluster zu mehreren verschiedenen Lauten entwickelten und erst relativ spät in das portugiesische Lexikon eingingen. Daher gehören Obstruent-r-Cluster zum Erbwortschatz und Obstruent-l-Cluster wurden erst später dem gehobenen Wortschatz hinzugefügt (Mateus & Andrade, 2000: 40, Veloso 2007a: 122).

- (ii) Nichtsdestotrotz ist der gemeinsame Ursprung beider Plosiv-Liquid-Gruppen im Klassischen Latein zu verankern. Beide Cluster stammen zum großen Teil aus lateinischen Clustern und zu einem kleineren Teil aus lateinischen CVC-Sequenzen, die nach der Tilgung unbetonter Vokale zu reinen Clustern wurden.
- (iii) Die Herausbildung von Clustern sowohl aus Clustern als auch aus CVC-Sequenzen fand in beiden Fällen teilweise in einem frühen Stadium der historischen Entwicklung statt, so dass die im Erbwortschatz vorhandenen Cluster zu beiden Ursprungsformen zurückverfolgt werden können. Dies ist außerdem an der parallelen Entwicklung von lateinischen Clustern und CVC-Sequenzen ersichtlich.

So wird bereits aus dem historischen Rückblick ersichtlich, dass die Trennung zwischen lexikalischen Clustern und CVC-Folgen in der Sprachentwicklung nicht stabil ist: Es wurden Lautwandelprozesse in mehreren Kontexten belegt, bei denen CVC-Sequenzen zu reinen Konsonantengruppen wurden. Allerdings: Für den umgekehrten Prozess, d.h. ein Lautwandel, bei dem ein Cluster zu einer Sequenz mit einem intermediären Vokal verändert wurde, wurden im romanischen Kontext – wie im nächsten Abschnitt diskutiert – keinerlei Hinweise gefunden.

2.1.2 Die Entwicklung der Obstruent-Liquid-Cluster im romanischen Sprachvergleich

Auch im romanischen Kontext weisen initiale Obstruent-r-Cluster eine stabile Entwicklung auf, denn die Cluster aus dem Klassischen Latein blieben in den größten romanischen Sprachen erhalten (KL *bracchīum* > Rum. *braț*, It. *braccio*, Sp. *brazo*, Kat. *braç*, Okz. *bratz*, Fr. *bras*, KL *crēdēre* > Rum. *cred*, It. *credere*, Sp. *creer*, Kat. *creure*, Okz. *creire*, Fr. *croire* und KL *frēno* > Rum. *frâu*, It. *freno*, Sp. *freno*, Kat. *fre*, Okz. *fren*, Fr. *frein* (Boyd-Bowman 1954: 110-113, Jensen 1999: 173-181).

In medialer Position waren die Obstruent-r-Cluster anfälliger für Lautwandelprozesse, die in gleichem Maße mediale Cluster und CVC-Sequenzen des Klassischen Lateins betrafen. Die Ähnlichkeit beider Formen ist an der parallelen Entwicklung erkennbar: Bei /pr/ wurde der stimmlose Plosiv zu einem stimmhaften Plosiv, zu einem stimmhaften labiodentalen Frikativ oder er wurde beibehalten (KL *caprā* > Rum. *capră*, It. *capra*, Sp. *cabra*, Kat. *cabra*, Okz. *cabra*, Fr. *chèvre* und KL *lēpōre* > Rum. *iepure*¹³, It. *liebre*, Sp. *lebre*, Kat. *llebre*, Okz. *lebre*, Fr. *lièvre*). Die Entwicklung von /tr/ und /kr/ verlief in manchen Sprachen parallel zu /pr/, außer im Okzitanischen und Französischen, denn dort wurde der Plosiv vokalisiert oder getilgt (KL *pētra* > Rum. *piatră*, It. *pietra*, Sp. *pedra*, Kat. *pedra*, Okz. *peira*, Fr. *pierre* und KL *lacrimā* > Rum. *lacrimă*, It. *lacrima*, Sp. *lágrima*, Kat. *llagrima*, Okz. *lairme*, Fr. *larme*).

Bei den stimmhaften Clustergruppen (/br/ und /gr/) wurden die stimmhaften Plosive in den meisten Fällen entweder vokalisiert oder sie blieben erhalten. Außerdem wurden bilabiale Plosive im Italienischen geminiert sowie /d/ im Französischen und /g/ im Italienischen getilgt, so dass folgende Variation entstand: KL *ŷebre* > Rum. , It. *febbre*, Sp. *fiebre*, Kat. / Okz. *lebre*, Fr. *fièvre*, KL *quadru* > Rum. *codru*, It. *quadro*, Sp. *cuadro*, Kat. *caire*, Okz. *creire*,

¹³ Im Rumänischen wurden in diesem Kontext keine Cluster gebildet.

Fr. *croire* und KL *nīgru* > Rum. *negru*, It. *nero*, Sp. *negro*, Kat. *negre*, Okz. *neir/ negre*, Fr. *noir* (Boyd-Bowman 1954: 47-49, Jensen 1999: 225-233).

Bei der Entwicklung der Obstruent-Liquid-Cluster fand mehr Variation statt. Initiale Obstruent-l-Cluster blieben im Katalanischen, Okzitanischen und Französischen erhalten und wurden in den weiteren Sprachen palatalisiert. Da die Palatalisierung im Rumänischen lediglich velare Plosive betraf, hat man angenommen, dass die Palatalisierung bei /kl/ und /gl/ begann und sich nicht auf weitere Kontexte im Rumänischen ausweitete. In den westromanischen Sprachen fand die Palatalisierung ausschließlich nach stimmlosen Plosiven statt. Die entstandenen Formen waren mit [ʎ] im Spanischen und mit [Cj] im Italienischen, so dass sich folgende weitere Alternanzen in der Romania ergaben:

Die medialen Obstruent-l-Cluster entwickelten sich parallel zu den initialen in dem Sinn, dass ebenso eine Palatalisierung stattfand. Mediale /kl/ und /gl/ verlief allerdings zu einem palatalen Lateral [ʎ] hin, das sich in medialer Position auf französische und katalanische Cluster ausweitete. Im Italienischen wurde der Lateral wie in initialer Position vokalisiert, der Unterscheid ist jedoch, dass der Plosiv geminiert wurde (KL *aurīcūla* > Rum. *ureche*, It. *orecchio*, Sp. *oreja*, Kat. *orella*, Okz. *aurella*, Fr. *oreille*). Es kann hierbei nicht zwischen ursprünglichen Clustern und CVC-Sequenzen im Klassischen Latein differenziert werden, die im Vulgärlatein durch Vokaltilgung zu einem Cluster wurden, da sie sich ähnlich entwickelten. Eine zusammenfassende Übersicht über die Entwicklung der initialen und medialen Obstruent-l-Cluster in den sechs besprochenen romanischen Sprachen wird in Tabelle 4 gegeben.

Aus der Entwicklung der Cluster vom Klassischen Latein bis hin zu den sieben größten romanischen Sprachen geht hervor, dass:

- (i) die Obstruent-l-Cluster sowohl im Portugiesischen als auch in den anderen romanischen Sprachen einer

größeren Anzahl von Lautwandelprozessen unterlagen als Obstruent-r-Cluster;

- (ii) in beiden Clustergruppen mediale Cluster anfälliger für Lautwandelprozesse waren als initiale Cluster. Aus Sicht der Produktion könnte dies mit der stärkeren Koartikulation und Überlappung beider Konsonantengesten in medialer Position im Vergleich zur initialen erklärt werden (Byrd 1993);
- (iii) wie im Portugiesischen die aktuellen romanischen Cluster zu einem großen Teil aus den im Klassischen Latein bereits vorhandenen Clustern entstanden, jedoch zu einem kleinen Teil auch aus entsprechenden CVC-Sequenzen.

Dieser Überblick zeigt, dass die diachrone Entwicklung der aktuellen Obstruent-r-Cluster aus dem Lateinischen stabiler war als die von Obstruent-l-Clustern, da Erstere an weniger Lautwandelprozessen beteiligt und in manchen Fällen das Ziel des Abbaus der lateralen Cluster waren (Rhotazismus, siehe Müller 2011 für eine aktuelle, ausführliche Diskussion). Dies war auch im Sardischen der Fall. Die entgegengesetzte Entwicklung (von /r/- zu /l/-Laut, der sogenannte Lombardismus) ist in der Romania ebenfalls belegt, trat allerdings viel seltener auf (siehe ebenso Müller 2011).

Dennoch waren die beiden initialen Clustergruppen (abgesehen vom Portugiesischen) bereits im Erbwortschatz aller weiteren romanischen Sprachen vorhanden. Die mediale Position zeigte sich anfälliger für Lautwandelprozesse als die initiale, da sich die Palatalisierung in medialer Position auf das Katalanische und Französische ausweitete (Jensen 1999, Koike & Macedo 1992, Williams 1961).

Die Tilgung unbetonter Vokale hatte die Entstehung reiner Konsonantenfolgen bereits im Vulgärlatein zur Folge. Diese im Vulgärlatein neu entstandenen Cluster entwickelten sich in den

romanischen Sprachen großteils parallel zu den aus dem Klassischen Latein übernommenen Clustern. Infolgedessen muss man annehmen, dass manche übernommenen und neuen Cluster sich bereits im Vulgärlatein kaum unterschieden und weitere Unterscheidungen schon zu einem frühen Zeitpunkt in der Übertragung hin zu den romanischen Sprachen sukzessive verloren gingen. Ein eher kontinuierlicher Prozess ist anhand mancher Ausnahmen ersichtlich, z.B. wurde KL *lēpōre* 'Hase' in allen besprochenen Sprachen außer im Rumänischen zu einem Cluster (Rum. *iepure*). Die Feststellung von solchen Ausnahmen schließt allerdings nicht die Neutralisierung des Unterschieds zwischen Clustern und CVC-Sequenzen zu einem sehr frühen Zeitpunkt in der Entwicklung der romanischen Sprachen aus.

2.1.3 Synchrone Variabilität bei Liquid-Clustern im Portugiesischen

Im gegenwärtigen Portugiesisch ist die Frequenz lexikalischer Obstruent-r-Cluster deutlich höher als die von Obstruent-l-Clustern (Velošo 2007a: 124). Dies ist eine logische Konsequenz aus der historischen Entwicklung, da Obstruent-r-Cluster ziemlich früh in das Lexikon aufgenommen wurden (Erbwörter), während Obstruent-l-Cluster erst relativ spät hinzugefügt wurden (gehobene Wörter). Hinzu kam noch, dass die Entwicklung der Obstruent-l-Gruppe bei den sogenannten halbgehobenen Wörtern in die r-Gruppe mündet, weshalb die in halbgehobenen Wörtern entstandenen Obstruent-r-Cluster einer Obstruent-l-Gruppe im Klassischen Latein entsprachen.

In der gesprochenen Sprache wurde bei der europäischen Varietät (EP) eine optionale Vokalepenthese (von [i]) bei der Obstruent-l-Gruppe festgestellt.

Por outro lado, corroborando ainda os dados de Freitas (1997) e uma nossa constatação empírica, ainda não referida, segundo a qual produções fonéticas como [pʰlātɐ] (para "planta") e [fʰlɔf] (para "flor"), p. ex., são relativamente comuns num registo coloquial mesmo em sujeitos adultos, uma outra manifestação deste tipo de acrescentos epentéticos merece

uma particular menção neste momento. Referimo-nos às produções metrificadas tradicionais, de raiz popular, supostamente oriundas de falantes detentores de níveis de instrução formal nulos ou pouco elevados, nas quais é detectável a introdução de uma vogal epentética entre as duas consoantes deste tipo de seqüências (Veloso 2007a: 125).

[Auf der anderen Seite bekräftigen die Daten [aus dem Erstspracherwerb] von Freitas (1997) und eine noch nicht erwähnte Feststellung von uns, dass solche phonetischen Produktionen wie zum Beispiel [pi'látɐ] (für "planta" 'Pflanze') und [fɨ'lor] (für "flor" 'Blüte/Blume') ziemlich üblich in der Umgangssprache, selbst bei Erwachsenen, sind. Ein anderer Ausdruck solcher epenthetischen Hinzufügungen verdient Erwähnung an dieser Stelle: Wir meinen traditionelle metrifierte Produktionen volkstümlichen Ursprungs, die angeblich von Sprechern mit keinem oder nur niedrigem Bildungsniveau stammen. Sie weisen die Einfügung eines epenthetischen Vokals zwischen beiden Konsonanten in diesen Sequenzentyp auf. Veloso 2007a: 125, Übersetzung CC].

Der zitierte Auszug fügt unterschiedliche Indizien für die Vokalepenthese in lateralem Kontext zusammen:

- (i) Die impressionistischen Beobachtungen des Autors, durch die die Epenthese im lateralen Kontext beschrieben wird ([pi'látɐ], planta 'Pflanze'; [fɨ'lor], flor 'Blüte/Blume'),
- (ii) Zitierung von Daten aus dem Erstspracherwerb (Freitas 1997: 173-174, 189),
- (iii) Beispiele aus einer volkstümlichen Kurzgedichtsammlung aus den achtziger Jahren aus Südpotugal (*Alentejo*).

Diese instabile Entwicklung und die späte Aufnahme der Obstruent-l-Cluster in das portugiesische Lexikon sowie die gegenwärtige Vokalepenthese werden so gedeutet, dass l-Cluster keine tautosilbischen Cluster des Portugiesischen seien. Mit anderen Worten, Obstruent-l-Cluster gehören nicht zu einem

Silbenanlaut, sondern jeder der Konsonanten bildet den entsprechenden Anlaut in unterschiedlichen Silben (Veloso 2007a: 127, 2006). Mit diese These steht der Autor allein in seinem Forschungsumfeld, das diese Cluster u.a. aufgrund ihre Sonorität und Ähnlichkeiten zur verwandten Sprachen als eindeutig tautosilbisch ansieht (u.a. Bisol 1999, Freitas 2003, Freitas & Santos 2001, Mateus & d'Andrade 2000, Netto 2001, Schmeiser 2009b, Vigário 2003).

Bei den zitierten Erstspracherwerbsdaten wird die Vokalepenthese als eine von fünf möglichen 'Rekonstruierungsstrategien'¹⁴ der Kinder beim Erwerb von Konsonantenclustern (*estratégias de reconstrução*) aufgeführt. Zu den verwendeten Strategien gehörten C₁-Tilgung, C₂-Tilgung, Cluster-Tilgung (C₁C₂), Vokalisierung von C₂ und Epenthese von [i] oder in seltenen Fällen [ɐ] (Freitas 1997: 174). Außerdem wurde die Epenthese ähnlich häufig bei Lateralen und bei Liquid-Lauten beschrieben (Freitas 1997: 173-174, 189). Somit wurde in den Spracherwerbsdaten kein Unterschied zwischen zentralen und lateralen Clustern festgestellt. Das dritte Argument (iii) umfasst Beispiele aus einer volkstümlichen, mündlich überlieferten Kurzgedichtsammlung aus den achtziger Jahren in Südportugal (*Alentejo*). Die mündliche Überlieferung und die geringere Bildung¹⁵ der Sprecher (*falantes*) sind Argumente dafür, dass die Gedichte von älteren Sprechern aus einer noch älteren Tradition stammen. Die Aussagekraft dieses Sammelbandes ist auch dadurch gegeben, dass deren Sprecher Beweise einer für die Sprachentwicklung natürlichen Übertragungsmethode (die mündliche Übertragung) sind und eine geringere Bildung aufweisen. Die

¹⁴ Die zitierte Autorin interpretiert die Produktion der Kinder als Versuch, die unmarkierte CV-Silbe zu rekonstruieren, weshalb die eigentlichen Produktionsbeschreibungen 'Rekonstruierungsstrategien' (*estratégias de reconstrução*) genannt werden (Freitas 1997: 174).

¹⁵ Da die Alphabetisierung in Portugal spätestens seit Ende der Salazar-Diktatur am 24. April 1975 und seit dem Beitritt des Landes zur CEE (aktuell: Europäische Union) am 1. Januar 1986 eine zunehmende Rolle spielte, wird hier impliziert, dass es sich eher um ältere Sprecher handelt.

geringere Bildung garantiert, dass diese Sprecher von orthographischem Wissen nicht stark beeinflusst waren. Daher besitzen diese Sprecher ein besonderes Empfinden für die intuitive Struktur der Sprache, unter anderen im Bereich der Silbenstruktur. Ein illustrierendes Beispiel wurde im Folgenden übernommen.

Não olhes p'ra mim, não olhes, / Qu'eu não sou o teu amor /
 Eu não sou como a figueira / Que dá fruto sem f[i]lor.
 Schau mich nicht an, schau nicht, /Denn ich bin nicht deine
 Liebe / Ich bin nicht wie der Feigenbaum, /Der Frucht ohne
 Blüte gibt. Delgado, 1980: 236; Nr. 2114, übernommen von
 Veloso 2007a: 126, Unterstreichung aus Veloso, Übersetzung
 CC.

Die Epenthese im zitierten Auszug belegt, dass sich im EP Liquid-Cluster ähnlich wie reine Plosivcluster¹⁶ verhalten, bei welchen die optionale Epenthese auf die Zugehörigkeit der Konsonanten zu zwei unterschiedlichen Anlauten hindeutet (d'Andrade & Viana 1993, Mateus & Andrade 1998, 2000). Aus theoretischer Sicht wird die mögliche Nichtrealisierung eines Vokals im Plosivkontext mit einem leeren Nukleus-Platzhalter erklärt und diese Annahme wird von Veloso (2007: 127) auf die Plosiv-Liquid-Cluster übertragen. Demnach können entweder beide Nuklei vollständig realisiert werden oder der erste Platzhalter kann leer bleiben (Veloso 2007a: 127). So wird die Variation bei Obstruent-l-Clustern nicht nur im EP, sondern auch in anderen Portugiesisch-Varietäten anhand der Realisierung oder Nicht-Realisierung des leeren Nukleus-Platzhalters erklärt. Eine solche Variation wurde auch in anderen Sprachen beschrieben und mit anderen, vielleicht besser geglückten Ansätzen modelliert, die im nächsten Unterpunkt diskutiert werden.

An der zitierten Passage ist außerdem interessant (und an den weiteren Beispielen in Veloso 2007a: 126-127), dass die Schreibweise sowohl auf eine Vokalepenthese als auch auf die Tilgung einiger Vokale hinweist, indem Vokale hinzufügt (*[f[i]lor]*)

¹⁶ Die Plosivcluster werden separat unter 4.4 diskutiert.

oder durch einen Apostroph angedeutet wurden (*p'ra, Qu'en*). So kann man an dieser Stelle festhalten, dass in diesen Texten von älteren Sprechern mit einem niedrigen Bildungsniveau Vokalepenthese und Tilgung in der europäischen Varietät koexistieren. Dies ist allerdings nicht die Deutung von Veloso (2007), der vor allem das Auftreten der Vokalepenthese dokumentieren möchte.

Pensamos que as produções de criações metrificadas populares que acabamos de transcrever constituem, entre outros, um indício considerável de que [...] também as sequências Obstruinte+Lateral são frequentemente produzidas com uma vogal "epentética" entre as duas consoantes. [Wir glauben, dass die volkstümlichen metrifizierten Konstruktionen, die wir gerade abgedruckt haben, unter anderem ein starkes Indiz dafür sind, dass auch die Sequenzen mit Obstruent + Lateral oft mit einem epenthetischen Vokal zwischen den Konsonanten produziert werden. (Veloso 2007a: 126-127, Kursiv im Zitat und Übersetzung CC)].

Aus dem Zitat lässt sich das Adverb *frequentemente* ('oft') entnehmen, das allerdings, abgesehen von der Gedichtsammlung, mit keinen weiteren empirischen Daten belegt wurde. Angesichts dessen kann seine Frequenz nicht zuverlässig eingeschätzt werden. In der weiteren Literatur sind ebenfalls ähnliche Beispiele von jüngeren Sprechern kaum nachweisbar. Man findet eher Beschreibungen der massiven Vokaltilgung im gegenwärtigen EP (Cunha 2011a, Martins 1975, Mateus & Martins 1982, Silva 2008, 2007, Vigário 1998). So traten die geschilderten Epenthese-Produktionen bei wenig alphabetisierten, älteren Sprechern in ländlichen Gegenden von Portugal (*Alentejo*) auf, während im übrigen Land einschließlich der Azoren (Silva 1998, 1997) eine durchgehende Tendenz zur Vokaltilgung vorherrscht. Womöglich sind beide Prozesse in einem Lautwandelkontext zu sehen, bei dem die Epenthese eine ältere Form belegt, die zumindest eine Zeit lang mit der Vokaltilgung koexistierte und in späteren Generationen verloren ging. Welche Rolle dabei die Vokaltilgung spielte und wie es zu der Koexistenz von Epenthesen und

Tilgungen bei dem angeführten Beleg kam, bleibt Untersuchungsgegenstand künftiger Forschung.

In der brasilianischen Varietät (BP) ist das Phonem /r/ variabel, in Konsonantenclustern kann der r-Laut als [r, ɾ, x] realisiert werden (Monaretto 2009: 143), in anderen Positionen ist die Variation noch größer [r, ɾ, ʀ, x, ʁ] (Monaretto 2009: 141-144). Bei der Produktion von Obstruent-r-Clustern im BP wurde ein intermediäres vokalisches Element beschrieben, das entweder als vokalisches Element (Silva 1996), als epenthetischer (Carvalho 2004) oder Svarabhakti-Vokal¹⁷. (Nishida 2009) bezeichnet wurde. Eine akustische Analyse dieses epenthetischen Vokals in tautosilbischen Clustern im BP zeigte, dass er nach Plosiven die gleichen Formantenstrukturen wie der folgende Vokal aufweist, während er nach Frikativen zentriert ist (Nishida 2009: 34-35, 47-48). Silva (1996: 69) stellte fest, dass /r/ sowohl als zweites Element eines Clusters als auch in Auslautposition zwischen Formantenstrukturen realisiert wird. Der Svarabhakti-Vokal kann daher links oder rechts vom r-Laut vorkommen und weist in beiden Fällen eine kürzere Dauer auf als der lexikalische Vokal selbst. Aussagekräftig bei der r-Produktion sei, dass sich /r/ unabhängig von der Position immer an zwei Vokalen stützt (Silva 1996: 68).

Außerdem wurden im EP einige interessante Konsonantentilgungen festgestellt. Plosiv-Liquid-Gruppen können im BP vereinfacht werden, indem die Liquide in unbetonten Silben ausgelassen werden. Dies wurde sowohl post- wie prätonisch beobachtet ([ˈlivru] vs. [ˈlivu]: *livro* ‘Buch’, [preˈsiza] vs. [peˈsiza]: *precisa*, 3.Person Singular ‘braucht’), jedoch nicht in primärbetonter Silbe ([ˈpratu], Cristófaros-Silva 2000). Bei der Analyse akustischer Aufnahmen von formellen und informellen Diskursen aus den Landeshauptstädten von Minas Gerais, Maranhão, São Paulo, Rio Grande do Sul und Rio de Janeiro stellten Lopes & Santos (1995) fest, dass diese Cluster zusätzlich in betonter Silbe vereinfacht werden können, wenn

¹⁷ Siehe 3.4.1 für detailliertere Informationen hierzu.

zwei komplexe (mehrgliedrige) Silbenanlaute aufeinander folgen (z.B. *problema*, *problema* ‘Problem’ und *frustado*, *frustrado* ‘frustriert’, Lopes & Santos 1995, siehe auch Madureira 1997, Ramos 2000).

Eine andere Studie von Cristófaró-Silva (2000) zeigte, dass auch heterosilbische Konsonantencluster mit einem der Auslautkonsonanten /N, R, S, L/ und einem weiteren Konsonanten im nächsten Silbenanlaut reduziert werden können. Diese Konsonantenreduktion erfolgt sowohl in betonter als auch in unbetonter Position und wurde bei unbetonten Silben sowohl vor als auch nach der Hauptbetonung festgestellt (*perunta*, *pergunta* ‘Frage’).

Nach der quantitativen Analyse von Freitas (1999) hängt der Anteil der Konsonantentilgung von der Position des Clusters relativ zur Primärbetonung ab. Sie stellte fest, dass 50% der Tilgungen in unbetonter Silbe vor der primärbetonten erfolgten, 33% in der primärbetonten Silbe und die übrigen 17 % nach der Hauptbetonung. Aus der Studie konnte kein eindeutiger Bezug zwischen Betonung und Tilgung herausgelesen werden. Daher ist mehr Forschung zu den modellierenden Faktoren dieser Konsonantentilgung nötig. Die Motivation dieser Tilgung ist ebenfalls unklar: Eines der stärksten Argumente dafür ist die Suche nach der universell unmarkierten CV-Silbe (Clements & Keyser 1983, McCarthy 1981, 1986, Prince & Smolensky 2004). Das getilgte Element kann allerdings nicht einheitlich erklärt werden: Bei den tautosilbischen Clustern wird der zweite Konsonant getilgt und somit der komplexe Anlaut vereinfacht, was für die CV-Theorie spricht. Bei den heterosilbischen Clustern wird jedoch der einfache Anlaut getilgt und der Auslaut zum Anlaut resilbifiziert. Eine weitere Möglichkeit wäre, dass der Sprecher die freie Wahl zwischen der Realisierung eines Clusters oder eines einfachen Konsonanten hat. Dies würde durch die Existenz beider Formen im Lexikon des Sprechers ermöglicht (Cristófaró-Silva 2000).

So können brasilianische Konsonantencluster /C₁C₂/ optional zu einem einzelnen Konsonanten vereinfacht werden. Die

Vereinfachung zu C_1 erfolgt sowohl bei tautosilbischen als auch bei heterosilbischen Clustern, die sich über eine Silbengrenze hinweg verteilen. Die Kontexte sind bei den tautosilbischen Clustern zwar stärker eingeschränkt, dennoch kann die Vereinfachung sowohl in unbetonter als auch in betonter Silbe vorkommen.

Insgesamt stellen die letzten Absätze ein vielschichtiges Bild der portugiesischen Konsonantencluster dar: Es wurde eine historische Entwicklung von Obstruent-Liquid-Clustern skizziert, bei der sich /r/-Cluster als Erbwörter des Portugiesischen herauskristallisierten, während /l/-Cluster erst später in das Lexikon aufgenommen wurden (gehobene Wörter). Im gegenwärtigen EP können die gehobenen Wörter mit einem intermediären [i] von älteren Erwachsenen produziert werden (Veloso 2007a), während v.a. die jüngeren Sprecher zur Vokaltilgung und damit Clusterbildung neigen. Es wird unterstellt, dass die spätere Aufnahme der /l/-Cluster in das portugiesische Lexikon einen Einfluss auf die gegenwärtige Epenthese im EP haben könnte. Beide Argumente dienen auf jeden Fall als Indiz dafür, dass Obstruent-Liquid-Gruppen keine Cluster sind, weil sie in der Tiefenstruktur des Portugiesischen einen intermediären Vokal besitzen, der an der Oberfläche entweder realisiert werden kann oder nicht (Veloso 2007a: 127). Sowohl Gruppen mit /l/ als auch mit /r/ können in der brasilianischen Varietät auf einen einzelnen Konsonanten reduziert werden (Cristófaró-Silva 2001, 2000, Lopes & Santos 1995). Wollte man alle beschriebenen Realisierungen als Evidenz für wichtige Phänomene in einem Sprachsystem (das Portugiesische) betrachten, müsste man dabei auf die Existenz von Konsonantenclustern verzichten. Vielmehr deuten die beschriebenen Studien auf Einzelbeobachtungen hin, deren Aussagekraft ohne eine solide empirische Grundlage schwer zu beurteilen ist.

2.1.4 Variation bei Obstruent-Liquid-Clustern im romanischen Vergleich

Das Vorkommen eines intermediären vokalischen Elementes in Clustern mit Liquiden [r, l] wurde im Sanskrit als Svarabhakti-Vokal bezeichnet. Svarabhakti bedeutet ursprünglich ‘der Vokal aus der mittleren Reihe’. Dieses vokalische Element tritt zwischen einem Liquid und einem nachfolgenden Konsonanten auf, hat eine deutlich kürzere Dauer als ein lexikalischer Vokal und sein Auftreten fügt dem Wort keine zusätzliche Silbe hinzu (Allen 1953: 73-80, Ali 2008). Man hat versucht, diesen Vokal anhand möglicher Kontexte zu beschreiben, nämlich dass er nur bei Clustern mit Sonorant vorkommt und die gleiche Qualität des nachfolgenden Vokals hat, dass er bei schnellen Sprechgeschwindigkeiten nicht auftritt und keine zusätzliche Silbe hervorruft. Die wichtigste Eigenschaft dieser Svarabhakti-Vokale ist, dass sie keine eigenständigen lexikalischen Segmente sind, sondern durch ein spezifisches Timing in der artikulatorischen Organisation entstandene Wahrnehmungen, die vom lautlichen Kontext und der Vokalqualität beeinflusst werden (Hall 2003: 1-3, 2006).

Svarabhakti-Vokale wurden bereits im Sanskrit (Allen 1953) und im späten Vulgärlatein (Schuchardt 1867: 421) sowie in einigen romanischen Sprachen und Varietäten (für das Spanische: Almeida & Dorta 1993: 106-107, Bradley 2001, 2002, 2004, 2006, Bradley & Schmeiser 2003: 1-3, Quillis 1993: 330, Lipksi 1995) beschrieben. Im Spanischen wurde der Svarabhakti-Vokal bei tautosilbischen Clustern zwischen Obstruent und Tap beschrieben, z.B. [ɲɟala'tera] für *Inglaterra* ‘England’, [eɟe'lesja] für *eglesia* ‘Kirche’, [co'ronika] für *cronica* ‘Chronik’ (Quillis 1993: 330, Menendez-Pidal 1926: 127). Seine Formantstruktur ist ähnlich wie die des nachfolgenden Vokals. Eine detaillierte Analyse anhand des Spanischen aus Teneriffa zeigte, dass F1 bei offenem nachfolgenden Vokal höher und bei nachfolgendem geschlossenen Vokal niedriger ist. Zusätzlich dazu sind F2 und F3 höher bei nachfolgenden vorderen und niedriger bei hinteren

Vokalen (Almeida & Dorta 1993: 107). Svarabhakti-Vokale haben tendenziell eine längere Dauer in initialer als in medialer Position und in betonten als in unbetonten Silben. Die Produktionsrichtung zeigt eine deutliche Tendenz zu längeren Svarabhakti-Vokalen in der *back-to-front* Richtung (z.B. /kr/) als umgekehrt (/pr/, Bradley & Schmeiser 2003: 2).

Ohala (1992: 323) erläutert, dass Realisierungen des Wortes *blow* lediglich durch Längung der Laterale als überzeugende Formen von *below* fungieren können, ohne dass ein Vokal hinzugefügt wird. In diesem Fall sei Silbizität ein perzeptives Konstrukt, das beim Hörer liege. Hörer interpretierten bei einem deutlich längeren Lateral ein zusätzliches Element, das die akustischen Ereignisse des wahrgenommenen akustischen Signals besser wiedergibt (Ohala 1992: 331, 1999, Steriade 1999, 2000). So hänge die Wahrnehmung eines Vokals oder vokalischen Elements lediglich mit der Konsonantenlänge zusammen.

Eine andere Erklärungsmöglichkeit für die Variation bei Konsonantenclustern ist der Einfluss einer graduellen zeitlichen Koordination der einzelnen Konsonantengesten auf die Sprachperzeption (Cunha & Harrington 2011b, Browman & Goldstein 1990a, b) im Rahmen der Artikulatorischen Phonologie (Browman & Goldstein 1986, 1988, 1990, ff.). Browman & Goldstein (1990a, b) zeigten, dass sich Wortpaare wie *bray* vs. *beret* vor allem anhand der Konsonantenkoordination unterscheiden lassen. Bei einem Perzeptionsexperiment nahmen Hörer bei den Stimuli mit sich überlappenden Konsonantengesten (/b/ und /ɪ/) *bray* wahr und bei den nicht-überlappenden Stimuli *beret*. Die Grundidee dahinter war, dass, wenn C_2 in einer $\langle C_1C_2 \rangle$ -Sequenz später auftritt, der größere Abstand zwischen den Konsonanten in der Sprachperzeption durch ein kurzes vokalisches Element wiedergegeben wird, ohne dass ein Vokal in die lexikalische Repräsentation des Wortes integriert werden muss. So würden folglich (lexikalisch) reine Konsonantenabfolgen (Cluster) mit einem intermediären Vokal wahrgenommen werden.

2.2 Reine Obstruentencluster im Portugiesischen

Im Portugiesischen können unter anderem reine Obstruenten¹⁸-Cluster auftreten (Mateus & Andrade 2000: 42, Bisol 1999). Dabei sind sie meistens entweder stimmlos oder stimmhaft, Cluster mit gemischter Stimmhaftigkeit kommen dagegen selten vor.

Clust- er	initial		medial	
	Port.	D-Glosse	Port.	D-Glosse
[pt]	<i>ptério</i>	‘Pterion’	<i>captar</i>	‘fangen’
[kt]	<i>ctenóforo</i>	‘Kammqualle’	<i>pacto</i>	‘Pakt’
[bt]			<i>obter</i>	‘erhalten’
[dk]	<i>bdélio</i>	‘Bedolachharz’	<i>adquirir</i>	‘erwerben’
[bd]			<i>abdómen</i>	‘Bauch’
[ps]	<i>psicologia</i> ¹⁹	‘Psychologie’	-	
[bs]			<i>absurdo</i>	‘Unsinn’
[bv]			<i>óbvio</i>	‘offen- sichtlich’
[bʒ]			<i>abjurar</i>	‘widerrufen’
[tz]			<i>quartzo</i>	‘Quarz’
[dv]			<i>advertir</i>	‘warnen’
[ks]			<i>axioma</i>	‘Axiom’

Tabelle 3: Obstruentencluster im Portugiesischen (von Mateus & d’Andrade 2000: 42-43 adaptiert).

¹⁸ Außerdem kommen Cluster aus Obstruent und Nasal vor (*tmese, ritmo, diafagma*). Diese wurden hier allerdings ausgeschlossen, da sie für die Fragestellung nicht ausschlaggebend sind.

¹⁹ Die Autoren listen das Beispiel unter den medialen Clustern auf. Offenbar ist den Autoren hier aber ein Fehler unterlaufen, der in unserer Darstellung korrigiert wurde.

Einige Cluster können initial vorkommen, sie sind allerdings in wortmedialer Position zahlreicher, wie sich aus den Beispielen in Tabelle 3 entnehmen lässt. Aus dieser Auswahl werden wir uns im Folgenden auf reine zweigliedrige Plosivcluster, d.h. Sequenzen von zwei Plosiven, beschränken.

2.2.1 Die Entwicklung reiner Plosivcluster

Die Entwicklung der reinen Plosivcluster aus dem Klassischen Latein bis hin zum Portugiesischen verlief wie bei den Obsruent-Liquid-Clustern innerhalb von unterschiedlichen Zyklen. Bei diesen sukzessiven Zyklen ist eine starke Tendenz zur Vereinfachung der Cluster zu finden, die sich erst sehr spät abschwächte, wodurch die Aufnahme einiger reiner Plosivcluster in das Portugiesische ermöglicht wurde. Diese Cluster traten im Klassischen Latein in medialer Position auf. Ihre Entwicklung wird im Folgenden am Beispiel der häufigen Cluster /pt/ und /kt/, die nicht selten in der diachronischen Entwicklung verwechselt wurden, skizziert.

Bei der Übertragung von Lexemen mit /pt/-Clustern wurden die folgenden vier Möglichkeiten beschrieben (Castro 1991, Koike & Macedo 1992, Mattos & Silva 1991, Williams 1968: 86)

- (i) Die ältere und am weitesten verbreitete Entwicklungsmöglichkeit innerhalb des Erbwortschatzes war, dass KL /pt/ durch regressive Assimilation zu /tt/ im Vulgärlatein (VL) und zu /t/ im Portugiesischen vereinfacht wurde (z.B. KL *aptāre* > port. *atar* 'binden, schnüren' und KL *scrīptum* > Port. *escrito* 'geschrieben', (Williams 1968: 86).
- (ii) In einer halbgehobenen Entwicklungsmöglichkeit wurde /pt/ über */ft/ im VL zu /ut/ im Altportugiesischen (aport.). Diese Entwicklung wurde zu einem späteren Zeitpunkt in den meisten

Fällen wieder zurückgenommen (KL *aptum* > AltPT *auto* > Port. *apto* ‘fähig’ und KL *adoptāre* > AltPT *adoutar* > Port. *adoptar* ‘adoptieren’).

- (iii) In den Fällen, in denen /pt/ dem Vokal /e/ folgte, durchlief es die gleiche Entwicklung wie /kt/, d.h. es wurde im Portugiesischen zu /it/. Bei manchen Wörtern ist allerdings wiederum eine Aufhebung dieses Wandels zu beobachten (KL *conceptum* > AltPT und Port. *conceito* ‘Begriff’ und KL *septimum* > AltPT *seitimo* > Port. *séptimo* ‘siebte’).
- (iv) In gehobenen Wörtern wurde /pt/ beibehalten (KL *capīāre* > Port. *captar* ‘fangen’)

Das Cluster /kt/ wurde zu zwei verschiedenen hohen Vokalen vereinfacht (Williams 1968: 95, Castro 1991, Mattos e Silva 1991):

- (i) Im Erbwortschatz fand die Vokalisierung des lateinischen /k/ zu /i/ statt. Der entstandene Vokal wurde allerdings nach /u/ getilgt (KL *nocte* > Port. *noite* ‘Nacht’ und KL *fructum* > AltPort. *fruito*²⁰, Port. *fruto* ‘Ertrag’).
- (ii) Nach /a/ oder /o/ wurde /kt/ – analog zu /pt/ – zu /ut/. Auch hier ist in manchen Fällen eine Aufhebung dieser Vokalisierung festzustellen (KL *doctorem* > Port. *doutor* ‘Doktor’ und KL *contractum* > AltPort. *contrauto* > Port. *contracto* ‘Vertrag’).
- (iii) In gehobenen Wörtern wurden die Cluster zwar übernommen, allerdings ist häufig eine Tilgung von /k/ anzutreffen, jedoch ohne eine nachgewiesene Regelmäßigkeit (KL *victimam* > Port. *victima* ‘Doktor’, KL *actum* > Port. *acto/ato* ‘Akt’ und KL *factum* > Port. *facto/fato* ‘Fakt’).

²⁰ Im Galicischen ist dies auch heute noch die Form des Lexems.

In der Übertragung der Cluster vom Lateinischen hin zum Portugiesischen wurde bei den Erbwörtern der ersten Plosiv entweder assimiliert oder getilgt, bei später aufgenommenen gehobenen Wörtern wurde optional der erste Plosiv getilgt. Die Übernahme der Cluster in das portugiesische Lexikon fand erst für gehobene Wörter statt, die keine mündliche Übertragung durchmachten, sondern direkt aus dem Klassischen Latein importiert wurden.

2.2.2 Synchroner Variabilität bei Obstruentenclustern

Plosivcluster weisen auch synchronisch Variabilität in portugiesischen Varietäten auf: Wenn ein Plosiv und ein anderer Konsonant außer einem Liquid aufeinanderfolgen, kann der erste Konsonant elidiert werden, wie z. B. bei port. [faktu] ~ [fatu] ‘Fakt’; [toraksiku] ~ [torasiku] ‘Thorax’. Diese Reduktion ist unregelmäßig, kommt selten vor und ist häufiger im BP als im EP anzutreffen. Bei Sequenzen von Obstruenten tritt im BP, in wenigen regionalen EP-Varietäten und im Erstspracherwerb ein epenthetischer Vokal auf (z.B. [kaptar] ~ [kapitar] ‘fangen’). Die Vokalepenthese wird aktiv verwendet und ist der bevorzugte Prozess, um Konsonantencluster der Sprache zu vereinfachen (vgl. Mateus & d’Andrade 2000: 89, 1998, Costa 2006).

Da Plosive der Klasse der Obstruenten angehören, ist es nicht ersichtlich, wie die Distribution der Konsonantentilgung vs. Vokalepenthese bei reinen Plosivclustern erfolgt. Nach Bisol (1999: 731) dient die Epenthese vor allem dazu, das Sprachsystem von unerlaubten Clustern zu befreien. Das hat zur Folge, dass im selben Sprachsystem unerlaubte Auslaute, die aus älteren Sprachstufen übernommen wurden, mit angepassten Strukturen koexistieren. In manchen Fällen alternieren Cluster und Sequenzen mit epenthetischem Vokal bei Produktionen desselben Sprechers (z.B. *capitar a simpatia* ‘Sympathie gewinnen’ (D2-SP 360: 1039) und *certo você teve que adaptar o horário deles* ‘ja, sie müssten sich an ihren Zeitplan anpassen’ (D2-SP 360: 371), Beispiele desselben Sprechers aus São Paulo aus dem Projekt NURC, Bisol 1999: 732). Die Tatsache, dass die Epenthese im BP aktiv verwendet wird, ist

heutzutage an der Lexikalisierung von englischen Lehnwörtern ersichtlich: Eng. *stress* > [is'tɾɛsi] *estresse* (BP) und ['strɛs] *stress* (EP) und Eng. *scanner* > [is'kanɛf] *escaner* (BP) und ['skanɛr] *scanner* (EP).

Beide Prozesse treten insgesamt öfter im BP als im EP auf, so dass beide Varietäten anhand der Alternanz zwischen durch Vokalepenthese oder Konsonantentilgung entstandenen CVC-Sequenzen im BP und reinen Plosivclustern im EP gekennzeichnet werden können (z.B. /kapItar/ und /fatu/ im BP vs. /kaptar/ und /faktu/ im EP).

2.2.3 Die theoretische Modellierung von Plosivclustern im Portugiesischen

Es gibt mehrere Argumente gegen die Auffassung, dass beide Konsonanten Teil eines Clusters sind und dafür sprechen, dass Plosivcluster im Portugiesischen heterosilbisch sind, d.h. durch eine Silbengrenze in zwei Silben aufgespaltet sind.

- (i) Diese Cluster kommen hauptsächlich in wortinterner Position vor, was dafür spricht, dass eine Silbengrenze zwischen beiden Konsonanten verläuft und C₁ zum Auslaut der vorhergehenden Silbe und C₂ zum Anlaut der folgenden Silbe gehört (Mateus & d'Andrade 2000, 1998, Lousada & Jesus 2006, Lousada et al. 2010).
- (ii) Sie richten sich nicht nach der Sonoritätshierarchie und werden von dem Silbifizierungsalgorithmus²¹ für das Portugiesische ausgeschlossen (Mateus & d'Andrade 2000: 62, 1998: 60-64, Veloso 2007a: 114, Vigário & Falé 1994: 476).

²¹ Silbifizierungsalgorithmus ist ein Bündel von Regeln, die alle möglichen Silbenkombinationen einer Sprache erfassen.

- (iii) Sie zählen zu den extrem seltenen Clustern der Sprache (Barbosa 1965: 180-181/211-212, 1994: 149, Câmara 1970: 56, Freitas 1997: 167, Veloso 2007a: 114, Vigário & Falé 1994: 469-471).
- (iv) Sie können mit einem epenthetischen Vokal ([i] in BP und [ĩ] in EP) zwischen beiden Konsonanten realisiert werden (Barbosa 1965: 212, Câmara 1970: 56-57, Vigário & Falé 1994: 477, Mateus & d'Andrade 1998: 16, 2000: 44, u. a.
- (v) Daten zum Erstspracherwerb von Freitas zeigten, dass Obstruentencluster mit einem epenthetischen Vokal produziert werden. Allerdings war ihre Frequenz mit 0,07% aller analysierten Daten im Korpus sehr gering (1997: 182), weshalb jede Schlussfolgerung mit Vorsicht zu behandeln ist.
- (vi) Diese Cluster entsprechen späteren Übernahmen aus dem Griechischen und dem Klassischen Latein in das portugiesische Lexikon, um spezifische Lücken im wissenschaftlichen und technischen Bereich zu schließen. Diese Kombinationen gehören daher nicht zum Erbwortschatz und können abweichende Lösungen für das implizite phonologische Wissen der Sprecher darstellen (Veloso 2007a: 115, Fußnote 5).
- (vii) Diese Sequenzen bereiteten Probleme bei der Implementierung eines Silbifizierungssystems für die Sprachsynthese, da die Abstammungsregeln nicht mit den auditiven Silbifizierungsmustern übereinstimmten (d'Andrade & Viana 1993a).
- (viii) Quantitative korpusbasierte Arbeiten betonen, dass das Portugiesische zu offenen Silben tendiert (Barbosa 1965, d'Andrade & Viana 1993a, b). Die letzte Arbeit zeigte, dass von den 17 möglichen Silbenformen die folgenden vier, nämlich CV, CVC,

V und CVV, für 90% des gesamten Wortschatzes verantwortlich sind, wobei nur eine dieser Silbenformen geschlossen ist.

Für die Analyse dieser Sequenzen in medialer Position könnte C_1 zum Auslaut der vorhergehenden und C_2 zum Anlaut der nächsten Silben hinzugefügt werden, allerdings bleiben so initiale Obstruentencluster außer Acht. Aus diesem Grund wurde die Annahme von extrasilbischen Konsonanten bevorzugt (X [+Kons], Mateus & d'Andrade 2000: 69). Diese Konsonanten werden mit keiner Silbe assoziiert, deswegen wird die Bildung eines leeren Nukleus postuliert, wodurch der extrasilbische Konsonant an das Skelett des Wortes angebunden werden kann, z.B. port. *pneu* – *pVneu* dt. 'Reifen' (Mateus & Andrade 2000: 69). Diese zurzeit sehr verbreitete Annahme (vgl. Cristófaros-Silva 1999, 2000, 2002, d'Andrade & Viana 1993, Mateus & d'Andrade 1998, 2000, Mateus et al. 2005, Veloso 2007a, 2003, u.a.) wurde bereits in den 70er Jahren vorgeschlagen:

(...) sendo ambas as consoantes pronunciadas, às vezes até em pares distintivos como em pacto «acordo»: pato «ave». A gramática tradicional e mesmo a fonética rigorosa de Gonçalves Viana vêem na primeira consoante uma consoante decrescente e a fronteira silábica entre elas. (...) Hoje, entretanto, parece-me a correta outra solução (...). Na realidade há entre uma e outra consoante a intercalação de uma vogal, que não parece poder ser fonemicamente desprezada, apesar da tendência a reduzir a sua emissão no registo formal da língua culta. Ela é /i/ na área do Rio de Janeiro e /e/ ([a] neutro em Portugal). [Beide Konsonanten werden ausgesprochen, manchmal sogar als Minimalpaare wie in *pacto* ('Pakt') und *pato* ('Ente'). Die traditionelle Grammatik und sogar die strenge Phonetik von Gonçalves Viana behaupten, dazwischen sei eine silbische Grenze. Mir scheint heute eine andere Lösung korrekt zu sein (...). In der Wirklichkeit gibt es zwischen dem einen und dem anderen Konsonanten einen intermediären Vokal, der trotz seiner seltenen Produktion in der gehobenen Sprache phonemisch nicht ignoriert werden darf. Es handelt sich um /i/ in Rio de Janeiro und /e/ (neutralisiertes [a]) in Portugal. Câmara, 1970: 56-57, Übersetzung CC]

Die Nicht-Realisierung dieses leeren Nukleus im EP und seine Realisierung durch das Hinzufügen eines epenthetischen Vokals im BP bringt die Variation zwischen beiden Varietäten zusammen. Für das BP wird vor einem ähnlichen theoretischen Hintergrund von demselben Prozess Gebrauch gemacht, allerdings kann in der lexikalischen Phonologie die Epenthese sowohl im Lexikon als auch postlexikalisch stattfinden (Bisol 1999).

Die Annahme eines leeren Nukleus ist zwar eine elegante Lösung für die Vereinfachung unerlaubter Sequenzen, allerdings wird dadurch nicht erklärt, weshalb Epenthesen mit Konsonantentilgungen variieren. Außerdem bleibt unklar, wieso das BP in größerem Ausmaß von beiden Prozessen betroffen ist als das EP. Im EP fehlt jeder Hinweis dafür, dass – außer bei Kindern während des Erstspracherwerbs (Freitas 1997) und älteren Sprechern in ländlichen Gegenden – solche epenthetischen Vokale realisiert werden. Verlässt man die Sprachsynchronie und verfolgt den Ursprung beider Realisierungen bis hin zum Klassischen Latein, fehlt erst recht jeder Hinweis auf eine Epenthese in der Entwicklung dieser Cluster, denn es traten ausschließlich regressive Assimilation und Vereinfachung, Vokalisierung von Konsonanten hin zu hohen Vokalen und Konsonantentilgung auf.

3 Die Herausbildung des portugiesischen Vokalsystems

Im Folgenden wird die Herausbildung des portugiesischen Vokalsystems skizziert. Aus Übersichtlichkeitsgründen und aufgrund der großen Variabilität der Vokale in ihrer Entwicklung wird die Herausbildung des portugiesischen Systems auf den iberoromanischen Kontext eingeschränkt. Ein spezieller Fokus liegt dabei auf den unbetonten Vokalen und den dazugehörigen phonologischen Prozessen, die dem Portugiesischen erneut eine außergewöhnliche Stellung im romanischen Vergleich verleihen. Davor werden die betonten oralen und die nasalen Vokalen vorgestellt.

3.1 Betonte orale Vokale

Der Verlust der Vokalquantität im Klassischen Latein verursachte bereits im Vulgärlatein eine Restrukturierung des gesamten Vokalsystems, wobei die Betonung die entscheidende Rolle bei der Vokaldiskriminierung übernahm (Jensen 1999, Krefeld, 1999). Demnach bekam bereits im Vulgärlatein der Gegensatz zwischen betonten von unbetonten Vokalen einen bedeutungsunterscheidenden Charakter, der bis zu den aktuellen romanischen Sprachen erhalten blieb. Das vulgärlateinische betonte Vokalsystem zeigte sich in seiner Entwicklung bis zu den iberoromanischen Sprachen ziemlich stabil: Die sieben betonten Vokale des Vulgärlateins /i, u, e, o, ε, ɔ, a/ blieben mit wenigen Ausnahmen²² erhalten.

²² Das kurze /ē/ wird im Galicischen und im Spanischen diphthongiert und das kurze /ō/ im Spanischen.

KL	VL	Katalanisch	BP	EP	Galisch	Spanisch
ǎ				a		
ā						
ĕ		ε				ie
ē				e		
ĭ						
ī				i		
ō			ɔ			ue
ō				o		
ŭ						
ū				u		

Tabelle 4: Die betonten Vokale vom Klassischen Latein bis zur Iberoromania.

Auch das Portugiesische (EP und BP) behielt die sieben betonten Vokale des Vulgärlateins mit den vier unterscheidenden Höhenstufen der Zungenposition bei (hohe [i, u], mittelhohe [e, o], mitteltiefe [ɛ, ɔ] und tiefe [a]). Außer bei dem tiefen Vokal ist die Struktur des Vokalraumes symmetrisch mit drei hinteren gerundeten [u, o, ɔ] und drei vorderen ungerundeten Vokalen [i, e, ɛ]. Akustisch weisen allerdings die hinteren Vokale etwas höhere F1-Werte auf als die vorderen Vokale (Escudero et al. 2009: 1391). Die betonten Vokale beider Varietäten unterscheiden sich jedoch durch weitere phonetische Eigenschaften: Insgesamt weisen BP-Vokale eine signifikant längere Dauer und einen größeren Vokalraum (sowohl für die 10 Männer als auch für die 10 Frauen der Studie) als die europäischen Vokale auf (vgl. Escudero et al. 2009: 1385). Die intrinsische F0 war bei den europäischen Vokalen höher und der vordere mitteltiefe Vokal /ɛ/ zeigte im EP sowohl einen höheren F1 als auch einen höheren F2 (vor allem bei den Frauen). Somit liegt /ɛ/ im EP im Vokalraum näher am vorderen mittelhohen Vokal /e/ als im BP, was ein künftiges Zusammenfallen dieser beiden vorderen Vokale im EP andeuten könnte (Costa 2004, Cruz-Ferreira 1995, Escudero et al. 2009: 1384, 1391).

3.2 Nasale Vokale

Bekanntlich hatte das Lateinische keine nasalen Vokale (Williams 1961). Die aktuellen portugiesischen Vokale entstanden aus eigentümlichen Schwächungsprozessen nasaler Konsonanten, die ziemlich früh in der Sprachgeschichte (ca. im 10. Jahrhundert) und ausschließlich im Galicisch-Portugiesischen stattfanden (Meier 1954).

Zum BP gehören folgende fünf nasale Vokale: [ĩ, ã, ẽ, õ, ă] (vgl. Barbosa & Albano 2004: 229) und zum EP ebenfalls fünf Vokale mit dem Unterschied, dass der tiefe Vokal zentralisierter ist: [ĩ, ã, ẽ, õ, ă] (vgl. Mateus & d'Andrade 2000: 46-48). Bei den nasalen Vokalen blieben drei der Höhenstufen der betonten Vokale mit Ausnahme der mitteltiefen ([ɛ, ɔ]) erhalten und sie können in betonter sowie in unbetonter Position vorkommen. Außerdem hat das Portugiesische als einzige romanische Sprache phonologische nasale Diphthonge wie in dem Wortpaar ersichtlich ['maw] *mau* 'böse' vs. ['māw] *mão* 'Hand'. Außer diesem Diphthong umfasst das Inventar noch [äe] wie in ['māj] *mãe* 'Mutter' und [õj] *põe* 'setzen, 3. Person Singular'.

Die Entstehung nasaler Diphthonge war diachronisch betrachtet ein eher komplexer und noch ungeklärter Prozess, der möglicherweise zwischen dem 10. und dem 12. Jahrhundert mit der Tilgung intervokalischer Nasale begann (Meier 1954) und sich im Laufe der Zeit auf weitere Kontexte ausbreitete. Die Entstehung nasaler Vokale kann auf die Reduzierung nasaler Konsonanten in den folgenden drei Kontexten (Oliveira 1999: 171-173, Parkinson 1983, Sampson 1999, Teixeira & Vaz 2001) zurückgeführt werden:

- (i) /VNV/ (Vokal-Nasal-Vokal): Die intervokalischen Nasale wurden nach der Nasalierung des vorausgehenden Vokals synkopiert (hermanu > irmão). Diese Synkope wurde ausschließlich für das

Galicisch-Portugiesische²³ beschrieben. Ihr Beginn wird in etwa auf das 10. Jahrhundert (evtl. bereits beginnend im 9. Jahrhundert) datiert und sie wird als im 12. Jahrhundert endgültig abgeschlossen eingeschätzt. Der Grund hierfür ist, dass es keine Belege für intervokalische Nasale in den ersten schriftlichen Texten (13. Jahrhundert) gibt.

- (ii) /VN#/ (Vokal-Nasal am Wortende): Die Nasale am Wortende führten ebenfalls zur Nasalisierung des vorhergehenden Vokals und wurden schließlich synkopiert. Dieser Prozess war zur Zeit der Entstehung der ersten Texte fast abgeschlossen (Sampson 1999).
- (iii) /VNC/ Vokal-Nasal-Konsonant: In mittelalterlichen Texten wurde die Synkope vor allem vor Frikativen beschrieben (*īnfanten* > *iffante*), diese Nasale wurde allerdings später wieder eingefügt.

Für einen ausführlichen Forschungsüberblick und eine synchrone artikulatorische Analyse der nasalen Vokale wird an dieser Stelle auf Oliveira (2009) und Martins (2014) verwiesen.

3.3 Unbetonte orale Vokale

Bereits im Vulgärlatein war die Betonung ein unterscheidendes Kriterium. Die fünf unbetonten oralen Vokale des Vulgärlateins werden im BP in prätonischer Position realisiert [i, u, e, o, a]. Allerdings werden diese Vokale in posttonischer Position stark zentriert und auf die folgenden vier Vokale reduziert [ɪ, ʊ, e̞, ø] (Barbosa & Albano 2004: 229, Everett 2011).

²³ Die Synkope intervokalischer Nasale – und auch intervokalischer Laterale /l/ – vollzog sich vor dem 13. Jahrhundert und fand nicht innerhalb des ‚Mozarabischen‘ Raums (heute Südportugal) statt. Daher blieben diese Segmente in manchen Begriffen arabischen Ursprungs und in der Toponymik des Alentejo und der Algarve bis heute erhalten (Williams 1961).

Im EP wurden diese Vokale unabhängig von der Position im Wort und von der Entfernung zur Hauptbetonung auf die folgenden vier Vokale [i, i̇, u, ɐ] reduziert. Das unbetonte System schrumpft in dieser Varietät auf drei hohe Vokale [i, i̇, u] und den tiefen Vokal [ɐ]. Leider ist keine aktuelle akustische Studie zu den unbetonten Vokalen dieser Varietät bekannt. Der aktuelle unbetonte Vokalismus wurde stark von der sogenannten Vokalerhöhung oder -reduzierung beeinflusst. Aus dem Vergleich zwischen betonten und unbetonten Vokalen geht hervor, dass die Vokalerhöhung zu einem Zusammenfall der mittleren Vokale führte, so dass sich das unbetonte System auf vier Vokale reduzierte: [i, i̇, u, ɐ] wie in [kazɐ – kɛziɲɐ] *casa, casinba* ‘Haus, Häuschen’ und [pɛdru – pidiɲu] *Pedro, Pedrinho* ‘Pedro, kleiner Pedro’ (Mateus & d’Andrade 2000, Mateus & Martins 1982, Freitas 2002).

Die früheste Datierung der Vokalerhöhung in der Literatur liegt im 17. Jahrhundert (Castro 1991). Mit der Vokal- und Konsonantenquantifizierung in geschriebenen Texten im Zeitraum von 1502 bis 1836 konnten Frota et al. (2008) eine starke Änderung im Sprachrhythmus auf ca. 1600 datieren. Die Änderung bestand nicht nur aus der Vokalerhöhung im EP, sondern führte darüber hinaus auch zu einem syntaktischen Wandel, der unter anderem einen Stellungswechsel der Klitika von proklitisch zu enklitisch zur Folge hatte. Diese Änderung des Sprachrhythmus war so zentral in der Entwicklung des Portugiesischen, dass sie als Grundstein der Ausdifferenzierung beider Varietäten betrachtet wird (Frota et al. 2008).

3.4 Einfügung und Auslassung von Vokalen

3.4.1 Vokalepenthese oder Svarabhakti Vokal im Portugiesischen

Bei der Produktion von Konsonantenclustern wurde ein intermediäres vokalisches Element beschrieben, das entweder als vokalisches Element (Silva 1996), als epenthetischer (Collischonn

2003, Carvalho 2004, Silveira & Seara 2009) oder als svarabhaktischer Vokal (Nishida 2009) bezeichnet wurde. Dieser Vokal kann sowohl in tauto- (Silva 1996, Nishida 2009) wie in heterosilbischen Clustern vorkommen (Veloso 2007a, Mateus & d'Andrade 2000, Bisol 1999, u.a.). In tautosilbischen Clustern zeigt der Vokal nach Plosiven die gleichen Formantenstrukturen wie der folgende Vokal, ist allerdings nach Frikativen zentriert (Nishida 2009: 34-35, 47-48). Dieser Vokal wird im BP als [i] oder [ɨ] realisiert, die durch die hohe Position der Zungenspitze bei der Verschlusslösung eines apikalen Konsonanten begünstigt werden (Barbosa & Albano 2004: 227). Eine andere Erklärungsmöglichkeit wäre, dass der epenthetische Vokal, abhängig vom darauffolgenden Vokal, sowohl [i] als auch [e] sein kann (Bisol 1999). Im EP ist der epenthetische Vokal seltener als im BP (vor allem bei älteren Sprechern aus ländlichen Gegenden und bei Kindern in der Erstspracherwerbsphase) und der beschriebene Vokal ist immer [ɨ] (Freitas 1997, Veloso 2007a, 2003).

Im portugiesischen Kontext wird der epenthetische Vokal mit einer neuen Silbe assoziiert (Bisol 1999, Girelli 1988, Mateus & d'Andrade 2000, Veloso 2007a), so dass aus lexikalischen Clustern CVC-Sequenzen entstehen. Es gibt allerdings ebenfalls die Möglichkeit, solche vokalischen Elemente als keine eigenständigen lexikalischen Segmente zu betrachten. In diesem Sinne entstehen sie durch ein spezifisches Timing in der artikulatorischen Organisation, das in der Perzeption als Vokal wahrgenommen wird, ohne dass ein Vokal in der lexikalischen Repräsentation vorhanden ist (z.B. Hall 2003: 1-3, Buchwald et al. 2007, Ohala 1992: 331).

3.4.2 Svarabhakti Vokale im romanischen Vergleich

Svarabhakti Vokale oder intermediäre vokalische Elemente wurden ebenso in anderen Sprachen dokumentiert. Sie kamen im Sanskrit (Allen 1953) sowie im späten Vulgärlatein (Schuchardt 1867: 421) und im romanischen Kontext vor, z.B. im Spanischen (Bradley & Schmeiser 2003: 1-3, Almeida & Dorta 1993: 106-107, Quillis 1993: 330).

Außerdem tritt das vokalische Element bei Clustern in diesem Kontext in zahlreichen weiteren europäischen Sprachen auf (für das Englische: Davidson & Stone 2004, das Irische: Ó Siadhail 1989: 21-22, das Gälische: Borgström 1949: 153, Bosch 1995, das Deutsche: Jannedy 1994: 113, Geng et al. 2010, das Niederländische: Warner et al. 2001: 395, Booij 1995: 127, das Polnische: Kul 2010, das Finnische: Harrikari 1999: 6) sowie in nicht-europäischen Sprachen (z.B. Hebräisch: McCarthy 1979: 34, Lakhota: Albright 1999 u.a.). Für einen aktuellen Überblick und die Analyse des svarabhakti Vokals in heterosilbischen Clustern im Sprachvergleich siehe Hall (2003, 2006).

3.4.3 Vokaltilgung im EP

Martins (1975) führte die erste empirische Studie zur Vokaltilgung durch. In einer akustischen Studie ließ sie 40 Sätze mit jeweils 20 Phonemen von einem einzelnen Sprecher aus Lissabon vorlesen. Die Sätze waren mit allen Segmenten des Portugiesischen ausgestattet und wurden auf zwei Ebenen transkribiert: auf der ersten Ebene mit allen phonologisch zu erwartenden und auf der zweiten Ebene mit den phonetisch realisierten Lauten. Aus dem Vergleich beider Ebenen wurde der Anteil der Tilgungen errechnet, indem die Anzahl der produzierten Vokale von der Anzahl der erwartbaren lexikalischen Vokale subtrahiert wurde. Die Ergebnisse zeigen, dass von den 394 auf der ersten Ebene erwarteten Vokalen 49 nicht realisiert wurden, was insgesamt einer Tilgung von 11% entspricht. Die nasalen Vokale [ĩ, u, e, o, a] und die Halbvokale [j,w] zeigten keine Tilgung (0%). Der am häufigsten elidierte Vokal war [u], an zweiter Stelle kam mit großem Abstand [ĩ], gefolgt von [a]. Leider wurde die Gesamtanzahl der einzelnen Vokale nicht angegeben, so dass ihr Tilgungsanteil nicht berechnet werden konnte.

In einer späteren Perzeptionsstudie gaben die Probanden bei 16 Minimalpaaren, unter anderem *prece* vs. *perce* 'gebet vs. untergehen' und *crer* vs. *querer* 'glauben vs. wollen', an, wie viele Silben sie bei jedem Stimulus wahrnahmen (2 vs. 3 und 1 vs. 2). Die Ergebnisse zeigten, dass die Teilnehmer *prece* vs. *perce*

identifizieren konnten (von den 24 Stimuli jedes Wortes wurden 19 Mal drei und 20 Mal zwei Silben erkannt). Für die velaren Stimuli (*crer* vs. *querer*) wurde als Antwort meist ‘eine Silbe’ ausgewählt, was *crer* entspricht, und zwar für beide intendierten Wörter (20 und 23 Antworten bei jeweils 24 Stimuli). Die Autoren folgern daraus, dass zum Zeitpunkt der Studie *crer* und *querer* in der Sprachperzeption nicht mehr unterschieden werden konnten (Mateus & Martins 1982: 177-179).

Weiter dokumentiert wurde die Tilgung im Portugiesischen im Rahmen einer Studie auf der Azoreninsel *Faial*. Für diese Untersuchung wurden etwa 22 Minuten Konversation einer 42-jährigen Sprecherin aus einem informellen Dialog zwischen ihr und einer anderen Sprecherin derselben Varietät aufgenommen. Es wurde die Anzahl der produzierten und der getilgten Vokale erhoben. Dabei zeigte /i/ den größten Anteil von Tilgungen sowohl in finaler als auch in nicht-finaler Position, gefolgt von /u/. Der Vokal /ɐ/ wurde hingegen in den meisten Fällen realisiert (Silva 1997: 303). Eine weitere Studie analysierte Monologe zum Thema ‘letzter Urlaub’ von vier weiblichen Sprecherinnen ohne Migrationshintergrund im Alter zwischen 22-24 Jahren die einen akademischen Abschluss erlangt haben (Cunha 2011a). Es wurden 3,5 bis 4,5 Minuten jeder Teilnehmerin transkribiert (insgesamt 15 Minuten). Wie bei Martins (1975) wurden die Daten auf zwei Ebenen transkribiert – eine für die phonologischen Segmente und eine für die phonetisch realisierten Laute. Der Anteil der Tilgung wurde aus dem Vergleich beider Ebenen errechnet. Die Gesamtergebnisse zeigen ein klar geteiltes Bild: Einerseits werden [i, u] in über 80% der phonologisch erwarteten Positionen getilgt, während [ɪ, ɐ, a, ɛ, ɔ] eine Tilgung von unter 10% aufweisen (Cunha 2011a).

3.4.4 Vokaltilgung in anderen Sprachen

Vokaltilgungen und Vokalschwächungen oder -entstimmungen wurden auch in zahlreichen weiteren romanischen Sprachen festgestellt, nämlich im Französischen (Pustka 2011, 2007, Torreira & Ernestus 2011, Krötsch 2004), im Italienischen

(Garrapa 2011), im Rumänischen (Krefeld 1987), im Katalanischen (z.B. Prieto 2004: 195-199) und in mehreren Spanisch-Varietäten aus Lateinamerika (in Mexiko: Serrano 2006, Lope Blanch 1963, Boyd-Bowman 1952, Matluk 1952; in Peru: 2010, 2009, 2008; in Bolivien: Gordon 1980 und in Ecuador: Lipski 1989, 1990, 1995).

Außerdem wurden sie in einer Reihe weiterer europäischer Sprachen beschrieben, beispielsweise im Englischen (Davidson 2006, Manuel et al. 1992) und im Deutschen (Szczepaniak 2007, Nübling et al. 2006: 32-34, Ramers & Vater 1995: 47-66).

3.4.5 Entstehung weiterer Konsonantencluster

Eine Folge der häufigen Vokaltilgung ist die Erhöhung der Anzahl der realisierten Konsonantencluster in der gesprochenen Sprache. Daraus geht die interessante Frage hervor, inwiefern sich lexikalische und nach Vokaltilgung entstandene (postlexikalische) Konsonantencluster unterscheiden, was nicht minder interessant für die synchrone Beschreibung und die Entwicklung der Sprache ist.

Für das Portugiesische zeigten Mateus & Martins (1982: 177-179) in einer Perzeptionsstudie, dass die Wortpaare *crer* und *querer* seltener differenziert werden können als *prece* vs. *perece*. So waren Unterschiede zwischen lexikalischen und postlexikalischen Clustern, oder anders ausgedrückt, zwischen lexikalischen Clustern vs. CVC-Sequenzen, bereits in den 80er Jahren teilweise neutralisiert. In einer akustischen Studie zu Plosivclustern im Vergleich zu den entsprechenden CVC-Sequenzen mit Wortpaaren wie *adaptaram* 'anpassen' vs. *atapetaram* 'Teppich legen' zeigten sich zwischen bilabialem und alveolarem Plosiv in beiden Lexemen keine Stimmhaftigkeit oder Formanten, allerdings war die Verschlussdauer des bilabialen Segments länger in /pít/ als in /pt/ (Andrade 1993).

3.4.6 Entstehung komplexer Konsonantengruppen

Die optionale verbreitete Tilgung der unbetonten Vokale /i, u/ in allen möglichen Wortpositionen hat einen raschen Anstieg der Konsonantengruppen in Anzahl und Komplexität zur Folge. So entstehen eine massive Anzahl von Homophonen wie ['kɾɐ] für *crer* 'glauben' und *querer* 'wollen' sowie reine Konsonantengruppen, die sich aus zwei bis sechs Konsonanten in Folge zusammensetzen können.

Diese Gruppen können aus allen möglichen Kombinationen der Konsonanten gebildet werden und treten in allen Wortpositionen auf. Eine kleine Auswahl an Beispielen in initialer Position aus der Literatur folgt zur Illustration:

	Cluster	Transkription ²⁴	Beispiel	D-Glosse
Zwei Kons.	[pd]	[pdír]	pedir	'bitten'
	[dp]	[dpɛdɛ́ɾ]	depender	'abhängen'
	[ds]	[dsifáf]	decifrar	'entziffern'
	[sd]	[sdiádu]	sediado	'ansässig'
	[sp]	[spɛrar]	separar	'trennen'
	[ʃp]	[ʃpadɐ]	espada	'Schwert'
	[sg]	[sgúru]	seguro	'Versicherung'
	[ʃg]	[ʃgár]	chegar	'ankommen'
	[mt]	[mtɛ́ɾ]	meter	'stecken'
	[tm]	[tmɛ́ɾ]	temer	'fürchten'

²⁴ Die Betonung wurde wie bei Mateus & d'Andrade (2000: 39) einheitlich durch einen graphischen Akzent (´) gesetzt, da in manchen der Konsonantengruppen die einzelnen Silben nicht definiert werden können.

Drei Kons.	[tʃ]	[tʃõn]	telefone	‘Telefon’
	[mʃs]	[mʃsér]	merecer	‘verdienen’
	[dvd]	[dvdór]	devedor	‘Schuldner’
	[dʃn]	[dʃnikár]	depenicar	‘naschen’
	[Rmm]	[Rmmurár]	rememorar	‘gedenken’
Vier Kons.	[dʃpg]	[dʃpgár]	despegar	‘versetzen’
	[dʃps]	[dʃpsér]	deperecer	‘untergehen’
Fünf Kons.	[dʃprg]	[dʃprgár]	despregar	‘annageln’
	[dʃprd]	[dʃprdisár]	desperdiçar	‘verschwenden’
Sechs Kons.	[dʃprvn]	[dʃprvnír]	despre- venido	‘unvor- bereitet’
	[dʃprʃt]	[dʃprʃtizjár]	despresti- giar	‘herab- würdigen’

Tabelle 5: Beispiele von komplexen Konsonantengruppen (aus Mateus & d’Andrade 2000: 43-44, Schwarzinger 2006: 39-42, mit eigenen Beispielen ergänzt).

Aufgrund ihrer Vielfaltigkeit und Häufigkeit haben diese Sequenzen die Aufmerksamkeit mehrerer Studien auf sich gezogen (Magalhães 2009, Schwarzinger 2006, Mateus & Andrade 2000: 43-46). Für weitere Beispiele und ihre Häufigkeit kann Schwarzinger (2006: 31-42) konsultiert werden. Die phonologische Analyse solcher Folgen wurde bislang anhand der lexikalischen Form mit dem Vokal durchgeführt, der phonetisch realisiert werden kann oder nicht (Mateus & d’Andrade 2000: 43-48). Allerdings verfehlt diese Analyse die Kernfrage der portugiesischen Phonologie und zwar, wie können beide portugiesischen Varietäten die gleiche phonologische Struktur besitzen, wenn sie sich so unterschiedlich anhören?

Mit dieser Frage haben sich auch anderen Forscher beschäftigt: Angesichts ihrer Frequenz in der gesprochenen Sprache, ihrer Häufigkeit in den Sprachen der Welt und ihrer Existenz bereits im Klassischen Latein argumentiert Freitas (1997) dafür, dass /ʃ/ plus Plosiv ([ʃgár] *chegar* ‘ankommen’ und [ʃkádø] *escada* ‘Leiter’) im EP lexikalische Cluster sind, in denen die historische Epenthese zurückgenommen wurde.

Unabhängig davon, wie diese Konsonantengruppen analysiert werden, sind sie im Hinblick auf ihre Vielfalt sehr interessant. Wie an den in der Tabelle 5 abgebildeten Beispielen ersichtlich, sind kaum Einschränkungen bei der Zusammenstellung dieser Gruppen vorhanden: Konsonantengruppen mit gemischter Stimmhaftigkeit sind im Gegensatz zu lexikalischen Clustern hier nicht selten der Fall. Es können beliebige Artikulationsorte gekoppelt werden ([ʃpádø] *espada* ‘Schwert’ vs. [spɛrar] *separar* ‘trennen’ vs. [ʃgár] *chegar* ‘ankommen’ vs. [sgúro] *seguro* ‘Versicherung’) und Konsonanten können in beiden Reihenfolgen aufeinander folgen ([pdír] *pedir* ‘bitten’ vs. [dpɛdér] *depender* ‘abhängen’). Außerdem können bei diesen Gruppen Artikulationsarten kombiniert werden, die in lexikalischen Clustern nicht vorkommen, und dies ebenso in beiden Richtungen (z.B. [tmér] *temer* ‘fürchten’ vs. [mtér] *meter* ‘stecken’).

Je größer die Anzahl von Elementen, desto geringer ist ihre Häufigkeit, so dass Konsonantengruppen mit fünf oder sechs Konsonanten als extrem selten einzustufen sind (Schwarzinger 2006: 31-42). Außerdem kann die Vokaltilgung zwar in allen Silbenpositionen vorkommen, so dass alle diese Formen theoretisch möglich sind. Allerdings sind empirische Belege eher selten und es fehlt jede genauere sowohl akustische als auch artikulatorische Beschreibung. Aufgrund ihrer besonderen Relevanz für die synchrone Sprachbeschreibung kann man nur hoffen, dass künftige Forschungsvorhaben diese Lücken schließen können.

4 Konsonantenkoordination mit Liquiden

4.1 Einführender Forschungsstand

Konsonantencluster können mit einem intermediären Vokal realisiert werden, wie beispielsweise [pilãtə] für *planta* ‘Pflanze’ im EP vgl. Veloso, 2003, [ĩngala'tera] für *Inglaterra* ‘England’, [ege'lesja] für *eglesia* ‘Kirche’, [co'ronika] für *cronica* ‘Chronik’ im Spanischen (vgl. Menendez-Pidal 1926: 127) oder [belo] für *blow* im Englischen (vgl. Browman & Goldstein 1990a, b, Ohala 1992: 323 1999, Price 1980). Allerdings wurde ebenso die gegensätzliche Tendenz festgestellt und zwar, dass manche lexikalisch unbetonten Vokale (vor allem das zentrierte Schwa /ə/) geschwächt, gekürzt oder getilgt werden können und infolgedessen lexikalische CVC-Sequenzen als reine Konsonantenabfolgen (CC) beschrieben werden. Diese Tendenz wurde in zahlreichen Sprachen festgestellt, ebenso im romanischen Kontext, z.B. im Französischen (Pustka 2011, 2007, Torreira & Ernestus 2011), im Italienischen (Garrapa 2011), im Rumänischen (Krefeld 1987), im Katalanischen (z.B. Prieto 2004: 195-199) und im Spanischen (Serrano 2006, Lope Blanch 1963, Boyd-Bowman 1952, Delforge 2009, 2008a, b²⁵).

Mit diesen Phänomenen werden Prozesse wie svarabhaktischer, epenthetischer Vokal oder Tilgung assoziiert, die diese Phänomene nur teilweise erfassen. Nun strebt das folgende Kapitel an, anhand von akustischen und artikulatorischen Daten aus zwei portugiesischen Varietäten die Realisierungen dieser Formen genauer zu beschreiben. Insbesondere ist von Interesse, inwiefern sich die Organisation von Konsonanten in lexikalischen Clustern und CVC-Sequenzen in der Produktion der Sprache

²⁵ Siehe 3.4 und Hall (2003) für einen aktuellen, ausführlichen Überblick.

unterscheidet und inwiefern sich dabei die Reduzierung des lexikalischen Kontrastes auf die Variabilität in der zeitlichen oder räumlichen Koordination der Konsonanten im Rahmen der Artikulatorischen Phonologie zurückführen lässt.

Nach einer Einführung in den theoretischen Rahmen und der Vorstellung der genaueren Fragestellungen und Hypothesen folgt unter Punkt 4.2 die Methode. Die Ergebnisse und ihre Diskussion werden in zwei Blöcken präsentiert: Der erste beschäftigt sich mit der zeitlichen intergestischen Koordination (4.3) und der zweite mit der räumlichen intragestischen Koordination der Zungenspitze (4.4). Abschließend werden Schlussfolgerungen gezogen und ein Ausblick gegeben.

4.1.1 Konsonantencluster und Sonoritätshierarchien

Die Koordination von Konsonantenclustern und CVC-Sequenzen beruht auf der zeitlichen und räumlichen Artikulation der einzelnen Segmente im Vokaltrakt und konkret darauf, wie genau ein Segment in Bezug auf die benachbarten Segmente produziert wird. Auf den ersten Blick befindet sich dies im Bereich der Sprachproduktion und der artikulatorischen Phonetik (was auch der Fall ist). Allerdings bildet die Sprachproduktion nicht nur einen Teilbereich der Phonetik, sondern ist ebenso die Grundlage für die phonologischen Repräsentationen im Rahmen der Artikulatorischen Phonologie.

Bestehende Zusammenhänge zwischen Phonetik und Phonologie sind keine rezente Feststellung: Bereits Trubetzkoy (1939) führte den Begriff des Phonems als Abgrenzung zu den tatsächlich realisierten phonetischen Sprachlauten ein und legte somit den Grundstein für die Herausbildung der Phonologie aus dem Bereich der Phonetik. Die Phonetik widmet sich seither der materiellen Seite der sprachlichen Kommunikation, d.h. wie Schallereignisse artikulatorisch zustande kommen, wie die akustischen Vorgänge der Schallproduktion und -übertragung entstehen und beschrieben werden können und wie sie über das Gehör verarbeitet werden. Die Phonologie andererseits abstrahiert

von dieser materiellen phonetischen Realisierung der Sprachlaute und betrachtet diese in ihrer Relation zu den anderen Sprachlauten innerhalb eines Sprachsystems, d.h. jeder abstrahierte Sprachlaut unterscheidet sich in seiner bedeutungstragenden Funktion von allen anderen Lauten eines Sprachsystems.

Trotz des klaren Zusammenhangs zwischen Phonetik und Phonologie ist der Bezug beider Disziplinen im Hinblick auf gewisse Phänomene nicht immer eindeutig: Ein gutes Beispiel dafür ist die Definition von Konsonantenclustern oder die Frage, warum bestimmte Konsonantensequenzen synchron und/oder diachron im Vergleich zu anderen Sequenzen bevorzugt werden. Eine verbreitete Erklärung bietet die sogenannte Sonoritätshierarchie, deren ursprüngliche Idee bis Jespersen (1904: 185-189), Sievers (1893) oder Whitney (1874) zurückverfolgt werden kann, so dass deren Ursprung älter als die Phonologie selbst ist. Diese Hierarchie wurde im Rahmen mehrerer Theorien weiterentwickelt: in der Autosegmentalen Phonologie (Selkirk 1986), in der CV-Phonologie (Clements & Keyser 1983, McCarthy 1981, 1986) und in der Optimalitätstheorie (McCarthy & Prince 1993, McCarthy 2003, Kager 1999). Manche Autoren haben sie leicht abgeändert oder ergänzt: Ein Beispiel ist die Skala der Konsonantischen Stärke von Vennemann (1978, 1987, 1988), wobei die Sonorität das eine Extrem und die Konsonantische Stärke das andere Extrem der Skala darstellen (Vennemann 1988, Restle & Vennemann, 2001: 1312).

Die Aufteilungen dieser Skala in stimmhafte und stimmlose Plosive und Frikative oder zentrale und laterale Liquide trugen dazu bei, dass eine größere Anzahl von möglichen Silbenstrukturen in den Sprachen der Welt ausgeschlossen oder präzisiert werden und einer größeren Anzahl von Lautwandelprozessen erklärt werden konnte (vgl. Vennemann 1988).

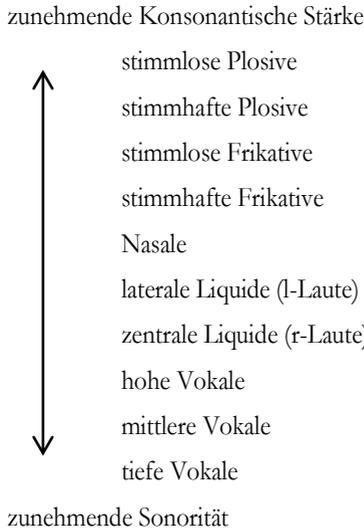


Abbildung 1: Skala der Sonorität / Konsonantischen Stärke (nach Restle & Venemann, 2001: 1312).

Dabei wird eine Silbe durch steigende Sonorität und sinkende konsonantische Stärke vom Anlaut bis zum Nukleus sowie fallende Sonorität und steigende Konsonantische Stärke vom Nukleus bis zum Auslaut definiert (Selkirk 1986, Mateus & Andrade 2000, Venemann 1988, Restle & Venneman 2001). Konsonantencluster besitzen demnach eine steigende Sonorität, wenn sie im Anlaut von Silben auftreten, und eine fallende Sonorität im Auslaut. Beispiele aus dem Deutschen hierfür wären die Verbform *brennst* oder das Substantiv *Pfalz*, bei denen die Anlautcluster /br/ und /pf/ eine steigende und die Auslautcluster /ts/ und /nst/ eine sinkende Sonorität aufweisen.

Solche Hierarchien sind hilfreich, um zahlreiche Konsonantencluster zu beschreiben, allerdings berücksichtigen sie häufige phonotaktische Begebenheiten nicht, beispielsweise, dass /w/ selten nach bilabialen und oft nach velaren Konsonanten auftritt (Ohala 1992: 321-323). Als Beispiel findet man im romanischen Raum häufig /wV/ nach velarem Konsonant (/k, g/)

(*quando, quanto, quatro, guarda* ‚wann, wieviel, vier, Wächter‘ im Portugiesischen, *cuando, cuanto, cuatro, guardia* im Spanischen, *quand, quant, quatre* im Französischen, *quando, quanto, quattro, guardia* im Italienischen), während /wV/ nach Bilabialen (/p, b/) selten und in weniger Sprachen vorhanden ist (*puerto* ‚Hafen‘ im Spanischen, *puer* ‚stinken‘ im Französischen, *poate* ‚kann‘ und *boala* ‚Krankheit‘ im Rumänischen). Selbst von der Sonoritätshierarchie ausgeschlossene Verbindungen wie beispielsweise Frikativ-Plosiv-Cluster sind u.a. in den germanischen Sprachen besonders verbreitet (im Englischen *stone, statue*, im Deutschen *Stein, Statue*, im Niederländischen *steen, standbeeld*, im Norwegischen *stein, statue*, usw.).

Es ist außerdem unklar, welche phonetische Einheit solchen Hierarchien zugrunde liegt (Price 1980). Ihr Erklärungsmodell ist in gewisser Weise zirkulär, da weder Sonorität noch Silbe empirisch überprüfbar sind und sich gegenseitig definieren. Als letzter Kritikpunkt sei hier angeführt, dass sie nicht in der Lage sind, einige Lautwandelprozesse bei Clustern zu erklären, z.B. warum lateinische /l/-Cluster im Gegensatz zu /r/-Clustern anfälliger für Lautwandelprozesse in der Entwicklung hin zum Romanischen waren (s. Jensen 1999: 173-181 für einen ausführlichen Vergleich), d.h. warum Obstruent-l-Cluster häufiger zu Obstruent-r-Clustern gewandelt wurden (Rhotazismus) als vice versa, d.h. /r/ zu /l/ (cf. Müller 2011: 127-148, für einen Überblick zu den romanischen Sprachen; Proctor 2009 für die Variation bei Liquiden im Spanischen und Russischen).

4.1.2 Die dynamische Grundlage der Artikulatorischen Phonologie

Ein alternatives Modell für die Definition von Silbe und Konsonantencluster ist die in den 80er Jahren entwickelte Artikulatorische Phonologie (Browman & Goldstein 1984, 1986, 1988, 1990, ff.). In diesem dynamischen Modell werden Sprachlaute durch die Bewegung des für einen bestimmten Laut relevanten Artikulators im Vokaltrakt in dem entsprechenden temporalen Abschnitt (Geste) beschrieben. Es geht hier also

darum, für jedes Segment den für seine Produktion relevanten Artikulator festzustellen, zum Beispiel wird die Zungenspitze als relevanter Artikulator für apikale und das Zungendorsum für velare Laute genutzt. Die Bewegung dieses Artikulators wird durch das Kontinuum der Gesamtäußerung eingeschränkt (Geste) und das erkannte Muster durch den Hörer abstrahiert. In diesem Zusammenhang werden Laute durch die entsprechende dynamische Bewegung definiert (Geste), ihr Verlauf zu einem Verlaufsmuster (*pattern*) abstrahiert und diese abstrahierten Muster als Grundlage der phonologischen Repräsentation des Segments betrachtet.

Gestures are units of action that can be identified by observing the coordinated movements of vocal tract articulators. That is, repeated observations of the production of a given utterance will reveal a characteristic pattern of constrictions being formed and released. The fact that these patterns of (discrete) gestures are similar in structure to the nonlinear phonological representations being currently postulated (Browman & Goldstein, 1989: 69).

In diesem Sinne sind Segmente nicht mehr statisch und linear zu betrachten (wie etwa in einem Spektrogramm), sondern dynamisch durch die Aktivität der verschiedenen Artikulatoren im Vokaltrakt. Für jedes Segment wird der dynamische Verlauf auf einen bestimmten Artikulator eingeschränkt. Aus der Beobachtung und Abstraktion von mehreren Wiederholungen eines Segments und dem Vergleich der einzelnen Verläufe (Gesten) entsteht ein Muster. Dieses dient als Grundlage für die phonologische Repräsentation dieses Segments. Die kritische Frage, inwiefern solche Gesten natürlich sind, lässt sich mit zwei starken Argumenten beantworten: Einerseits definiert das Modell die Geste als eine dynamische und zielorientierte Bewegung, die nur eine geringe Variabilität ermöglicht, da sie sonst Gefahr liefe, das Ziel zu verfehlen. Andererseits sind solche Gesten bei Kindern im Vorspracherwerb (in der sogenannten Lala-Phase) bereits vorhanden (Browman & Goldstein 1989: 72).

Da die Gesten einen grundlegenden Mechanismus der Sprachproduktion nachbilden, der in der Produktion aller Laute unentbehrlich ist, ist das gestenbasierte Modell der Sprache in sich sehr mächtig: Es bietet einen optimalen Rahmen für die synchrone Beschreibung von Variation innerhalb einer Sprache oder einer Varietät sowie zwischen Sprachen oder Varietäten. Des Weiteren findet es ebenfalls Anwendung in der Analyse von pathologischer Sprache und in der historischen Sprachforschung insbesondere zum Lautwandel.

The use of gestures, rather than constellations of gestures as in Root nodes, as basic units of description makes it possible to characterize a variety of language patterns in which gestural organization varies. Such patterns range from the misorderings of disordered speech through phonological rules involving gestural overlap and deletion to historical changes in which the overlap of gestures provides a crucial explanatory element (Browman & Goldstein, 1989: 72).

In diesem Zusammenhang sind die vielen Arbeiten über Cluster mit Lateralen ein gutes Beispiel dafür, wie im Rahmen der Artikulatorischen Phonologie Variation erklärt werden kann. Obwohl Cluster aus Plosiv und Lateral (z.B. /pl/ in *Planet* oder /kl/ in *Klausur*) den angesprochenen Sonoritätsskalen folgen, zeigen sie sowohl synchronisch als auch diachronisch Variation: Diese Cluster werden mit klarem²⁶ Lateral in den Sprachen Englisch (Hardcastle 1985, Gibbon et al. 1993, Marin & Pouplier 2008, 2010), Deutsch (Gibbon et al. 1993, Bombien et al. 2006, Bombien 2011), Italienisch und Schwedisch (Gibbon et al. 1993),

²⁶ Die Unterscheidung zwischen klarem und dunklem Lateral wird an die englische Terminologie von *clear* /l/ und *dark* /ɫ/ angelehnt (Recasens 1996, Recasens et al. 2010). Von der Unterscheidung *velarisiert* vs. *nicht velarisiert* wird in anderen Arbeiten ebenso Gebrauch gemacht. In der vorliegenden Arbeit wird allerdings Ersteres bevorzugt, da dunkle Laterale sowohl durch die Erhöhung des Zungendorsums zum weichen Gaumen (Velarisierung) hin als auch zum Pharynx hin (Pharyngalisierung) produziert werden können (Recasens 1996: 64) und wir uns hier nicht auf eine Variante festlegen möchten.

Französisch (Fougeron 1999, Kühnert & Hoole 2006) und Griechisch (Nicolaidis 2001) produziert, jedoch mit dunklem /ɫ/ in verschiedenen katalanischen Dialekten (Gibbon et al. 1993, Recasens & Espinosa 2005), ebenfalls im Englischen (Marin & Pouplier 2008, 2010) und im europäischen Portugiesisch (Martins et al. 2011, Oliveira et al. 2011, 2012, Teixeira et al. 2011). Historisch wurde das lateinische Cluster /pl/ zu [ʃ] im Portugiesischen (z.B. das lateinische *pluvia* wurde im Portugiesischen zu [ʃuvɐ] *chuva* ‘Regen’) und zu [ʎ] im Spanischen (z.B. *pluvia* > [ʎuβja] *lluvia* ‘Regen’). Spätere Übernahmen aus dem Lateinischen unterlagen einem anderen Wandel zum Portugiesischen (auch zum Sardischen, zum Okzitanischen, usw., vgl. Schmeiser 2009a, Müller 2011). Im Portugiesischen wurde das lateinische /pl/ zu [pʃ] (Rhotazismus, z.B. *placēre* > [pʃɛzɐ] *prazer* ‘Vergnügen’, Boyd-Bowman 1954: 110-112, Williams 1968: 74-75). Akustische und artikulatorische Analysen beider Lateraltypen zeigten, dass das klare /l/ einen kurzen Verschluss und einen alveolaren Artikulationsort aufweist, während dunkle laterale /ɫ/ einen signifikant längeren Verschluss und einen dentalen oder weiter frontierten Artikulationsort haben (Recasens & Farnetani 1994: 198, Recasens & Espinosa 2005: 20). Der Grund für Letzteres ist, dass die Verengung des Zungendorsums stärker bei dunklen Lateralen ausgeprägt ist als bei klaren (Sproat & Fujimura 1993: 307) und die Erhöhung des Zungendorsums eine Frontierung der Zungenspitze hin zur dentalen Region (oder sogar weiter anterior) hervorbringt (Recasens 2009: 478).

4.1.3 Definition einer Geste

Die Bewegung eines Artikulators impliziert eine inhärente Geschwindigkeit, d.h. ein Artikulator, wie jedes Objekt in Bewegung, besitzt eine Geschwindigkeit, die ungleich null ist. Ist die Geschwindigkeit null, bewegt sich der Artikulator bzw. das Objekt nicht. Standpunkte der Bewegung sind sehr hilfreich für die Eingrenzung der Gesten, da jede Geste zwischen zwei Standpunkten stattfindet.

Der Gestenverlauf besteht somit aus einem Anfangspunkt A_1 , einer Zielkonfiguration B_1 sowie dem Rückweg von dieser Zielkonfiguration B_2 zum Anfangspunkt A_2 ²⁷. Ähnlich wie ein Fußballspieler, der von Pfosten A_1 zu Pfosten B_1 läuft und anschließend wieder zurück von B_2 zu A_2 . Beim Anfangs- und Endpunkt A_1 und A_2 ist die Geschwindigkeit gleich null, und somit grenzt man zwischen diesen beiden Null-Geschwindigkeiten den Verlauf ein. Man kann den Gesamtverlauf in Hinweg und Rückweg einteilen: Der Abschnitt von A_1 zu B_1 entspricht der Verschlussgeste und der Abschnitt von B_2 zu A_2 der Öffnungsgeste.

Diese Bewegungsverläufe im Vokaltrakt können räumlich, aber auch zeitlich beschrieben werden, denn alle A- und B-Punkte entsprechen definierten Zeitpunkten. Die zeitliche Beschreibung der Verläufe hat den großen Vorteil, dass sie auch bei unterschiedliche Artikulatoren anwendbar ist, da alle Gesten die gleiche interne Struktur von Anfangspunkt A_1 zu Zielpunkt B_1 haben und dann von Zielpunkt B_2 zum Ausgangspunkt der Geste A_2 zurück. Da die Zeit linear ist und nicht rückwärts läuft, entsprechen die festgelegten räumlichen Punkte vier unterschiedlichen Zeitpunkten (Gafos et al. 2010: 657; Kühnert et al. 2006).

- (i) START (A_1) – der Zeitpunkt, ab dem sich der Artikulator zum Verschluss hin bewegt;
- (ii) TARGET (B_1) – der Zeitpunkt, ab dem der Verschluss gebildet wird;
- (iii) LÖSUNG (*Release*, B_2) – der Zeitpunkt, ab dem der Verschluss gelöst wird;
- (iv) ENDE (A_2) – der Zeitpunkt, an dem der Artikulator zur Anfangsposition gelangt.

²⁷ Warum jeder Punkt zwei Benennungen zu haben scheint, wird im folgenden Absatz deutlich werden

Die so definierten zeitlichen Markierungspunkte der Geste (*landmarks*) grenzen die interne Struktur der artikulatorischen Gesten und die entsprechenden phonologischen Repräsentationen ab. In diesem Sinne entsteht der labiale Plosiv /p/ in der Sprachproduktion durch die Schließungsbewegung beider Lippen, die den Verschluss bilden, und die darauffolgende Öffnungsbewegung. Das Phonem /p/ wird folglich ab dem Zeitpunkt definiert, ab dem die Lippen anfangen, sich in Richtung Verschluss zu bewegen (Geschwindigkeit ist null vor der Geste), über den Zeitpunkt, an dem sie den Verschluss bilden und den Verschluss wieder lösen, bis zu dem Zeitpunkt, an dem die Lippen die Öffnungsbewegung abschließen und sich nicht mehr bewegen (die Geschwindigkeit ist am Ende der Geste wieder null).

Nach der Definition der Gesten als Grundeinheit kann das Modell sich den Beziehungen zwischen Gesten widmen, also die Frage beantworten, in welchem Verhältnis die Gesamtgeste zu ihren Bestandteilen steht, z.B. wie lange der Verschluss in Bezug auf die Gesamtgeste dauert. Darüber hinaus werden die Verhältnisse verschiedener Gesten zueinander analysiert, beispielsweise, zu welchem Zeitpunkt innerhalb einer CV- oder CCV-Sequenz der Vokal im Vergleich zum angrenzenden Konsonanten beginnt. Derartige Fragen kann man mithilfe der Koordination von Gesten beantworten.

4.1.4 Koordination von CVC-Sequenzen und Clustern

Durch die Beteiligung verschiedener gekoppelter Artikulatoren nach dem dynamischen Modell von Saltzman (*Task Dynamics*, Saltzman & Kelso 1987, Saltzman & Munhall 1989) werden Laute nicht isoliert produziert, sondern mit gleichzeitigen oder überlappenden Bewegungen, in Abhängigkeit davon, welche Artikulatoren beteiligt sind und welche Kategorie von Segment erzielt wird. Bei der Produktion von Konsonanten und Vokalen (CV-Sequenzen) stellte Öhman (1967) fest, dass beide Bewegungen gleichzeitig beginnen. Der Grund dafür ist, dass die Vokalproduktion keine starke Konstriktion des Vokaltrakts bedingt und die akustischen Korrelate des Vokals vom Konsonant

nicht komplett verhindert oder verdeckt werden. Daher kann bei Vokalen und Konsonanten die parallele Übertragung maximal werden (Mattingly 1981), d.h. ein Konsonant und ein Vokal können gleichzeitig produziert und übertragen werden, ohne dass sie sich gegenseitig beeinträchtigen (maskieren). Auf diese Weise kann man die maximale Menge an Information übertragen werden (Effizienz der Produktion), ohne die Wahrnehmung der einzelnen Segmente zu behindern.

Aufgrund dieser maximalen Übertragung in der Produktion der Sprache wird in der Artikulatorischen Phonologie jede vokalische Geste mit dem linken, vorhergehenden Konsonanten in einer CV-Sequenz assoziiert und beide miteinander koordiniert (Browman & Goldstein 1988). Unter Koordination versteht man ein zeitliches Verhältnis zwischen beiden Elementen, das zustande kommt, indem die Elemente von gekoppelten (aber mehr oder weniger unabhängigen) Artikulatoren produziert werden. Wie zuvor beschrieben, kann das zeitliche Verhältnis zwischen Konsonanten und Vokalen bis zu 100% überlappen, d.h. die Geste des Konsonanten und des Vokals können zum gleichen Zeitpunkt beginnen und enden (maximale Überlappung). Die Assoziierung des Vokals mit dem linken, vorhergehenden Konsonanten in einer CV-Struktur ist keine neue Erfindung der Artikulatorischen Phonologie und wurde bereits bei der Autosegmentalen Phonologie (Selkirk 1986), der CV-Phonologie (Clements & Keyser 1983, McCarthy 1981) und der Optimalitätstheorie (Prince & Smolensky 2004, 1997, Kager 1999) verwendet.

Die Grundzüge dafür basieren bezüglich des Modells der Artikulatorischen Phonologie auf dem Zusammenspiel zwischen Produktion und Perzeption in der Sprache. In diesem Sinne wird in der Artikulatorischen Phonologie vorhergesagt, dass jeder der Vokale mit dem linken, vorhergehenden Konsonanten koordiniert wird. So erwarten wir, dass auch in einer CVC(V)-Sequenz jeder der Vokale mit dem links stehenden Konsonanten koordiniert wird (CV.CV) und die Überlappung beider Segmente (Vokal und

Konsonant) maximal sein kann. Wenn beide Konsonanten mit unterschiedlichen Vokalen koordiniert werden, wird keine Überlappung beider Konsonanten erwartet.

Die Koordination von Konsonantenclustern ist generell komplexer, da sie der Vielfältigkeit der Konsonanten und ihrer Kombinationsmöglichkeiten Rechnung tragen muss. Die Konsonanten sind an sich eine mannigfaltige Kategorie, die aus der Kombination verschiedener Parameter wie Artikulationsart (Plosiv, Frikativ, Nasal, usw.), Artikulationsort (labial, dental, alveolar, palatal, usw.) und Stimmhaftigkeit (stimmhaft, stimmlos) entsteht. Die Kombination von Elementen einer so vielfältigen Kategorie lässt sich aufgrund der Verschiedenheit der einzelnen Elementen und der großen Anzahl von Kombinationsmöglichkeiten schwer generalisieren. Hinzu kommen zusätzliche Gemeinsamkeiten und Unterschiede zwischen Sprachen oder Sprachstufen in der historischen Entwicklung. Die Fragen, warum bestimmte Cluster nur in manchen Sprachen auftreten, andere insgesamt häufiger in den Sprachen der Welt vorkommen, historisch stabiler sind oder beides (öfter vorkommen und stabiler sind), machen die Clusterforschung auf der einen Seite besonders interessant, auf der anderen Seite aber auch zu einer größeren Herausforderung.

Innerhalb der aktuellen Forschung zu Konsonantenclustern im Rahmen der Artikulatorischen Phonologie können zwei Forschungsrichtungen unterschieden werden, die sich vor allem methodisch voneinander unterscheiden: Die eine beschäftigt sich mit der (räumlich-) temporalen Organisation der Gesten zueinander (Überlappung einzelner Konsonantengesten), die andere dagegen eher mit dem Abstand der Konsonanten des Clusters zu einem weiter rechts definierten Ankerpunkt (der sogenannte *C-Center-Effekt*). Sie teilen sich die dynamisch-basierte Definition von Gesten und die Auffassung, dass Gesten und ihre Koordination Teil der phonologischen Repräsentation des Segmentes sind. Konkret geht es bei der Ersteren um die Fragen, wie die Konsonantengesten in Bezug zueinander koordiniert werden, inwieweit sich beide Konsonantengesten überlappen und

welche Faktoren die Überlappung modellieren können (Byrd 1996a, b, Chitoran et al. 2002, Gafos 2002, Gafos et al. 2010, Hoole et al. 2009, Kühnert et al. 2006 u.a.).

Ein möglicher Grund für die Vielzahl der Arbeiten zu Konsonantenclustern in diesem Bereich könnten die vielseitigen Übertragungsmöglichkeiten der einzelnen Elementen eines Clusters sein: Die angesprochene parallele Übertragung von Konsonant und Vokal (Mattingly 1981) funktioniert bei reinen Konsonantenclustern je nach Artikulationsart anders, da sich Konsonanten je nach Stärke des Verschlusses mehr oder weniger gegenseitig blockieren können. Je fester der Verschluss für die Artikulation des Konsonanten, desto stärker kann ein Konsonant einen anderen beeinträchtigen (maskieren). Stellt man sich eine Maskierungsskala (ähnlich der Sonoritätsskala) vor, liegt die geringste Maskierung bei einer CV-Sequenz, gefolgt von Sequenzen aus Plosiv-Liquid und die stärkste Maskierung bei einer reinen Plosivsequenz. Für Letztere wird vorhergesagt, dass die Überlappung möglichst gering ausfällt, damit beide Plosive für den Hörer perceptiv wiederherstellbar (*recoverable*) bleiben. So stehen bei Clustern zwei Kräfte im Konflikt zueinander: Einerseits hat der Sprecher den Anspruch, durch die parallele Übertragung segmentaler Informationen (Mattingly 1981) möglichst viel gleichzeitig zu übertragen, damit das Produzieren für ihn so effizient wie möglich ist (auch Ökonomie-Prinzip bei Martinet (1960) genannt). Andererseits muss der Sprecher die produzierte Information für den Hörer wiederherstellbar, d.h. zugänglich machen, damit die Kommunikation erfolgreich ist. Ist die Übertragung zu schnell, kann das Gesagte bei dem Hörer mit Problemen ankommen.

Ein weiteres Argument für die Interaktion zwischen Produktion und Perzeption ist der Einfluss der Produktionsrichtung bei der Koordination von Clustern. Chitoran et al. (2002) zeigten, dass georgische Plosivcluster, die aus einem vorderen C_1 gefolgt von einem hinteren C_2 (*front-to-back* Cluster) bestanden, mehr Überlappung aufwiesen als die Cluster in der umgekehrten

Produktionsrichtung (*back-to-front*). Diese Hypothese konnte ebenso für französische Cluster bestehend aus Obstruent und Liquid (z.B. /pl/ wie in *plaque*) oder Obstruent und Nasal (z.B. /pn/ wie in *pneu* ‘Reifen’, siehe Kühnert et al. 2006: 329) und für reine Plosivcluster im Marokkanischen Arabisch für einen von zwei Sprechern (Gafos et al. 2010) bestätigt werden. Ein möglicher Grund für diesen Produktionsrichtungs-Effekt (*place order effect*) ist, dass ein bilabialer Laut selbst bei stärkster Überlappung der beteiligten Konsonanten für den Hörer in der Produktion besser wiederherstellbar bleibt im Vergleich zu einem weiter hinten produzierten Laut wie beispielsweise /k/.

Allerdings wird ein geringerer Einfluss der Perzeption auf die aus Obstruent und Liquid bestehenden französischen Cluster angenommen, da Liquide andere Laute nicht im gleichen Ausmaß maskieren können wie Plosive (Mattingly 1981, Kühnert et al. 2006). Daraus ergibt sich eine andere Erklärungsmöglichkeit für diesen Effekt, nämlich die Aktivität unabhängiger Artikulatoren in der *front-to-back* Richtung (Kühnert et al. 2006). Beispielsweise bei /pl/ im Vergleich zu /kl/ entsteht das erste Cluster /pl/ durch die Bewegung von Lippen und Zungenspitze, die sich unabhängig voneinander und somit zeitgleich bewegen können, wodurch eine stärkere Überlappung ermöglicht wird. Im Gegensatz dazu ist die Zungenspitze bei /kl/ nicht frei beweglich, da der erste Konsonant /k/ mit dem Zungendorsum produziert wird und beide Artikulatoren, also Zungendorsum und Zungenspitze, miteinander verbunden sind. Die Relativierung des Einflusses der Perzeption auf die Clusterproduktion schließt auf keine Weise den Einfluss der Perzeption vollständig aus, sondern es soll hervorgehoben werden, wie komplex Konsonantencluster eigentlich sind und dass sie sich nicht allein aus der Beziehung zwischen Produktion und Perzeption modellieren lassen.

Die Position des Clusters innerhalb eines Wortes bestimmt ebenso in gewissem Maße die Koordination der Konsonanten miteinander: Wort-mediale Cluster zeigen mehr Überlappung als wort- initiale in verschiedenen Sprachen (Byrd 1996 für Englisch, Chitoran et al. 2002 für Georgisch, Gafos et al. 2010, 2011 und

Shaw et al. 2009a, b, 2011 für Marokkanisches Arabisch). Mögliche Gründe dafür sind zum einen das Fehlen von akustischen Übergängen von einem vorangehenden Vokal für die initialen Konsonanten, vor allem wenn sie ganz am Anfang einer Äußerung stehen (Chitoran et al. 2002). Zum anderen sind initiale Segmente besonders wichtig für die Wortspeicherung im mentalen Lexikon (Marslen-Wilson 1987) und werden deshalb besonders deutlich produziert.

In den letzten Abschnitten wurde ersichtlich, dass sich ein Teil der Arbeiten primär mit der zeitlichen und räumlichen Organisation der Konsonanten in einem Cluster auseinandersetzt. Dabei wird untersucht, welche Rolle verschiedene Faktoren für die Koordination der Konsonanten spielen, wie z.B. die Artikulationsstelle, die Produktionsrichtung, die Wiederherstellbarkeit des Segments bei der Perzeption, die motorischen Mechanismen bei der Produktion oder die Position innerhalb eines Wortes.

Im Vordergrund der Arbeiten, die sich hauptsächlich mit dem *C-Center-Effect* bei den Konsonantenclustern beschäftigen, steht dagegen primär die artikulatorische Grundlage von phonologischen Einheiten wie Segment und Silbe. In diesem Ansatz wird der Abstand zwischen dem Mittelpunkt des Konsonanten oder der Konsonantensequenz (*C-Center*) in Bezug auf einen weiter rechts lokalisierten, folgenden Ankerpunkt verglichen. Der Ankerpunkt für die Analysen ist entweder der vokalische Silbennukleus oder der konsonantische Anlaut der darauffolgenden Silbe. Browman & Goldstein (1988) kamen zu der Hypothese, dass die Bestandteile komplexer Konsonantencluster im Anlaut miteinander koordiniert werden (der sogenannte *C-Center-Effekt*), während die Konsonanten eines komplexen Auslautclusters sequenziell hinzugefügt werden. Dieser Effekt kann durch eine messbare Verschiebung des Konsonanten in Bezug auf einen weiter rechts positionierten Anker ermittelt werden. Wenn weitere Konsonanten links hinzugefügt werden (also C zu CC wird), so bleibt der Abstand zwischen dem Anker

und dem *C-Center* (der Mitte des konsonantischen Anteils des Anlauts) relativ stabil, unabhängig davon, ob der Anlaut aus einem oder mehreren Konsonanten besteht.

Anlautcluster werden global mit dem darauffolgenden Vokal koordiniert, so dass der Mittelpunkt des gesamten Clusters (der *C-Center*) einen stabilen zeitlichen Abstand zum Anker aufweist, der unabhängig von der Komposition des konsonantischen Anteils der Sequenz ist (ein einzelner Konsonant oder ein Cluster mit zwei oder mehreren Elementen). Dieser Effekt wurde von Browman & Goldstein 1988, Honorof & Browman 1995, Marin & Pouplier 2010 für englische Konsonantenanlaute, von Hermes et al. 2008 für italienische, Shaw et al. 2009a, b, Shaw 2011 für arabische und Marin 2011 für rumänische Cluster nachgewiesen.

Im Gegensatz dazu werden Auslautcluster sequenziell hinzugefügt, d.h. im Auslaut findet keine Verschiebung des Mittelpunkts des Konsonantenclusters in Bezug auf den Anker statt. Es tritt kein *C-Center*-Effekt auf, sondern die Konsonanten werden nacheinander hinzugefügt (siehe Byrd 1996a, 1996b, Browman & Goldstein 1988, 1990a, 2000, Marin & Pouplier 2008, 2010 u.a.). Es wurden allerdings Ausnahmen zu dieser Generalisierung aufgestellt: Manche Sprachen zeigen auch im Anlaut keinen *C-Center*-Effekt, da selbst bei einer reinen Konsonantensequenz im Anlaut keine Verschiebung des Abstands zwischen Anker und *C-Center* auftritt, sondern der Abstand zwischen Anker und dem rechten, folgenden Konsonanten unverändert bleibt. Dies wurde unabhängig von der Konsonantenanzahl für Tashlhiyt Berber (*mun* vs. *smun* vs. *tsmun*, vgl. Goldstein et al. 2007a, b) und Marokkanisches Arabisch (*bulha* vs. *shulha* vs. *ksbulha*, vgl. Shaw et al. 2009) festgestellt. Eine weitere Arbeit zeigte, dass zumindest für Tashlhiyt Berber das Nicht-Vorhanden-Sein des *C-Center*-Effekts daran liegt, dass diese Sprache keine Konsonantencluster, sondern lediglich einzelne Konsonanten im Anlaut erlaubt. Die reinen Konsonantensequenzen kommen zustande, weil die Sprache sowohl vokalische als auch konsonantische Silbennuklei erlaubt (Hermes et al. 2011).

Die Koordination von aufeinanderfolgenden CV-Sequenzen (CVCV) miteinander wurde in der Artikulatorischen Phonologie kaum untersucht. Ausgehend von der bestehenden Forschung wird jedoch erwartet, dass Konsonanten in einer alternierenden $C_1VC_2(V)$ -Sequenz anders koordiniert werden als in einem reinen Konsonantencluster: In alternierenden Sequenzen erwartet man, wie oben angesprochen, dass beide Vokale mit dem links stehenden vorhergehenden Konsonanten koordiniert werden. Die Hauptgründe dafür sind, dass die parallele Übertragbarkeit in der Produktion und die gute Wiederherstellbarkeit in der Perzeption eine maximale Überlappung von Vokal und Konsonant begünstigen. Es wird allerdings keine Überlappung von C_1 und C_2 erwartet und auch kein C-Center-Effekt von C_1 in Bezug auf einen Ankerpunkt, da beide Konsonanten jeweils mit unterschiedlichen Vokalen koordiniert werden.

Abschließend lässt sich für den Fall der in dieser Arbeit interessierenden reinen Konsonantencluster festhalten, dass die Koordination beider Konsonanten komplex ist und mit verschiedenen Methoden untersucht wird: Die Überlappung beider Konsonantengesten scheint von unterschiedlichen Faktoren besonders beeinflusst zu sein: von der Übertragbarkeit in der Sprachproduktion und der Wiederherstellbarkeit in der Perzeption, der Produktionsrichtung und der Position des Clusters im Wort. Die Koordination des Clusters in Bezug auf den C-Center-Effekt scheint vor allem von der Position des Clusters in der Silbe (Anlaut vs. Auslaut) beeinflusst zu werden. Man erwartet einen C-Center Effekt bei Anlaut- aber nicht bei Auslautclustern.

4.1.5 Anpassungen in der gesprochenen Sprache

Die etablierte gestische Unterscheidung zwischen alternierender CVC-Sequenz und reinen Konsonantenclustern (CC) basiert auf der Annahme, dass beispielsweise das englische Wort *beret* ('Baret, Baskenmütze') durch einen anderen lexikalischen Eintrag als das Wort *bray* ('Schrei') repräsentiert wird und dass die Koordination der Segmente in abstrahierter Form auch der Koordination ihrer

Repräsentationen entspricht (Browman & Goldstein 1990a, 1992, Byrd 1996c, Gafos et al. 2010). Die gleiche Erwartung trifft auf das deutsche Wort *beraten* im Vergleich zu *braten* oder das portugiesische Wort *querer* ('wollen') im Vergleich zu *crer* ('glauben') zu, die unterschiedlichen Lexemen entsprechen und verschiedene abstrakte Repräsentationen hervorrufen.

In der gesprochenen Sprache gibt es allerdings die Tendenz, das unbetonte Schwa zu lenisieren oder zu tilgen (Davidson 2006, Kohler 2001, Oda 2007). So wird *beret* als [bɛɪ] produziert (Browman & Goldstein 1990a, 1992) und daraus entstehen neue Konsonantensequenzen in der Produktion. Diese graduelle Tilgung wird von kontextuellen Faktoren wie dem Lautkontext, der Sprechgeschwindigkeit und dem Sprechstil beeinflusst (Barry 1992, Davidson 2006, Manuel et al. 1992, Fokes & Bond 1993, Fougeron & Steriade 1997) und von der Entstehung von in der Sprache möglichen Konsonantenclustern begünstigt (*believe* > [bl]ieve, *potential* > [pt]ential für Englisch, Hooper 1978, Glowacka 2001). Sie kann jedoch ebenso neue Cluster hervorbringen (*semester* > [sm]ester ‚Semester‘ oder *fatigué* > [ft]igué, ‚müde‘, z.B. Davidson 2006, Fougeron & Steriade 1997 u.a). Akustische Studien zeigen, dass die so entstandenen (postlexikalischen) Konsonantensequenzen sich bei der Konsonantendauer oder bei der Aspiration von C₂ von lexikalischen Clustern signifikant unterscheiden (Davidson 2006, Fokes & Bond 1993, Fougeron & Steriade 1997, Manuel et al. 1992). Die phonetischen Unterschiede am Sprachsignal können entweder mit einem probabilistischen Modell oder mit Variabilität in der Koordination erklärt werden (Browman & Goldstein 1990, 1992). Bei einem probabilistischen Modell im Sinne der Exemplar-Theorie (Bybee 2006, 2008, Goldinger 1998, Johnson 1997, 2006, Pierrehumbert 2000, 2001, 2002, 2006) wird angenommen, dass der Sprecher mehrere Varianten in seinem mentalen Lexikon zur Verfügung hat und sie je nach kontextuellem Bedarf einsetzt. Die andere Erklärungsmöglichkeit ist, dass sich beide Formen lediglich in der Koordinierung der Gesten unterscheiden, ohne Änderung der lexikalischen Repräsentation.

From the gestural point of view, the relationship between the lexical characterization of a word and its characterization in connected speech is much simpler and more highly constrained. We propose that most of the phonetic units (gestures) that characterize a word in careful pronunciation will turn out to be observable in connected speech, although they may be altered in magnitude and in their temporal relation to other gestures. In faster, casual speech, we expect gestures to show decreased magnitudes (in both space and time) and to show increasing temporal overlap. We hypothesize that the types of casual speech alternations observed (segment insertions, deletions, assimilations, and weakenings) are consequences of these two kinds of variation in the gestural score. (Browman & Goldstein 1991b: 343-344).

In diesem Sinne behalten reduzierte und unreduzierte (volle) Formen die gleichen Gesten, allerdings mit reduziertem Umfang oder stärkerer Überlappung bei der reduzierten Form. Diese Hypothese ist überprüfbar, indem man die Koordination der reduzierten Form von *beret* und *bray* artikulatorisch vergleicht. Dabei gibt es zwei Möglichkeiten, die Hypothese zu bestätigen: Entweder besitzt *beret* eine zusätzliche Zungenbewegung für die Produktion des Schwas oder die Gesten von *beret* und *bray* unterscheiden sich lediglich in ihrer Überlappung. Browman & Goldstein (1990a: 107) konnten keine zusätzliche Zungenbewegung feststellen, die mit der Produktion des Schwas assoziiert werden konnte. In diesem Kontext (/bər/) scheint das Schwa gestisch unterspezifiziert zu sein und keine zusätzliche Bewegung hervorzurufen. Aus diesem Grund konnte der lexikalische Unterschied zwischen *bray* und *beret* nur anhand von Überlappungsunterschieden modelliert werden.

In einem Perzeptionsexperiment, bei dem lediglich die Überlappung eines Kontinuums in Schritten von 10 ms. variiert wurde, zeigten die Ergebnisse, dass Hörer die Abfolge von /b/ und /r/ ohne Überlappung der einzelnen Konsonanten als Anfang von *beret* und die gleiche Abfolge mit Überlappung als Anfang von *bray* wahrnahmen (Browman & Goldstein 1990a: 106). Somit konnte das Experiment die Hypothese bestätigen, dass *bray* und *beret* eine ähnliche gestische Struktur mit den

gleichen Gesten besitzen und der lexikalische Unterschied zwischen beiden Wörtern an der zeitlichen Organisation der initialen Konsonanten liegt: In *bray* sind /b/ und /r/ enger miteinander koordiniert, d.h. sie zeigen eine stärkere Überlappung, während der bilabiale Plosiv /b/ in *beret* weiter entfernt von /r/ und dem Vokal liegt. Das erste Wort zeigte ebenso eine C-Center-Organisation, was für das zweite nicht der Fall war (Browman & Goldstein 1990a, 1992).

Da sich die lexikalischen Unterschiede in solchen Wortpaaren lediglich in der zeitlichen Koordination manifestierten und diese beim Sprechen mit dem Sprechstil, der Sprechgeschwindigkeit, usw. variieren können, besteht die Gefahr, dass die lexikalischen Unterschiede in der gesprochenen Sprache neutralisiert werden oder ganz verloren gehen.

4.1.6 Fragestellung und Hypothesen

Inspiziert durch die Unterscheidung zwischen *bray* vs. *beret* (Browman & Goldstein 1988, 1991a, 1992) und zwischen *sport* vs. *support* (Manuel et al. 2002) im Englischen, liegt der Schwerpunkt des aktuellen Kapitels auf Koordinationsunterschieden zwischen reinen Konsonantenclustern (bestehend aus Plosiven und Liquiden) und den entsprechenden CVC-Sequenzen in zwei portugiesischen Varietäten. Allerdings ist der portugiesische Ausgangspunkt aus zwei Gründen noch interessanter:

- (i) Es werden zwei verschiedene Varietäten (Brasilianisches Portugiesisch vs. Eu-ropäisches Portugiesisch, im Folgenden BP vs. EP) verglichen. Das Hauptinteresse liegt darin begründet, dass EP Vokalschwächungen und -tilgungen aufweist und BP nicht. Somit kann man zusätzlich die Annahmen der Artikulatorischen Phonologie für realisierte Vokale testen.

- (ii) Im EP wird kein zentrales Schwa (/ə/) getilgt, sondern der hintere hohe gerundete Vokal [u] und der zentrierte hohe Vokal /i/. Mit dem dreifachen Vergleich (C₁C₂ vs. <C₁eC₂> vs. <C₁uC₂>) kann man einen graduellen Einfluss beider geschwächten Vokale auf die Unterscheidung zwischen Clustern und CVCs untersuchen.

Bei den Token wurden der zweite Konsonant (C₂: /r/ vs. /l/), der Artikulationsort des ersten Konsonanten (C₂: labial vs. velar) und die Silbenbetonung (betont vs. unbetont) variiert. Da sich die meisten Arbeiten mit der zeitlichen Koordination beschäftigen, werden sich auch die folgenden Hypothesen zu den einzelnen Faktoren darauf konzentrieren.

4.1.6.1 Stimulustyp: Cluster vs. <C₁eC₂> vs. <C₁uC₂>

Im Englischen werden sowohl vordere als auch hintere Vokale in unbetonter Silbe als Schwa realisiert (z. B. *bæret* /bə'ræ/ und *support* /sə'pɔrt/), das in der gesprochenen Sprache getilgt werden kann (Davidson 2006, Manuel et al. 2002). Im EP wird kein zentrales Schwa (/ə/) getilgt, sondern der hintere hohe Vokal [u] und der zentrierte hohe Vokal /i/. In den gleichen Positionen kommen im BP [e, ʊ] (Barbosa & Albano 2004: 229) vor. Das Hauptinteresse an diesem Faktor liegt im dreifachen Vergleich zwischen Konsonantenclustern (CC, die keinen intermediären Vokal aufweisen) vs. <C₁uC₂> (Sequenzen mit einem hinteren Vokal, d.h. [ʊ] im BP und [u] im EP) vs. <C₁eC₂> (Sequenzen mit einem vorderen oder zentrierten Vokal, d.h. [e] im BP und [i] im EP).

Die Verwechselbarkeit beider Vokale wurde bereits in einer Perzeptionsstudie von Mateus & Martins (1982: 177-180) belegt. Dabei hatten 24 Teilnehmer die Aufgabe, Minimalpaare mit dem Zielvokal am Wortende (*parto* vs. *parte* 'brechen, 1. und 3. Person Singular Präsens' oder *caso* vs. *case* 'heiraten, 1. und 3. Person Singular Präsens') zu identifizieren. Die Ergebnisse zeigten eine gewisse Neutralisierung beider Vokale in wort-finaler Position:

Insgesamt wurden 20% an Substitutionen beider Vokale mit einem starken Einfluss der jeweils vorausgehenden Konsonanten festgestellt. Der Anteil der Substitutionen war ähnlich in beiden Richtungen (21% für $i > u$ und 19% $u > i$), allerdings wurde /i/ vor allem nach palatalen Konsonanten [ɲ, ʎ] mit /u/ verwechselt und /u/ nach dem bilabialen Plosiv [b] mit /i/ (Mateus & Martins 1982: 179). Trotz des ähnlichen Substitutionsanteils beider Vokale und der ähnlichen Instabilität beider Konsonanten erwarten wir, dass das zentralisierte /i/ dank seiner artikulatorischen Ähnlichkeit mit dem zentralen Schwa ([ə] in *beret*, Browman & Goldstein 1990a, 1992) eher artikulatorisch unterspezifiziert sein könnte als /u/. Daher die Vorhersage, dass Cluster vs. $\langle C_1eC_2 \rangle$ mit vorderem Vokal stärker neutralisiert werden als Cluster vs. $\langle C_1uC_2 \rangle$ mit hinterem Vokal.

4.1.6.2 Sprechervarietäten: BP und EP

Trotz des gemeinsamen Ursprungs und der teilweise gemeinsamen Sprachentwicklung zeigen beide Varietäten synchron unterschiedliche unbetonte Vokalsysteme (Barbosa & Albano 2004) und eine unterschiedliche Rhythmik (vgl. Barbosa, 2002, Bisol 1990, 2000, Bisol & Brescancini 2002, Bisol et al. 2008, u.a.). Besonders interessant für diese Studie ist die Tendenz im EP, unbetonte Vokale zu schwächen und zu tilgen, was im BP nur unter sehr eingeschränkten prosodischen Bedingungen²⁸ vorkommt. Daher können Wortpaare wie *querer* ('wollen') und *crer* ('glauben') im BP deutlich unterschieden werden (/ke'ref/ vs. /'kref/), während sie im EP nahezu Homophone (/ 'kref/) sind.

So sind insgesamt ausgeprägtere Unterschiede zwischen Konsonantenclustern und CVC-Sequenzen im BP als im EP zu erwarten. Konkret erwarten wir eine deutliche Vokalgeste für die CVC-Sequenzen im BP und im EP lediglich Unterschiede in der zeitlichen Organisation der Konsonanten (ähnlich wie im Englischen *beret* vs. *bray*, vgl. Browman & Goldstein 1990a, 1992).

²⁸ Unbetonte Vokale werden in post-tonischer und prä-finaler Position getilgt, wenn daraus ein legaler Cluster entsteht, z. B. [a'bɔbra] anstatt /a'bɔbɔra/ *abóborá* 'Kürbis' (Bisol 2000, 2010).

4.1.6.3 Artikulationsart: zentrale vs. laterale Liquide²⁹

Ein alveolarer Tap wird mit einer schnelleren und ballistischen Bewegung der Zungenspitze und einem kürzeren Verschluss produziert (Proctor 2009) als Laterale³⁰. Aufgrund der längeren Geste und des längeren Verschlusses bei Lateralen wird geringere Gestenüberlappung bei /r/ erwartet. Ein möglicher Grund ist die größere Überlappung der Lateralgeste, ohne die Perzeption zu beeinträchtigen. Wäre die kürzere Tap-Geste von C₁ überlappt, wäre der Tap womöglich nicht mehr wiederherstellbar in der Perzeption. So ist anzunehmen, dass Laterale eine längere C₂-Konstriktionsdauer aufweisen und die Zungenspitzengeste später mit C₁ koordiniert wird als bei dem Tap.

4.1.6.4 Artikulationsort: bilabialer vs. velarer C₁

Eine frühere Perzeptionsstudie zeigte, dass der Unterschied zwischen Cluster und <C₁eC₂> in Wortpaaren wie *prece* ('Gebet') vs. *perce* ('Meinung, Gutachten') und *crer* ('glauben') vs. *querer* ('wollen') nach dem velaren Konsonanten stärker neutralisiert wurde als nach dem bilabialen (Mateus & Martins 1982: 180). Bei dieser Studie hatten 24 Studierende aus Lissabon die Aufgabe, zu

²⁹ In beiden Varietäten ist der zentrale Liquid als alveolarer Tap /r/ realisiert (regional auch Trill /r/ sowie andere Varianten möglich, vgl. Mateus & Andrade 2000, Silva 1996), während die Laterale als klares /l/ im BP (Silva 1996, Pagan 2003, Pagan & Wertzner 2007) und als dunkles /ɫ/ im EP realisiert werden (Andrade 1999, Martins et al. 2011, Oliveira et al. 2011). Da diese Unterschiede nicht Gegenstand der aktuellen Arbeit sind, werden sie nur am Rand beachtet.

³⁰ Sowohl die Dauer der Zungenspitzengeste (Sproat and Fujimura 1993: 307) als auch der Verschluss sind länger bei dunklen als bei hellen Lateralen (Recasens & Farnetani 1994: 198 für den Vergleich von italienischen hellen und katalanischen dunklen Lateralen und Recasens & Espinosa 2005: 12 für den Vergleich von valencianischem hellen und mallorquinischem dunklen /ɫ/). Der Artikulationsort war eher alveolar bei hellen und dental oder weiter posterior bei dunklem /l/ (Recasens & Espinosa 2005: 4-7, 20). Der Grund hierfür könnte sein, dass die Erhöhung des Zungendorsums für die hintere Artikulation, die zeitlich zuerst stattfindet, die Zungenspitze in der vorderen Region verflacht (dental oder weiter anterior, vgl. Recasens 2009: 478). So werden kürzere Konstriktionen bei den Lateralen im BP als im EP erwartet, die allerdings nicht Gegenstand der Arbeit sind.

jedem gehörten Stimulus die Anzahl der Silben zu notieren. Die Hörer sollten zum auditiven Stimulus die Anzahl der wahrgenommenen Silben angeben. Anhand der getesteten sechs Wortpaare mit vier Wiederholungen zeigen die Ergebnisse, dass die Token der bilabialen C₁ selten verwechselt wurden (5 von 48 Token), während die Token der velaren C₁ häufiger als mit einem Vokal produziert wahrgenommen wurden: Alle Token von *querer* (/ki.'rer/) und 22 Wiederholungen von *erer* (/.'krer/) wurden von den Versuchspersonen als zweisilbige Wörter wahrgenommen. Dies macht deutlich, dass Hörer das letzte Wortpaar nicht mehr unterschieden und nur die Variante mit dem intervokalischem Vokal wahrnahmen.

Aus heutiger Sicht ist die Richtung der Neutralisierung unerwartet, da mehrere Studien eine steigende Vokaltilgung feststellten (Cunha 2011a, Martins 1975, Mateus & Martins 1982, Silva 1998, 1997). Bei der Interpretation dieses Experiments könnte man die Ergebnisse mit der schlechten Wiederherstellbarkeit der Konsonanten in einem reinen Cluster in der *back-to-front* Richtung (Chitoran et al. 2002) assoziieren, die dazu führt, dass Cluster in dieser Produktionsrichtung weiter auseinander synchronisiert werden. Aufgrund der häufig auftretenden Vokaltilgung (Martins 1975) haben sich die Hörer daran gewöhnt und kompensieren es, indem sie für die weiter synchronisierten Konsonanten einen intermediären Vokal in der Perzeption einsetzten. Dabei ist allerdings unklar, ob der Produktionsrichtungseffekt, nach dem *front-to-back* Plosivcluster (/pt/) mehr Überlappung zeigen als Cluster in der umgekehrten Produktionsrichtung (/tp/, Chitoran et al. 2002), auch bei Plosiv-Liquid-Clustern auftritt. Der Grund dafür ist, dass Laterale andere Konsonanten nicht so stark maskieren können wie Plosive (Mattingly 1981).

Leider kann man jedoch bei der aufgeführten Opposition von bilabialem vs. velarem C₁ diesen Effekt und den Einfluss unabhängiger Artikulatoren (Lippen und Zungenspitze) bei den bilabialen Token im Vergleich zu verbundenen Artikulatoren (Zungendorsum und -spitze) bei den velaren Gegenstücken nicht voneinander trennen (Kühnert et al. 2006). Daher wird

ausgeprägte Überlappung bei den bilabialen Token, sei es aufgrund der Aktivität unabhängiger Artikulatoren oder aufgrund der perceptiven Wiederherstellbarkeit der Konsonanten in der *front-to-back*-Richtung, erwartet. Außerdem werden längere Konstriktionsdauern bei den velaren C₁ vorhergesagt, da das Zungendorsum träger und langsamer ist als die Lippen.

4.1.6.5 Betonung: betonte vs. unbetonte Kondition

Die akustische Dauer von betonten Silben ist länger als die von unbetonten (Browman & Goldstein 1990a: 103). Artikulatorisch entspricht die akustische Dauer entweder Änderungen an der zeitlichen Koordination (*Timing*) oder an der Steifheit der Gesten (*stiffness*, Browman & Goldstein 1990a: 103, Kelso et al. 1985). Beckman et al. (1990) zeigten, dass der Unterkiefer in betonten Silben größere Amplituden aber keine signifikant geringeren maximalen Geschwindigkeiten (*peak velocities*) aufweist. Demnach sollten betonte Silben einen vergleichsweise größeren räumlichen intergestischen Abstand (*phasing*) zwischen den Konsonanten in Anspruch nehmen als unbetonte Silben. Die Vorhersage lautet, dass Cluster und CVC-Sequenzen in unbetonter Position einen geringeren zeitlichen und räumlichen Umfang aufweisen als die in der betonten Umgebung.

Bei der Analyse werden Konsonantencluster in betonter und unbetonter Kondition verglichen. Bei den CVC-Sequenzen vergleichen wir Token, in denen C₂ (und V₂ in einer /C₁V₁C₂V₂/-Sequenz) entweder in betonter oder unbetonter Position (eine Silbe vor der betonten Position) vorkommen (z.B. prada vs. pradada bei den Clustern und perada vs. perador bei den CVC-Sequenzen³¹). Damit Vokaltilgung auftreten kann, steht der Zielvokal (V₁) in unbetonter Silbe.

Kurz zur Erinnerung, es wird im Weiteren untersucht, inwiefern sich die Koordination von portugiesischen Konsonantenclustern

³¹ Die Zielcluster und Ziel-CVC wurden *kursiv* gesetzt und die primär betonte Silbe unterstrichen.

(Plosiv und Liquid) vs. $\langle C_1eC_2 \rangle$ - vs. $\langle C_1uC_2 \rangle$ -Sequenzen unterscheidet und welche Rolle dabei Sprechervarietät, Artikulationsart, Artikulationsort und Betonung spielen. Wie in Bombien (2011) wird hier erwartet, dass die Betonung einen Effekt auf den unmittelbar vorhergehenden Konsonanten hat und der C_2 -Verschluss bei betonten Silben länger ist als bei unbetonten.

Die Methode umfasst synchronisierte akustische und artikulatorische EMA-Daten von vier verschiedenen Clustern (/pr, pl, kr, kl/) von fünf EP- und vier BP-Sprechern. Es folgt die Analyse der Koordination von Konsonantenclustern und CVC-Sequenzen. Dafür werden folgende Messverfahren verwendet:

- (i) Die zeitliche intergestische Koordinierung (*Timing*) mit der Organisation des Konstriktionsplateaus und der Synchronisierung von C_2 in Bezug auf C_1 . Bei Ersterem geht es um die Dauer der einzelnen Konstriktionsplateaus sowie den intermediären Abstand und bei Letzterem um die Synchronisierung beider Konsonanten zueinander.
- (ii) Die räumliche intragestische Dynamik der Zungenspitzenbewegung wird ausgehend von ihrem Umfang (*magnitude*) und ihrer maximalen Geschwindigkeit (*peak velocity*) analysiert. Hier soll festgestellt werden, inwiefern Cluster vs. $\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$ signifikante Unterschiede im Umfang und der Geschwindigkeit der Zungenspitzenbewegung aufweisen.

Abschließend folgt gegen Ende des Kapitels eine allgemeine Diskussion der Ergebnisse, ihre Positionierung gegenüber dem aktuellen Forschungsstand und ein Ausblick auf die weiteren Kapitel.

4.2 Methode

Für die Beantwortung dieser Fragestellungen wurden akustische und artikulatorische Daten mithilfe des 5D-Artikulatographen AG500 (Hoole & Zierdt 2010, Zierdt 2007) der Firma Carstens Medizinelektronik GmbH³² am Institut für Phonetik und Sprachverarbeitung in München erhoben.

4.2.1 Elektromagnetische Artikulographie

Die Elektromagnetische Artikulographie (EMA) ermöglicht es, Positionen und Bewegungsabläufe beim Sprechen anhand von auf den Artikulatoren angebrachten Sensoren im dreidimensionalen Raum zu verfolgen und zu messen.

EMA-Koordinaten ergeben sich aus der Lösung von komplexen, nicht-linearen Gleichungen. Die Eingangsdaten hierfür sind die demodulierten, elektrischen Signale, die durch sechs magnetische Feldtransmitter (Sender) in den Sensor induziert werden (Hoole & Zierdt 2010). Jeder der Sender produziert ein magnetisches Wechselfeld mit unterschiedlicher Frequenz. Dieses magnetische Wechselfeld ruft eine Wechselspannung in jedem Sensor hervor und ermöglicht so die Ermittlung der Distanzen von jedem Sensor zu den sechs Sendern. Mit diesen Distanzen kann man dann die drei Koordinaten XYZ und die zwei Winkel Azimut und Elevation für die Messung, Speicherung und Visualisierung der Sensorenpositionen bestimmen. Die drei Positionskoordinaten (X, Y, Z) sind kartesische Koordinaten und die zwei Winkelkoordinaten (Azimut, Elevation) sind sphärisch und geben an, in welche Richtung die Sensoren zeigen. Sie werden wie folgt definiert (Hoole & Zierdt 2010: 332):

- Die X-Achse verläuft horizontal von vorne nach hinten (anterior-posterior).
- Die Y-Achse verläuft von rechts nach links.

³² Siehe <http://www.articulograph.de>.

- Die Z-Achse verläuft vertikal von oben nach unten (superior-inferior).
- Azimut ist ein Winkel in der XY-Ebene, der den positiven, gegen den Uhrzeigersinn gedrehten Winkel eines Punktes von der Z-Achse weg bezeichnet (ähnlich wie ein geographischer Längengrad (longitude) auf einem Globus).
- Elevation ist ein Höhenwinkel, der die Verschiebung weg von der XY-Ebene beschreibt (ähnlich wie ein geographischer Breitengrad (latitude) auf einem Globus).

Aufgrund dieser fünf Koordinaten und der Komplexität der damit aufgenommenen Daten ist es angemessener zu behaupten, dass der verwendete Artikulograph (AG500) eher ein 5-dimensionales als ein 3-dimensionales (3D) System ist (Hoole & Zierdt 2010). Da die relevanten Artikulatoren (Zunge, Lippen und Unterkiefer) sich nicht nur aktiv bewegen, sondern auch passiv durch Bewegungen des gesamten Kopfes beeinflusst werden, muss eine Kopfbewegungskorrektur im Rahmen der Weiterverarbeitung durchgeführt werden. Hierfür werden Sensoren an statischen Positionen (oberhalb der Nase, am Oberkiefer und hinter beiden Ohren) befestigt, so dass daraus die Kopfbewegungen bestimmt und aus den Daten der relevanten Artikulatoren entfernt werden können.

Mit EMA können Positionen und Bewegungen zu jedem Zeitpunkt der Aufnahme bestimmt werden, wobei auch verschiedene Artikulatoren untersucht werden können, wie z.B. Kiefer- und Lippenbewegungen. Dies bietet eine umfangreichere Betrachtungsmöglichkeit von Artikulatoren und deren Zusammenspiel beim Sprechen im Vergleich zu anderen Messmethoden, wie beispielsweise der Elektropalatographie (EPG), die lediglich die Zunge bzw. ihren Kontakt zum Gaumen betrachtet.

Im Vergleich zu älteren 2D-Systemen (EMMA - *Elektromagnetic Midsagittal Artikulography*) verwenden die neuen 5D-Systeme keinen Helm, wodurch die Aufnahmen weniger anstrengend für die Versuchspersonen sind und damit länger dauern können. Außerdem müssen die Sensoren nicht mehr genau mediosagittal befestigt sein, so dass kaum Daten verloren gehen. Das neue System verwendet die gleiche Anzahl an Sensoren (bis zu 12), hat jedoch mehr Sender zur Verfügung (6), die eine größere Menge von Information pro Sensor ermöglichen. Die Verwendung von zwei Winkeln für die Datenberechnung ermöglicht eine zusätzliche Erfassung der Zungenkontur und steigert die Stabilität und Robustheit der Daten.

Darüber hinaus hat sich EMA als Untersuchungsmethode in vielen anderen Studien (Bombien 2011, Bombien et al. 2006, 2010, Gafos et al. 2010, Hoole et al. 2009, Marin 2011, Marin & Pouplier 2008, 2010 u.a.) bewährt und ermöglicht somit eine gute Vergleichbarkeit der Ergebnisse aus der vorliegenden Untersuchung zum Portugiesischen mit anderen Studien bzw. Sprachen.

4.2.2 Versuchspersonen

Es wurden Daten von fünf Muttersprachlern des EP und vier Muttersprachlern des BP aufgenommen. Die Teilnehmer waren im Alter zwischen 22-29 Jahren und wurden nach bestimmten Regionen ausgewählt: Alle EP-Sprecher kamen aus Porto und Umgebung in Portugal und alle BP-Sprecher aus der Stadt São Paulo in Brasilien. Sie waren entweder Studenten im Austauschjahr oder hatten ihr Studium bereits abgeschlossen und waren aus privaten Gründen in Deutschland. Keiner der Probanden lebte länger als 12 Monate im Ausland (also außerhalb ihres Heimatlandes Portugal).

4.2.3 Sprachmaterial

Für die Aufnahmen wurden die in der Liste der aufgenommenen Stimuli mit den Zielclustern und Sequenzen (hier in *kursiv* dargestellt) und den betonten Silben (als unterstrichene Silben

markiert). aufgelisteten Zielwörter in den Trägersatz „Ele leu ____ bem“ (‘Er las ____ richtig’) eingebettet.

Cluster	r-Stimuli	<i>cr<u>a</u>ta</i> <i>cr<u>a</u>t<u>a</u>da</i> <i>pr<u>a</u>da</i> <i>pr<u>a</u>d<u>a</u>da</i>	‘Person’ ‘Streit’ ‘Eigennamen’ ‘Wiesen’
	l-Stimuli	<i>cl<u>a</u>ta</i> <i>cl<u>a</u>t<u>a</u>do</i> <i>pl<u>a</u>da</i> <i>pl<u>a</u>t<u>a</u>da</i>	‘Eigennamen’ ‘E. Genitiv’ ‘Eigennamen’ ‘E. Genitiv’
CVC- Sequenzen	r-Stimuli	<i>qu<u>e</u>rat<u>o</u>cone</i> <i>qu<u>e</u>rat<u>o</u>so</i> <i>cu<u>r</u>ado</i> <i>cu<u>r</u>ador</i> <i>pe<u>r</u>ada</i> <i>pe<u>r</u>ador</i> <i>pu<u>r</u>atos</i> <i>Pu<u>r</u>at<u>ã</u>o</i>	‘Keratoconus’ ‘Keratose’ ‘geheilt’ ‘Vormund’ ‘Binnenmuss’ ‘Lager’ ‘Mehlsorte’ ‘Ort’
	l-Stimuli	<i>qu<u>e</u> l<u>a</u>ta</i> <i>qu<u>e</u> l<u>a</u>t<u>a</u>da</i> <i>cu<u>l</u>atra</i> <i>cu<u>l</u>atr<u>a</u>no</i> <i>pe<u>l</u>ada</i> <i>pe<u>l</u>ador</i> <i>pu<u>l</u>ada</i> <i>pu<u>l</u>ador</i>	‘wie frech!’ ‘wie frech!’ ‘Gewehrkolben’ ‘G. Genitiv’ ‘geschält’ ‘Schäler’ ‘gesprungen’ ‘Springer’

Tabelle 6: Liste der aufgenommenen Stimuli mit den Zielclustern und Sequenzen (hier in *kursiv* dargestellt) und den betonten Silben (als unterstrichene Silben markiert).

Die fünf EP-Sprecher produzierten jeweils 8 Wiederholungen von jedem Satz und die vier BP-Sprecher jeweils 9 Wiederholungen. Die Sätze wurden in randomisierter Reihenfolge präsentiert und von einem Computerbildschirm abgelesen. Die Präsentation der Stimuli wurde manuell gesteuert, so dass inkorrekte Äußerungen (z.B. Versprecher) wiederholt und korrigiert werden konnten. Die

aufgenommene Anzahl von Äußerungen (Token) war für das EP insgesamt 960 Stimuli (24 Wörter x 5 Sprecher x 8 Wiederholungen) und für das BP insgesamt 864 Stimuli (24 Wörter x 4 Sprecher x 9 Wiederholungen).

4.2.4 Akustische und artikulatorische Stimuli

Bei den Aufnahmen wurden die Messsensoren an Ober- und Unterlippe, am Unterkiefer und an drei Positionen auf der Zunge (Zungenspitze, -mitte und -dorsum) angebracht. Für die Kopfbewegungskorrektur wurde jeweils ein zusätzlicher Sensor hinter dem rechten und dem linken Ohr, an der Nase und am Oberkiefer fixiert. Da das Prompt-Programm zur Stimulus-Präsentation die gesamte Aufnahme steuert, sind die akustischen und artikulatorischen Daten synchronisiert. Die Bewegungsdaten wurden mit einer Abtastrate von 200 Hz und das Audiosignal mit einer Abtastrate von 24 kHz aufgenommen.

4.2.4.1 Weiterverarbeitung der physiologischen Rohdaten

Die Weiterverarbeitung der physiologischen Rohdaten³³ wurde in Matlab (The MathWorks Inc., Version 7.11.0.584, R2010b) vorgenommen. Es folgt nun die Beschreibung weniger Grundschritte der Weiterverarbeitung, um diese zu illustrieren: Nach der Filterung der Rohdaten werden die Startpositionen und die Konturverläufe der Sensoren auf rekursive Weise anhand zweier unterschiedlicher Algorithmen (kalman von Korin Richmond und tapad von Andreas Zierdt) ermittelt. Die sehr unterschiedlichen Algorithmen werden parallel verwendet und miteinander verglichen, um die sichersten Ergebnissen zu erzielen.

Es folgt die Anpassung der Amplitude durch eine Regressionsanalyse. Das Ziel der Regressionsanalyse ist es, das Residuum von jedem Sensor-Sender-Paar aus den Signalen aller Sender vorherzusagen, um die Zuverlässigkeit der Ergebnisse zu erhöhen. Das Verfahren wird durch den Vergleich aller Sensoren

³³ Die hierfür benutzen Programme wurden mir von Philip Hoole zur Verfügung gestellt.

zu einem Kompensationssensor für die Daten jedes Senders durchgeführt. Die folgende Kopfbewegungskorrektur subtrahiert die (unbeabsichtigten) Bewegungen des gesamten Kopfes von den artikulatorischen Verläufen. Mit diesem Schritt werden also Bewegungen der Artikulatoren, die durch die Bewegung des gesamten Kopfes entstanden sind, herausgerechnet. Dies erfolgt durch die Definition eines anatomischen, sprecherbasierten Referenzobjekts und die Anpassung des Koordinatensystems für jeden Sprecher, nach welchem dann alle Daten dieses Sprechers angepasst werden (*remapping*).

Parallel zu den Bewegungsverläufen wurden horizontale, vertikale und tangentielle Geschwindigkeiten, ebenso wie Lippenöffnung aus der Euklidischen Distanz zwischen Ober- und Unterlippe berechnet (Bombien 2011: 75-76). Die akustischen Daten wurden nach der Aufbereitung in Matlab mit dem Münchener Automatischen Segmentierungssystem (MAUS, Schiel 1999, Kisler et al. 2012, Reichel 2012) automatisch phonemisch segmentiert, in das Emu-System (Harrington 2010) importiert und manuell (nach-)korrigiert.

4.2.4.2 Artikulatorische Segmentierung

Die artikulatorische Segmentierung erfolgte ebenfalls in Emu mit einem halbautomatischen Algorithmus, der die artikulatorischen Gesten aufgrund ihres Verlaufs und ihrer absoluten Geschwindigkeit in die sieben folgenden Markierungspunkte (*landmarks*) segmentiert:

- START – Anfang der Geste
- PVEL1 – Maximale Geschwindigkeit der Schließungsgeste
- TARGET – Beginn der Konstriktion/des Verschlusses
- MAX – Maximale Konstriktion
- ÖFFNUNG – Beginn der Lösung der Konstriktion/des Verschlusses

- PVEL2 – Maximale Geschwindigkeit der Öffnungsgeste
- ENDE – Ende der Geste.

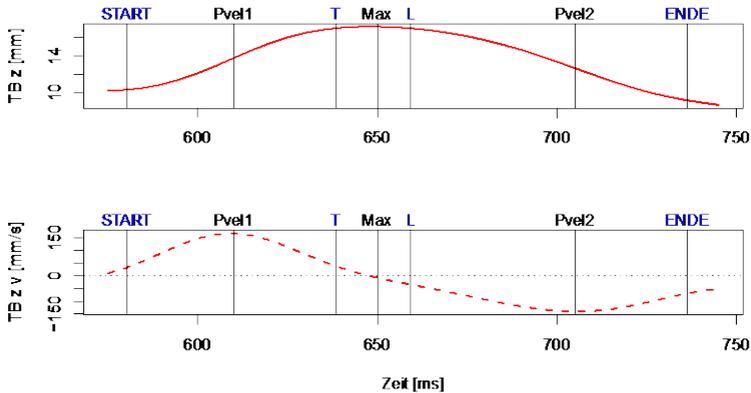


Abbildung 2: Markierungspunkte des Zungendorsums (TB) für den Konsonanten /k/ im Wort *crata* von S2.

Abbildung 2 zeigt die Markierungspunkte START, PVEL1, TARGET, MAX, ÖFFNUNG, PVEL2, ENDE, illustriert durch den Verlauf des Zungendorsums (TB) für den Konsonanten /k/ im Wort *crata* von S2 (erste Wiederholung). Oben wird der Verlauf des Zungendorsums (TB) in Millimetern und unten die vertikale Geschwindigkeit des Zungendorsums in Millimetern/Sekunde dargestellt. Der abgebildete Verlauf des Zungendorsums wird über die vertikale Geschwindigkeit dieses Artikulators definiert: Die zwei maximalen Geschwindigkeiten (PVEL1 und PVEL2) entsprechen der maximalen Geschwindigkeit des Zungendorsums in der Schließungs- (positive Geschwindigkeit) und in der Öffnungsphase (negative Geschwindigkeit) der Geste. MAX entspricht der maximalen Konstriktion und gleichzeitig der Null-Geschwindigkeit des Zungendorsums (Geschwindigkeit = 0).

Die anderen Markierungspunkte START, TARGET, ÖFFNUNG und ENDE wurden definiert auf Basis eines 20% Schwellenwerts der vertikalen Geschwindigkeit für das Zungendorsum in dem oben

abgebildeten Beispiel und der Spitzen-Tangentialgeschwindigkeit für die apikalen Laute (vgl. Bombien 2011: 75-78). Für unterschiedliche Artikulatoren wurden unterschiedliche Geschwindigkeiten verwendet. Der velare Konsonant /k/ wurde, wie in obigem Beispiel verdeutlicht, durch die Bewegungen des am Zungendorsum (*tongue back* – TB) fixierten Sensors erhoben. Der Verlauf wurde innerhalb beider Null-Geschwindigkeiten der vertikalen Position des Zungendorsums markiert. Da für die vorliegende Arbeit für /k/ die Anhebung des Zungendorsums interessant ist, wird dementsprechend lediglich die vertikale Geschwindigkeit³⁴ für die Verlaufskontur herangezogen.

Die apikalen Konsonanten /r, l/, die mit der Zungenspitze produziert werden, wurden entsprechend mit dem an der Zungenspitze fixierten Sensor erhoben. Da für die Produktion dieser Konsonanten sowohl die vertikalen (*inferior* – *superior*) als auch die horizontalen (*anterior* – *posterior*) Zungenbewegungen relevant sind, wurde für die Bewegung der Zungenspitze (*tongue tip* – TI) die Tangentialgeschwindigkeit aus der horizontalen und vertikalen Geschwindigkeit berechnet und anschließend etikettiert. Für den bilabialen Plosiv /p/ wurde die Lippenöffnung (*lip aperture* – LA), genauer gesagt die Euklidische Distanz zwischen Ober- und Unterlippe, berechnet. Die Gestenverläufe wurden in einem zweidimensionalen Raum berechnet, da (wie bei der Zungenspitze) lediglich die vertikale (*inferior* – *superior* auf der Z-Ebene) und horizontale (*anterior* – *posterior* auf der X-Ebene) Bewegung relevant ist und die laterale Bewegung ausgeschlossen werden kann (vgl. Bombien, 2011: 75-76).

4.2.5 Statistische Auswertung

Für die statistische Auswertung wurde eine Varianzanalyse mit Messwiederholung (*RM-Anova*, *repeated measures Anova*) mit einem linearen *Mixed Model* in R (R Development Core Team 2011) verglichen. Die Varianzanalyse wurde mit dem EZ-Paket

³⁴ Generell ist die horizontale Bewegung (*anterior-posterior*) des Zungendorsums auf der X-Achse für die Retraktion wichtig. Sie ist in Bezug auf die aktuelle Fragestellung allerdings nicht relevant.

(*ez-library*, Lawrence 2011) durchgeführt. Die unabhängigen Variablen waren Sprechervarietät (BP vs. EP), Stimulustyp (dreistufig: C_1C_2 vs. $\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$), C_1 (bilabial vs. velar), C_2 (/l/ vs /r/) und Betonung (betonte vs. unbetonte Kondition). Die Sprechervarietät war der einzige Faktor, bei dem sich die Sprecher unterschieden (*between subjects factor*), denn jeder Sprecher gehört entweder der einen oder der anderen Kategorie an (jeder Teilnehmer war entweder EP- oder BP-Sprecher, nicht beides). Für jede Faktorstufe der übrigen vier Faktoren (C_1 , C_2 , Betonung und Vokal, *within-subject factors*) wurde der Mittelwert aus den acht Wiederholungen pro Versuchsperson berechnet. Jeder der vier Faktoren wurde mit der Sprechervarietät (EP vs. BP) kombiniert und in einer separaten Varianzanalyse mit Messwiederholung geprüft. Da der Faktor Vokal dreistufig war, wurde dabei für Sphärizität (*Sphericity*³⁵) korrigiert.

Das *Mixed Model* wurde ebenfalls in R mithilfe der Pakete für lineare gemischte Effekte (*lme4-library*, Bates et al. 2011) und für parametrische Modellierung (*multcomp-library*, Hothorn et al. 2008) zur Anwendung des anschließenden Tukey Post-Hoc-Tests durchgeführt. Die unabhängigen Variablen waren die gleichen wie bei der Varianzanalyse und Sprechervarietät wurde als randomisierter Faktor miteinbezogen. Die abhängigen Variablen werden bei den jeweiligen Verfahren weiter unten im Text definiert.

Im Vergleich zur *Anova* mit Messwiederholungen haben *Mixed Models* die Vorteile, dass die Daten nicht ausgewogen sein müssen und nicht pro Sprecher gemittelt werden muss, was in beiden Fällen den Datenverlust verringert. Daher basiert die Analyse mit dem *Mixed Model* auf einer größeren Datenmenge und ist insgesamt robuster. Der Nachteil dieses Verfahrens ist, dass keine Wahrscheinlichkeiten ausgegeben werden, weil die Freiheitsgrade im Nenner nicht berechnet werden können. Um dies auszugleichen, wird der Einfluss der einzelnen Faktoren und der

³⁵ Siehe Baayen (2008), Delgaard (2008) und Johnson (2011) für weiterführende Lektüre

einzelnen Interaktionen separat berechnet. Im Falle, dass die Interaktionen signifikant sind, werden Tukey Post-Hoc-Tests mithilfe des *Multcomp*-Pakets berechnet. (Für alternative Verfahren siehe Reubold et al. 2010, Baayen, 2008: 263-321) Da die Analyse mit dem *Mixed Model* insgesamt robuster war und um Wiederholungen in der Darstellung zu vermeiden, werden lediglich die Ergebnisse des *Mixed Models* präsentiert.

4.3 Die zeitliche intergestische Koordinierung

Für die zeitliche Koordinierung beider Konsonanten wurden zwei Verfahren angewendet, das eine beschäftigt sich mit der zeitlichen Organisation beider Konstriktionsplateaus, das andere mit der Synchronisierung von C₂ in Bezug auf das C₁-Plateau. Mit diesen zwei Verfahren wird überprüft, inwiefern zeitliche Unterschiede zum einem eher mit der Länge der Konstriktionsplateaus oder dem Abstand dazwischen und zum anderen mit der Überlappung von C₂ in Bezug auf C₁ zusammenhängen.

4.3.1 Definition der Variablen

Es wurde die Dauer der einzelnen Plateaus sowie des intermediären Abstands und die Überlappung beider Konsonantengesten berechnet. Das Plateau (Konstriktionsdauer) ist definiert als das Intervall zwischen TARGET und LÖSUNG, dividiert durch die Gestendauer (ENDE-START) und damit auch normalisiert. Der Abstand zwischen den Plateaus (PLAT-DIST) wurde als das Intervall zwischen Ende des ersten und Anfang des zweiten Konstriktionsplateaus definiert, was durch die Gesamtdauer beider Gesten normalisiert wurde.

- $C_1\text{-PLATEAU} = (C_1\text{-LÖSUNG} - C_1\text{-TARGET}) / (C_1\text{-ENDE} - C_1\text{-START})$.
- $C_2\text{-PLATEAU} = (C_2\text{-LÖSUNG} - C_2\text{-TARGET}) / (C_2\text{-ENDE} - C_2\text{-START})$.

- $PLAT-DIST = (C_2TARGET - C_1LÖSUNG) / (C_2-ENDE - C_1-ANFANG)$

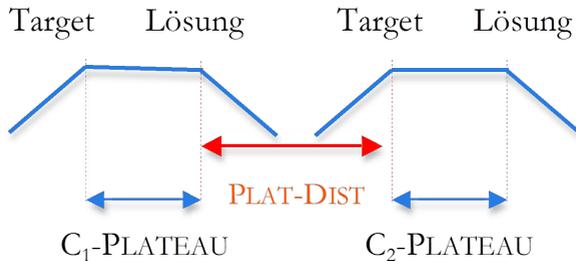


Abbildung 3: Schematische Darstellung der Plateau-Distanz.

- $ÜBERLAPPUNG = 1 - (C_2-START - C_1-TARGET) / (C_1-PLATEAU)$

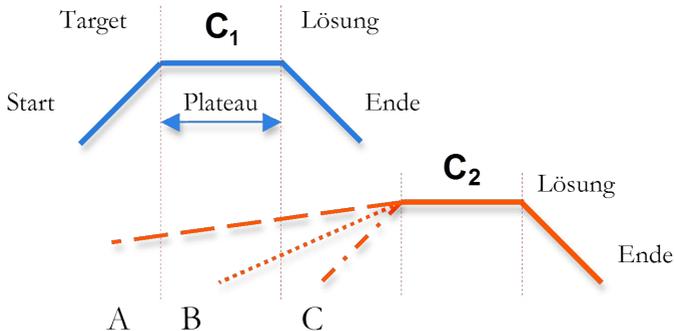


Abbildung 4: Schematische Darstellung der Überlappung/Synchronisierung der Gesten.

Für die Untersuchung der Synchronisierung und Überlappung beider Konsonanten wurde das zeitliche Intervall zwischen dem Target des ersten Konsonanten ($C_1-TARGET$) und dem Beginn des zweiten Konsonanten ($C_2-START$) berechnet. Dieser Index wurde durch die Plateaudauer von C_1 ($C_1-PLATEAU$) dividiert und

wiederum von 1 subtrahiert. Somit kann man prüfen, ob und wie der Gestenanfang von C_2 mit dem Verschluss des ersten Konsonanten (C_1) synchronisiert wird bzw. überlappt.

Da bei diesem Verfahren die Überlappung (C_2 -START - C_1 -TARGET) durch das C_1 -PLATEAU dividiert wurde, ist die Überlappung in Bezug auf das C_1 -PLATEAU berechnet worden. Daher ist bei der Interpretation der Ergebnisse Folgendes zu beachten (Gafos et al. 2010):

- Ein Überlappungswert zwischen 0 und 1 bedeutet, dass C_2 während des C_1 -Plateaus (im Bereich B der Grafik) beginnt, was als übliche Überlappung definiert wird (Ausgangspunkt).
- Bei einem negativen Überlappungswert beginnt C_2 nach dem Plateau(-ende) von C_1 (im Bereich C) und es findet keine Überlappung beider Gesten statt.
- Bei einem positiven Wert, der größer als 1 ist, beginnt C_2 vor dem Plateau(-anfang) von C_1 (im Bereich A) und es findet eine stärkere Überlappung statt.

Dieses Messverfahren wurde bereits in anderen Arbeiten verwendet, beispielsweise in Chitoran et al. (2002), Byrd & Choi (2010) und Gafos et al. (2010). Der Index wird von 1 subtrahiert (wie bei Gafos et al. 2010), damit größere Werte einer größeren Überlappung entsprechen (was bei den anderen Autoren entgegengesetzt zu interpretieren ist).

4.3.2 Ergebnisse

Folgende Abbildungen zeigen die Dauer des intermediären Abstands für die BP- (links) und die EP-Daten (rechts). Die Obstruent-r-Token befinden sich in der oberen Hälfte der Abbildungen und die Obstruent-l-Token in der unteren.

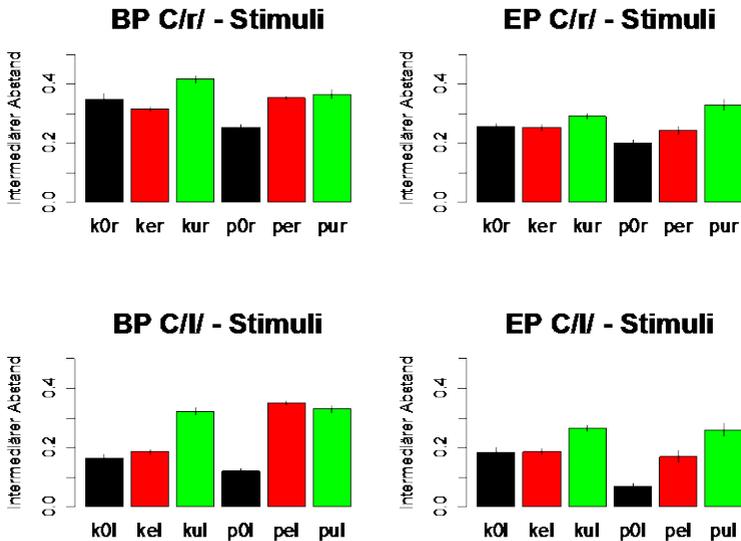


Abbildung 5: Dauer des intermediären Abstands (Mittelwerte über vier BP und fünf EP Sprecher mit Standardabweichungen)³⁶.

Der Abstand zwischen Plateaus war in beiden Varietäten signifikant ($\chi^2[2] = 495.8$, $p < 0.001$) und deutlich kleiner bei Clustern als bei beiden Sequenztypen mit intermediären Vokalen. Sequenzen mit dem hinteren Vokal /u/ zeigten durchgehend größere Abstände als die mit /e/. Wenn man die Weite der Koordinierung mit dem Abstand korrelieren würde, dann ergäbe

³⁶ Die Konsonantencluster wurden vereinheitlicht und parallel zu den anderen Stimulustypen mit drei Elementen dargestellt. Dafür wurden sie mit einer Null zwischen den Konsonanten verzeichnet, die für einen leeren Vokal(-platz) stehen soll. Die Dauer der Plateaus wurde nicht abgebildet.

sich bei $\langle C_1uC_2 \rangle$ die weiteste Koordinierung, gefolgt von $\langle C_1eC_2 \rangle$ und am engsten bei Clustern.

In der deskriptiven Statistik ist eine leichte Tendenz zu längeren intermediären Abständen im BP als im EP vorhanden ($\chi^2[1] = 5.9$, $p < 0.05$). Die größeren Abstände im BP bedeuten, dass die entsprechenden Konstriktionen weiter voneinander entfernt liegen als im EP und folglich die Koordination der Plateaus im BP weiter ist als im EP. Im Letzteren liegen die Gesten näher aneinander und die Koordination ist dementsprechend enger als im BP. Die Artikulationsart von C_2 übte ebenso einen starken Einfluss auf die Abstandsdauer aus ($\chi^2[1] = 190.8$, $p < 0.001$): /r/-Laute zeigten insgesamt deutlich längere Abstände als Laterale und somit wurden /r/-Token weiter auseinander koordiniert als /l/-Token. Die starken Interaktionen zwischen Stimulustyp * Sprechervarietät: $\chi^2[2] = 41.7$, $p < 0.001$, Stimulustyp * C_2 : $\chi^2[2] = 12.5$, $p < 0.01$ und C_2 * Sprechervarietät $\chi^2[1] = 15.7$, $p < 0.001$ kamen zustande, da sich die Dauer von $\langle C_1er \rangle$ ($p < 0.01$) und $\langle C_1ur \rangle$ ($p < 0.05$) zwischen den Varietäten unterschied, jedoch nicht die Dauer der lateralen Gegenstücke ($\langle C_1el \rangle$ und $\langle C_1ul \rangle$). Außerdem hatte die C_2 einen signifikanten Einfluss auf alle Stimulustypen im BP (alle $p < 0.001$), im EP jedoch nur bei Clustern ($p < 0.001$) und $\langle C_1eC_2 \rangle$ ($p < 0.01$).

C_1 und Betonung übten keinen signifikanten Einfluss auf die Dauer der intermediären Abstände aus. Es gab allerdings bei beiden Varietäten eine signifikante Interaktion zwischen Stimulustyp und C_1 . Tukey Post-Hoc-Tests zeigten, dass sich in beiden Varietäten $\langle C_1eC_2 \rangle$ und $\langle C_1uC_2 \rangle$ nach dem bilabialen C_1 unterscheiden, wobei der Abstand bei $\langle C_1eC_2 \rangle$ deutlich länger ist als bei $\langle C_1uC_2 \rangle$. Die Unterschiede nach dem velaren C_1 waren nicht signifikant. Die durchgeführte Analyse mit dem Mixed Model zeigt einen leichten Effekt der Sprechervarietät ($\chi^2[1] = 3.9$, $p < 0.05$) auf die Dauer des C_1 -Plateaus, indem sie in den BP-Daten länger ist als in den EP-Daten. Stimulustyp und Artikulationsart hatten keinen signifikanten Einfluss auf die Plateaudauer und es gab keine signifikanten Interaktionen.

Die pro Sprechvarietät durchgeführten Analysen zeigen einen hochsignifikanten Effekt des Artikulationsorts in beiden Varietäten ($\chi^2[1] = 34.0$, $p < 0.001$ in den BP- und $\chi^2[1] = 45.6$, $p < 0.001$ in den EP-Daten), der an der längeren Dauer bei bilabialen als bei velaren C_1 ersichtlich ist. In den BP-Daten gibt es außerdem eine signifikante Interaktion zwischen Stimulustyp und Artikulationsort ($\chi^2[2] = 23.4$, $p < 0.001$), die bedeutet, dass der Artikulationsort nicht den gleichen Einfluss auf die verschiedenen Stimulustypen hatte. Tukey Post-Hoc-Tests zeigten, dass die Dauer des C_1 -Plateaus signifikant länger bei bilabialen als bei velaren Clustern war ($p < 0.001$), jedoch nicht bei den CVC-Sequenzen. Die Betonung hatte keinen signifikanten Einfluss auf die Plateau-Dauer in beiden Varietäten und wird aufgrund ihres insgesamt marginalen Einflusses sowie aus Platzgründen hier nicht aufgeführt.

Anders als bei der Dauer des C_1 -Plateaus zeigte die Artikulationsart einen hochsignifikanten Einfluss auf die Dauer des C_2 -Plateaus ($\chi^2[1] = 490.8$, $p < 0.001$), das deutlich länger bei Lateralen als bei Taps war. Die Plateaudauer wurde ebenso vom Stimulustyp beeinflusst ($\chi^2[2] = 78.9$, $p < 0.001$): Insgesamt zeigen $\langle C_1eC_2 \rangle$ längere C_2 -Plateaus als die anderen Stimulustypen. Die Interaktion zwischen den Faktoren war allerdings nicht signifikant. Die fehlende Interaktion bedeutet, dass beide Faktoren unabhängig voneinander signifikant waren und sich nicht gegenseitig beeinflussten. D.h., die Stufen des Faktors Artikulationsart unterschieden sich zwar signifikant voneinander, zeigten aber die gleiche Tendenz bei den verschiedenen Stimulustypen. Diese unterschieden sich wiederum untereinander, zeigten jedoch die gleichen Tendenzen für die Stufen der Artikulationsart.

Die Sprechervarietät hatte keinen signifikanten Einfluss auf die Plateaudauer, eine signifikante Interaktion zwischen Sprechervarietät und Artikulationsart ($\chi^2[1] = 16.0$, $p < 0.001$) kam zustande, weil $\langle C_1eC_2 \rangle$ mit Lateral deutlich länger waren als die anderen Stimulustypen im EP ($\langle C_1eC_2 \rangle$ vs. Cluster: $p < 0.001$

und $\langle C_1eC_2 \rangle$ vs. $\langle C_1eC_2p \rangle$ ($p < 0.001$), jedoch nicht im BP. Hingegen wiesen Taps längere C_2 -Plateaus bei $\langle C_1eC_2 \rangle$ als bei $\langle C_1uC_2 \rangle$ und als bei Clustern (beide $p < 0.001$) im BP auf, aber nicht im EP. Die Dauer des C_2 -Plateaus war in beiden Varietäten signifikant länger bei betonten als bei unbetonten Token ($\chi^2[1] = 7.1$, $p < 0.001$ im BP- und $\chi^2[1] = 24.8$, $p < 0.001$ im EP). Außerdem hatte der Artikulationsort einen leichten Einfluss auf die Plateaudauer ($\chi^2[1] = 12.0$, $p < 0.001$), der durch die tendenziell längere Plateaudauer bei bilabialen im Vergleich zu velaren Token und bei betonten im Vergleich zu unbetonten Token zum Ausdruck kam.

So kann man zusammenfassend festhalten, dass die Dauer des C_1 -Plateaus länger in den BP- als in den EP-Daten ist und in beiden Varietäten deutlich länger nach bilabialen als nach velaren C_1 . Die weiteren Faktoren hatten keinen eindeutigen Effekt auf die Plateau-Dauer. Die Dauer der C_2 -Plateaus variierte mit der Artikulationsart und dem Stimulustyp: Sie ist länger bei Lateralen als bei Taps und bei $\langle C_1eC_2 \rangle$ als bei Sequenzen mit hinterem Vokal ($\langle C_1uC_2 \rangle$) oder Clustern. Es wurde ein schwacher Effekt der Betonung und des C_1 -Artikulationsorts bei EP-Token festgestellt, bei dem bilabiale C_1 längere C_2 -Plateaus auslöste. Der Einfluss der Sprechervarietät war marginal.

Die Länge der intermediären Abstands und die Koordinierung beider Plateaus fiel unterschiedlich bei Clustern vs. $\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$ in beiden Varietäten aus. Beide Varietäten tendierten zu einer vergleichsweise weiteren Koordinierung beider Plateaus bei $\langle C_1uC_2 \rangle$ und einer engeren bei Konsonantenclustern, BP-Token wurden allerdings insgesamt weiter koordiniert als EP-Token und Plosiv-r-Token weiter als Plosiv-l-Token. Der Artikulationsort von C_1 und die Betonung hatten keinen signifikanten Einfluss auf die Koordinierung beider Plateaus.

In Abbildung 6 werden die Synchronisierungsmittelwerte über fünf EP-Sprecher und vier BP-Sprecher mit den entsprechenden Standardabweichungen veranschaulicht, aus denen die Überlappung abgeleitet wurde. Die Plosiv-r-Token befinden sich

parallel zu den vorherigen Abbildungen in der oberen Hälfte und die Plosiv-l-Token in der unteren Hälfte der Abbildung. Die BP- wurden links und die EP- Token rechts dargestellt.

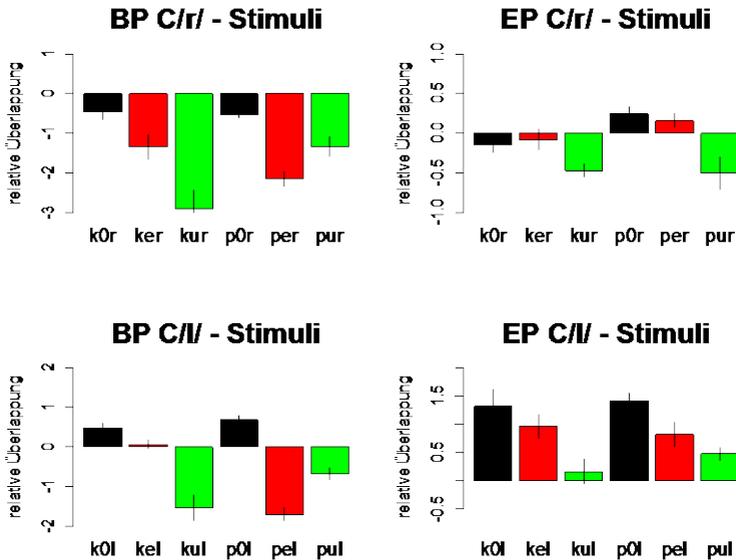


Abbildung 6: Übersicht der Synchronisierung der Gesten nach Liquid-Gruppen und Varietäten getrennt (Mittelwerte über fünf EP- und vier BP-Sprecher mit Standardabweichungen)³⁷.

Synchronisierung meint, zu welchem Zeitpunkt in C_1 der Beginn von C_2 bei Clustern und CVC-Sequenzen erfolgt und anschließend den Bezug zur Überlappung herstellt. Beginnt C_2 während des C_1 -Plateaus, findet man Überlappung der Gesten (Ausgangspunkt, *baseline*). Beginnt C_2 vor Beginn des C_1 -Plateaus, erfolgt vergleichsweise mehr Überlappung als im Ausgangspunkt. Falls C_2 nach Ende des C_1 -Plateaus beginnt, findet keine

³⁷ Auch hier wurden Konsonantencluster vereinheitlicht und parallel zu den anderen Stimulustypen mit drei Elementen dargestellt. Dafür wurden sie mit einer Null zwischen den Konsonanten verzeichnet, die für einen leeren Vokal(-platz) steht.

Überlappung statt und die Synchronisierung der Gesten wird als weit eingeordnet.

Aus den Gesamtergebnissen wird ersichtlich, dass die Konsonantengesten nach Stimulustyp (3-stufig: C_1C_2 vs. $\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$) anders synchronisiert werden und somit übe der Stimulustyp einen hochsignifikanten Einfluss auf die Verteilung aus ($\chi^2[2] = 150.6$, $p < 0.001$): Die höchsten positiven Werten ergaben sich bei den Konsonantenclustern (C_1C_2). Positive Werte stehen für Überlappung der Gesten und infolgedessen zeigen Cluster mehr Überlappung der Gesten als die anderen Stimulustypen. Bei den negativen Werten, die keiner Überlappung der Gesten entsprechen, lagen die Werte der Cluster nahe an Null, was einem geringeren Abstand zwischen beiden Gesten entspricht als bei CVC-Sequenzen. Insgesamt haben Cluster also eine vergleichsweise engere Koordinierung als die anderen Token. Die Sequenzen zeigen insgesamt das umgekehrte Bild: größere negative Werte als Cluster, die für einen größeren intergestischen Abstand stehen sowie kleine positive Werte, die eine geringere Überlappung bedeuten. Somit begann C_2 später in C_1 und es überlappte nicht oder zu einem kleineren Anteil.

Die Sprechervarietät hatte einen signifikanten Einfluss auf die Synchronisierung der Token ($\chi^2[1] = 8.7$, $p < 0.01$) mit größeren negativen Werten im BP und größeren positiven Werten im EP. Dies erklärt die Daten allerdings nur teilweise aufgrund der weiteren Effekte und Interaktionen: Die hochsignifikante Interaktion zwischen Sprechervarietät und Stimulustyp ($\chi^2[2] = 27.8$, $p < 0.01$), kam zustande weil, sich $\langle C_1eC_2 \rangle$ ($p < 0.001$) und $\langle C_1uC_2 \rangle$ ($p < 0.01$) zwischen den Varietäten unterscheiden, jedoch nicht die Cluster. Insgesamt waren die Token mit Tap weiter koordiniert ($\chi^2[1] = 143.0$, $p < 0.01$) als die mit Lateral: Die Überlappungsmittelwerte der Plosiv-r-Token zeigten größere negative oder kleinere positive Werte als Plosiv-l-Token und die Gesten wiesen dementsprechend entweder keine oder eine geringere Überlappung auf. Im Vergleich dazu zeigten Plosiv-l-

Token größere positive Werte (mehr Überlappung) und kleinere negative Werte (kürzerer Abstand zwischen den Konsonanten).

Die fehlende Interaktion zwischen Sprechervarietät und Artikulationsart wird in der deskriptiven Statistik daraus ersichtlich, dass die gleichen Tendenzen bei den verschiedenen Artikulationsarten und Sprechervarietäten zum Ausdruck kamen: Plosiv-r-Token überlappten weniger als Plosiv-l-Token, sowohl im BP als auch im EP. Die von BP-Sprechern produzierten Token überlappten weniger als die von EP-Sprechern, dies war sowohl für /r/- als auch für Plosiv-l-Token der Fall.

Der Artikulationsort hatte keinen signifikanten Effekt bei der Synchronisierung der Konsonanten, es tritt allerdings eine signifikante Interaktion zwischen Stimulustyp und Artikulationsort in den BP-, aber nicht in den EP-Daten auf. Für die BP-Token zeigten Post-Hoc-Tests einen signifikanten Einfluss des Artikulationsortes bei $\langle C_1eC_2 \rangle$ ($p < 0.001$) und $\langle C_1uC_2 \rangle$ ($p < 0.01$), jedoch nicht bei Clustern. Die negativen Werte waren größer und somit die Überlappung geringer bei $\langle C_1eC_2 \rangle$ nach bilabialem und bei $\langle C_1uC_2 \rangle$ nach velarem Konsonanten.

Die Betonung³⁸ ($\chi^2[1] = 7.2$, $p < 0.01$ in den BP- und $\chi^2[2] = 4.9$, $p < 0.05$ in den EP-Daten) hatte ebenfalls einen signifikanten Einfluss auf die Überlappung. Im Vergleich zu den betonten Token zeigen unbetonte kleinere negative (-10) oder größere positive Überlappungsmittelwerte (+8, siehe Anhang A). Daher war die Koordinierung der Gesten enger in der unbetonten als in der betonten Kondition.

Zusammenfassend gesagt zeigte dieses Messverfahren, dass Cluster vs. $\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$ jeweils eine unterschiedliche Koordinierung aufweisen: Die Konsonanten hatten eine engere Synchronisierung bei Clustern und eine weitere bei $\langle C_1uC_2 \rangle$. Die

³⁸ Dieser Faktor ist um der Übersichtlichkeit willen nicht auf der letzten Abbildung zu sehen, siehe Cunha, *submitted* für die vollständige Analyse.

Sequenz $\langle C_1uC_2 \rangle$ zeigte intermediäre Werte. Sie wurde von Sprechervarietät, C2-Artikulationsart und Betonung signifikant beeinflusst, wobei die Interaktion zwischen Stimulustyp und Artikulationsort im BP ebenso eine Rolle spielte. Insgesamt ergab sich bei den BP-Token eine vergleichsweise geringere Überlappung der Gesten und daher eine weitere Synchronisierung als bei den EP-Token, ebenso wie bei Plosiv-r-Token im Vergleich zu Plosiv-l-Token. So könnten die Stimulustypen in ein Kontinuum eingeordnet werden, in dem die betonte Sequenz $\langle C_1ur \rangle$ im BP die weiteste und die unbetonten Plosiv-Liquid-Cluster im EP die engste Koordinierung aufweisen. Bei den unbetonten Token ergab sich insgesamt eine engere Koordinierung als bei den betonten. Im BP war die Überlappung außerdem von der Interaktion zwischen Stimulustyp und Artikulationsort geprägt, wobei $\langle C_1eC_2 \rangle$ weniger Überlappung nach bilabialem und $\langle C_1uC_2 \rangle$ nach velarem C₁-Artikulationsort aufwiesen.

4.4 Räumliche intragestische Koordination (*Phasing*)

Im vorherigen Abschnitt beschäftigten wir uns mit der Koordination beider Konsonanten. Im Folgenden liegt nun der Schwerpunkt auf den räumlichen Eigenschaften der Zungenspitze (Jancke 1997, Roon 2007), die an der Produktion beider Liquide und beider Vokale der CVC-Sequenzen beteiligt ist. Dabei kann man untersuchen, inwiefern festgestellte Unterschiede zwischen Clustern vs. $\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$ auf die räumliche Organisation zurückzuführen sind und es werden die gleichen Faktoren miteinbezogen.

4.4.1 Definition der Variablen

Für die Analyse der Dynamik der Zungenspitze wurden die maximale Geschwindigkeit und der räumliche Verlauf der vertikalen Bewegung der Zungenspitze (ITY-Umfang & TTY-Geschwindigkeit) wie folgt berechnet.

- TTY-GESCHWINDIGKEIT (MM/MS). Dies entspricht dem höchsten Wert der Geschwindigkeit der vertikalen Zungenspitzenbewegung (IT-Y).
- TTY-UMFANG (MM) = TTY-MAX - TTY-MIN. Dies entspricht der Differenz zwischen der höchsten und der niedrigsten Position der vertikalen Zungenspitzenbewegung (IT-Y).

Aus diesen zwei Messverfahren wird ersichtlich, inwiefern sich die räumliche Organisation der Konsonanten bei Clustern und CVC-Sequenzen im Hinblick auf die maximale Geschwindigkeit und den Umfang der Zungenspitzenbewegung unterscheidet.

4.4.2 Ergebnisse

Es mussten einige /l/-Token von EP-Sprechern von der Auswertung ausgeschlossen werden, da manche Zungenspitzenbewegungen aufgrund ihres flachen Verlaufs nicht erfasst werden konnten. Wie Oliveira et al. (2011) anhand von MRI-Daten³⁹ feststellten, wird im EP der dunkle Lateral /ɫ/ von manchen Sprechern mit dem Zungenblatt (laminal) produziert. Eine solche Bewegung kann nur ungenau durch den Sensor an der Zungenspitze (apikal) erfasst werden. Eine laminale Bewegung könnte eventuell mit dem Verlauf des Sensors an der Zungenmitte verfolgt werden, dies würde allerdings die Vergleichbarkeit der Token in Frage stellen (Zungenspitze bei BP-Sprechern vs. Zungenmitte bei EP-Sprechern) und wenig zur aktuellen Fragestellung beitragen. Daher wurde dies für die vorliegende Arbeit nicht durchgeführt.

Die maximale Geschwindigkeit der Schließgeste der Zungenspitzenbewegung wird vom Sprecher besser kontrolliert

³⁹ *Magnetic resonance imaging*, dt. Magnetresonanz Imaging. Dies ist ein bildgebendes Untersuchungsverfahren zur Darstellung des Körperinneren eines Menschen, ohne diesen dabei mit Röntgenstrahlen zu belasten (z.B. <http://www.rodia.ch/default.asp?id=2010000%20&siteid=1&langid=de>).

als die Öffnungsgeste (Byrd & Saltzman 1998). Die Mittelwerte über fünf EP-Sprecher und vier BP-Sprecher mit den entsprechenden Standardabweichungen wurden in Abbildung 7 dargestellt. Wie in den vorherigen Abbildungen befinden sich die Plosiv-r-Token in der oberen und die Plosiv-l-Token in der unteren Hälfte der Abbildung, wobei die BP-Token links und die EP-Token rechts stehen.

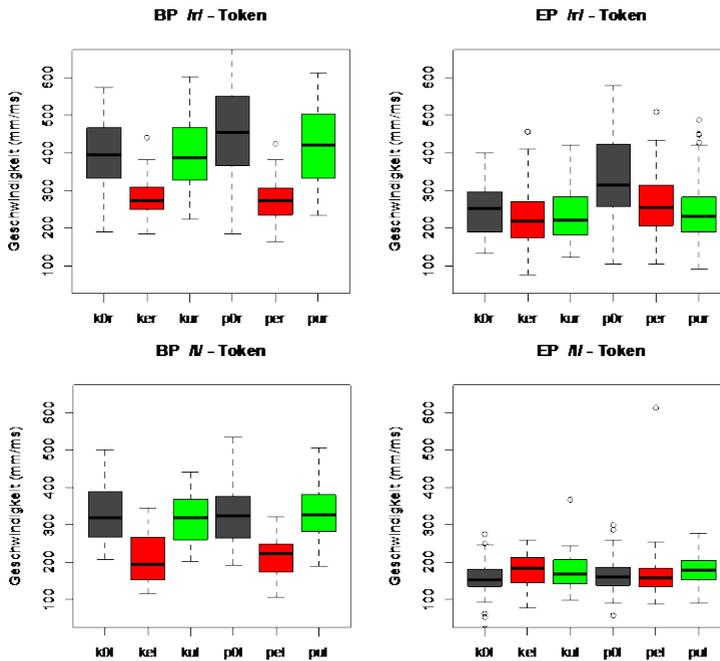


Abbildung 7: Maximale Geschwindigkeit des vertikalen Verlaufs der Schließungsgeste der Zungenspitze (ITY, Mittelwerte über fünf EP- und vier BP-Sprecher mit Standardabweichungen).

Die maximale Geschwindigkeit der Schließgeste variierte insgesamt stark nach Stimulustyp ($\chi^2[2] = 298.0$, $p < 0.001$) und Sprechervarietät ($\chi^2[2] = 11.2$, $p < 0.001$) mit einer insgesamt höheren Geschwindigkeit in BP- als in EP-Token. Noch aussagekräftiger war allerdings die hochsignifikante Interaktion

zwischen Stimulustyp und Sprechvarietät ($\chi^2[2] = 214.8$, $p < 0.001$), die sich dadurch kennzeichnet, dass $\langle C_{1u}C_2 \rangle$ und Cluster eine deutlich höhere Geschwindigkeit im BP als im EP aufweisen ($p < 0.001$ für beide Stimulustypen). Für $\langle C_{1e}C_2 \rangle$ ergaben sich keine Geschwindigkeitsunterschiede zwischen den Varietäten.

Die für jede Varietät getrennt durchgeführten *Mixed Models* zeigen einen deutlichen Einfluss des Stimulustyps auf die maximale Geschwindigkeit bei den BP- ($\chi^2[2] = 384.0$, $p < 0.001$), jedoch nicht bei den EP-Daten ($p = 0.09$). Im BP wies $\langle C_{1e}C_2 \rangle$ durchgehend eine vergleichsweise geringere Geschwindigkeit auf als die anderen Stimulustypen.

Die Artikulationsart beeinflusste ebenso auf eine hochsignifikante Weise die Geschwindigkeit der Schließgeste, in der Hinsicht, dass Plosiv-l-Token mit einer niedrigeren Geschwindigkeit produziert wurden als die entsprechenden r-Token ($\chi^2[1] = 532.3$, $p < 0.001$). Die hochsignifikante Interaktion mit dem Stimulustyp ($\chi^2[2] = 31.4$, $p < 0.001$) kam zustande, indem im EP $\langle C_{1e}C_2 \rangle$ und $\langle C_{1u}C_2 \rangle$ von Clustern mit zentralen Approximanten unterschieden werden können (beide $p < 0.001$), jedoch nicht mit Lateralen. Der Einfluss des Artikulationsorts ($\chi^2[1] = 7.9$, $p < 0.001$ für BP- und $\chi^2[1] = 14.6$, $p < 0.001$ für EP-Daten) ist an der durchgehend höheren Geschwindigkeit bei bilabialen als bei velaren Token erkennbar. Die Betonung hatte in keiner der Varietäten einen signifikanten Einfluss auf die untersuchten Phänomene

Der räumliche Umfang in Abbildung 8 entspricht dem Abstand in Millimetern zwischen der höchsten und der niedrigsten Position der vertikalen Zungenspitzenbewegung.

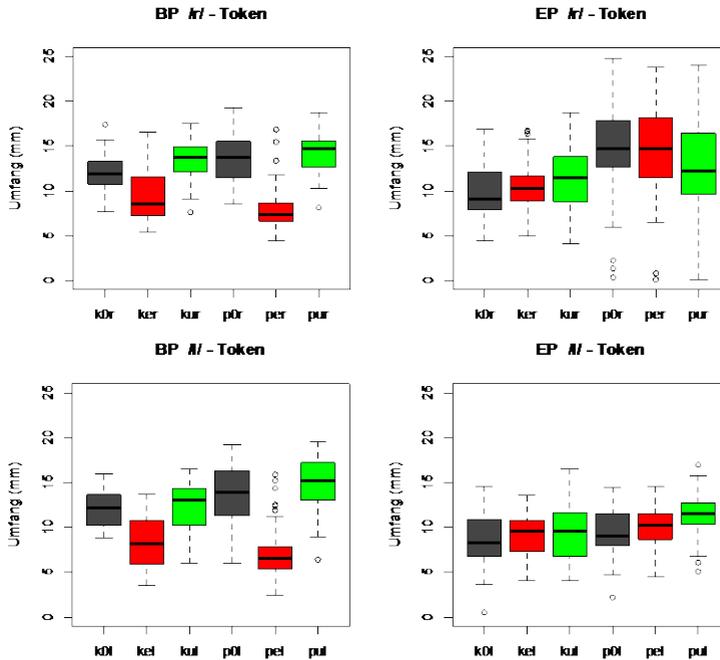


Abbildung 8: Räumlicher Umfang der Zungenspitzenbewegung (Mittelwerte über fünf EP- und vier BP- Sprecher mit Standardabweichungen).

Die Plosiv-r-Token wurden in der oberen und die Plosiv-l-Token in der unteren Hälfte der Abbildung dargestellt, wobei die BP-Token links und die EP-Token rechts stehen. Der räumliche Umfang der Zungenspitzenbewegung wurde vom Stimulustyp stark beeinflusst ($\chi^2[2] = 206.8$, $p < 0.001$), allerdings zeigten nach Varietät getrennte Analysen, dass der Einfluss des Stimulustyps erheblich größer im BP ($\chi^2[2] = 570.4$, $p < 0.001$) als im EP ($\chi^2[2] = 8.4$, $p < 0.05$) war. Die Sprechervarietät an sich hatte keinen signifikanten Einfluss auf den Gestenumfang ($p = 0.6$), es wurde jedoch eine hochsignifikante Interaktion zwischen Stimulustyp und Sprechervarietät festgestellt ($\chi^2[2] = 247.8$, $p < 0.001$). Tukey Post-Hoc-Tests zeigten, dass sich der Umfang bei Clustern vs.

$\langle C_1eC_2 \rangle$ im BP unterschied ($p < 0.001$), jedoch nicht im EP. Außerdem war der Umfang im EP bei Clustern vs. $\langle C_1uC_2 \rangle$ unterschiedlich ($p < 0.01$), was allerdings im BP nicht zutraf. Die Prüfstatistik bestätigte also den Eindruck aus der Abbildung: Der räumliche Zungenspitzenumfang verändert sich im EP nur zwischen Cluster vs. $\langle C_1uC_2 \rangle$ und unterscheidet sich im BP nur bei Clustern vs. $\langle C_1eC_2 \rangle$. Bei Letzterem ergab sich im BP ein deutlich geringerer Umfang als bei den anderen Token. Die Artikulationsart hat einen positiven Einfluss auf den Gestenumfang der Zungenspitze ($\chi^2[1] = 99.1$, $p < 0.001$), vor allem bei den EP- ($\chi^2[1] = 126.0$, $p < 0.001$) im Vergleich zu den BP-Daten ($\chi^2[1] = 5.5$, $p < 0.05$). In den EP-Daten zeigen Laterale einen geringeren Umfang als Taps. In den BP-Daten ist der Gestenumfang deutlich kleiner bei $\langle C_1eC_2 \rangle$ im Vergleich zu $\langle C_1uC_2 \rangle$ bei den Token mit dem zentralen Approximanten ($p < 0.001$), aber nicht mit Lateral. Der Zungenspitzenumfang war geringer nach einem velaren als nach einem bilabialen C_1 ($\chi^2[1] = 14.0$, $p < 0.001$ bei BP und $\chi^2[1] = 91.9$, $p < 0.001$ bei EP), wurde jedoch von der Betonung nicht beeinflusst.

Die wichtigen Ergebnisse sind, dass die maximale Geschwindigkeit der Schließgeste im BP bei $\langle C_1uC_2 \rangle$ und bei Clustern höher war als bei $\langle C_1eC_2 \rangle$, jedoch nicht im EP. Die r-Gesten wurden mit einer höheren Geschwindigkeit produziert als die l-Gesten, ebenso wie die BP- im Vergleich zu den EP-Gesten. Eine etwas höhere Geschwindigkeit wurde ebenso von dem bilabialen Artikulationsort der Plosive begünstigt, die Betonung beeinflusste sie allerdings in keiner Hinsicht. Der räumliche Umfang der Zungenspitzenbewegung wurde zwar vom Stimulustyp beeinflusst, allerdings war seine Auswirkung größer auf die BP- als auf die EP-Token. Die Sprechervarietät hatte keinen Einfluss auf den Umfang im Allgemeinen, sondern sich lediglich zwischen Clustern vs. $\langle C_1eC_2 \rangle$ im BP und Clustern vs. $\langle C_1uC_2 \rangle$ im EP. Der Einfluss der Artikulationsart und des Artikulationsorts auf den Umfang war vor allem im EP erheblich, indem Laterale einen geringeren Umfang zeigen als Taps und velare C_1 einen geringeren Umfang begünstigten als Bilabiale.

4.4.3 Korrelation beider Faktoren

Darüber hinaus wurde eine starke Korrelation zwischen der maximalen Geschwindigkeit und dem Umfang der Zungenspitzengeste festgestellt. In Abbildung 9 wurde der räumliche Umfang auf die Y-Achse und die maximale Geschwindigkeit auf die X-Achse aufgetragen.

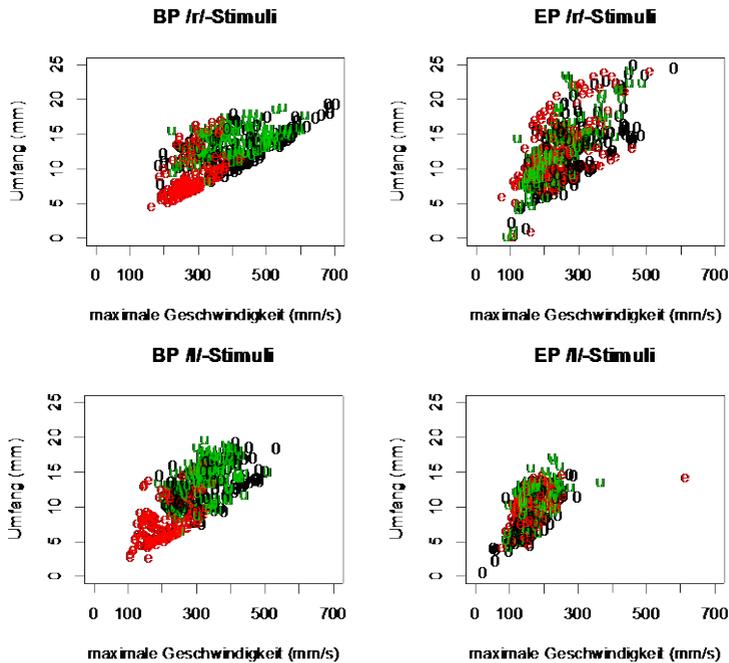


Abbildung 9: Korrelation des Umfangs und der maximalen Geschwindigkeit der Zungenspitzengeste (schwarze /o/ stehen für Cluster, rote /e/ für $\langle C_1eC_2 \rangle$ und grüne /u/ für $\langle C_1uC_2 \rangle$, Mittelwerte über fünf EP- und vier BP-Sprecher mit Standardabweichungen).

Anhand dieser Abbildung ist noch einmal ersichtlich, dass sich die Stimulustypen bei beiden Verfahren eher in den BP- als in den EP-Daten unterscheiden, denn nach Farben getrennte Gruppen sind eher in der linken (BP) als in der rechten Hälfte (EP) der

Abbildung zu erkennen. Es ist außerdem eindeutig erkennbar, dass Laterale niedrigere Geschwindigkeiten und Gesten mit einem kleineren Umfang aufweisen als alveolare Taps. Die Konzentration der Werte in beiden Verfahren vor allem im EP ist stärker ausgeprägt bei den r-Token als bei den Lateralen. Dies bestätigt das Ergebnis der Statistik, dass Geschwindigkeit und Umfang größer bei Taps als bei Liquiden sind. Die Geschwindigkeit und der Umfang der Geste sind größer im BP als im EP, da die Konzentration der Werte auf beiden Achsen im BP höher liegt als bei EP. Bei den europäischen Taps können die Stimulustypen aufgrund der ähnlichen Geschwindigkeiten und Umfänge der Gesten nicht mehr differenziert werden.

4.4.4 Diskussion der wichtigsten Ergebnisse

Der Stimulustyp übte keinen signifikanten Einfluss auf die Dauer des C_1 -Plateaus, aber auf die Dauer des C_2 -Plateaus aus, indem sie länger bei $\langle C_1eC_2 \rangle$ als bei $\langle C_1uC_2 \rangle$ oder Clustern war. Dies kann mit Koartikulationseffekten in Verbindung gebracht werden, in der Hinsicht, dass Plosive resistenter gegen Koartikulation sind als Liquide (z.B. Farnetani & Recasens 1997, Recasens 2011, 2001, Recasens & Pallarès 1999).

Die Dauer des intermediären Abstands und die darauffolgende Koordinierung der Plateaus sowie die Synchronisierung und Überlappung beider Konsonantengesten waren je nach Stimulustyp unterschiedlich: Der intermediäre Abstand war am längsten bei $\langle C_1uC_2 \rangle$, gefolgt von $\langle C_1eC_2 \rangle$ und am kürzesten bei Clustern. Folglich ergab das Verfahren eine weitere Koordinierung der $\langle C_1uC_2 \rangle$ und $\langle C_1eC_2 \rangle$ -Token, während Cluster eine enge Koordinierung hatten. Ein ähnliches Ergebnis zeigte die Analyse der Synchronisierung, bei der Cluster ebenso enger koordiniert wurden als $\langle C_1eC_2 \rangle$ und $\langle C_1uC_2 \rangle$. Zwischen $\langle C_1eC_2 \rangle$ und $\langle C_1uC_2 \rangle$ gab es zwar Unterschiede, aber ihre Koordinierung war insgesamt weiter als die von Clustern. Die übereinstimmenden Ergebnisse zeigen vor allem, wie robust die engere Koordination der Cluster im Vergleich zu den beiden CVC-Sequenzen ist.

Die Sprechervarietät wirkte sich signifikant auf die Dauer des C₁-Plateaus und auf den intermediären Abstand sowie auf die Synchronisierung beider Konsonantengesten aus. Die längeren intermediären Abstände im BP bedeuten, dass die entsprechenden Konstriktionen weiter voneinander entfernt liegen als im EP und folglich die Koordination der Plateaus im BP weiter ist als im EP. Im Letzteren liegen die Gesten näher aneinander und die Koordination ist dementsprechend enger als im BP. Auch die Synchronisierung zeigte, dass die beiden Konsonanten in BP-Token signifikant weiter koordiniert werden als in EP-Token. Dies war angesichts der mehrfach beschriebenen Vokaltilgung im EP (Martins 1975, Silva 1997, 1998, Cunha 2011a) zu erwarten.

Außerdem war die Dauer des C₂-Plateaus länger bei Lateralen als bei Taps in beiden Varietäten. Auch dies war zu erwarten, da ein Tap mit einer schnelleren und ballistischen Bewegung der Zungenspitze sowie einem kurzen Verschluss produziert wird und die Konstriktionsdauer von /t/ deutlich kürzer war als die von /l/. Bei Lateralen kann dagegen der zentrale Verschluss beliebig lange gehalten werden (Proctor 2009) und die Luft strömt über die Seitenränder der Zunge aus. Die Plosiv-r-Token wiesen insgesamt deutlich längere intermediäre Abstände und eine breitere Synchronisierung mit weniger Überlappung beider Konsonantengesten auf als die Plosiv-l-Token. Daher waren die Konsonanten bei den Plosiv-r-Token weiter auseinander koordiniert als bei den Plosiv-l-Token. Alles deutet darauf hin, dass die beiden Liquide anders koordiniert werden, und zwar beginnen die Taps zu einem späterem Zeitpunkt im Vergleich zu den Lateralen. Dies hängt möglicherweise mit der kurzen Konstriktion von Taps zusammen, die erst nach dem Plosiv einsetzt, damit sie vom Plosiv nicht verdeckt wird.

Ob das erste Element der Token (C_i) labial oder velar war, hatte keinen aussagekräftigen Einfluss auf die Konsonantenkoordination. Bei dem aktuellen Datensatz konnte die von Chitoran et al. (2002) und Kühnert et al. (2006) beschriebene Tendenz zur stärkeren Überlappung bei *front-to-back* Clustern im Vergleich zu Clustern mit der umgekehrten Richtung (*back-to-front*)

nicht bestätigt werden. Allerdings wurden signifikante Interaktionen in beiden Koordinationsverfahren bei den CVC-Sequenzen festgestellt: Der intermediäre Abstand war in beiden Varietäten länger bei $\langle C_1eC_2 \rangle$ als bei $\langle C_1uC_2 \rangle$ und länger bei bilabialem als bei velarem C_1 . Dies spricht dafür, dass $\langle C_1eC_2 \rangle$ und $\langle C_1uC_2 \rangle$ stärker in bilabialem als in velarem Kontext neutralisiert sind. Das Synchronisierungsverfahren zeigt jedoch keine signifikanten Unterschiede zwischen den Artikulationsorten im EP auf. Im BP hingegen ergab sich eine starke Interaktion von Artikulationsort und Vokal, wobei $\langle C_1eC_2 \rangle$ weniger Überlappung nach bilabialem und $\langle C_1uC_2 \rangle$ nach velarem C_1 -Artikulationsort aufwiesen.

Die Betonung hatte einen marginalen Effekt auf die Dauer des C_1 -Plateaus sowie auf die intermediären Abstände, indem betonte Token etwas längere Dauern als unbetonte zeigten. Dies hängt teilweise mit der Verwendung der absoluten Dauer im Verfahren zusammen und könnte möglicherweise durch eine Zeitnormalisierung minimiert werden. Allerdings zeigte sich der gleiche marginale Effekt auf die Synchronisierung beider Konsonantengesten, bei dem Token enger in der unbetonten als in der betonten Kondition koordiniert wurden.

Bezüglich der Dynamik der Zungenspitzenegeste zeigten die Ergebnisse, dass sich sowohl die maximale Geschwindigkeit als auch der räumliche Umfang insbesondere im BP unterschieden: Beide Verfahren wiesen insgesamt geringere Werte bei $\langle C_1eC_2 \rangle$ als bei Clustern oder $\langle C_1uC_2 \rangle$ auf. Der Grund hierfür ist die mittlere Zungenhöhe des vorderen Zungenteils für die Produktion von /e/. Diese intermediäre Höhe verringert im BP den Abstand zwischen der Zungenspitze und den Alveolen für die (anschließende) Produktion von C_2 und dies betrifft auch den Umfang der vertikalen Position der Zungenspitze. Die Geschwindigkeit wird womöglich entsprechend angepasst.

Bei den hinteren Vokalen findet die Zungenerhöhung im Bereich des Zungendorsums statt und hat daher einen geringeren Einfluss auf die Zungenspitze. Dies würde erklären, warum Sequenzen mit

dem hinteren Vokal (<C₁uC₂>) und Cluster einen ähnlichen Umfang aufweisen. Das Gleiche gilt für den hinteren Vokal im EP. Der andere analysierte Vokal ist hoch und zentriert (/i/). Im Hinblick auf die Tatsache, dass dieser Vokal ähnliche Geschwindigkeiten und Umfänge wie die Cluster aufwies, wäre interessant zu prüfen, inwiefern bei diesem Vokal ein Target vorhanden ist, d.h. ob der Vokal eine zusätzliche Geste hervorruft oder nicht (wie für das Schwa /ə/ im Englischen diskutiert wurde, Browman & Goldstein 1990). Der Einfluss des Stimulustyps auf die EP-Daten war nicht signifikant bei der Geschwindigkeit und gering beim Gestenumfang. Dies bringt zum Ausdruck, dass die räumliche Organisation der Gesten bei allen drei Stimulustypen im EP stark neutralisiert ist. Diese Neutralisierung war angesichts der verbreiteten Vokaltilgung (Cunha 2011a, Silva 1998, 1997, Martins 1975, u.a.) nicht unerwartet.

Zentrale Approximanten wurden mit einer höheren Geschwindigkeit produziert und erreichten einen höheren Umfang als Laterale, was angesichts der schnelleren ballistischen Bewegung der Zungenspitze für die Produktion der Taps im Vergleich zu den Lateralen vorhersehbar war. Bilabiale Plosive in der C₁-Position begünstigten eine hohe Geschwindigkeit und einen größeren Umfang der Geste als die velaren Gegenstücke. Dies könnte einerseits mit der Unabhängigkeit der Artikulatoren bei den bilabial-apikalen Token (z.B. /pr/ – /pl/) zusammenhängen und andererseits mit der langsamen, trägen Bewegung des Zungendorsums, die die Zungenspitze beeinträchtigen kann.

4.5 Zusammenfassung und Ausblick

Das Ziel dieses Kapitel war es, ausgehend von akustischen und artikulatorischen Daten aus zwei portugiesischen Varietäten, zu untersuchen, inwiefern sich die Organisation von Konsonanten in Clustern und CVC-Sequenzen in der Produktion der Sprache unterscheidet und ob dabei mögliche Neutralisierungen in der lautlichen Realisierung auf Variabilität in der Koordination der Konsonanten im Rahmen der Artikulatorischen Phonologie zurückzuführen ist.

Das Kapitel zeigte, dass sich die zeitliche Koordination von Konsonanten unterschied, je nachdem, ob sie Teil eines Clusters oder einer CVC-Sequenz waren: In einem Cluster waren die Konsonantengesten enger und in einer CVC-Sequenz weiter koordiniert. Dabei ist allerdings zu berücksichtigen, dass Plosiv-r-Cluster in beiden Varietäten und in beiden angewendeten Verfahren deutlich weiter koordiniert waren als die Plosiv-l-Cluster. Die grundsätzlich weite Koordinierung der r-Laute mag der Grund sein, weshalb im Falle des Auftretens eines kurzen intermediären vokalischen Elements dieses bei r-Clustern oft mit nicht-phonematischen, vokalischen Elementen – den sogenannten svarabhaktischen Vokalen – assoziiert wird (Nishida 2009, Bradley 2001, 2004, 2006, Bradley & Schmeiser 2003, Almeida & Dorta 1993) und bei l-Clustern von epenthetischen Vokalen mit einem silbischen Charakter die Rede ist (Velo 2006, 2007b). Dies würde außerdem erklären, weshalb l-Cluster bei einer kleineren Änderung der Weite der Koordination leichter mit CVC-Sequenzen verwechselt werden als r-Cluster. Diese Hypothesen müssten allerdings mithilfe weiterer Daten überprüft werden.

Die zeitlichen Koordinationsunterschiede zwischen Clustern und CVC-Sequenzen könnten durch den Einfluss weiterer Faktoren minimiert werden. In unseren Daten hatte der Vokal im Stimulustyp einen signifikanten Einfluss, wobei hintere Vokale weiter koordiniert waren als nicht-hintere Vokale (außer nach bilabialen C_1 im BP bei dem Synchronisierungsverfahren). Dies würde erklären, warum Sequenzen mit einem nicht-hinteren Vokal ($\langle C_1eC_2 \rangle$) anfälliger für Vokaltilgungen zumindest im EP sind (Cunha 2011a, Silva 1997, 1998) und im Kapitel 4 werden wir überprüfen, ob diese Token auch in der Perzeption häufig mit Clustern verwechselt werden.

Die Sprechervarietät spielte eine wichtige Rolle bei der Unterscheidung zwischen Clustern und CVC-Sequenzen, indem die Koordinationsunterschiede in den BP-Daten stärker ausgeprägt waren als in den EP-Daten. Es ist infolgedessen nicht unerwartet, dass es zu mehr Verwechslungen zwischen CVC-

Sequenzen und Clustern im EP als im BP kommt, worauf ebenfalls im Kapitel 6 eingegangen wird.

Die räumliche Organisation der Zungenspitze wies zudem deutliche Unterschiede in der Produktion des apikalen Lauts bei Clustern vs. $\langle C_1eC_2 \rangle$ im BP auf, allerdings nicht im EP. Bei Letzterem unterschieden sich die Stimulustypen nicht aufgrund ihrer maximalen Geschwindigkeit und nur geringfügig aufgrund ihres Umfangs. Dies führte dazu, dass die Zungenspitzenbewegungen beider CVC-Sequenzen im EP denen von Clustern sehr ähneln. Sollte sich die Perzeption vor allem an der räumlichen und nicht an der zeitlichen Organisation der Konsonanten orientieren, würde dies vorhersagen, dass im EP nicht nur $\langle C_1eC_2 \rangle$, sondern beide CVC-Sequenzen mit Clustern verwechselt würden.

Im BP hatten nicht beide CVC-Sequenzen den gleichen Einfluss auf die Analyseverfahren: Die räumliche Organisation der Konsonanten wies größere Unterschiede zwischen Clustern und $\langle C_1eC_2 \rangle$ als zwischen Clustern und $\langle C_1uC_2 \rangle$ auf. Da die Ergebnisse bei der zeitlichen Koordinierung genau umgekehrt waren, könnte man anhand von BP-Beispielen untersuchen, ob es zu mehr Verwechslungen zwischen Clustern und $\langle C_1eC_2 \rangle$ oder zwischen Clustern und $\langle C_1uC_2 \rangle$ in der Perzeption kommt und somit, ob die Perzeption sich eher an der zeitlichen oder an der räumlichen Organisation der Konsonanten orientiert. In der Zungenspitzenbewegung gab es weniger Unterschiede bei allen drei Stimulustypen mit Lateralen als mit zentralen Approximanten. Dies lässt für die Perzeption vorhersagen, dass im EP Stimuli mit Lateralen leichter zu verwechseln sind als bei r-Token.

Der Artikulationsort und die Betonung hatten keinen stabilen Effekt auf die Fragestellung. Bezüglich des Artikulationsorts könnte dies mit der geringeren Relevanz des assoziierten Produktionsrichtungseffekts in Plosiv-Liquid-Token in Verbindung stehen, der aufgrund der schwachen Maskierung anderer Konsonanten durch Liquide nicht zum Ausdruck kommt.

Sollte dies der Fall sein, würde der Artikulationsort eine wichtige Rolle im nächsten Kapitel spielen. Dort werden die gleichen Messverfahren auf die Koordination von wortmedialen Plosiven in Clustern und CVC-Sequenzen angewendet.

5 Die Koordination von Plosiven

5.1 Einführender Forschungsstand

Basierend auf der inter- und intragestischen Koordination von Konsonanten liegt der Fokus dieses Kapitels auf der genaueren Organisation von Plosiven bei der Produktion in reinen Clustern und in CVC-Sequenzen. Reine Plosivcluster haben im Portugiesischen eine eingeschränkte Distribution, sie sind aber äußerst interessant, da sie synchron besonders variable und diachron instabil sind. Diese Cluster lassen sich nicht mithilfe der Sonoritätshierarchie (z.B. Vennemann 1998) beschreiben, sie treten jedoch auf. Eine detaillierte Analyse könnte zu einem besseren Verständnis dieser Cluster beitragen sowie aufschlussreiche Erkenntnisse bezüglich ihrer synchronen Variation, ihrer seltenen initialen Position im Wort, ihrer geringen lexikalischen Frequenz nicht nur im Portugiesischen (Bisol 1999, Mateus & d'Andrade 2000, Veloso 2007a), sondern auch in zahlreichen anderen romanischen (Jensen 1999) und weiteren Sprachen (Maddieson 1993) und nicht zuletzt bezüglich ihrer äußerst ausgeprägten diachronen Instabilität liefern.

Aus Sicht der Sprachproduktion wird bei Plosivclustern mit einer starken Beeinträchtigung des ersten Plosivs durch den zweiten Plosiv gerechnet, welche durch die nicht parallele Übertragung (Ladefoged & Maddieson 1996) der einzelnen Segmente zustande kommt. Infolgedessen nimmt man bisher bei diesen Clustern im Grunde eine geringere Überlappung ihrer Bestandteile an als bei anderen gemischten Konsonantenclustern, wie beispielsweise den im vorherigen Kapitel besprochenen Plosiv-Liquid-Clustern. In bereits erforschten Sprachen war der Einfluss zusätzlicher Faktoren, wie etwa ihre Position im Wort oder ihr tauto- oder heterosilbischer Charakter von Bedeutung (siehe Gafos et al. 2010, Shaw et al. 2009, 2011 für Marokkanisches Arabisch, Shaw et al. 2009, 2011, Chitoran et al. 2002 für das Georgische).

Für eine Aufklärung der auftretenden Variation in Plosiv-Token eignen sich die analysierten portugiesischen Varietäten ausgesprochen gut, da diese Cluster unterlagen im Portugiesischen diachron zahlreichen Lautwandelprozessen und unterliegen in der synchronen gesprochenen Sprache⁴⁰ einer dreifachen Variation:

- (i) Plosivcluster tendieren im BP überwiegend dazu, mit einem intermediären Vokal realisiert zu werden, so dass sich beispielsweise folgende Wörter – abgesehen vom Anlaut – nicht unterscheiden: [kapi'tar] *captar* 'fangen' und [tapi'tar] *tapetar* 'Teppich legen'. Im EP trifft diese Tendenz eher bei älteren Sprechern und Kindern im Erstspracherwerb zu: ([kɛpi'tar] *captar* 'fangen' und [tɛpi'tar] *tapetar* 'Teppich legen' vgl. Bisol 1999, Mateus & Andrade 1998, 2000, Veloso 2006, 2007b, Vigário 2003).
- (ii) Alternativ kann der erste Plosiv, insbesondere im BP, getilgt werden: [fatu/faktu] *facto* 'Fakt' (vgl. Mateus & Andrade 1998, 2000, Vigário, 2003).
- (iii) Da beide Phänomene optional sind, können ebenso Cluster mit zwei Plosiven vorkommen.

So werden Plosivcluster im Portugiesischen entweder mit einem intermediären Vokal, als Clusterfolge bestehend aus zwei Konsonanten oder als einzelner Konsonant realisiert. Im BP ist die häufige Realisierung mit intermediärem Vokal einer CVC-Sequenz ähnlich. Zusätzlich werden im EP auch zwischen Plosiven lexikalisch unbetonten Vokale (/i,u/) getilgt und infolgedessen Plosiv-CVC-Sequenzen als reine Konsonanten-

⁴⁰ Die Variation beider Plosive in einem Cluster ist zwar in der gesprochenen Sprache besonders auffällig, allerdings ist sie längst nicht auf das Mündliche beschränkt. Ein gutes Beispiel für die Auswirkung dieser Variation von der gesprochenen auf die geschriebene Sprache ist die seit 1990 diskutierte und ratifizierte Rechtschreibreform, die vor allem die Schreibweise von Plosivclustern neu regelt und äußerst umstritten ist (z.B. Emiliano 1990a, b, 2008, Moura 2008).

abfolgen (CC) realisiert. In diesem Sinne unterscheiden sich im EP sowohl [tɐp'tar] *tapetar* 'Teppich legen' als auch [kɐp'tar] *capotar* 'umkippen' mit Ausnahme des Anlauts nicht von [kɐp'tar] *captar* 'fangen' (Mateus & Martins 1982). Das hat zur Folge, dass im EP mehrere Wortpaare beinahe homophon werden. Angesichts dieser Verringerung der lexikalischen Unterschiede in beiden Varietäten und ausgehend von der vielfachen Variation zwischen Clustern und CVC-Sequenzen wird auch in diesem Kapitel mithilfe von physiologischen Daten untersucht, inwiefern lexikalische Unterschiede in der Konsonantenkoordination verankert sind und inwiefern sich die beschriebene Variation durch unterschiedliche Koordination der Plosive im Sinne der Artikulatorischen Phonologie (Browman & Golstein, 1986, 1988, 1990, 2000) erklären lässt.

Vor der Präsentation und Diskussion der Ergebnisse wird zunächst auf einige Besonderheiten der Übertragung von Plosiven hingewiesen (5.1.1) und es werden die genauen Fragestellungen (5.1.2) sowie die zugehörigen Hypothesen (5.1.3) erläutert.

5.1.1 Die Übertragung von Plosiven

Die Übertragung von Plosiven in reinen Konsonantenclustern einerseits und in CVC-Sequenzen andererseits wird bereits in der Theorie unterschiedlich modelliert: Der Grund dafür ist, dass die akustischen Korrelate des Vokals den Konsonanten nicht behindern. Der Vokal wird ohne eine starke Konstriktion produziert (Öhman 1967), so dass die Bewegung für den Plosiv wie bei jedem Konsonant gleichzeitig mit der Bewegung für den Vokal beginnen kann, ohne dass dabei die Wahrnehmung beider Segmente beeinträchtigt wird. Daher kann bei der Produktion eines Konsonanten und eines Vokals in Folge die parallele Übertragung maximal sein.

Dies ist allerdings bei Konsonantenclustern nicht der Fall, da Konsonanten andere Konsonanten unterschiedlich stark maskieren können. Bei Plosiven ist die Maskierung am stärksten (Mattingly 1981), so dass ein Plosiv in der C₂-Position eines

Clusters (/C₁C₂/) den ersten Konsonanten komplett verdecken kann. Dies trifft unabhängig vom Konsonanten in der ersten Position zu und kann jede Artikulationsart der Konsonanten betreffen. Infolgedessen kann ein Plosiv in der C₂-Position auch einen anderen Plosiv komplett verdecken, so dass der erste Plosiv nicht mehr wahrgenommen werden kann. Dies hat in der Konsonantenkoordination zur Folge, dass Plosive in der C₂-Position weiter von C₁ entfernt koordiniert werden müssen als Liquide oder Nasale in dieser Position, damit beide Segmente in der Perzeption wiederherstellbar bleiben (Gafos et al. 2010, Chitoran et al. 2002). Der Grund liegt darin, dass ein Liquid oder ein Nasal als C₂ den C₁-Plosiv nicht so stark beeinträchtigen kann, wie es bei einem Plosiv als C₂ der Fall ist (Kühnert et al. 2006). Mit anderen Worten, ein Plosiv als C₂ kann den C₁-Plosiv so massiv verdecken, dass kein C₁-Verschluss wahrnehmbar oder im Spektrogramm sichtbar ist und infolgedessen nur C₂ in der Perzeption wiederherstellbar ist. Ein Liquid oder ein Nasal als C₂ können den C₁-Plosiv nicht so stark beeinträchtigen. Demnach können Liquide und Nasale parallel übertragen werden, allerdings ist dies nicht so stark ausgeprägt wie bei Vokalen. Bei Plosiven ist hingegen die parallele Übertragbarkeit minimal (vgl. Mattingly 1981), d.h., die Laute werden selbst in einer Plosivgruppe nacheinander übertragen, damit ihre Wahrnehmung gewährleistet werden kann.

Für die Konsonantenkoordination im Rahmen der Artikulatorischen Phonologie (Browman & Goldstein 1986, 1988, ff.) hat dies zur Folge, dass sich die Plosivgesten in einem Cluster nicht stark überlappen können und ihre Koordinierung der Konsonanten breiter ist als bei anderen bzw. gemischten Konsonantenclustern. Somit wird gewährleistet, dass beide akustischen Verschlusslösungen realisiert und wahrgenommen werden. Außerdem werden die Produktion und die Perzeption von Plosivclustern von weiteren Faktoren beeinflusst: Chitoran et al. (2002) zeigten, wie bereits mehrfach erwähnt, einen zusätzlichen Effekt der Produktionsrichtung bei der Wahrnehmung von Plosivclustern. Georgische Plosivcluster, die

aus einem vorderen C_1 gefolgt von einem hinteren C_2 (*front-to-back* Cluster) bestanden, zeigten mehr Überlappung als Cluster in der umgekehrten Produktionsrichtung (*back-to-front*). Dies ist zu erwarten, wenn die *front-to-back* Produktionsrichtung in der Perzeption besser wiederherstellbar ist als die *back-to-front* Richtung.

Des Weiteren ging aus einem Vergleich von Plosivclustern sowohl in verschiedenen Wortpositionen als auch in den möglichen tauto- und heterosilbischen Kontexten im Marokkanischen Arabisch hervor, dass mehr Überlappung bei wortmedialen Plosiven als in den anderen Silbenpositionen sowie in hetero- als in tautosilbischen Clustern auftritt (Gafos et al. 2010:686).

5.1.2 Fragestellung und theoretischer Rahmen

Wie aus den oberen Abschnitten hervorgeht, muss die Koordination von Plosiven den Besonderheiten und Einschränkungen ihrer Produktion Rechnung tragen, so dass selbst bei aufeinanderfolgenden Plosiven keine große Überlappung der Konsonantengesten erwartet wird, damit die Wiederherstellbarkeit beider Elemente in der Perzeption gewährleistet werden kann.

In den analysierten Sprechervarietäten wurden allerdings bislang weniger Unterschiede zwischen Clustern und CVC-Sequenzen festgestellt, sondern vielmehr varietätsspezifische Besonderheiten: Im BP werden lexikalische Cluster und CVC-Sequenzen mit einem intermediären Vokal beschrieben, während im EP – mit Ausnahme vereinzelter älterer Sprecher (Veloso 2007a) – beide aufgrund der häufigen Tilgung der lexikalisch unbetonten Vokalen /i,u/ als Konsonantengruppe auftreten. Beide lexikalischen Formen werden im BP mit einem intermediären Vokal produziert, während sie im EP beide als Konsonantengruppe auftreten. Dies hat zur Folge, dass trotz der angenommenen unterschiedlichen Koordinierung von Clustern und CVC-Sequenzen die Unterschiede zwischen lexikalischen Kategorien innerhalb einer

Varietät kleiner ausfallen als die Variation zwischen den beiden Varietäten.

Diese unterschiedlichen Realisierungen wurden bislang mit phonologischen Prozessen wie Vokalepenthese, Konsonanten- und Vokaltilgung beschrieben. Die Vokalepenthese und die Konsonantentilgung wurden mittels Extrasilbizität begründet: Da Plosive aufgrund ihrer fehlenden Sonorität nicht in ein Cluster integriert werden können und ebenso wenig zum Auslaut der letzten Silbe gehören, weil aufgrund von phonotatischen Beschränkungen Plosive keine Auslaute im Portugiesischen sind, bleibt C_1 in einem C_1C_2 -Cluster extrasilbisch. Diese Segmente werden dann entweder getilgt oder durch Vokalepenthese zum Anlaut einer neuen Silbe resilbifiziert (Mateus & d'Andrade 1998, 2000, Bisol 1999). Daraus entstehen optimierte Strukturen, die der allgemeinen Tendenz zur Bildung einer CV-Silbe folgen. Bei der Vokaltilgung bleiben die Silben erhalten, da der Vokal jederzeit hinzugefügt werden kann und somit in der Repräsentation des Wortes fortbesteht (vgl. Mateus & d'Andrade 1998, 2000).

Man kann allerdings – wie im vorherigen Kapitel geschehen – diese Produktionsunterschiede mit der zeitlichen Koordination der artikulatorischen Gesten analog zu Browman & Goldstein (1986, 1988, 1990, 2000) im Rahmen der artikulatorischen Phonologie in Verbindung bringen. Es stellt sich daher die Frage, ob sich die Koordination lexikalischer Plosivcluster im Vergleich zu $\langle C_1et \rangle$ und $\langle C_1ut \rangle$ innerhalb und/oder zwischen den Varietäten überhaupt unterscheidet oder ob sich ihre Variabilität in der gesprochenen Sprache durch eine unterschiedliche Koordinierung der Plosive erklären lässt. Inwiefern dabei die Koordination von Plosiven Variabilität zulässt, die zu Variationen und Instabilitäten aus synchroner und diachroner Sicht führt, gehört zu den zentralen Fragestellungen, mit denen sich das aktuelle Kapitel beschäftigt wird.

5.1.3 Hypothesen und Vorhersagen

Aufgrund der starken Beeinträchtigung anderer Konsonanten durch die Plosive in der Produktion und mit dem Ziel, die Wahrnehmung beider Plosive zu gewährleisten beginnen Sprecher den zweiten Plosiv so spät, dass die Plosive nacheinander produziert werden und sich somit nicht mehr überlappen. Sollten sich allerdings zwei Plosive stark überlappen, kann C_1 in der Perzeption nicht mehr wahrgenommen werden. Somit gewinnt die perzeptive Wiederherstellbarkeit der Segmente bei Plosiven eine ausschlaggebende Rolle für die Koordination der Plosivcluster.

Dieser Effekt wird allerdings von anderen Faktoren eingeschränkt. Für das Marokkanische Arabisch zeigten Gafos et al. (2010), dass eine stärkere Überlappung von Plosiven in heterosilbischen als in tautosilbischen Clustern und auch in medialer im Vergleich zu initialer Position auftritt. Angenommen, das Portugiesische weist die gleichen Tendenzen auf, so wird ebenfalls bei den portugiesischen Plosivclustern mehr Überlappung beider Konsonantengesten erwartet, da sie sowohl heterosilbisch sind als auch in medialer Position vorkommen. Davon ausgehend lauten unsere übergreifenden Vorhersagen wie folgt:

- Die Konsonantenkoordinierung berücksichtigt die mögliche Wiederherstellbarkeit in der Perzeption und daher werden Plosive weit koordiniert.
- Dort allerdings, wo die perzeptive Wiederherstellbarkeit ohnehin geschwächt ist, also in medialen, heterosilbischen Plosivclustern, unterbleibt die „Rettungsmaßnahme“ der weiten Koordination, so dass in diesen Fällen Überlappung oder sogar starke Überlappung beider Plosivgesten festgestellt werden kann.

So wird im Folgenden untersucht, inwiefern sich die Koordination von medialen, heterosilbischen, lexikalischen Plosivclustern vs.

<C₁et> vs. <C₁ut> (C= /p,k/) in den portugiesischen Varietäten unterscheidet und inwiefern die große Variabilität auf die unterschiedliche Koordination der einzelnen Konsonanten zueinander zurückzuführen ist. Dabei wird die Rolle des Artikulationsorts des ersten Konsonanten (labial vs. velar) und der Betonung (betont vs. unbetont) miteinbezogen. Gegen Ende des Kapitels wird diskutiert, inwiefern die Koordination von Plosiven an sich Variabilität zulässt, die sowohl zu Variationen und Instabilitäten aus synchroner, aber auch aus diachroner Sicht führen kann. Dabei werden wie im letzten Kapitel die zeitliche intergestische (*Timing*) und die intragestische Koordinierung analysiert. Dafür werden folgende Messverfahren verwendet:

- (i) Die zeitliche intergestische Koordinierung (*Timing*). Hierbei werden die zeitliche Organisation beider Konstriktionsplateaus sowie die Synchronisierung von C₂ in Bezug auf C₁ berücksichtigt. Bei Ersterem geht es um den Abstand zwischen beiden Konstriktionsplateaus und bei Letzterem um die Überlappung von /t/ und C₁.
- (ii) Die räumliche intragestische Dynamik der Zungenspitzenbewegung wird hinsichtlich Umfang (*magnitude*) und maximaler Geschwindigkeit (*peak velocity*) untersucht. Dabei wird geprüft, inwiefern Cluster vs. <C₁et> vs. <C₁ut> signifikante Unterschiede im Umfang und in der Geschwindigkeit der Zungenspitzenbewegung hervorrufen.

Im Folgenden sollen die genauen Hypothesen formuliert werden.

5.1.3.1 Sprachvarietäten: BP und EP

Die diachrone Variabilität wurde in der historischen Entwicklung dieser Formen hin zum Portugiesischen ausführlich beschrieben, bei der diese Cluster mehreren Lautwandelprozessen unterlagen: Der erste Konsonant wurde entweder regressiv assimiliert und vereinfacht (KL *aptāre* > attar > Port. *atar* 'binden, schnüren')

oder vokalisiert (KL *conceptum* > Port. *conceito* ‘Begriff’, Jensen 1999, William 1961). Die synchrone Variabilität der Plosivcluster, bei der lexikalische Plosivcluster /pt/ und CVC-Sequenzen /pit, put/ im EP als [pt] realisiert werden, während im BP bei Clustern und CVC-Sequenzen ein intermediärer Vokal [pit] produziert wird (Bisol 1999, Mateus & d’Andrade 2000, Mateus et al. 2005), wurde bereits mehrmals angesprochen. Alternativ dazu kann insbesondere im BP der erste Plosiv (C₁) in reinen Plosivclustern gefilgt werden (z.B. [fatu] für *facto* ‘Fakt’, vgl. Mateus & Andrade 1998, 2000, Vigário, 2003), so dass unter synchroner Betrachtung Plosivcluster entweder mit einem intermediären Vokal oder als reine Cluster oder als einzelne Konsonanten (C₂) beschrieben werden.

Angesichts der synchronen Beschreibungen wird eher eine weite Koordinierung beider Plosivgesten im BP erwartet, die bei Hörern manchmal dazu führt, dass sie die weite Koordinierung durch ein intermediäres vokalisches Element wiedergeben, ohne dass ein Vokal in der Repräsentation des Wortes vorhanden wäre. Im Falle einer umgekehrt entsprechend engen Koordinierung der CVC-Sequenzen im EP könnten die Hörer teilweise den Vokal und im Extremfall sogar den ersten Plosiv nicht mehr wahrnehmen.

5.1.3.2 Stimulustyp: Cluster vs. <C₁et> vs. <C₁ut>

Im Hinblick auf die bereits beschriebene unterschiedliche Koordination von Clustern und CVC-Sequenzen wird eine weitere Koordination eher bei CVC-Sequenzen als bei reinen Clustern erwartet.

5.1.3.3 Artikulationsort: bilabialer vs. velarer C₁

Da sich der sogenannten Produktionsrichtungseffekt (Chitoran et al. 2002) auf Plosiv-Token besonders stark auswirken könnte, wird in dieser Studie mehr Überlappung bei den *front-to-back* als bei den *back-to-front* Token erwartet.

5.1.3.4 Betonung: betonte vs. unbetonte Kondition

Auch in diesem Kapitel werden Plosivcluster in betonter und unbetonter Silbe miteinander verglichen. Da die untersuchte Vokaltilgung ausschließlich in unbetonter Position auftritt, wurde die Betonungskondition für CVC-Sequenzen mithilfe von C₂ definiert. Dabei werden also Sequenzen mit C₂ in betonter (z.B. *computar*, ‘rechnen’) und in unbetonter Silbe (z.B. *computador*, ‘Rechner’) verglichen. Der Zielvokal musste dabei, wie gesagt, in beiden Konditionen in einer unbetonten Silbe vorkommen, um Tilgung zu ermöglichen.

Segmente in unbetonter Position zeigen mehr Koartikulation als in betonter Position (Beckman et al. 1992, Byrd 1993, Iskarous & Kavitskaya 2010, Tabain 2000), was mit der steigenden Überlappung der Konsonantengesten in diesem Kontext erklärt wurde (z.B. Browman & Golstein 1991a). Infolgedessen wird ebenso mehr Überlappung bei den unbetonten als bei den betonten Clustern vorhergesagt. Für die CVC-Sequenzen wird der gleiche Effekt erwartet, allerdings etwas schwächer, da in der betonten Kondition lediglich C₂ und nicht das gesamten Zielcluster betont ist.

In Bezug auf die Betonung werden außerdem signifikante Unterschiede in der intragestischen Koordinierung vorhergesagt. Dies liegt darin begründet, dass die längere akustische Dauer von betonten im Vergleich zu unbetonten Silben (z.B. Browman & Goldstein 1990a: 103) artikulatorisch sowohl von der zeitlichen intergestischen Koordination als auch von der räumlichen Steifheit der Gesten (*stiffness*, Browman & Goldstein 1990a:103, Kelso et al. 1985, 1988) beeinflusst wird. Beckman et al. (1990) zeigten, dass der Unterkiefer in betonten Silben einen größeren Gestenumfang, aber keine signifikant geringeren maximalen Geschwindigkeiten (*peak velocities*) aufwies. So muss man annehmen, dass betonte Silben einen vergleichsweise größeren räumlichen intergestischen Abstand zwischen den Konsonanten in Anspruch nehmen als unbetonte Silben. Daher wird davon ausgegangen, dass Cluster und CVC-Sequenzen in unbetonter

Position einen geringeren zeitlichen und räumlichen Gestumfang aufweisen als die in der betonten Umgebung.

Zunächst folgt eine knappe Wiederholung der Methode unter Angabe der relevanten Plosiv-Token (5.2) und im Anschluss daran die Analysen der intergestischen Koordination (5.3) und der räumlichen Organisation der Zungenspitzen-geste (5.4) zu den Plosiv-Token. Abschließend werden die wichtigen Schlussfolgerungen zusammengefasst und die entstandenen Vorhersagen für das folgende Kapitel präsentiert (5.5).

5.2 Methode

Die akustischen und artikulatorischen Daten zur Beantwortung unserer Fragestellungen wurden im Rahmen des in Kapitel 4 beschriebenen Experiments aufgenommen. Die Durchführung und die Methode waren daher die gleichen wie in Kapitel 4, ihre ausführliche Beschreibung kann unter 4.2. nachgelesen werden.

5.2.1 Sprachmaterial

Zur Erstellung unserer Stimuli wurden die zu analysierenden Cluster und CVC-Sequenzen in Zielwörter integriert, die in Tabelle 7 aufgelistet werden. Sie wurden in folgenden Trägersatz „Ele leu ___ bem“ („Er las ___ richtig“) eingebettet, randomisiert und von einem Computerbildschirm abgelesen.

Die fünf EP-Sprecher produzierten jeweils 8 Wiederholungen von jedem Satz und die vier BP-Sprecher jeweils 9 Wiederholungen. Die Präsentation der Stimuli wurde manuell gesteuert, so dass inkorrekte Äußerungen (z.B. Versprecher) wiederholt und korrigiert werden konnten. Die aufgenommene Anzahl von Äußerungen (Token) betrug für das EP insgesamt 480 Stimuli (zwölf Wörter x fünf Sprecher x acht Wiederholungen) und für das BP insgesamt 432 Stimuli (zwölf Wörter x vier Sprecher x neun Wiederholungen).

Cluster	<i>compact<u>a</u>do</i>	‘kompakt’
	<i>compact<u>a</u>do<u>r</u></i>	‘Verdichter’
	<i>cap<u>t</u>ado</i>	‘empfangen’
	<i>cap<u>t</u>ado<u>r</u></i>	‘Empfänger’
C₁eC₂	<i>que <u>t</u>ado⁴¹</i>	‘wie doof’
	<i>que <u>t</u>ad<u>a</u>do</i>	‘wie doof’
	<i>tap<u>e</u>tado</i>	‘mit Teppich’
	<i>tap<u>e</u>tado<u>r</u></i>	‘Teppichverleger’
C₁uC₂	<i>cut<u>a</u>no</i>	‘Eigennamen’
	<i>cut<u>a</u>no<u>s</u>o</i>	‘E., Genitiv’
	<i>comp<u>u</u>tado</i>	‘ausgerechnet’
	<i>comp<u>u</u>tado<u>r</u></i>	‘Rechner’

Tabelle 7: Liste der aufgenommenen Plosiv-Token mit den Zielclustern und Sequenzen (hier in kursiv dargestellt) und den betonten Silben (als unterstrichene Silben markiert).

5.2.2 Statistische Auswertung

Die statistische Auswertung wurde parallel zur Analyse im vorherigen Kapitel mit einer Varianzanalyse mit Messwiederholung (*RM-Anova*, *repeated measures Anova*) und einem linearen *Mixed Model* in R (R Development Core Team 2011) durchgeführt. Die unabhängigen Variablen waren Sprechervarietät (BP vs. EP), Stimulustyp (dreistufig: Cluster vs. <C₁et> vs. <C₁ut>), C₁ (bilabial vs. velar), C₂ (Lateral vs. /r/) und Betonung (betonte vs. unbetonte Kondition). Die abhängigen Variablen werden im Zuge der weiter unten definierten Messverfahren erläutert. Da die Ergebnisse der Analyse mit dem *Mixed Model* insgesamt robuster waren, werden ausschließlich diese präsentiert. Eine detaillierte Beschreibung der beiden Verfahren und der angewendeten Analyse kann 4.2.5 entnommen werden.

⁴¹ Da das Sprachsystem hier eine Lücke aufweist, wurde die Tabelle mit lexikalisierten festen Ausdrücken aufgefüllt.

5.3 Die zeitliche intergestische Koordinierung

Kurz zur Erinnerung: Das erste Verfahren analysiert die zeitliche Organisation beider Konstriktionsplateaus und der Synchronisierung des Zeitpunkts des Beginns von C₂ in Bezug auf das C₁-Plateau. Mit den beiden Verfahren wird die Koordinierung sowohl anhand des intermediären Abstands beider Plateaus als auch der Überlappung der C₂- mit der C₁-Geste festgestellt. Der Vergleich beider Verfahren bringt robustere Ergebnisse hervor.

5.3.1 Definition der Variablen

Wie in Kapitel 4 wird im Folgenden die Dauer der einzelnen Plateaus sowie des intermediären Abstands berechnet. Hier wird lediglich die Grundlage des Messverfahrens zur Erinnerung wiederholt, für weitere Details siehe Punkt 4.2. im vorherigen Kapitel.

- $C_1\text{-PLATEAU} = (C_1\text{-LÖSUNG} - C_1\text{-TARGET}) / (C_1\text{-ENDE} - C_1\text{-START})$. Dies entspricht der Differenz aus C₁-Lösung und C₁-Target dividiert durch die gesamte Gestendauer.
- $C_2\text{-PLATEAU} = (C_2\text{-LÖSUNG} - C_2\text{-TARGET}) / (C_2\text{-ENDE} - C_2\text{-START})$. Dies entspricht der Differenz aus C₂-Lösung und C₂-Target dividiert durch die gesamte Gestendauer.
- $\text{PLAT-DIST} = (C_2\text{TARGET} - C_1\text{LÖSUNG}) / (C_2\text{-ENDE} - C_1\text{-ANFANG})$
- $\text{ÜBERLAPPUNG} = 1 - (C_2\text{-START} - C_1\text{-TARGET}) / (C_1\text{-PLATEAU})$ Dies entspricht der Differenz aus C₂-Start und C₁-Target, dividiert durch C₁-Plateau, anschließend wird der Wert von 1 subtrahiert.

5.3.2 Ergebnisse

In den folgenden Abbildungen wurde die Dauer des intermediären Abstands abgebildet. In der oberen Hälfte befinden sich in diesem Kapitel die betonten und in der unteren die unbetonten Token.

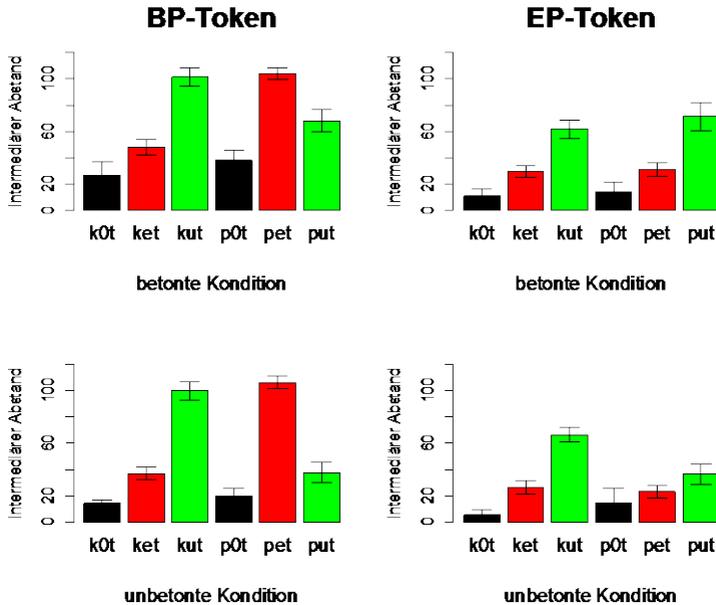


Abbildung 10: Dauer des intermediären Abstands (Mittelwerte über vier BP und fünf EP Sprecher mit Standardabweichungen⁴²).

Die Abstände zwischen den Plateaus waren ausnahmslos kürzer im EP ($\chi^2[1] = 13.1$, $p < 0.01$) als im BP und somit die hier analysierten Plosive bei allen EP-Token enger koordiniert als bei den BP-Token. Der Stimulustyp hatte den Haupteffekt

⁴² Die Konsonantencluster wurden vereinheitlicht und parallel zu den anderen Stimulustypen mit drei Elementen dargestellt. Dafür wurden sie mit einer Null zwischen den Konsonanten verzeichnet, die für einen leeren Vokal(-platz) stehen soll. Die Dauer der Plateaus wurde nicht abgebildet.

($\chi^2[2] = 237.4, p < 0.001$), wurde allerdings durch die signifikanten Interaktionen zwischen Stimulustyp und Sprechervarietät ($\chi^2[2] = 35.5, p < 0.001$) sowie zwischen Stimulustyp und Artikulationsort ($\chi^2[2] = 183.6, p < 0.001$) eingeschränkt. Aus den Tukey Post-Hoc-Tests ging hervor, dass der Artikulationsort einen signifikanten Einfluss auf die Plateaudauer der beiden CVC-Sequenzen ausübte ($p < 0.001$ für $\langle C_{1et} \rangle$ und $\langle C_{1eC_2} \rangle$), aber nicht der Cluster ($p = 0.8$). Die fehlende Interaktion zwischen Stimulustyp und Artikulationsort im EP ist daran ersichtlich, dass die Stimulustypen in beiden Artikulationsorten die gleiche Tendenz aufwiesen, und zwar war die Dauer des intermediären Plateaus sowohl bei velaren als auch bei bilabialen Token am kürzesten bei Clustern, gefolgt von $\langle C_{1et} \rangle$ und schließlich $\langle C_{1ut} \rangle$ mit dem längsten Plateau.

Im BP waren Cluster in beiden Artikulationsorten eng koordiniert, da sie sowohl nach velarem als auch nach bilabialem C_1 einen kürzeren intermediären Abstand aufwiesen. Bei den CVC-Sequenzen kam die Interaktion zum Ausdruck, indem $\langle C_{1et} \rangle$ vergleichsweise kürzere intermediäre Abstände nach /k/ und längere nach /p/ zeigte. Demnach war die Koordination von $\langle C_{1et} \rangle$ enger nach /k/, aber weiter nach /p/ als die von $\langle C_{1ut} \rangle$ und die Koordination von $\langle C_{1ut} \rangle$ verhielt sich genau umgekehrt: weiter nach /k/ und enger nach /p/. Die Betonung hatte einen größeren Einfluss auf die Abstandsdauer im BP als im EP, aber die Tendenz zu kürzeren Abständen bei unbetonten als bei betonten Token war in beiden Varietäten signifikant ($\chi^2[1] = 13.2, p < 0.001$ für BP und $\chi^2[1] = 9.3, p < 0.01$ für EP).

Die Unterschiede in der Dauer des C_1 -Plateaus lassen sich am besten durch den signifikanten Einfluss des Stimulustyps ($\chi^2[2] = 47.0, p < 0.001$) und der Sprechervarietät ($\chi^2[1] = 5.9, p < 0.05$) erklären. Die C_1 -Plateaus waren im BP länger als im EP und in beiden CVC-Sequenzen ($\langle C_{1et} \rangle$ $\langle C_{1ut} \rangle$) ergaben sich deutlich längere Plateaus als in Clustern. Bei den europäischen Token wurde eine signifikante Interaktion zwischen Stimulustyp und Artikulationsort festgestellt, die an längeren Plateaus bei $\langle C_{1et} \rangle$

nach velaren als nach bilabialen Plosiven lagen. Die weiteren Faktoren oder Interaktionen hatten keinen signifikanten Einfluss auf die Plateaudauer.

Der Stimulustyp ($\chi^2[2] = 34.9, p < 0.001$) hatte einen signifikanten Einfluss auf die Dauer des C_2 -Plateaus und die Sprechervarietät zeigte an sich keinen Einfluss, allerdings wurde eine Interaktion zwischen beiden Faktoren festgestellt. Diesbezüglich zeigten Post-Hoc-Tests eine längere Dauer von $\langle C_{1et} \rangle$ im Vergleich zu Clustern ($p < 0.001$) und zu $\langle C_{1ut} \rangle$ ($p < 0.001$) im BP und keine signifikanten Unterschiede für EP. Die Dauer des C_2 -Plateaus wurde außerdem in beiden Varietäten von der Betonung ($\chi^2[1] = 168.5, p < 0.001$ im BP und $\chi^2[1] = 37.8, p < 0.001$ im EP) erheblich beeinflusst, wobei die Plateaudauer vor allem in den BP-Daten bei betonten deutlich länger als bei unbetonten Token ist.

Der Effekt des C_1 -Artikulationsorts ist ebenso ausgeprägter in den BP- als in den EP-Daten ($\chi^2[1] = 44.0, p < 0.001$ im BP und $\chi^2[1] = 13.0, p < 0.001$ im EP), weil velare Konsonanten in beiden Varietäten signifikant längere Plateaus zeigten als die Bilabiale. Im BP gab es außerdem eine signifikante Interaktion zwischen Stimulustyp und Artikulationsort ($\chi^2[1] = 36.1, p < 0.001$), die dadurch zustande kam, dass die Dauer des C_2 -Plateaus deutlich länger nach dem velaren Plosiv in $\langle C_{1et} \rangle$ als in Clustern ($p < 0.001$) war und in Clustern deutlich länger im Vergleich zu $\langle C_{1ut} \rangle$ ($p < 0.001$). Der gleiche Effekt von längeren C_2 -Plateaus nach dem velaren C_1 trat ebenfalls bei beiden CVC-Sequenzen auf ($p < 0.001$ in beiden Fällen). Bei den bilabialen C_1 wurden keine Dauerunterschiede bzw. keine Effekte festgestellt. Im EP interagierte der Artikulationsort sowohl mit dem Stimulustyp ($\chi^2[2] = 10.9, p < 0.01$) als auch mit der Betonung ($\chi^2[1] = 4.1, p < 0.05$). Der Grund für diese Interaktionen lag darin, dass C_2 -Plateaus länger waren bei $\langle C_{1et} \rangle$ im Vergleich zu $\langle C_{1ut} \rangle$ nach velaren ($p < 0.05$) aber nicht nach labialen C_1 . Cluster ($p < 0.01$) und $\langle C_{1ut} \rangle$ ($p < 0.01$) waren länger in betonter als in unbetonter Position. Dies traf allerdings nicht auf $\langle C_{1et} \rangle$ zu.

Abbildung 11 veranschaulicht die Synchronisierungsmittelwerte von fünf EP-Sprechern (links) und vier BP-Sprechern (rechts). In der oberen Hälfte befinden sich die betonten, in der unteren die unbetonten Token.

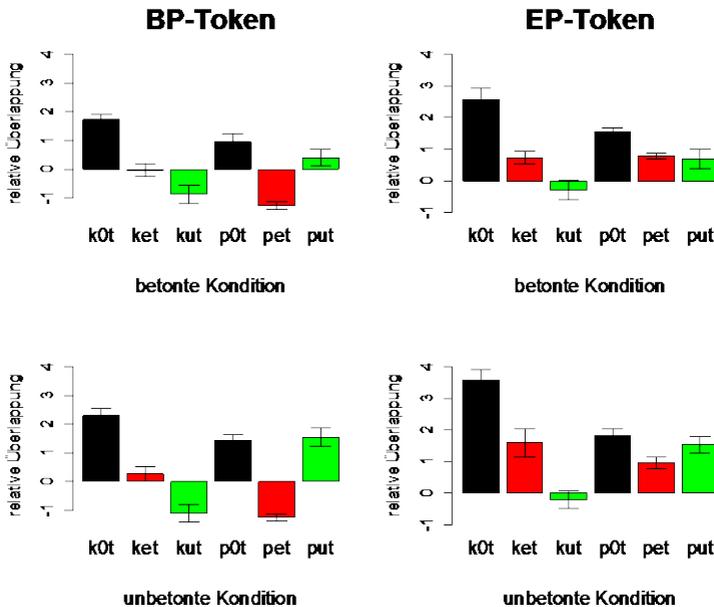


Abbildung 11: Übersicht der Synchronisierung der Gesten nach Liquid-Gruppen und Varietäten getrennt (Mittelwerte über fünf EP- und vier BP-Sprecher mit Standardabweichungen)

Betrachtet man die Gesamtergebnisse, können Sprechervarietät ($\chi^2[1] = 13.1$, $p < 0.01$) und Stimulustyp ($\chi^2[2] = 226.1$, $p < 0.001$) sowie die Interaktion ($\chi^2[2] = 20.9$, $p < 0.01$) zwischen beiden die meiste Variabilität bei der Synchronisierung beider Plosive erklären. In der Koordination der Plosive wurden die gleichen Tendenzen in beiden Varietäten ersichtlich, allerdings fand ausnahmslos mehr Überlappung der Gesten im EP als im BP statt. Cluster zeigten in beiden Varietäten positive Werte > 1 , und

somit eine starke Überlappung und eine enge Koordinierung der Konsonanten.

Die Koordination der CVC-Sequenzen wurde von der Interaktion zwischen Artikulationsort und Stimulustyp stark geprägt ($\chi^2[2] = 119.9$, $p < 0.001$ für BP und $\chi^2[2] = 62.7$, $p < 0.001$ bei EP), indem die Plosive in $\langle C_{1et} \rangle$ mehr Überlappung nach dem velaren als nach dem bilabialen C_1 zeigten und sich in $\langle C_{1ut} \rangle$ die Gesten nach /p/ überlappten, aber nicht nach /k/. Infolgedessen war die Koordination von $\langle C_{1et} \rangle$ eng nach velarem und weit nach bilabialem Plosiv, während die Koordination von $\langle C_{1ut} \rangle$ genau die entgegengesetzte Tendenz zeigte, und zwar waren die Plosive in $\langle C_{1ut} \rangle$ eng koordiniert nach dem bilabialen und weit auseinander nach dem velaren Plosiv. Die Überlappung war ausnahmslos stärker bei unbetonten als bei betonten Token und dementsprechend waren unbetonte Token enger koordiniert als betonte.

An dieser Stelle waren die Ergebnisse dieses Überlappungsverfahrens also ähnlich wie die des intermediären Abstands der Konsonanten: Die Synchronisierung beider Konsonanten fiel je nach Sprechervarietät unterschiedlich aus, indem Plosive enger im EP als im BP koordiniert wurden. Der Stimulustyp hatte ebenso einen hochsignifikanten Einfluss auf die Plosivkoordination, die von einer starken Interaktion mit dem Artikulationsort geprägt war: Die Koordination von $\langle C_{1et} \rangle$ war eng nach velaren und weit nach bilabialen Plosiven und $\langle C_{1ut} \rangle$ zeigte eine enge Koordination nach bilabialen und eine weite nach velaren Plosiven. Außerdem begann C_2 in unbetonten Token früher innerhalb von C_1 als in betonten in beiden Varietäten.

5.3.3 Diskussion der wichtigsten Ergebnisse

Bezüglich beider Konstriktionsplateaus übte bei den Plosiv-Token der Stimulustyp einen signifikanten Einfluss sowohl auf die Länge des C_1 - als auch des C_2 -Plateaus aus, und zwar waren beide Plateaus tendenziell länger bei $\langle C_{1et} \rangle$ als bei den anderen Token. Diese Tendenz war im BP für beide Plateaus signifikant, jedoch

nicht im EP. In Letzterem war $\langle C_{1et} \rangle$ ausschließlich nach velarem C_1 /k/ länger. Die Längung beider Plateaus wurde möglicherweise von der zentralen Lage der Zunge für die /e/-Produktion im BP begünstigt. Die Dauer des C_1 -Plateaus war sowohl bei /p/ als auch bei /k/ deutlich kürzer bei den reinen Plosiv-Token im Vergleich zu den im vorherigen Kapitel analysierten Plosiv-Liquid-Token. Dies könnte mit der medialen Position der Konsonanten(-cluster) in diesem Kapitel zusammenhängen, deren Verschluss aufgrund von möglichen Transitionen zu vorhergehenden Konsonanten weniger prominent sein kann als bei initialen Plosiven (Byrd 1993).

Das C_1 -Plateau wurde von der Sprechervarietät beeinflusst, indem es im BP länger als im EP war, allerdings spielte die Sprechervarietät keine Rolle für die Dauer des C_2 -Plateaus. Der Einfluss des Artikulationsorts führt vor allem nach velarem C_1 zu Längenunterschieden in den Plateaus der verschiedenen Stimulustypen, die nach dem bilabialen C_1 nicht zum Ausdruck kommen. Dies kann mit der starken Überlappung der Konsonanten in der *front-to-back* Produktionsrichtung (Chitoran et al. 2002) in Verbindung gebracht werden, wodurch die Unterschiede der verschiedenen Stimulustypen minimiert wurden, d.h. die Koordination von CVC-Sequenzen und Clustern war ähnlich. Die Betonung hatte den Effekt, dass die Plateaudauern geringer bei unbetonten als bei betonten Token ausfielen, die – wie vorhergesagt – mit der stärkeren Koartikulation der Segmente in unbetonter Position (Byrd 1996) zusammenhängen.

Der Haupteffekt bezüglich der Stimulustypen ist, dass bei beiden angewandten Verfahren die Koordinierung beider Plosive eindeutig enger bei Clustern als bei CVC-Sequenzen war. Dies war aufgrund der vielen Studien, die Überlappung und enge Koordinierung bei Konsonanten in einem Cluster feststellten (z.B. Bombien et al. 2013, Gafos et al. 2010, Hoole et al. 2009 u.v.a.) zu erwarten. Bei den CVC-Sequenzen fand insbesondere in den BP-Token eine starke Interaktion zwischen Stimulustyp und Artikulationsort statt: Die Plosive in $\langle C_{1et} \rangle$ waren nach velarem

C₁ eng koordiniert, während dies für <C₁ut> nach bilabialem C₁ der Fall war. Ein ähnliches Ergebnis war ebenso bei der Synchronisierung der unbetonten EP-Token vorhanden, indem die Überlappung der Plosive so stark bei <C₁ut> nach dem bilabialen Plosiv war, dass sich die Koordination von der CVC-Sequenz /put/ und dem Cluster /pt/ kaum unterschieden. Der intermediäre Abstand zwischen beiden Plosiven war allerdings deutlich größer bei /put/ als bei /pt/. Dies kommt dadurch zustande, dass die /t/-Geste in /put/ zwar früher innerhalb von C₁ beginnt, insgesamt allerdings einen langsamen flacheren Anstieg bis zum Erreichen des Plateaus aufweist. Dieser längere (und auch raumgreifendere) Verlauf spiegelt sich in den längeren intermediären Abständen wider.

Die zeitliche Koordination der Cluster wurde ebenfalls stark von C₁ und dem damit assoziierten Produktionsrichtungseffekt bestimmt, allerdings nicht in der erwarteten Richtung: Die Konsonantengesten überlappten sich unerwarteterweise eindeutig mehr bei *back-to-front* Clustern (/kt/) und zeigten zudem signifikant kleinere intermediäre Abstände zwischen den Plateaus als bei *front-to-back* /pt/. Es geht somit eindeutig hervor, dass die Produktionsrichtung einen erheblichen Effekt auf die intergestische Koordination der Plosive hat, indem /kt/ erheblich enger koordiniert wird als /pt/.

Die Sprechervarietät zeigte sich als signifikanter Haupteffekt bei beiden Verfahren, bei denen die Plosivgesten im EP ausnahmslos enger koordiniert waren als im BP. Die Übereinstimmung der Ergebnisse beider Verfahren bekräftigt die Stabilität dieses Effekts. Angesichts der beschriebenen Vokaltilgungen im EP (Martins 1975, Silva 1997, 1998, Cunha 2011a) sowie der zahlreichen Studien zu Rhythmus-Unterschieden in den beiden Varietäten (Abaurre & Sandalo 2007, Abaurre & Galves 1998, Barbosa 2006, 2002, Frota et al. 2002, Frota & Vigário 2001, 2000, Reich 2004, Vigário et al. 2003) waren die varietätsspezifischen Unterschiede nicht unerwartet.

Die Betonung hatte einen erheblichen Einfluss auf die intergestische Koordination, indem Token in unbetonter Kondition deutlich enger koordiniert wurden als betonte Token. Dies ist möglicherweise eine Folge der deutlich stärkeren Hypoartikulation der Segmente (Lindblom 1990) bzw. der damit assoziierten stärkeren Koartikulation der Segmente in nicht prominenten und vorhersagbaren Kontexten, wie es bei unbetonten Silben oft der Fall ist (z.B. Fowler 2005). Ob dies auch bei den untersuchten Clustern und CVC-Sequenzen im Portugiesischen zutrifft, kann mit dem im nächsten Punkt beschriebenen Verfahren anhand des Umfangs der Gesten überprüft werden.

5.4 Räumliche intragestische Koordination

Nach der Analyse der intergestischen Koordination zwischen beiden Plosiven verlagert sich der Schwerpunkt der Analyse auf die räumliche Organisation der Zungenspitze für die Produktion von /t/, um die Variabilität bedingt durch Stimulustyp, Sprechervarietät, C_1 und Betonung zu beschreiben.

5.4.1 Definition der Variablen

Für die Analyse der Dynamik der Zungenspitzenegeste wurden die maximale Geschwindigkeit und der räumliche Verlauf der vertikalen Zungenspitzenbewegung (T^*TY -Umfang & T^*TY -Geschwindigkeit) erfasst, die hier zur Erinnerung knapp wiederholt werden⁴³.

- T^*TY -Geschwindigkeit (mm/ms). Dies entspricht dem höchsten Wert der Geschwindigkeit der vertikalen Bewegung der Zungenspitze (T^*Y).

⁴³ Für nähere Informationen hierzu siehe 4.4.

- TTY-Umfang (mm) = TTY-Max - TTY-Min. Dies entspricht der Differenz zwischen der höchsten und der niedrigsten Position der vertikalen Bewegung der Zungenspitze (T-Y).

5.4.2 Ergebnisse

Abbildung 12 zeigt die Mittelwerte der maximalen Geschwindigkeit der Schließgeste der Zungenspitze. Links befinden sich die BP- und rechts die EP-Token, in der oberen Hälfte der Abbildung die Token in der betonten Kondition und in der unteren Hälfte die in der unbetonten Kondition.

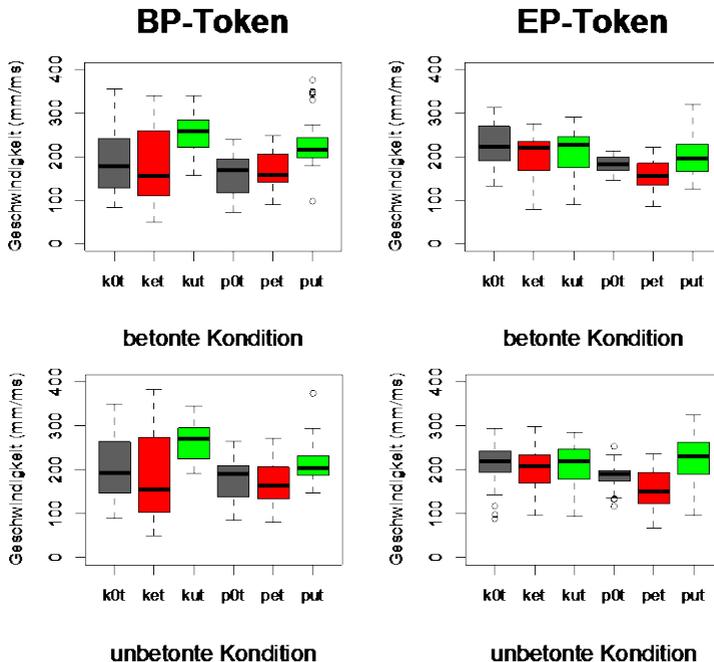


Abbildung 12: Maximale Geschwindigkeit des vertikalen Verlaufs der Schließungsgeste der Zungenspitze (TTY, Mittelwerte über fünf EP- und vier BP-Sprecher mit Standardabweichungen).

Der Stimulustyp hatte einen signifikanten Einfluss auf die maximale Geschwindigkeit der Schließgeste ($\chi^2[2] = 152.6$, $p < 0.001$), indem die Geschwindigkeit niedriger bei $\langle C_{1et} \rangle$ als bei den anderen Token war. Diese Tendenz findet sich in beiden Varietäten und ist der Grund, weshalb kein Effekt der Sprechervarietät auftrat. Allerdings war die Interaktion zwischen den beiden genannten Faktoren signifikant ($\chi^2[2] = 48.3$, $p < 0.001$). Tukey Post-Hoc-Tests ergaben signifikante Geschwindigkeitsunterschiede zwischen $\langle C_{1ut} \rangle$ und Clustern ($p < 0.001$) sowie zwischen $\langle C_{1ut} \rangle$ und $\langle C_{1et} \rangle$ ($p < 0.001$) im BP, indem $\langle C_{1ut} \rangle$ höhere Geschwindigkeiten zeigte als die anderen Stimulustypen. Die Geschwindigkeit unterschied sich allerdings nicht bei $\langle C_{1et} \rangle$ vs. Cluster. Im EP war sie ebenfalls höher bei $\langle C_{1ut} \rangle$ als bei Clustern ($p < 0.001$) und diese wiederum höher als bei $\langle C_{1et} \rangle$ ($p < 0.001$), jedoch gab es keine Geschwindigkeitsunterschiede zwischen $\langle C_{1ut} \rangle$ und $\langle C_{1et} \rangle$.

Für die Sprechervarietät getrennt durchgeführte *Mixed Model*-Analysen ergaben einen Haupteffekt des C_1 -Artikulationsorts in beiden Varietäten ($\chi^2[1] = 37.9$, $p < 0.001$ in den BP- und $\chi^2[1] = 32.4$, $p < 0.001$ in den EP-Daten) und eine signifikante Interaktion zwischen Artikulationsort und Stimulustyp im EP ($\chi^2[1] = 32.4$, $p < 0.001$), aber nicht im BP. Tukey Post-Hoc-Tests zeigten eine signifikant höhere Geschwindigkeit der apikalen Geste bei Clustern nach velaren als nach bilabialen C_1 -Plosiven im EP, was jedoch nicht auf die anderen Stimulustypen (d.h. auf beiden CVC-Sequenzen) zutraf. Im BP war die /t/-Geste schneller nach den velaren als nach den bilabialen C_1 . Der Einfluss der Betonung war größer im EP ($p < 0.01$) als im BP ($p < 0.05$), wobei allerdings in beiden Varietäten die maximale Geschwindigkeit höher in betonten als in unbetonten Token war.

Der räumliche Umfang zwischen der höchsten und der niedrigsten Position der vertikalen Zungenspitzenbewegung wurde in für die BP-Daten links und für die EP-Daten rechts dargestellt.

Der Stimulustyp übte einen hochsignifikanten Einfluss auf den räumlichen Umfang der Zungenspitzenbewegung aus ($\chi^2[2] = 374.5$, $p < 0.001$), allerdings zeigten nach Varietät getrennte Analysen, dass dieser größer bei den BP-Daten ($\chi^2[2] = 280.6$, $p < 0.001$) als bei den EP-Daten war ($\chi^2[2] = 126.7$, $p < 0.001$). Der Gestenumfang war allerdings in beiden Varietäten deutlich größer bei $\langle C_{1ut} \rangle$, gefolgt von Clustern und am geringsten bei $\langle C_{1et} \rangle$. Die Sprechervarietät hatte an sich keinen signifikanten Einfluss auf den Gestenumfang und ihre Interaktion mit dem Stimulustyp ($\chi^2[2] = 30.3$, $p < 0.001$) kam durch die unterschiedlich starken Ausprägungen der Varietäten im Umfang bei den einzelnen Stimulustypen zustande.

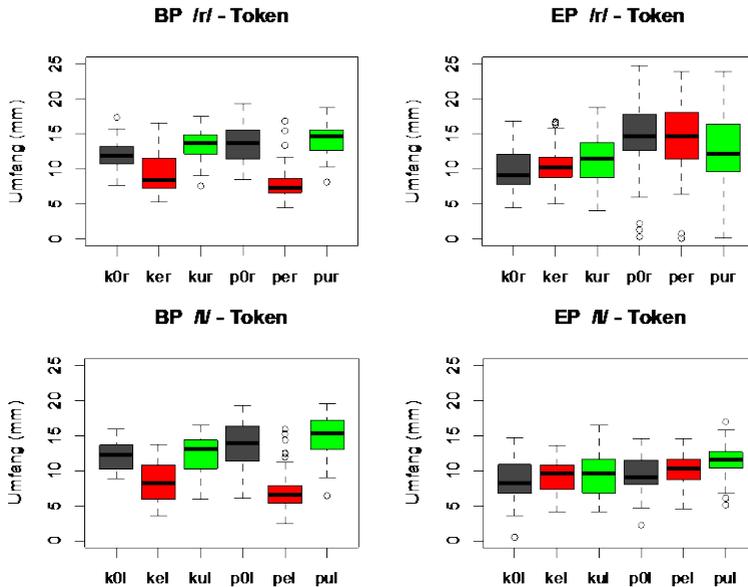


Abbildung 13: Räumlicher Umfang der Zungenspitzenbewegung (Mittelwerte über fünf EP- und vier BP- Sprecher mit Standardabweichungen).

Sowohl der C_1 -Artikulationsort ($\chi^2[1] = 12.5, p < 0.001$) als auch die Betonung ($\chi^2[1] = 4.3, p < 0.05$) zeigten im EP signifikante Effekte und es ergab sich eine signifikante Interaktion zwischen Artikulationsort und Stimulustyp ($\chi^2[1] = 81.0, p < 0.001$). Der Umfang von $\langle C_{1ut} \rangle$ war größer nach bilabialen als nach velaren Plosiven, was jedoch nicht auf die anderen Stimulustypen zutraf. Die Betonung sorgte für eine signifikante Tendenz zu umfangreicheren Gesten in der betonten im Vergleich zur unbetonten Kondition. Im BP hatte weder der C_1 -Artikulationsort noch die Betonung einen signifikanten Einfluss auf den Umfang.

5.4.2.1 Korrelation beider Faktoren

Die Korrelation zwischen maximaler Geschwindigkeit der vertikalen Zungenspitzenbewegung und dem Umfang der apikalen Geste wird in Abbildung 14 dargestellt. Die Plosiv-Token wurden nach Sprechervarietät und C_1 -Artikulationsort aufgeteilt. Die maximale Geschwindigkeit wurde auf die X-Achse und der räumliche Umfang auf die Y-Achse aufgetragen.

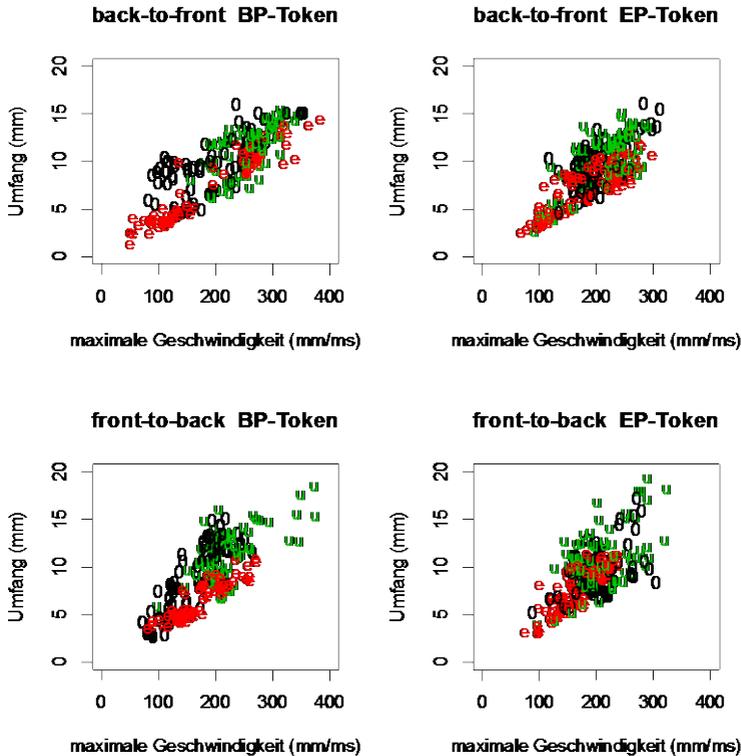


Abbildung 14: Korrelation des Umfangs und der maximalen Geschwindigkeit der Zungenspitzengeste nach Stimulustyp (schwarze /0/ stehen für Cluster, rote /e/ für $\langle C_{1et} \rangle$ und grüne /u/ für $\langle C_{1ut} \rangle$).

Anhand dieser Abbildung kommt stärker zum Ausdruck, dass sich die Stimulustypen eher in den BP- als in den EP-Daten unterscheiden, denn nach Farben getrennte Gruppen sind eher in der linken (BP) als in der rechten Hälfte (EP) der Abbildung zu erkennen. Dennoch hatte der Artikulationsort ebenfalls einen signifikanten Effekt auf die Verteilung, der allerdings in den beiden Varietäten unterschiedlich ausfiel: Wie in der letzten Abbildung zu sehen ist, ergibt sich in den BP-Daten eher bei der

front-to-back Produktionsrichtung eine Konzentration der Werte nach Stimulustypen. Daher wiesen die Stimulustypen in den /pt/-Token weniger signifikante Unterschiede in der Geschwindigkeit und im Umfang auf als in der *back-to-front* Richtung. Außerdem waren die verschiedenen Stimulustypen hier insgesamt stärker neutralisiert als bei /kt/, denn die einzelnen Farbgruppen sind schwerer voneinander abzugrenzen.

In den EP-Daten zeigen die Ergebnisse der Korrelation beider Verfahren insgesamt geringere Streuungen als in den BP-Daten, allerdings liegt die stärkere Konzentration der Werte im EP bei den *back-to-front* Token, die im EP deutlich dichter beieinander liegen als die /pt/-Token. Dieses Ergebnis spiegelt sich ebenfalls im signifikanten Einfluss der Produktionsrichtungen auf die Analyseverfahren wider.

5.4.2.2 Diskussion der wichtigsten Ergebnisse

Aus den Gesamtergebnissen kann man entnehmen, dass die Sprechervarietät per se keinen eindeutigen Einfluss auf die räumliche Organisation der Zungenspitzenbewegung hatte. Dennoch treten mehr signifikante Effekte des Stimulustyps in den BP- als in den EP-Daten auf. Diese Effekte sorgten dafür, dass sich die Dynamik der /t/-Geste im BP deutlicher nach Stimulustyp unterschied als im EP. Im EP waren die Werte des Umfangs und der Geschwindigkeit bei den CVC-Sequenzen ähnlich wie bei den lexikalischen Clustern, insbesondere <C₁et> wies keine signifikanten Unterschiede zu Clustern auf. Das Ergebnis bestätigt die starke Neutralisierung der lexikalischen Unterschiede zwischen Clustern und <C₁et> und in etwas geringerem Maße der Unterschiede zwischen Cluster vs. <C₁ut> aufgrund der häufigen Vokaltilgung der entsprechenden unbetonten Vokale.

Trotz der varietätsspezifischen Unterschiede verhält sich die Dynamik der Zungenspitze nach Stimulustyp in beiden Verfahren entgegengesetzt: Die /t/-Geste bei <C₁ut> wurde mit der höchsten Geschwindigkeit produziert und erreichte gleichzeitig den größten Umfang. Die /t/-Geste bei der anderen CVC-

Sequenz $\langle C_1et \rangle$ zeigte im Gegensatz dazu die niedrigste Geschwindigkeit und den niedrigsten Umfang in beiden Varietäten. Cluster zeigten intermediäre Werte und waren vor allem im EP kaum von $\langle C_1ut \rangle$ zu unterscheiden. In diesem Zusammenhang hatte der unterschiedliche Vokal im Stimulustyp einen eindeutigen Einfluss auf die Gestendynamik, indem hintere Vokale schnellere und umfangreichere apikale Gesten begünstigten als die nicht-hinteren Vokale.

Die Betonung zeigte einen schwachen Effekt, der allerdings geringer ausfiel als erwartet. Es wurden schnellere Gesten in der betonten als in der unbetonten Kondition produziert, allerdings war der Einfluss der Betonung auf den Gestenumfang ausschließlich im EP signifikant und zeigte keinen Effekt im BP. Die Gesten waren im EP wie erwartet umfangreicher in betonter als in unbetonter Kondition. Der Einfluss des C_1 -Artikulationsorts auf die Dynamik der Zungenspitzenbewegung war eher marginal. Der Artikulationsort und der damit verbundene Produktionsrichtungseffekt hatten lediglich eine geringe Auswirkung auf die räumliche Organisation der Gesten. Dies liegt möglicherweise daran, dass der Vokal einen größeren Effekt ausübt als der vorangehende Konsonant.

5.5 Schlussfolgerungen und Ausblick

Mithilfe von akustischen und artikulatorischen Sprachdaten von Sprechern aus Porto in Portugal und São Paulo in Brasilien beschäftigt sich dieses Kapitel mit folgenden Fragen:

- (i) inwiefern sich die Koordination von lexikalischen Plosivclustern im Vergleich zu den CVC-Sequenzen $\langle C_1et \rangle$ und $\langle C_1ut \rangle$ anhand ihrer zeitlichen und räumlichen Organisation in beiden Varietäten unterscheidet;

- (ii) ob sich die synchrone Variabilität dieser Formen in den oben genannten Varietäten in der gesprochenen Sprache durch eine unterschiedliche Koordinierung der Plosive im Sinne der Artikulatorischen Phonologie (Browman & Golstein, 1986, 1988, 1990) erklären lässt.

Die Gesamtergebnisse zeigen wie vorhergesagt, dass Cluster im Vergleich zu $\langle C_1et \rangle$ vs. $\langle C_1ut \rangle$ anhand der zeitlichen Koordination eindeutig unterschieden werden können, indem Cluster enger koordiniert waren als beide CVC-Sequenzen, sowohl in Form von mehr Überlappung als auch von kleineren intermediären Abständen. Die Vorhersage basierte auf zahlreichen Arbeiten zu Konsonantenclustern in einer Vielzahl von Sprachen (Byrd 1993, Bombien 2011, Hardcastle 1995, Marin & Pouplier 2010, Recasens & Espinosa 2005 u.v.a.) und konkret auf Arbeiten zu Plosivclustern (Chitoran et al. 2002, Gafos et al. 2010, Shaw et al. 2009a, b), wofür eben diese Tendenzen beschrieben wurden. Die erneute Bestätigung dieser Ergebnisse bekräftigt, wie robust die Organisation von Konsonanten in Clustern eigentlich ist.

Bezüglich der Dynamik der Zungenspitze hatten erheblich mehr Signifikanzen des Stimulustyps im BP zur Folge, dass sich die Dynamik der /t/-Geste im BP deutlicher nach Stimulustyp unterschied als im EP. Insgesamt war die /t/-Geste schneller und umfangreicher bei $\langle C_1ut \rangle$ in beiden Varietäten als bei den anderen Stimulustypen. Im Gegensatz dazu war sie bei $\langle C_1et \rangle$ am langsamsten und am wenigsten umfangreich. Diese beiden Verläufe unterschieden sich deutlich stärker voneinander im BP als im EP. In Letztem unterscheiden sie sich nach velaren Plosiven nicht signifikant voneinander. Die deutlichen Unterschiede bei den räumlichen Verläufen im BP – sowohl zwischen beiden CVC-Sequenzen als auch im Vergleich zu den Clustern – spricht eindeutig für die Existenz von vokalischen Targets im BP, die mit der beschriebenen Realisierung lexikalischer Vokale im BP übereinstimmen. Im EP waren die Unterschiede zwischen den Stimulustypen bei den räumlichen

Verläufen insbesondere nach velaren Konsonanten stark neutralisiert und die Verläufe insbesondere von Clustern und <C₁et> so ähnlich, dass kein vokalisches Target anhand des Verlaufs zu erkennen war. Dies spricht dafür, dass die Unterschiede im räumlichen Verlauf der Zungenspitze zwischen den Stimulustypen vor allem nach den velaren Konsonanten stärker neutralisiert sind und sich die Stimulustypen eher in der zeitlichen Organisation beider Gesten zueinander unterscheiden.

Wie erwartet wurden erhebliche Koordinationsunterschiede zwischen beiden Varietäten festgestellt, indem BP ein weiteres intergestisches Timing aufwies als EP. Dieser Effekt war über die Gesamtdaten betrachtet sehr robust und trat in beiden Messverfahren auf. Man kann daher an dieser Stelle festhalten, dass sich die Organisation der Konsonanten eindeutig zwischen den beiden Varietäten unterscheidet: Konsonanten werden im BP weiter koordiniert als im EP. Die starke Überlappung der Konsonanten im EP könnte der Grund für die häufigen Vokaltilgungen (Cunha 2011a, *in press*, Cunha & Harrington 2011a, Silva 1998, 1997, Vigário 1998) in dieser Varietät sein. Wenn Konsonanten einen großen Druck auf die benachbarten Segmente ausüben, könnten wenig prominente Segmente graduell kürzer werden und im Extremfall getilgt werden.

Die zeitliche Koordination wurde ebenfalls stark vom C₁-Artikulationsort und dem damit assoziierten Produktionsrichtungseffekt beeinflusst: Im Gegensatz zur Vorhersage überlappten sich die Konsonantengesten eindeutig mehr bei *back-to-front* Clustern (/kt/) und zeigten ebenso signifikant kleinere intermediäre Abstände zwischen den Plateaus als bei *front-to-back* /pt/. Es geht somit eindeutig hervor, dass /kt/ trotz der laut Produktionsrichtungseffekt postulierten schlechteren perzeptiven Wieder-herstellbarkeit erheblich enger koordiniert wurde als /pt/.

Bei beiden CVC-Sequenzen wurde der Produktionsrichtungseffekt von der Interaktion mit dem Stimulustyp geprägt, indem die Plosive in <C₁et> vor allem im BP mehr Überlappung nach

velaren und in <C₁ut> nach bilabialen zeigten. Möglicherweise ist der Artikulationsort des Plosivs weiter hinten bei hinteren Vokalen (Flemming 2003). Außerdem ergab sich eine starke Überlappung beider Konsonanten in /put/ in beiden Varietäten, die in Bezug auf das intergestische Timing nicht von lexikalischen Clustern zu unterscheiden war.

Die Betonung hatte ebenso einen eindeutigen Einfluss auf die intergestische Koordination, denn Token in unbetonter Kondition wurden deutlich enger koordiniert als betonte Token. Die Dynamik der Zungenspitze für die /t/-Geste belegt im EP insgesamt schnellere und umfangreichere Gesten in der betonten als in der unbetonten Kondition. Dies kann eine Folge der stärkeren Koartikulation der Segmente in unbetonter Position sein. Im BP fand sich ebenso eine höhere Geschwindigkeit in betonter als in unbetonter Kondition, allerdings unterschieden sich die Stimulustypen nicht im Umfang. Möglicherweise könnte dies mit einer gewissen Isochronie der Segmente und der Silben im BP (Barbosa 2006) in Verbindung gebracht werden. So ist beispielsweise in Akzentsprachen der Unterschied zwischen betonten und unbetonten Silben nicht sehr stark und dies könnte im BP ähnlich sein (Auer, 1993).

Als Auftakt des Kapitels wurde ausdrücklich auf die wichtige Vorhersage der Sprachperzeption für die Übertragung von Plosivclustern in der Produktion hingewiesen. Diese besagt, dass sich Plosive nicht stark überlappen können, ohne die Wiederherstellbarkeit beider Segmente für den Hörer in Gefahr zu bringen. Trotz dieser Vorhersage zeigten die Ergebnisse eine erhebliche zeitliche Überlappung beider Konsonantengesten in Clustern. Das Synchronisierungsverfahren, bei dem der Anfang von C₂ in Bezug auf C₁ analysiert wurde, ließ erkennen, dass der Beginn von C₂ deutlich vor dem C₁-Plateau lag und somit eine sehr enge Koordinierung der Gesten miteinander vorlag bzw. eine starke Überlappung. Dabei entsteht akustisch entweder eine lange Verschlussphase mit einem deutlichen *Burst* (Verschlussprägung

mit Geräusch assoziiert) oder in extremen Fällen ist nur noch eine kurze Verschlussphase und ein *Burst* zu erkennen.

Bei der ersten geschilderten Situation vermutet man anhand der langen Verschlussphase zwei Plosive im Signal, die aufgrund des fehlenden charakteristischen *Bursts* schwer zu hören sind. Bei der zweiten Situation gibt es umgekehrt keinerlei Hinweise auf die Realisierung beider Plosive im akustischen Signal, aber man kann sie beispielsweise anhand von vorhandenen artikulatorischen Verläufen identifizieren. Anhand dieser weiteren Ergebnisse können dann in Kapitel 6 die übergreifenden Vorhersagen aus 5.1.3 zur Rolle der Perzeption abschließend geprüft werden.

6 Die Perzeption von Clustern und CVC-Sequenzen

6.1 Einführender Forschungsstand

Nachdem Unterschiede zwischen Clustern und CVC Sequenzen in der Sprachproduktion anhand der inter- und intragestischen Koordination der Konsonanten in zwei Portugiesisch-Varietäten untersucht wurden, liegt die Hauptmotivation dieses Kapitels darin, zu überprüfen, ob (1) die Differenzierung von lexikalischen Clustern und lexikalischen CVC-Sequenzen in Bezug auf möglicherweise neutralisierte Formen in der Perzeption erhalten bleibt, und (2) inwiefern dabei der Zusammenhang zwischen vom Sprecher intendierten⁴⁴ und vom Hörer wahrgenommenen Token durch Koordinationsunterschiede erklärt werden kann.

Zu diesem Zweck wird mit einer Auswahl von lexikalischen Clustern und entsprechenden lexikalischen CVC-Sequenzen mit intermediären <e> und <u> aus den vorangehenden Kapiteln ein Perzeptionsexperiment durchgeführt. Angenommen, es gibt in einer <C₁C₂>-Sequenz einen durchschnittlichen Umbruchpunkt, ab dem Hörer zwei Konsonanten wahrnehmen, so lautet unsere Hypothese wie folgt: Bei einer breiteren Koordination der Konsonanten führt der größere Abstand zwischen den Konsonantengesten zu der Wahrnehmung eines kurzen vokalischen Elements. Im Gegensatz dazu sind die Hörer bei einer engen Koordination nicht mehr in der Lage, zwei Konsonanten wahrzunehmen, sondern lediglich einen. Dieser Ansatz hat den Vorteil, die große Variabilität, die normalerweise

⁴⁴ Der Begriff „intendiert“ bezieht sich in dieser Arbeit, wie bereits bemerkt, nicht auf kognitive Prozesse beim Sprechen, sondern bezeichnet einfach das Token, das die Probanden laut Stimuluspräsentation produzieren sollten.

bei Prozessen wie Vokalepenthese, Vokal- und Konsonantentilgungen beschrieben wird, im gleichen theoretischen Rahmen erklären zu können.

6.1.1 Perzeption in der Sprachwissenschaft

Die Sprachwahrnehmung oder -perzeption als Teilbereich der Phonetik beschäftigt sich mit der Verarbeitung von Sprachlauten durch das Gehör im Gehirn. Für die aktuelle Fragestellung ist weniger die Übertragung der Laute durch das Gehör (auditive Phonetik) relevant, sondern ihre Übermittlung und Verarbeitung im Gehirn (perzeptive Phonetik). Bei diesem Prozess wird das Sprachsignal in Zeichen überführt, wie beispielsweise Phoneme, Silben, Wörter usw. und muss dabei der großen Variabilität des Sprechens Rechnung tragen. Beim Sprechen fließen z. B. sprecherspezifische Merkmale (z.B. Alter, Geschlecht, Körpergröße und Form des Artikulationsstrakts), soziale und kontextuelle Begebenheiten (Dialekt oder dialektale Färbung, Sprachstil und Sprechgeschwindigkeit) sowie der Lautkontext mit in die Variation ein. Diese große Variabilität sowohl zwischen Sprechern als auch innerhalb eines Sprechers (Groswald 2009) kann eine Ambiguität bei der Überführung von Sprachsignalen in Zeichen verursachen. Daher ist der große Anspruch der perzeptiven Phonetik, bei der extremen Variabilität der Sprachproduktion die nötige Invarianz für die Überführung des Signals in ein bestimmtes Zeichen/Phonem herauszustellen.

Diese Invarianz wird in der *Motor Theory* (Liberman et al. 1967, Liberman & Mattingly 1985) auf die abstrakten motorisch-artikulatorischen Gesten zurückgeführt, d.h. auf das artikulatorische Bewegungsmuster eines bestimmten Lautes (Liberman & Mattingly 1985). Im Hyper- & Hypoartikulationsmodell (H & H) von Lindblom (1990, 1996) wird die Wahrnehmung eher interaktiv verstanden. In diesem Sinne interagieren Sprecher und Hörer in einem Kommunikationsmodell, in dem jede Äußerung in einem Kontinuum von Hyper- zu Hypoartikulation produziert wird. Dabei bringt der Hörer sein Weltwissen zur Kommunikation mit und passt es in jedem

Gespräch neu an. Ein Wort oder eine Lautkette tendiert zur Hyperartikulation in neuen oder unvorhersehbaren Kontexten und zur Hypoartikulation in vorgegebenen, absehbaren Kontexten (Lindblom et al. 2011, Cho 2005, Cho et al 2005). Dies resultiert in einer deutlichen, gepflegten Aussprache und weniger Koartikulation bei der Hyperartikulation (Lindblom et al. 2007, Wright 2003). Bei Wiederholungen und absehbaren Kontexten hingegen, die meistens mit einem geringeren Informationsgehalt einhergehen, findet man eher mehr Koartikulation und eine weniger deutliche Aussprache. Diese Tatsache ist ein möglicher Grund, weshalb Vokalschwächungen und Tilgungen eher in unbetonter Position vorkommen (Davidson et al. 2004, Manuel et al. 1995).

Der Zusammenhang zwischen Produktion und Perzeption und die Vorteile der Einbeziehung der Perzeption in die phonologische Theorie gewannen auch in der Phonologie, und in der Sprachwissenschaft im Allgemeinen, eine zunehmende Relevanz. Dabei sollte die Analyse von sprachlichen Phänomenen aus dem Zusammenspiel von Produktion und Perzeption der Sprache hervorgehen (Krefeld, 2004), da Sprache aus Produktion und Perzeption besteht und sich beide Teilbereiche wie „zwei Seiten derselben kommunikativen Medaille“ (Krefeld & Pustka 2010: 9) ergänzen. Daher muss eine vollständige Analyse der Sprache ebenso beide Bereiche beinhalten. Dabei ist die Verwendung von Perzeptionsexperimenten unentbehrlich, um die „perzeptiv basierten Repräsentationen“ zu erschließen:

Doch nur letztere [Perzeptionsexperimente, C.C.] konfrontieren die Sprecher-Hörer mit konkreten Produktionsdaten und ermöglichen es so, die perzeptiv basierten Repräsentationen zu identifizieren und so von den außersprachlich basierten zu trennen (Krefeld & Pustka 2010: 16). Für eine Optimierung der perzeptiven Repräsentationen der Sprecher können Perzeptionsexperimente idealerweise mit Befragungen, Fragebögen oder anderen Methoden kombiniert werden. In der vorliegenden Arbeit liegt der Schwerpunkt auf dem Zusammenspiel zwischen

Produktion und Perzeption der Sprache und daher auf dem Vergleich zwischen Produktions- und Perzeptionsdaten.

6.1.2 Aktueller Stand der Forschung

Konsonantencluster können im Portugiesischen mit einem intermediären Vokal realisiert werden: Dieses Phänomen – die sogenannte Vokalepenthese – betrifft in der Gegenwart /l/-Cluster im EP ([pɪlātɐ] für *planta*, ‘Pflanze’ vgl. Veloso, 2003) und Obstruentencluster sowohl im BP als auch im EP (z.B. BP [kapi'tar] vs. EP [kɛpi'tar] für *captar* ‘fangen’, vgl. Bisol 1999, Mateus & Andrade 1998, 2000, Vigário 2003, Veloso 2007a und erstes Kapitel dieser Arbeit). Bei reinen Plosivclustern kann alternativ der erste von zwei Konsonanten, insbesondere im BP, getilgt werden (z.B. [fatu] für *facto* ‘Fakt’, vgl. Mateus & Andrade 1998, 2000, Vigário, 2003). Da Plosive Teil der Obstruenten sind, können bei Plosivclustern im Portugiesischen sowohl Vokalepenthese als auch C₁-Tilgung vorkommen.

Das Auftreten intermediärer Vokale bei Konsonantenclustern mit Liquiden wurde auch in anderen Sprachen beschrieben, z.B. Ingalaterra statt *Inglaterra* ‘England’, egelesia statt *eglesia* ‘Kirche’, coronica statt *cronica* ‘Chronik’ im Spanischen (vgl. Menendez-Pidal 1926: 127) oder [blo] vs. [belo] für *blow* im Englischen (vgl. Ohala 1992: 323, Browman & Goldstein 1990, Price 1980). Eine mögliche Erklärung für die letzteren Aussprachevarianten ist, dass die längere Dauer des Sonoranten nach dem bilabialen Plosiv in *blow* bei den Hörern ein anderes perzeptives Konstrukt induziert, das besser mit dem akustischen Signal ohne Vokal übereinstimmt (Ohala 1992: 331). Mit anderen Worten, Hörer interpretieren den verlängerten Sonoranten als Abfolge von Schwa und Sonorant ([ə]). Dies kann eintreten, weil z. B. VC-Transitionen weniger wichtig für die Wahrnehmung der einzelnen Segmente sind als CV-Transitionen (Ohde & Sharf 1977) oder weil Schwa wie z.B. im Englischen ohne ein vokalisches Target („*targetless schwa*“, Browman & Goldstein 1990) produziert wird.

Eine andere, gleichberechtigte Erklärungsmöglichkeit für diese Variation bei den Konsonantenclustern (z.B. [blo] vs. [belo]) ist der Einfluss einer graduellen zeitlichen Koordination der einzelnen Konsonantengesten auf die Sprachperzeption (Cunha & Harrington 2011b) im Rahmen der Artikulatorischen Phonologie (Browman & Goldstein 1986). Die Grundidee dahinter ist, dass, wenn C_2 in einer $\langle C_1C_2 \rangle$ -Sequenz später auftritt, der größere Abstand zwischen den Konsonanten in der Sprachperzeption durch ein kurzes vokalisches Element wiedergegeben wird, ohne dass ein Vokal in die lexikalische Repräsentation des Wortes integriert werden muss. So würden also lexikalisch reine Konsonantenabfolgen (Cluster) mit einem intermediären Vokal wahrgenommen werden.

Allerdings wurde ebenfalls die gegensätzliche Tendenz festgestellt: Manche lexikalisch unbetonten Vokale (vor allem das zentrierte Schwa /ə/) können geschwächt, gekürzt oder getilgt werden und infolgedessen können lexikalische CVC-Sequenzen als reine Konsonantenabfolgen (CC) beschrieben werden. Dies wurde im Englischen (Davidson 2006, Manuel et al. 1992) und im Deutschen (Szczepaniak 2007, Nübling et al. 2006: 32-34, Ramers & Vater 1995: 47-66), aber auch in zahlreichen romanischen Sprachen festgestellt, nämlich im Französischen (Pustka 2011, 2007, Torreira & Ernestus 2011), im Italienischen (Garrapa 2011), im Rumänischen (Krefeld 1987), im Katalanischen (z.B. Prieto 2004: 195-199), im EP (Cunha 2011a, Silva 1998, 1997, Mateus & Martins 1982, Martins 1975) und in mehreren Spanisch-Varietäten aus Lateinamerika (in Mexiko: Serrano 2006, Lope Blanch 1963, Boyd-Bowman 1952; in Peru: Delforge, 2009, 2008 a, b, Hundley 1986, 1983; in Bolivien: Gordon 1980 und in Ecuador: Lipski 1989, 1990, 1995).

Im EP wird kein zentriertes Schwa getilgt, sondern die unbetonten hohen Vokale (/u, i/), so dass sowohl Sequenzen mit intermediärem /i/ ($\langle C_1eC_2 \rangle$), wie in *tapetar* [təpi'tar] 'Teppich legen', als auch Sequenzen mit intermediärem /u/ ($\langle C_1uC_2 \rangle$), wie in *computar* [kõ'putar] 'ausrechnen' als reines Konsonantencluster

(CC wie in [tʁ'ptar] und [kõ'ptar]) realisiert werden können. Bei der Modellierung dieser Alternanz zwischen CVC-Vollformen und reduzierten Formen (CC) stellt sich die Frage, inwiefern sich reduzierte Formen und Vollformen unterscheiden. Auch hier könnte man das oben erwähnte Modell der Artikulatorischen Phonologie anwenden, um der Frage nachzugehen, inwiefern verschiedene Koordinationsmuster der Konsonanten die Sprachwahrnehmung beeinflussen. In diesem Fall wäre das graduelle Modell in der entgegengesetzten Richtung denkbar: Die Hypothese würde lauten: Wenn C_2 in einer $\langle C_1VC_2 \rangle$ Sequenz früher auftritt und sich beide Konsonanten mehr überlappen, wird der Abstand zwischen C_1 und C_2 verkürzt. Diese Abstandsverkürzung würde in der Sprachperzeption durch die Schwächung oder Tilgung des vokalischen Elements wiedergegeben werden.

Aus der vorgestellten Datenbesprechung entsteht der Eindruck, dass lexikalische Cluster als reine Konsonantensequenzen oder mit einem (epenthetischen) intermediären Vokal beschrieben werden und lexikalische CVC-Sequenzen ohne oder mit intermediärem Vokal auftreten können. Frühere Studien über Konsonantenkoordination von englischen Clustern und CVC-Sequenzen wiesen jedoch darauf hin, dass Probanden bei der Überlappung beider Konsonantengesten eher ein Cluster und bei Nicht-Überlappung eher eine CVC-Sequenz wahrnahmen (Browman & Goldstein 1990a, b). Bei portugiesischen Clustern und CVC-Sequenzen mit intermediären /e, u/ ($\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$) wurden in den Kapiteln 4 und 5 dieser Arbeit zum Teil kleine, aber konsistente Unterschiede bei der Koordinierung der Konsonantengesten festgestellt (siehe Ergebnisse und Diskussion in diesen Kapiteln), so dass jeder Stimulustyp auch in Abhängigkeit weiterer Faktoren in der Produktion unterschieden werden konnte.

Genau an diesem Punkt setzt das vorliegende Kapitel an. Die Hauptmotivation dieses Kapitels liegt darin, am Beispiel der zwei portugiesischen Varietäten EP und BP, in denen beide Phänomene auftreten, zu überprüfen, ob lexikalische Cluster und

lexikalische CVC-Sequenzen in der Perzeption identifiziert werden können. In diesem Zusammenhang wird auch der Frage nachgegangen, inwiefern dabei der Zusammenhang zwischen intendierten und wahrgenommenen Token durch Koordinationsunterschiede erklärt werden kann. Sollte sich die Alternanz zwischen lexikalischen Clustern und Clustern mit einem intermediären epenthetischen Vokal (CəC) sowie lexikalischen CVC-Sequenzen und deren reduzierten Formen (CC) vor allem durch die zeitliche Koordinierung ihrer Konsonanten unterscheiden, wäre eine leichte Verwechslung denkbar.

Zu diesem Zweck wird zunächst für eine Auswahl von lexikalischen Clustern (bestehend aus Plosiv und Liquid oder aus zwei Plosiven) und für die entsprechenden lexikalischen CVC-Sequenzen mit intermediären <e> und <u> aus den vorangehenden Kapiteln untersucht, ob die Probanden Cluster vs. <C₁eC₂> vs. <C₁uC₂> identifizieren können.

6.1.3 Kategoriale Wahrnehmung

Kategoriale Wahrnehmung bezeichnet einen Effekt in der Sprachperzeption, wonach einzelne Stimuli in der Wahrnehmung scharf voneinander abgegrenzten Kategorien zugeordnet werden. Dies betrifft insbesondere kontinuierliche Stimuli (z.B. entlang eines Kontinuums zwischen /i/ und /u/), die nicht graduell, sondern kategorial wahrgenommen werden. Verschiedene Perzeptionsexperimente zeigten, dass Hörer eine Reihe von Einzelreizen, bei denen ein akustischer Parameter in regelmäßigen Abständen variiert wurde, nicht als kontinuierlich empfanden, sondern in scharf abgegrenzte Kategorien einteilten (Liberman et al. 1957, Fujisaki & Kawashima 1968, 1970, Pastore 1987 für den Grenzeneffekt, Pisoni 1973, 1975 für das auditive Gedächtnis).

Bei dem für diese Arbeit durchgeführten Perzeptionsexperiment wurde kein Kontinuum präsentiert, sondern einzelne, nicht manipulierte akustische Stimuli, die einer von vier vorgegebenen Optionen zugeordnet werden sollten. In dieser Arbeit wird also das Paradigma der kategorialen Wahrnehmung – obwohl es eine

gängige Aufgabestellung in der Perzeptionsforschung ist – nur in abgewandelten Form angewendet⁴⁵ (für Genaueres siehe Punkt 6.2.1).

6.2 Fragestellung

6.2.1 Identifikation des Stimulustyps

Das erste Ziel des Experiments war es, zu untersuchen, ob Hörer in der Lage sind, in der eigenen und in der fremden portugiesischen Varietät den Stimulustyp, d.h. lexikalische Cluster vs. lexikalische $\langle C_1eC_2 \rangle$ vs. lexikalische $\langle C_1uC_2 \rangle$ zu identifizieren.

Hierbei fließen bisherige Forschungserkenntnisse und die Ergebnisse aus den Produktionskapiteln der vorliegenden Arbeit in die zwei folgenden Hypothesen mit ein: Entweder werden kleine Unterschiede in der Koordination der Konsonanten von den Hörern nicht wahrgenommen und daher nehmen sie den Unterschied zwischen Cluster vs. $\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$ nicht wahr. Oder bestimmte phonetische Details ermöglichen es den Hörern, den Stimulustyp zu identifizieren (unvollständige Neutralisierung, Port & O'Dell 1985). Dabei soll der Einfluss weiterer Faktoren, nämlich der Sprechervarietät, des Artikulationsorts von C_1 und der Artikulationsart von C_2 , der Primärbetonung und der Hörervarietät, miteinbezogen werden.

6.2.2 Wahrnehmung von Koordinationsunterschieden

Das zweite Ziel des Experiments war, zu testen, ob Hörer zu jedem präsentierten Stimulustyp (Cluster vs. $\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$) jeweils ein, zwei oder drei Elemente wahrnehmen und inwiefern dies mit der Enge der Konsonantenkoordination in der Produktion zusammenhängt. Dafür stellten wir folgende Vorhersagen auf:

⁴⁵ Für eine Übersicht über die klassischen Experimente zur kategorialen Wahrnehmung siehe Liberman et al. (1957), Repp (1982).

- (i) Wenn C_2 während des C_1 -Plateaus beginnt, überlappen sich beide Konsonantengesten, was dazu führt, dass die Probanden ein Cluster (CC) wahrnehmen (Ausgangspunkt).
- (ii) Wenn C_2 in der Produktion später auftritt und nach dem C_1 -Konstriktionsplateau beginnt, überlappen sich beide Konsonantengesten nicht mehr. Dies hat zur Folge, dass Hörer einen intermediären Vokal (CVC) wahrnehmen, unabhängig davon, ob ein Vokal von den Sprechern intendiert wurde oder nicht.
- (iii) Wenn C_2 in der Produktion früher auftritt und vor dem C_1 -Plateau beginnt, findet mehr Überlappung statt. In diesem Fall wird die Verschlusslösung von C_1 möglicherweise von C_2 verdeckt, so dass Hörer dazu tendieren, nicht mehr zwei, sondern nur einen Konsonanten wahrzunehmen (C).

Dahinter steckt allerdings die Annahme, dass sich einzelne Konsonanten, Cluster und CVC-Sequenzen ausschlaggebend in der Koordinierung der einzelnen Konsonantengesten unterscheiden und anderen phonetischen Details weniger relevant für die Wahrnehmung hinsichtlich der Anzahl der Elemente sind. Die drei Vorhersagen werden nochmals zur Veranschaulichung in schematischer Form wiederholt.

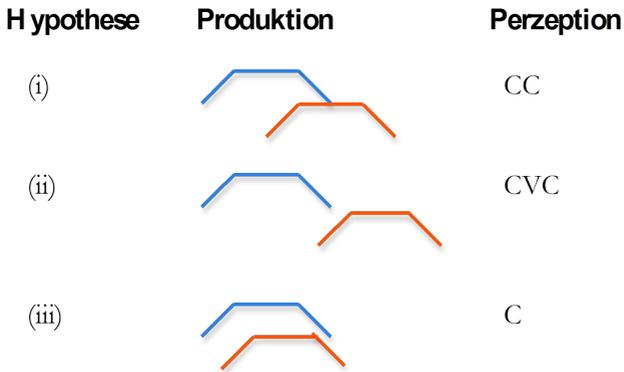


Abbildung 15: Schematische Darstellung der Hypothesen für den Zusammenhang zwischen Produktion und Perzeption von Clustern und CVC-Sequenzen.

Bei der Überprüfung der Vorhersagen zum Stimulustyp wird die Abhängigkeit von weiteren Faktoren, nämlich von dem Artikulationsort von C_1 , der Varietät von Sprechern und Hörern und der Betonungskondition in Betracht gezogen. Bei den Interaktionen zwischen den Faktoren gilt es allerdings zu bedenken, dass der Schwerpunkt dieser Untersuchung auf dem Vergleich von Clustern vs. CVC-Sequenzen liegt. Aus diesem Grund werden Interaktionen von Clustern mit jeder der CVC-Sequenzen geprüft (d.h. Cluster vs. $\langle C_{1et} \rangle$ und Cluster vs. $\langle C_{1ut} \rangle$), Vergleiche beider CVC-Sequenzen miteinander ($\langle C_{1et} \rangle$ vs. $\langle C_{1ut} \rangle$) jedoch ausgeklammert. Die untersuchten Faktoren sind in dieser Hinsicht analog zu dem Produktionsexperiment.

6.2.3 Hypothesen und Vorhersagen

6.2.3.1 Sprechervarietät: BP vs. EP

Mit dem Faktor *Sprechervarietät* soll getestet werden, ob die dialektale Ausprägung der Sprecher (BP vs. EP) einen Einfluss auf die Fragestellung hat, d.h. inwiefern sie die Identifikation des Stimulustyps oder die Wahrnehmung der Koordinationsunterschiede beeinflusst.

Beide analysierten Sprachvarietäten haben einen gemeinsamen Ursprung, allerdings werden sie mit unterschiedlichen Tendenzen assoziiert: EP tendiert dazu, hohe unbetonte Vokale in der gesprochenen Sprache zu tilgen (Cunha 2011a, Mateus & d'Andrade 2000, Silva 1998, 1997, Vigário 1998, Martins 1975), während BP-Sprecher alle Vokale realisieren und dazu tendieren, in Obstruentensequenzen einen intermediären Vokal zu produzieren (Vokalepenthese, z.B. Mateus & D'Andrade 2000, Bisol 1999). Davon ausgehend wird erwartet, dass die für Plosivcluster beschriebene Vokalepenthese die Identifikation der Token bei Plosiven beeinträchtigt, jedoch nicht bei Plosiv-Liquid-Token. Im EP wird eine gewisse Abschwächung der lexikalischen Unterschiede erwartet, so dass alle Stimulustypen tendenziell als reine Konsonantencluster (CC) wahrgenommen werden.

Bezüglich der Konsonantensynchronisierung konnte in den Kapiteln 3 und 4 dieser Arbeit festgestellt werden, dass EP-Token insgesamt enger als BP-Token koordiniert werden, indem sie eine größere Überlappung als BP-Token aufweisen. Angenommen, dass Produktion und Perzeption übereinstimmen, so wird vorhergesagt, dass bei BP-Stimuli mehr intermediäre Vokale und bei EP-Stimuli mehr einzelne Konsonanten wahrgenommen werden. Außerdem erwarten wir eine Interaktion zwischen Sprechervarietät und Stimulustyp, da die Vokale [u] im BP und [ɨ] im EP eine engere Synchronisierung beider Konsonanten zulassen. Daher sollten diese in der Perzeption stärker zu Konsonantenclustern neutralisiert werden, als die Sequenzen mit [e] im BP und mit [u] im EP.

6.2.3.2 Stimulustyp: Cluster vs. <C₁eC₂> vs. <C₁uC₂>

Mit dem Faktor Stimulustyp wird überprüft, inwiefern Cluster vs. <C₁eC₂> vs. <C₁uC₂> identifiziert werden können, bzw. ob die Wahrnehmung von ihrer Koordination beeinflusst wird. Aus der bestehenden Literatur ist bekannt, dass hohe unbetonte Vokale im EP neutralisiert (Mateus & Martins 1982) und getilgt werden können (Cunha 2011a, Mateus & d'Andrade 2000, Silva 1998, 1997, Vigário 1998, Martins 1975). Im BP kommt dies nicht vor,

dort werden alle Vokale realisiert. Daher wird nur für EP eine gewisse Abschwächung des lexikalischen Unterschieds erwartet, wobei alle drei Stimulustypen zu reinen Konsonantenclustern (CC) neutralisiert werden sollten.

Im BP wurde hingegen die Tendenz beschrieben, reine Obstruentenfolgen mit einem intermediären Vokal zu realisieren (Vokalepenthese, z.B. Mateus & D'Andrade 2000, Bisol 1999). Davon ausgehend wird erwartet, dass die beschriebene Vokalepenthese bei Obstruentenfolgen die Identifikation des Stimulustyps lediglich bei Plosivclustern beeinträchtigt und dementsprechend alle anderen Stimulustypen im BP identifiziert werden können.

6.2.3.3 Artikulationsort: bilabialer vs. velarer C₁

Der C₁-Artikulationsort wird aufgrund seiner Relevanz für die Modellierung von Clustern mit dem Produktionsrichtungseffekt assoziiert (Chitoran et al. 2002). Dieser bereits thematisierte Effekt besagt, dass sich die Konsonantengesten in *front-to-back* Clustern (/pt/) aufgrund der leichteren Wiederherstellbarkeit ihrer Elemente mehr überlappen können als in der umgekehrten Richtung (*back-to-front*, z.B. /kt/). Für Plosivcluster im Georgischen zeigten Chitoran et al. (2002), dass *front-to-back* Cluster eine konsistent größere Überlappung aufwiesen als die Cluster in der umgekehrten Richtung. Der Grund dafür ist möglicherweise, dass selbst bei einer größeren Überlappung der Plosive sich die *front-to-back* Cluster dank ihrer zusätzlichen visuellen Unterstützung leichter bei der Perzeption wiederherstellen lassen als die *back-to-front* Cluster (z.B. /tp/). Kühnert et al. (2006) stellten die gleiche Tendenz für französische Plosiv-Liquid-Cluster fest. Allerdings fügen die Autoren selbst hinzu, dass Liquide als zweiter Konsonant den ersten nicht so stark maskieren können wie Obstruenten. Daher wäre der Einfluss der perceptiven Wiederherstellbarkeit als Argument fraglich. Der mögliche Grund, weshalb die Argumente übereinstimmen, sei eher die Aktivität von unabhängigen Artikulatoren (z.B. Lippen und Zungenspitze in /pt/) bei der Produktion von *front-to-back* Clustern im Vergleich zu

verbundenen Artikulatoren (z.B. Zungendorsum und -spitze in /kt/) bei den *back-to-front*-Gegenstücken. Bei zwei unabhängigen Artikulatoren, z.B. Lippen und Zungenspitze in /pt/ oder /pr/, kann sich die Zungenspitze gleichzeitig mit dem Lippenverschluss zu den Alveolen bewegen, wo der zweite Konsonant /t/ oder /r/ produziert wird. Bei /kt/ oder /kr/ sind zwei Bereiche des Zungenartikulators aktiv (Zungendorsum und -spitze), die sich nicht frei bewegen können.

Unabhängig davon, welchen Ansatz man bevorzugt, sind beide in der Hinsicht stimmig, dass *front-to-back* Stimuli mehr überlappen können, ohne die Wahrnehmung der einzelnen Elemente zu beeinträchtigen. Angesichts der Tatsache, dass *back-to-front* Plosivcluster realiter mehr Überlappung zeigen, erwarten wir vor allem für die Plosivstimuli einen geringeren Anteil von Identifikation des Stimulustyps in der *back-to-front*-Richtung als in der *front-to-back* Richtung. Bezüglich des Einflusses der Koordination auf die Perzeption erwarten wir aufgrund der größeren Überlappung bei den Produktionsdaten mehr Wahrnehmungen von einzelnen Konsonanten bei *back-to-front* Token (/kt/) als in der umgekehrten Richtung.

6.2.3.4 Betonung: betonte vs. unbetonte Kondition

Wie bereits in den vorherigen Kapiteln erwähnt, werden bei den CVC-Sequenzen (<C₁eC₂> vs. <C₁uC₂>) Stimuli verglichen, in denen C₂ entweder in betonter (primärbetonte Silbe) oder in unbetonter Position (eine Silbe vor der Primärbetonung) steht (z.B. prada vs. pradada bei den Clustern und perada vs. perador bei den CVC-Sequenzen⁴⁶). Der Zielvokal (V=/e, u/) muss in einer unbetonten Silbe vorkommen, da nur unbetonte Vokale Tilgungsprozessen unterliegen.

Mit der Betonungskondition wird der Einfluss der Betonung auf die Identifikation des Stimulustyps und auf die Anzahl der wahrgenommenen Elemente getestet. Aufgrund der längeren

⁴⁶ Die Zielcluster und Ziel-CVC wurden kursiv gesetzt und die primärbetonte Silbe unterstrichen.

akustischen Dauer von betonten Silben (Browman & Goldstein 1990a: 103), die artikulatorisch entweder einer weiteren Koordination beider Konsonanten oder einer geringeren Steifheit der Gesten entspricht (*stiffness*, Browman & Goldstein 1990a: 103, Kelso et al. 1985), wird erwartet, dass Cluster in betonten Silben leichter identifizierbar sind und auch die Anzahl der wahrgenommenen Elemente eher dem Stimulustyp entspricht. Für CVC-Sequenzen zeigte Martins (1995), dass die Anzahl der erkannten Stimuli abnimmt, je weiter sie von der primärbetonten Silbe entfernt liegen. Bei beiden CVC-Sequenzen wird daher erwartet, dass die Stimuli schlechter identifizierbar sind, je weiter sie von der Primärbetonung entfernt sind.

6.2.3.5 Hörervarietät (BP vs. EP)

Zuletzt wird untersucht, inwiefern die Hörervarietät, d.h. die Sprachvarietät des Probanden (EP vs. BP), bei der Fragestellung eine Rolle spielt. Dahinter steckt die Annahme, dass die unterschiedliche Produktion zur Speicherung verschiedener Formen in Sinne der *Exemplar Theory* (Johnson 1997, 2006, 2011, Pierrehumbert 2000) führen kann. Demzufolge hätten Brasilianer z.B. mehrere Formen von breit koordinierten Plosivclustern in ihrem Inventar, die sich in minimalen phonetischen Details unterscheiden. Portugiesen hingegen hätten eher eine große Anzahl von Formen mit eng synchronisierten Konsonanten gespeichert, die sich ebenfalls lediglich in kleinen Details unterscheiden.

Sollten Produktion und Perzeption übereinstimmen, würden EP-Hörer die Cluster mit epenthetischem Vokal eher als CVC-Sequenz einstufen, während BP-Hörer in der Lage sein könnten, sie anhand von phonetischen Details von den lexikalischen CVC-Sequenzen zu unterscheiden. Bei den EP-Token wird erwartet, dass die EP-Hörer den etwas längeren Abstand zwischen den Konsonanten bei CVC-Sequenzen im Vergleich zu Clustern in ihrer Varietät mit einem kurzen intermediären Vokal assoziieren, während dies für BP-Hörer nicht wahrnehmbar ist. Der Grund hierfür ist möglicherweise, dass keine zentrierten, kurzen, Schwa-ähnlichen Vokale im BP-Inventar vorhanden sind. Daher wird bei

beiden CVC-Sequenzen aus dem EP erwartet, dass BP-Sprecher die kürzeren europäischen Vokale weniger gut identifizieren als europäische Hörer.

6.3 Methode

6.3.1 Stimuli

Anstelle von synthetischen graduellen Stimuli wurden hier natürliche Stimuli von jeweils einem BP und einem EP Sprecher verwendet. Für die Beantwortung der Frage nach varietäten-spezifischer Produktion wurden die in Tabelle 8 gelisteten Wörter mit /Ct/- und /Cr/-Clustern (C=/p, k/) ausgewählt. Wie im Produktionsexperiment befanden sich alle /r/-Token in wortinitialer und die Plosiv-Token in wortinterner Position.

Bei den Aufnahmen wurden die Zielwörter in folgende Trägersätze eingebaut, bei denen der Satzakzent variiert wurde. Beim ersten Trägersatz wurde das Zielwort in akzentuierter (*O Pedro leu ___ bem*, ‘Pedro las ___ richtig’) und beim zweiten in deakzentuierter Position (*O Pedro leu ___ bem*, ‘Pedro las ___ richtig’) vorgelesen. Jede Satzfolge wurde viermal pro Sprecher in randomisierter Reihenfolge wiederholt. Die Aufnahmen fanden am Institut für Phonetik und Sprachverarbeitung der LMU München statt⁴⁷.

Da es sich bei den Aufnahmen um eine Leseaufgabe handelte, war es möglich, die intendierten Stimuli bei der Produktion festzustellen, d.h. welches Wort der Sprecher gerade vorlas. Somit war es also möglich, die intendierten akustischen Stimuli festzuhalten, die bei der Auswertung des Experiments dann den dazugehörigen Perzeptionsantworten gegenübergestellt wurden.

⁴⁷ Die technischen Details waren dieselben wie die in Kapitel 4 beschriebenen (siehe 4.2). Diese Informationen wurden aus den Perzeptionskapiteln ausgeschlossen.

r-Token	Cluster	<i>crer</i>	‘glauben’
		<i>cridin<u>h</u>o</i>	‘gläubig’
		<i>prece</i>	‘Gebet’
		<i>pre<u>ce</u>r</i>	Logatom
C₁eC₂		<i>querer</i>	‘wollen’
		<i>queridin<u>h</u>o</i>	‘Liebling’
		<i>per<u>ce</u></i>	‘Meinung’
		<i>per<u>ce</u>r</i>	‘untergehen’
C₁uC₂		<i>cur<u>a</u>do</i>	‘geheilt’
		<i>cur<u>a</u>dor</i>	‘Vormund’
		<i>pur<u>a</u>tos</i>	‘Mehlsorte’
		<i>purat<u>ã</u>o</i>	‘frech’
Plosiv-Token	Cluster	<i>compact<u>a</u>do</i>	‘kompakt’
		<i>compact<u>a</u>dor</i>	‘Verdichter’
		<i>capt<u>a</u>r</i>	‘empfangen’
		<i>capt<u>a</u>dor</i>	‘Empfänger’
C₁eC₂		<i>que t<u>a</u>do⁴⁸</i>	‘wie doof’
		<i>que tad<u>a</u>do</i>	‘wie doof’
		<i>tap<u>e</u>tar</i>	‘Teppich verlegen’
		<i>tap<u>e</u>tador</i>	‘Teppichverleger’
C₁uC₂		<i>cut<u>a</u>no</i>	Eigenname
		<i>cut<u>a</u>no<u>s</u>o</i>	E., Genitiv
		<i>com<u>p</u>utar</i>	rechnen’
		<i>com<u>p</u>utador</i>	‘Rechner’

Tabelle 8: Liste der portugiesischen Zielwörter mit deutscher Übersetzung (hier in kursiv dargestellt) und den betonten Silben (als unterstrichene Silben markiert).

⁴⁸ Da das Sprachsystem hier eine Lücke aufweist, wurde die Tabelle mit lexikalisierten, festen Ausdrücken aufgefüllt.

6.3.2 Perzeptionsexperiment

Für das Perzeptionsexperiment wurden die akustischen Stimuli aus der zweiten Wiederholung von zwei Probanden (einmal EP, einmal BP) mit MAUS (Schiel, 1999) vorsegmentiert und anschließend in Emu (Harrington 2010) manuell korrigiert. Die Zielsilben wurden zwischen dem vorangehenden Vokal-Offset und dem nächsten Vokal-Onset extrahiert. Diese wurden bei den /r/-Token mit Minimalpaaren ergänzt, da diese im Portugiesischen existieren. Eine Auflistung der exakten Stimuli folgt in Tabelle 9:

	velar	bilabial
Plosiv-Token	/kt/	/pt/
	/ket/	/pet/
	/kut/	/put/
r-Token (I)	/kr/	/pr/
	/ker/	/per/
	/kur/	/pur/
r-Token (II)	crer	prece
	cridinho	precer
	querer	perece
	queridinho	perecer

Tabelle 9: Liste der Stimuli für das Perzeptionsexperiment.

Diese Stimuli waren mit weiteren Stimuli bei einer größeren Studie getestet worden (insgesamt fünf Blöcke: K-Plosiv, P-Plosiv, K-Liquid, P-Liquid und Minimalpaare, wovon wir hier die ersten zwei auswerten). Die Analyse berücksichtigt für die Plosivstimuli sechs Wörter x zwei Kontexte x zwei Sätze x zwei Bedingungen x zwei Sprecher (= 96 Stimuli, insgesamt 144 Stimuli).

Die Stimuli wurden in ein Online-Experiment eingefügt, bei dem jeder Proband (Hörer) sich durch das Anklicken eines

Lautsprecher einen Stimulus anhören sollte und jeden akustischen Stimulus einer von vier Antwortmöglichkeiten zuordnen musste. Die vier Optionen waren erstens die orthographische Form des entsprechenden Clusters, zweitens C_2 sowie drittens und viertens die entsprechenden CVC-Folgen mit jeweils /e/ oder /u/ in der V-Position. Die Probanden sollten zu jedem der präsentierten Stimuli eine der vier Optionen auswählen: Einzelkonsonant (C_1/C_2^{49}), Cluster ($\langle C_1C_2 \rangle$), Sequenz mit /e/ ($\langle C_1eC_2 \rangle$) oder Sequenz mit /u/ ($\langle C_1uC_2 \rangle$). Für die /r/-Stimuli hatten die Probanden $\langle Cr \rangle$, $\langle C \rangle$, $\langle Cer \rangle$, $\langle Cur \rangle$ und für die Plosivcluster $\langle Ct \rangle$, $\langle t \rangle$, $\langle Cet \rangle$, $\langle Cut \rangle$ (wobei $C=[p, k]$) zur Auswahl. Die Antwortmöglichkeiten wurden orthographisch präsentiert. Es handelt sich also um ein sogenanntes *forced-choice*-Experiment mit einer Identifikationsaufgabe, d.h. der Proband musste eine der vier Optionen identifizieren und auswählen⁵⁰. Es wird keine Interaktion zwischen Hörer- und Sprechervarietät erwartet, da die ausgeschnittenen randomisierten Stimuli mit keinerlei Hinweisen auf deren regionale Herkunft eingeführt wurden.

Die Plosiv-Liquid-Stimuli wurden von 37 EP- (11m/26w) und 32 BP-Probanden (20m/12w) und die Plosiv-Stimuli von 40 EP- (13m/27w) und 37 BP-Probanden (23m/13w) beurteilt⁵¹. Alle Teilnehmer waren entweder Studenten der Unicamp oder der Universität Porto und waren im Alter zwischen 23 und 36 Jahren. Die Teilnehmer stammten aus den gleichen Städten wie beim Produktionsexperiment: aus Porto in Portugal und aus Campinas

⁴⁹ Ausgehend von den Ergebnissen von Cristófaró Silva (2000) wird C_1 bei den Plosiv-Liquid-Clustern und ausgehend von Chitoran et al. (2002) C_2 bei reinen Plosiv-Clustern erwartet.

⁵⁰ Bei diesem Aufgabentyp ist der Proband gezwungen, eine der angegebenen Optionen auszuwählen und es war nicht möglich, zu einem früheren Zeitpunkt bzw. Stimulus zurückzukehren und die Antwort zu revidieren.

⁵¹ Die Anzahl der Teilnehmer ist unterschiedlich, weil bei allen Teilnehmern, die das Experiment nicht zuende durchführten aber die beiden Blöcke mit den Plosivstimuli abschlossen, die Antworten zu den Plosiv-Token berücksichtigt wurden.

in Brasilien. Nach eigenen Angaben litten die Teilnehmer an keinen Sprach-, Hör- oder Lesestörungen. Das Experiment fand mit Kopfhörern und in einem ruhigen, geschlossenen Raum statt.

6.3.3 Statistische Auswertung

Die Antworten wurden mit einem *Generalized Linear Mixed Model* (GLMM) in R ausgewertet. Ein GLMM ist eine Art Regression, bei der die Probandenantworten (abhängige Variable) aus den möglichst besten Kombinationen der getesteten Faktoren eingeschätzt werden. Dieses Verfahren ist von zentraler Bedeutung, um für eine konkrete Stichprobe die Parameter herauszufinden, die mit der größten Wahrscheinlichkeit (*maximum likelihood*) für ihr Zustandekommen verantwortlich sind (Baayen, 2008: 263-321). Das Modell kombiniert zwei Arten von Faktoren: Feste (*fixed*) Faktoren werden geprüft, während die randomisierten Faktoren von der Analyse ausgeklammert werden. Letzteres ist wichtig, um die Variabilität der Teilnehmer (Sprecher oder Hörer) oder bestimmter Stimulustypen (z.B. wenn für jede geprüfte Kategorie mehrere Wörter vorhanden sind) herauszurechnen⁵². In der Analyse wurde der Probandenname als randomisierter Faktor einbezogen. Dies hatte den Vorteil, dass die Variabilität zwischen den Probanden innerhalb einer Varietät ausgeschlossen wurde. Die festen Faktoren waren Sprechervarietät und Stimulustyp, Artikulationsort von C₁, Betonung und Hörervarietät.

Bei diesem Verfahren wurde ein GLMM mit maximal drei Prüffaktoren erst mit Interaktion und dann ohne Interaktion der einzelnen Faktoren durchgeführt. Bei dem dreistufigen Faktor *Stimulustyp* (Cluster vs. <C₁et> vs. <C₁ut>) wurden die Stufen anders angeordnet, damit alle Abstände bei dem Binomialtest berücksichtigt werden konnten. Beide GLMM-Analysen wurden mit einer ANOVA (*analysis of variance*) verglichen, um festzustellen, welcher der GLMM-Tests (mit oder ohne Interaktion) die Ergebnisse besser modellierte. Im Falle einer Interaktion wurden

⁵² Zur GLMM siehe z.B. Baayen (2010), Delgaard (2008), Johnson (2011), Vasishth & Broe 2010.

zusätzliche *Tukey Post-Hoc*-Tests mit den kombinierten Faktoren durchgeführt, um die einzelnen Interaktionen zwischen den Faktoren herauszustellen.

Im Folgenden werden die Ergebnisse zu beiden Fragestellungen präsentiert. Zunächst wird der Frage nachgegangen, inwiefern Hörer den Stimulustyp identifizieren konnten (6.3.1) und danach der mögliche Zusammenhang zwischen Koordination und Perzeption (6.3.2) untersucht.

6.4 Ergebnisse

6.4.1 Identifikation des Stimulustyps

Bei dem Perzeptionsexperiment hatten die Probanden die Aufgabe, zu jedem präsentierten Stimulus eine von vier Antwortmöglichkeiten auszuwählen. Das erste Ziel des Experiments war es, zu untersuchen, inwiefern die Probanden den Stimulustyp der präsentierten Stimuli identifizieren konnten. Es geht hierbei nicht darum, ob der Sprecher tatsächlich einen intermediären Vokal produzierte, sondern welches Wort der Sprecher im Produktionsexperiment vorlesen sollte. Die Wörter wurden für die Auswertung in drei Typen eingeteilt: ein Wort mit lexikalischem Cluster (C_1C_2 , z.B. *prada* 'Prada'), ein Wort mit intermediärem /e/ ($\langle C_1eC_2 \rangle$, z.B. *perada* 'Birnenmus') oder ein Wort mit intermediärem /u/ ($\langle C_1uC_2 \rangle$, z.B. *puratos*, 'Mehlsorte').

Da Plosiv-Liquid-Token in wortinitialer Position und Plosiv-Token medial vorkamen, wurden beide Gruppen getrennt analysiert und die unterschiedlichen Artikulationsarten nicht berücksichtigt. Geprüft wurde die Sprechervarietät und der Stimulustyp sowie der Artikulationsort von C_1 , Betonung und Hörervarietät, außerdem die Interaktionen zwischen Stimulustyp und Artikulationsort von C_1 , zwischen Stimulustyp und Betonung und schließlich zwischen Stimulustyp und Hörervarietät für die getrennten Daten nach Artikulationsart (d.h. Plosiv-Liquid- vs. Plosiv-Plosiv-Token). Die anderen Interaktionen sind für dieses Verfahren nicht maßgebend und wurden daher ausgelassen.

6.4.1.1 Plosiv-Liquid-Token

Abbildung 16 zeigt den Anteil der Perceptionsantworten zu den Plosiv-Liquid-Token, die entweder mit den vom Sprecher intendierten Stimulustypen übereinstimmten (ja) oder dies nicht taten (nein). Die bei der Produktion intendierten Stimuli werden vertikal auf der y-Achse abgebildet. Dabei werden Cluster wie bei der Produktion mit einer intermediären Null (0) gekennzeichnet, die bedeutet, dass kein intermediärer Vokal vorkommt.

Die Antworten der Teilnehmer wurden proportional berechnet und durch Balken auf der horizontalen Skala wiedergegeben. In der linken Spalte werden die BP- und in der rechten Spalte die EP-Sprechdaten dargestellt. Im oberen Balkenpaar werden die betonten Token und im unteren die unbetonten Token abgebildet.

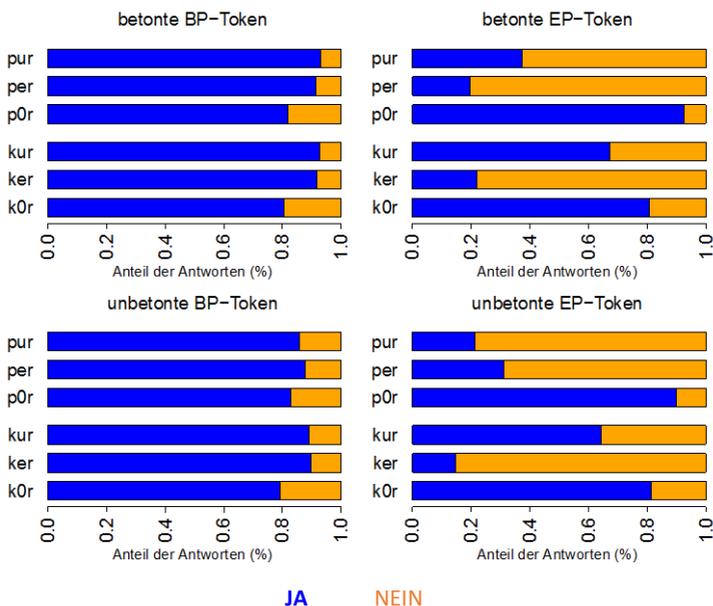


Abbildung 16: Übereinstimmung der entsprechenden Perceptionsantworten und intendierten Produktionsstimuli bei Plosiv-Liquid-Token (Antworten von 32 BP- und 37 EP-Hörern).

Die Antwortverteilung dieser Token wurde von der Sprechervarietät ($z=2.9$, $p<0.01$ ⁵³) und dem Stimulustyp bei Clustern vs. $\langle C_1er \rangle$ ($z=6.5$, $p<0.001$), Clustern vs. $\langle C_1ur \rangle$ ($z=6.1$, $p<0.001$), jedoch nicht bei den Sequenzen mit Vokal beeinflusst. Der Einfluss der Sprechervarietät ist an der deskriptiven Statistik deutlich, indem die Übereinstimmungen deutlich größer in den BP- als in den EP-Sprechdaten waren. Die Sprechervarietät hatte sogar eine größere Auswirkung auf beide CVC-Sequenzen ($p<0.001$ für $\langle C_1er \rangle$ und $\langle C_1ur \rangle$) als auf Cluster ($p<0.05$).

Die Unterschiede bezüglich Stimulustypen fielen deutlicher bei den EP- als bei den BP-Daten aus, wobei aber keine deutliche Tendenz zwischen Stimulustypen und Sprechervarietät vorhanden war: Cluster und $\langle C_1er \rangle$ zeigten beim BP den höheren Erkennungsanteil und $\langle C_1ur \rangle$ wurde häufiger beim EP als beim BP identifiziert. Diese Unterschiede zwischen Häufigkeiten der Stimulustypen und Varietäten spiegelte sich in der hochsignifikanten Interaktion zwischen den Faktoren ($\chi^2[7]=589.8$, $p<0.001$). Bei den EP-Daten beeinflusste jeder der Stimulustypen die Antwortverteilung ($p<0.001$ für alle drei Kombinationen), während sich bei den BP-Daten lediglich Cluster vs. $\langle C_1er \rangle$ ($p<0.001$) sowie Cluster vs. $\langle C_1ur \rangle$ ($p<0.001$) unterschieden, jedoch nicht für beide Vokale ($p=1.0$).

Bei den von Brasilianern produzierten Plosiv-Liquid-Token (links in der Graphik) war der Anteil der identifizierten Cluster (ja) kleiner als der Anteil der identifizierten CVC-Sequenzen ($\langle C_1er \rangle$ und $\langle C_1ur \rangle$), sowohl bei den bilabialen als auch bei den velaren C_1 und in betonter sowie in unbetonter Kondition. C_1 und Betonung hatten keinen signifikanten Einfluss auf die Antwortverteilung.

⁵³ Der in der Abbildung ersichtlich große Einfluss der Sprechervarietät kommt im präsentierten statistischen Test aufgrund der starken Interaktion geschwächt zum Ausdruck. Ohne die Berücksichtigung der Interaktion wäre der Einfluss der einzelnen Faktoren auf die Antwortverteilung deutlicher (Sprechervarietät: $z=34.1$, $p<0.001$ und Stimulustyp: $z=21.2$, $p<0.001$ für Cluster vs. $\langle C_1eC_2 \rangle$ und $z=11.9$, $p<0.001$ für Cluster vs. $\langle C_1uC_2 \rangle$).

Dies war allerdings nicht der Fall bei der Hörervarietät ($z=2.5$, $p<0.05$) sowie der Interaktion zwischen Stimulustyp und Hörervarietät ($\chi^2[7]=17.7$, $p<0.001$). Die Tukey Post-Hoc-Tests zeigten hier, dass die Stimulustypen Cluster und $\langle C_{1er} \rangle$ einen Einfluss auf die Antwortverteilung beider Hörergruppen hatten ($p<0.001$ bei BP- und $p<0.01$ bei EP-Hörern), während Cluster vs. $\langle C_{1ur} \rangle$ nur bei BP-Hörern ($z=7.1$, $p<0.001$), jedoch nicht bei EP-Hörern ($p=0.5$) Unterschiede bewirkten. EP-Hörer haben Cluster und $\langle C_{1ur} \rangle$ zu ähnlichen Anteilen identifiziert.

Der C_1 -Artikulationsort war signifikant bei den Antworten zu den EP-Token ($z=5.8$, $p<0.001$). Wie aus Abbildung 17 ersichtlich wird, war bei den europäischen Token (rechts in der Graphik) der Anteil der Übereinstimmungen am größten bei Clustern, gefolgt von $\langle C_{1ur} \rangle$ und am geringsten bei $\langle C_{1er} \rangle$. Diese Tendenz zeigte sich für alle betonten Token, aber nicht für den bilabialen C_1 in unbetonter Position, wobei das intermediäre /u/ weniger Übereinstimmungen als das intermediäre /e/ aufwies. Dies erklärt die signifikante Interaktion zwischen Stimulustyp und C_1 ($\chi^2[7]=170.0$, $p<0.001$). Tukey Post-Hoc-Tests zeigten einen größeren Einfluss von C_1 bei $\langle C_{1ur} \rangle$ ($p<0.001$) und bei Clustern ($p<0.001$) als bei $\langle C_{1er} \rangle$ ($p<0.05$). Die Betonung an sich war wenig aussagekräftig, aber ihre Interaktion mit dem Stimulustyp ergab einen kleinen Effekt bei $\langle C_{1ur} \rangle$ -Token ($p<0.05$), aber nicht bei den anderen. Die Hörervarietät beeinflusste leicht die Antwortverteilung ($z=2.2$, $p<0.05$) und zwar bei der Wahrnehmung von $\langle C_{1er} \rangle$ ($p<0.5$), jedoch nicht bei $\langle C_{1ur} \rangle$ ($p=0.9$) und Clustern ($p=0.2$).

6.4.1.2 Plosiv-Token

Bei den Plosiv-Token zeigte sich tendenziell ein größerer Einfluss der einzelnen Faktoren: Die Sprechervarietät beeinflusste die Antwortverteilung ($z=8.2$, $p<0.001$), indem BP-Token insgesamt häufiger identifiziert werden konnten als EP-Token.

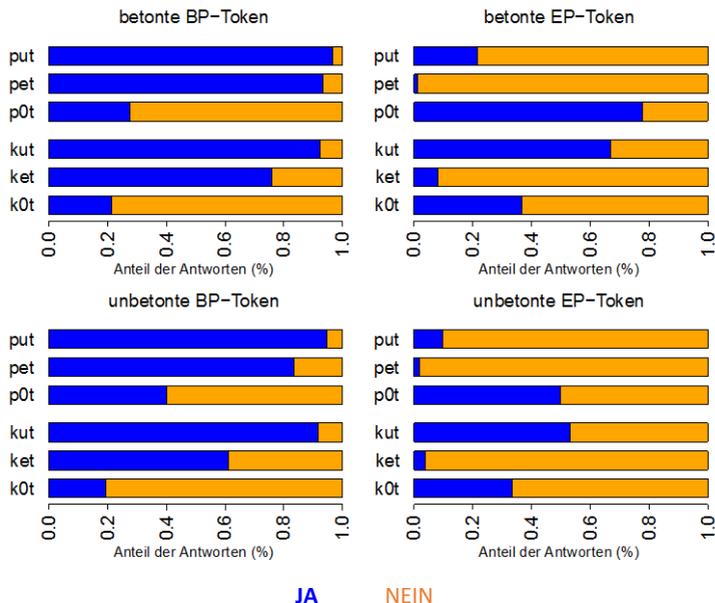


Abbildung 17: Anteil der entsprechenden Perzeptionsantworten und intendierten Produktionsstimuli bei reinen Plosiv-Token (Antworten von 37 BP- und 40 EP-Hörern).

Der Stimulustyp zeigte ebenso einen signifikanten Effekt, der bei Clustern vs. $\langle C_{1ut} \rangle$ ($z=20.1$, $p<0.001$) und Clustern vs. $\langle C_{1et} \rangle$ ($z=18.1$, $p<0.001$) deutlich wurde und etwas geringer zwischen beiden Sequenzen ($\langle C_{1et} \rangle$ vs. $\langle C_{1ut} \rangle$: $z=7.5$, $p<0.001$) ausfiel. Bei den BP-Token war sein Einfluss größer bei $\langle C_{1ut} \rangle$ vs. Cluster ($z=16.1$, $p<0.001$) als bei $\langle C_{1et} \rangle$ vs. Cluster ($z=12.9$, $p<0.001$). Der Einfluss von $\langle C_{1et} \rangle$ vs. $\langle C_{1ut} \rangle$ auf die Verteilung war kleiner, aber ebenso signifikant ($z=7.2$, $p<0.001$).

Bei den EP-Token (rechts in der Graphik) hatte der Stimulustyp insgesamt einen geringeren Einfluss, und in einer anderen Reihenfolge: $\langle C_{1et} \rangle$ vs. Cluster: $z=8.3$, $p<0.001$, $\langle C_{1et} \rangle$ vs. $\langle C_{1ut} \rangle$: $z=7.5$, $p<0.001$ und $\langle C_{1ut} \rangle$ vs. Cluster: $z=7.0$, $p<0.001$.

Diese Diskrepanzen zwischen Sprechervarietäten und Stimulustyp werden in der hochsignifikanten Interaktion zwischen beiden Faktoren wiedergegeben ($\chi^2[7]=890.7$, $p<0.001$).

Bei den BP-Daten hatte der Artikulationsort einen signifikanten Einfluss auf die Antwortverteilung ($=4.2$, $p<0.001$), aber die Identifikation des Stimulustyps wies die gleiche Tendenz nach bilabialen wie nach velaren C_1 auf, weshalb keine Interaktion zwischen Stimulustyp und C_1 ($p=0.06$) vorhanden war. Die Betonungskondition beeinflusste die Verteilung nicht, zeigte allerdings eine signifikante Interaktion mit dem Stimulustyp ($\chi^2[7]=13.0$, $p<0.01$), ebenso wie die Hörervarietät ($\chi^2[7]=11.1$, $p<0.01$). $\langle C_{1et} \rangle$ wurde schlechter in der unbetonten als in der betonten Kondition identifiziert. Allerdings war der Unterschied zwischen Clustern und $\langle C_{1et} \rangle$ sowie Clustern und $\langle C_{1ut} \rangle$ in beiden Betonungskonditionen signifikant ($p<0.001$ für beide Paare in beiden Betonungskonditionen). Die Wahrnehmung der BP-Token unterschied sich lediglich bei $\langle C_{1et} \rangle$ nach Hörergruppen ($p<0.05$), indem EP-Hörer weniger Token identifizierten als BP-Hörer.

C_1 beeinflusste die Antworten zu den EP-Token ($z=7.8$, $p<0.001$) und es gab dabei eine Interaktion zwischen Stimulustyp und C_1 , so dass bilabiale Cluster ($p<0.001$) und $\langle C_{1ut} \rangle$ ($p<0.001$) häufiger identifiziert wurden als velare. Die Betonung zeigte ebenso bei den EP-Daten einen signifikanten Effekt ($z=4.6$, $p<0.001$). Die Hörervarietät zeigte keine Effekte, aber eine signifikante Interaktion zwischen Stimulustyp und Hörervarietät ($\chi^2[7]=8.8$, $p<0.01$) kam zustande, da BP-Hörer Cluster und $\langle C_{1ut} \rangle$ deutlich besser unterschieden ($p<0.05$) als EP-Hörer ($p=0.3$).

6.4.1.3 Zwischenfazit

So bringen die Ergebnisse zum Ausdruck, dass Plosiv-Liquid-Token durch eine starke Interaktion zwischen Sprechervarietät und Stimulustyp modelliert wurden, was bedeutet, dass die Sprechervarietät nicht dieselbe Auswirkung auf die Wahrnehmung der einzelnen Stimulustypen hatte: Die brasilianischen Token

konnten relativ gut identifiziert werden, allerdings hatten $\langle C_{1er} \rangle$ und $\langle C_{1ur} \rangle$ ähnliche Anteile an Übereinstimmungen, während Cluster im Vergleich häufiger verwechselt wurden. Die Antworten wurden weder vom Artikulationsort noch von der Betonung beeinflusst, und die Hörervarietät zeigte, dass EP-Hörer Cluster und $\langle C_{1er} \rangle$ besser identifizierten als BP-Hörer, jedoch nicht $\langle C_{1ur} \rangle$, bei dem sich die Hörergruppen nicht unterschieden. Bei den europäischen Plosiv-Liquid-Token war die Tendenz genau entgegengesetzt: Cluster wurden mit den größten Anteilen als Cluster erkannt, während $\langle C_{1er} \rangle$ die größte Diskrepanz zwischen intendierten Stimuli und Perzeptionsantworten zeigte. Der C_1 -Artikulationsort hatte einen signifikanten Effekt auf $\langle kur \rangle$ und die Betonung auf $\langle C_{1ut} \rangle$, aber nicht auf die anderen Token. Die Hörervarietät zeigte, dass EP-Hörer $\langle C_{1er} \rangle$ besser identifizierten als BP-Hörer und sich beide Hörergruppen bei den weiteren Token nicht unterschieden.

Token mit zwei Plosiven wurden durch eine noch stärkere Interaktion zwischen Sprechervarietät und Stimulustyp gekennzeichnet: Die brasilianischen Sequenzen konnten relativ gut identifiziert werden und Cluster deutlich schlechter, weshalb Cluster vs. $\langle C_{1et} \rangle$ und Cluster vs. $\langle C_{1ut} \rangle$ einen deutlichen Einfluss auf die Antwortverteilung hatten. Die Antworten wurden von C_1 und der Hörervarietät beeinflusst, indem Token mit dem bilabialen C_1 häufiger identifiziert wurden als mit dem velaren. Die Hörergruppen unterschieden sich bei der Wahrnehmung von $\langle C_{1et} \rangle$, wobei EP-Hörer weniger Token identifizierten als BP-Hörer. Die Betonung übte einen Einfluss auf $\langle C_{1et} \rangle$ aus, jedoch nicht auf die weiteren Stimulustypen.

Bei den EP-Daten war die Tendenz im Vergleich zu den BP-Daten genau entgegengesetzt: Cluster wurden mit den größten Anteilen als reine Plosive identifiziert. Der Identifikationsanteil der Cluster war deutlich niedriger als bei den Plosiv-Liquid-Token. $\langle C_{1er} \rangle$ zeigte eine so große Diskrepanz zwischen intendierten Stimuli und Perzeptionsantworten, dass weniger als 10% der Token identifiziert wurden. C_1 und die Betonung beeinflussten die Antwortverteilung, indem der Stimulustyp in der betonten

Kondition und nach dem velaren C_1 etwas besser identifiziert werden konnten als in der unbetonten Kondition und nach dem bilabialen C_1 . Cluster und $\langle C_1ut \rangle$ wurden häufiger von BP- als von EP-Hörern identifiziert. Womit die Stimulustypen verwechselt wurden und inwiefern dies mit ihrer Koordination zusammenhängt, ist das Thema des nächsten Unterpunkts.

1.1.1 Wahrnehmung von Koordinationsunterschieden

Die Perzeptionsantworten wurden des Weiteren in Bezug auf die intendierten Stimuli abgebildet, und zwar im Hinblick auf die Anzahl der wahrgenommenen Elemente, d.h. ob eins bei einzelnen Konsonanten (C), zwei bei Clustern (CC) und drei bei CVC-Sequenzen (CVC) wahrgenommen wurden. Dies wurde in Verbindung mit den Koordinationsunterschieden in den vorherigen zwei Kapiteln gebracht. Die Perzeptionsantworten wurden zu den Produktionstypen für die getrennten Artikulationsarten in Abhängigkeit von Sprechervarietät, Artikulationsort von C_1 , Betonungskondition sowie Hörervarietät ausgewertet.

Plosiv-Liquid-Cluster

Bei den entsprechenden Blöcken des Perzeptionsexperiments entschieden die Teilnehmer, ob sie zum jeweiligen Stimulus einen Konsonanten ($\langle C_1 \rangle$), ein Cluster ($\langle C_1r \rangle$), oder eine CVC-Sequenz mit /e/ oder mit /u/ ($\langle C_1er \rangle$, $\langle C_1ur \rangle$), (wobei $C=[p, k]$) hörten. Aus dem Gesamtbild geht hervor, dass bei den Plosiv-Liquid-Clustern keine einzelnen Konsonanten perzipiert und für über 80% der Stimuli zwei Elemente wahrgenommen wurden. Da unser statistisches Modell die Antworten paarweise auswertet und an dieser Stelle der wahrgenommene Vokal nicht maßgebend ist, wurden für die weitere Analyse die zwei Antwortmöglichkeiten mit Vokal ($\langle C_1er \rangle$, $\langle C_1ur \rangle$) zu einer Kategorie zusammengefügt (CVC). Daher prüfen wir im Folgenden, inwiefern Hörer zweigliedrige (CC) oder dreigliedrige Elemente (CVC) wahrnahmen.

Die Gesamtergebnisse wurden in Abbildung 18 wie folgt dargestellt: Die bei der Produktion intendierten Stimuli wurden vertikal auf der y-Achse und die Perzeptionsantworten in einer proportionalen Skala horizontal auf der x-Achse abgebildet. In der linken Spalte werden die BP- und in der rechten Spalte die EP-Sprechdaten dargestellt. Das obere Balkenpaar zeigt die betonten Token und das untere Balkenpaar die unbetonten.

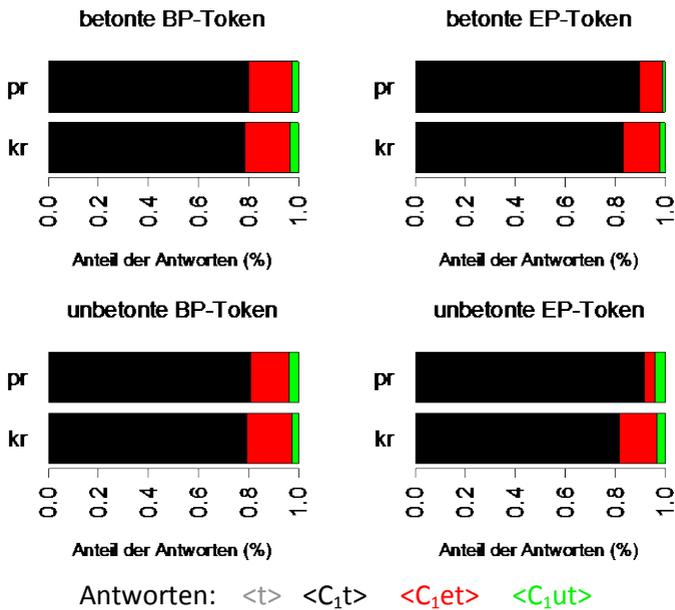


Abbildung 18: Antworten zu den produzierten Plosiv-Liquid-Clustern (anteilig)

Die deskriptive Statistik zeigte einen größeren Anteil von zweigliedrigen Elementen (<C₁r>) bei EP- als bei BP-Token und bei /pt/ als bei /kt/-Token. Die GLMM-Analyse ergab Haupteffekte von Sprechervarietät ($z=5.0$, $p<0.001$) und Artikulationsort ($z=3.7$, $p<0.001$). Die Interaktion zwischen Sprechervarietät und Artikulationsort war ebenfalls signifikant ($\chi^2[4]=11.3$, $p<0.01$) und die Tukey Post-Hoc-Tests zeigten, dass

die Sprechervarietät einen signifikanten Einfluss bei /pt/ sowohl in betonter ($z=4.0$, $p<0.01$) als auch in unbetonter ($z=3.7$, $p<0.01$) Kondition, jedoch nicht bei /kt/-Clustern hatte.

Obwohl die für die Produktions- und Perzeptionsexperimente verwendeten Wörter teilweise voneinander abwichen⁵⁴, wurden starke Parallelen zwischen den Typen beider Datensets festgestellt. Trotz der großen Variabilität zwischen Sprechern in der Produktion zeigten die Ergebnisse in Kapitel 4 eine weitere Koordinierung der Gesten im BP in der Form, dass C_2 nach Ende des C_1 -Plateaus bei beiden /r/-Clustern begann. Der so entstandene Abstand zwischen beiden Konsonanten in der Produktion könnte der Grund sein, weshalb in der Perzeption 20% der in der Produktion intendierten Cluster mit einem intermediären Vokal wahrgenommen wurden. Dennoch wurden über 80% der intendierten Cluster auch als $\langle C_1r \rangle$ (zweigliedriges Element) wahrgenommen.

Die Übereinstimmung zwischen Produktion und Perzeption ist allerdings bei den europäischen Token noch eindeutiger: Bei den in Kapitel 4 analysierten Stimuli begann C_2 von /pr/ während des C_1 -Plateaus und somit waren *front-to-back*-Cluster (/pr/) enger koordiniert als /kr/. Bei Letzteren wurde wie bei den BP-Clustern in der Produktion ein Abstand zwischen beiden Konsonanten festgestellt. Dies kann eindeutig mit dem größeren Anteil der $\langle C_1r \rangle$ -Antworten bei den EP-/pr/-Clustern in Verbindung gebracht werden. Damit zeigte der C_1 -Artikulationsort (velar vs. labial) einen signifikanten Einfluss auf die Antworten bei den europäischen, jedoch nicht bei den brasilianischen Token. So zeigte die Prüfstatistik, dass sich der Anteil wahrgenommener zweigliedriger Elemente ($\langle C_1r \rangle$) lediglich bei den bilabialen /pr/-Clustern im EP sowohl in der Produktion als auch in der

⁵⁴ In den Produktions- und Perzeptionskapiteln werden die gleichen Cluster und CVC-Sequenzen analysiert. Allerdings stimmen die Wörter, in die die r-Stimuli eingebaut wurden, bei Clustern und $\langle C_1er \rangle$ nicht überein. Vgl. Methode des Kapitels 4 (4.2).

Perzeption unterschied. Alle anderen Kombinationen trugen zu keinen Unterschieden in der Verteilung von korrekten CC- und CVC-Antworten bei.

Da sich die Perzeption bei den Produktionsrichtungen an Koordinationsunterschieden zu orientieren schien, scheinen die Probanden gut zu dem Produktionsrichtungseffekt zu kompensieren. Der Grund hierfür ist möglicherweise, dass Liquide andere Konsonanten nicht stark maskieren (Mattingly 1981) und die Hörer daher Unterschiede in der Konsonantenkoordination mit kategorialen Unterschieden assoziieren.

CVC-Sequenzen mit Liquida

In Abbildung 19 wurden die intendierten CVC-Sequenzen vertikal auf der y-Achse und die Perzeptionsantworten in einer proportionalen Skala horizontal auf der x-Achse in der gleichen Reihenfolge wie in der vorherigen Abbildungen dargestellt. Da die Hörer auch zu den intendierten CVC-Sequenzen nur drei der vier Möglichkeiten auswählten (einzelner Konsonant wurde nur einmal angeklickt), wurden die zwei Antwortmöglichkeiten mit Vokal ($\langle C_{1er} \rangle$, $\langle C_{1ur} \rangle$) zusammengefügt (CVC) und ausgewertet, ob Probanden zwei- (CC) oder dreigliedrige (CVC) Elemente wahrnahmen.

Obwohl beide Sprecher bei den Aufnahmen die gleichen Stimuli mit den orthographischen Vokalen $\langle e, u \rangle$ als Grundlage hatten, macht die Abbildung deutlich, dass bei den BP-Token (links in der Graphik) die intendierten CVC-Sequenzen auch entsprechend wahrgenommen wurden (mit Vokal), während bei den EP-Token (rechts in der Graphik) die zweigliedrigen Konsonantencluster (C_{1r}) überwogen. Auch die GLMM-Analyse ergab einen hochsignifikanten Einfluss der Sprechervarietät ($z=29.1$, $p<0.001$) auf die Antwortverteilung von CC und CVC.

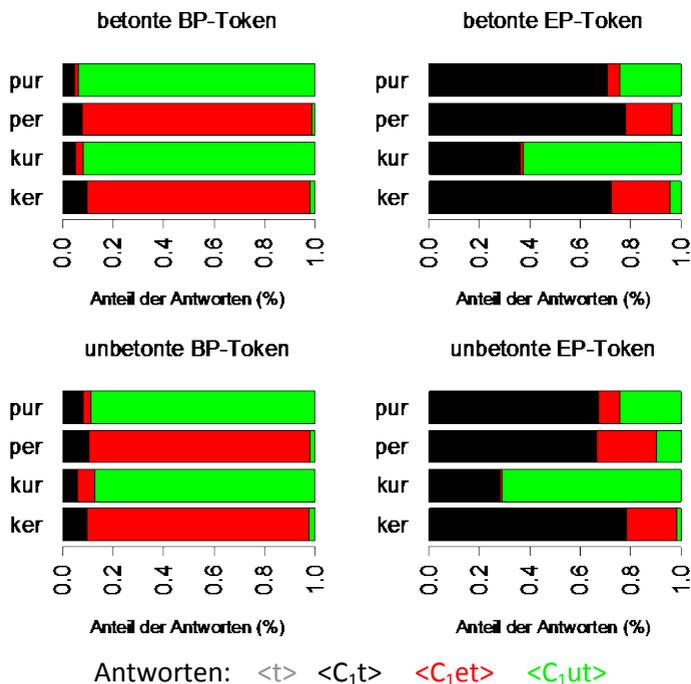


Abbildung 19: Anteilige Antworten zu den produzierten CVC-Sequenzen mit Liquid

Die Prüfstatistik zeigte außerdem einen leichten Einfluss der Betonung ($z=3.0$, $p<0.01$), der allerdings nicht eindeutig zu erkennen ist. C_1 zeigte keinen Einfluss ($p=0.2$), was aufgrund der großen Divergenzen zwischen <ker> und <kur> zustande kam und sich durch die folgenden Interaktionen modellieren lässt: Es gab hochsignifikante Interaktionen der drei Faktoren ($\chi^2[9]=45.3$, $p<0.01$) und die Tukey Post-Hoc-Tests zeigten einen hochsignifikanten Einfluss des Artikulationsorts auf die Antworten zu den EP-Token in betonter ($p<0.001$) und unbetonter Kondition ($p<0.01$). Bei den BP-Stimuli war dies nicht der Fall. Die Sprechervarietät hatte ebenfalls einen hochsignifikanten Einfluss auf beide Betonungsbedingungen

($p < 0.001$) und beide Artikulationsorte ($p < 0.001$). Die Betonung war beim bilabialen Artikulationsort in beiden Varietäten signifikant ($p < 0.05$ im BP und $p < 0.001$ im EP), jedoch nicht beim velaren Artikulationsort ($p = 1$). Auch bei den CVC-Sequenzen waren trotz unterschiedlicher Wörter in den Produktions- und Perzeptionsdaten gewisse Parallelen erkennbar: Bei der Synchronisierung der Konsonanten in den vom BP-Sprecher produzierten CVC-Sequenzen begann C_2 später, nach Ende des C_1 -Plateaus, und somit fand keine Überlappung bzw. eine weite Synchronisierung beider Konsonantengesten statt (siehe Diskussion in Kapitel 4). Dies passt zur Wahrnehmung eines intermediären Vokals in über 90% der Stimuli im Perzeptionsexperiment.

Bei den EP-Sprechdaten zeigten die Sequenzen mit hinterem Vokal ($\langle C_1ur \rangle$) größere negative Werte oder kleinere positive Werte als bei $\langle C_1er \rangle$, was bedeutet, dass bei ersteren weniger Überlappung auftrat. In der Perzeption nahmen Hörer bei $\langle kur \rangle$ einen größeren Anteil von Vokalen wahr, bei $\langle pur \rangle$ allerdings nicht, obwohl sie in der Produktion eine ähnliche Synchronisierung aufwiesen. So stellt sich heraus, dass die Koordination allein die Konsonantenwahrnehmung von Plosiv-Liquid-Token nicht erklären kann. Der C_1 -Artikulationsort hatte in der Produktion keinen Einfluss auf die Konsonantensynchronisierung. Bei der Wahrnehmung ist allerdings unklar, ob der hintere Vokal oder die Produktionsrichtung ausschlaggebend bei $\langle kur \rangle$ war. Die Betonungskondition hatte sowohl bei der Produktion als auch bei der Perzeption einen marginalen Einfluss auf die Antwortverteilung und die Hörervarietät spielte keine Rolle bei der Antwortverteilung bei Plosiv-Liquid-Token ($p = 0.2$).

Das Experiment zeigte, dass in den brasilianischen Sprechdaten Cluster als zweigliedrige und CVC-Sequenzen als dreigliedrige Elemente wahrgenommen werden konnten. Dies war allerdings bei den europäischen Stimuli nicht der Fall, da sowohl intendierte Cluster als auch intendierte CVC-Sequenzen (außer $\langle kur \rangle$) als zweigliedrige Elemente wahrgenommen wurden. Der Grund dafür konnte teilweise in der Koordination beider Konsonanten in

der Produktion gesehen werden. Bei <kur> war unklar, ob das abweichende Antwortverhalten mit dem hinteren Vokal, mit dem velaren Konsonanten oder mit beiden zusammenhängt. Für die Modellierungen feinerer Unterschiede zwischen Produktion und Perzeption bei den Plosiv-Liquid-Clustern sind allerdings weitere Untersuchungen nötig. Diese Fragen werden im Folgenden für die reinen Plosivcluster zum Untersuchungsgegenstand.

Plosivcluster

Die Gesamtergebnisse der Plosivcluster werden analog zu den vorherigen Abbildungen dargestellt. Bei den entsprechenden zwei Blöcken des Perzeptionsexperiments mussten die Teilnehmer entscheiden, ob sie den jeweiligen Stimulus als Einzelkonsonanten (<t>), Cluster (<C₁t>), <C₁et> oder <C₁ut> wahrnahmen.

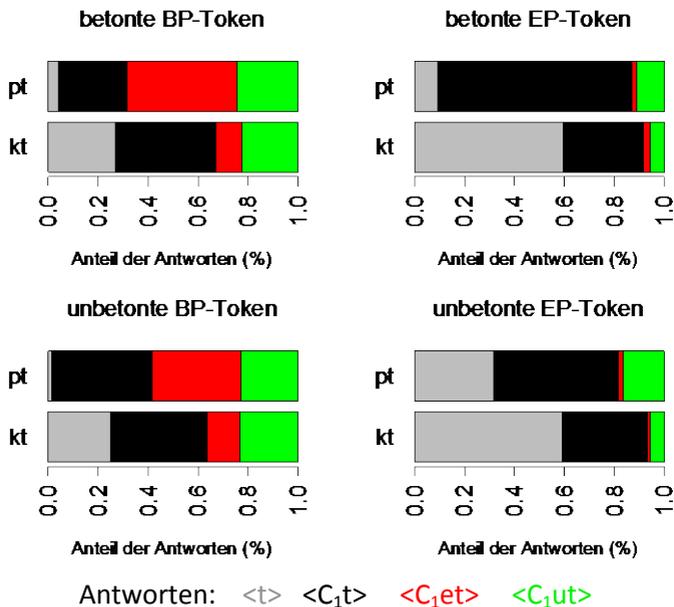


Abbildung 20: Anteilige Antworten zu den intendierten reinen Plosivclustern.

Die Ergebnisse zeigten deutlich mehr Wahrnehmungen von Einzelkonsonanten bei den EP-Stimuli als bei den BP-Stimuli. Dies trifft auch auf die velaren C_1 im Vergleich zu den bilabialen C_1 zu. Bei $\langle C_1et \rangle$ und $\langle C_1ut \rangle$ ist das Bild genau umgekehrt: Es wurden deutlich mehr Vokale bei den BP- als bei den EP-Stimuli und bei den bilabialen als bei den velaren C_1 wahrgenommen. Da unser statistisches Modell die Antworten paarweise auswertet, wurden für die weitere Analyse die vier Antwortmöglichkeiten $\langle t \rangle$, $\langle C_1t \rangle$, $\langle C_1et \rangle$ oder $\langle C_1ut \rangle$ (wobei $C = [p, k]$) wie folgt in zwei Gruppen eingeteilt:

- (i) $\langle t \rangle$ vs. alle anderen Antwortmöglichkeiten. Mit dieser Unterteilung kann man überprüfen, inwiefern die Probanden eher einen oder mehrere Konsonanten (abhängige Variable) in Abhängigkeit von Sprechervarietät, Artikulations-ort von C_1 , Primärbetonung und Hörervarietät wahrnahmen.
- (ii) Im zweiten Schritt wurden die Sequenzen mit ($\langle C_1et \rangle$, $\langle C_1ut \rangle$) vs. ohne Vokal ($\langle t \rangle$, $\langle C_1t \rangle$) untersucht. Somit kann überprüft werden, ob sich der Anteil der wahrgenommenen intermediären Vokale (abhängige Variable) unterschied.

/t/ vs. andere Antwortmöglichkeiten

In diesem Abschnitt werden die /t/-Antworten allen anderen gegenübergestellt. Insgesamt zeigten die Daten einen signifikanten Einfluss der Sprechervarietät ($z=9.0$, $p<0.001$) in der Form, dass bei den EP-Stimuli mehr einzelne Konsonanten als bei den BP-Stimuli perzipiert wurden. Der C_1 -Artikulationsort hatte ebenfalls einen signifikanten Einfluss auf die Antwortverteilung ($z=5.3$, $p<0.001$): Bei velaren Token im Vergleich zu bilabialen Token wurden mehr einzelne Konsonanten wahrgenommen. Die Betonungskondition hatte keine Auswirkung, aber es gab eine hochsignifikante Interaktion zwischen den Faktoren ($\chi^2[4]=32.0$, $p<0.001$). Dabei ergaben die Tukey Post-Hoc-Tests einen

signifikanten Einfluss der Betonung auf bilabiale Cluster im EP ($p < 0.001$), jedoch nicht auf velare Cluster.

Der C_1 -Artikulationsort hatte einen signifikanten Einfluss auf die Antwortverteilung in beiden Betonungsbedingungen (jeweils $p < 0.001$) und in beiden Varietäten (jeweils $p < 0.001$). In der Produktion dieser Stimuli (siehe Kapitel 3) hatte sich gezeigt, dass bei velaren Clustern C_2 bereits vor dem C_1 -Plateau begann und somit eine größere Überlappung beider Konsonantengesten stattfand bzw. eine engere Koordinierung beider Konsonanten. Bei einer starken Überlappung wird die Verschlusslösung von C_1 von C_2 verdeckt und das ist möglicherweise der Grund dafür, weshalb Hörer lediglich C_2 wahrnehmen. Außerdem stellten Chitoran et al. (2002) fest, dass Plosive in der *back-to-front* Produktionsrichtung für den Hörer schwerer erkennbar sind als in der umgekehrten Richtung.

Die Sprechervarietät hatte ebenso einen signifikanten Einfluss auf die Antwortverteilung für beide Artikulationsorte und dies kann mit der Konsonantenkoordinierung in der Produktion in Verbindung gebracht werden. C_2 begann entweder während oder nach dem C_1 -Plateau im BP, weshalb BP insgesamt eine signifikant weitere Koordinierung als EP aufwies. Bei EP begann C_2 vor oder zu Beginn des C_1 -Plateaus, weshalb sich beide Gesten stark überlappten. Die parallelen Perzeptionsantworten zeigten, dass bei manchen der eng koordinierten Cluster im EP die Hörer nicht mehr in der Lage waren, zwei, sondern lediglich einen Konsonanten wahrzunehmen.

Der Einfluss der Hörervarietät wurde für die Sprechdaten getrennt nach BP vs. EP überprüft. Für die brasilianischen Produktionsstimuli zeigten GLMM-Tests keinen signifikanten Einfluss der Hörergruppen auf die Wahrnehmung der brasilianischen /t/-Stimuli und keine signifikanten Interaktionen. Für die europäischen Produktionsstimuli wurde die Antwortverteilung von /t/ von der Hörervarietät beeinflusst ($z = 2.7$, $p < 0.01$) und es gab eine signifikante Interaktionen

zwischen Hörervarietät, Artikulationsort und Betonung ($\chi^2[4] = 26.5$, $p < 0.001$). Die Tukey Post-Hoc-Tests zeigten, dass die Variable Hörervarietät lediglich bei der Wahrnehmung unbetonter /kt/-Cluster signifikant ausfiel ($p < 0.05$), wobei EP-Hörer mehr einzelne Konsonanten wahrnahmen als BP-Hörer.

Anhand der geprüften Faktoren wurde ersichtlich, dass die Verteilung der /t/-Antworten am besten durch die Interaktion von Sprechervarietät und C₁-Artikulationsort modelliert werden konnte. Dabei ergab sich eine Reihenfolge, bei der /kt/ aus den EP-Sprechdaten am häufigsten in der Perzeption als einzelner Konsonant wahrgenommen wurde, gefolgt von /kt/ im BP, /pt/ im EP und schließlich /pt/ im BP mit dem geringsten Anteil an einzelnen Konsonanten. Die Rolle der Betonung blieb eher marginal. Bei der ausgeprägten Gewichtung der Sprechervarietäten in diesen Daten war zu erwarten, dass die unterschiedliche Produktion der Plosivcluster zu Diskrepanzen in der Wahrnehmung führen würde. Diese ergaben sich allerdings nicht in der erwarteten Richtung, wonach wir mehr Einzelkonsonantenanteile im BP im Vergleich zu EP-Hörern erwartet hatten.

/CVC/ vs. /C(C)/

Im Folgenden werden die Antworten mit Vokalen (/C₁et/, /C₁ut/) den Antworten ohne Vokal (/t/ und /C₁t/) gegenübergestellt. Die Ergebnisse der GLMM-Tests mit Sprechervarietät, Artikulationsort und Betonung als Prüf-Faktoren zeigten insgesamt einen signifikanten Einfluss der Sprechervarietät ($z = 7.3$, $p < 0.001$) und des Artikulationsortes von C₁ ($z = 7.5$, $p < 0.001$), aber nicht der Betonung auf die Verteilung der dreigliedrigen Elemente (CVC) bei der Wahrnehmung. Dies ist in Abbildung 20 ersichtlich: Der Anteil der dreigliedrigen Antworten (/C₁et/ und /C₁ut/) ist größer in den BP- als in den EP-Daten und bei bilabialen als bei velarem Konsonanten.

Bei der Synchronisierung begann C₂ bei Clustern mit Bilabialen im BP nach Ende des C₁-Plateaus. Somit überlappten sich beide Konsonantengesten nicht, sondern es gab einen Abstand zwischen den Konsonanten des Clusters. Dieser wurde erkennbar mit einem Vokal in der Perzeption assoziiert. Ob dabei auch die bessere perzeptive Wiederherstellbarkeit von Plosivclustern in dieser Produktionsrichtung (vgl. Chitoran et al. 2002) eine Rolle spielte, konnte nicht eindeutig festgestellt werden, da diese Token in der Produktion ebenfalls mehr überlappten. Es gab eine signifikante Interaktion zwischen den Faktoren ($\chi^2[4]=9.4$, $p<0.01$), die zwar durch die unterschiedliche Stärke des Einflusses gekennzeichnet war, aber in allen Kombinationen über dem Schwellwert ($p<0.001$) lag. Die Hörergruppen zeigten keinen Einfluss auf die Antwortverteilung in beiden Varietäten.

Folgen mit Plosiv

Auch hier wurden die Daten parallel zu den zuletzt gezeigten abgebildet und die Analyse in zwei Schritten durchgeführt:

- (i) <t> vs. alle anderen
- (ii) CVC vs. C(C).

Obwohl beide Versuchspersonen dieselben Token als Sprechvorlage hatten, ist in Abbildung 21 offensichtlich, dass bei den BP-Stimuli der Anteil der wahrgenommenen Vokale viel größer war als bei den EP-Stimuli. Bei letzteren überwog die Wahrnehmung von einzelnen Konsonanten und Clustern.

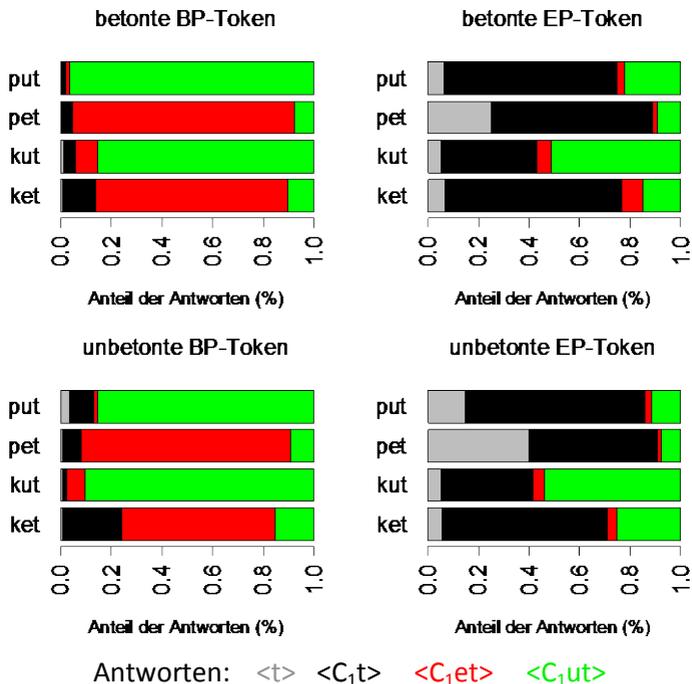


Abbildung 21: Anteilige Antworten zu den intendierten Plosiv-CVC-Sequenzen

/t/ vs. andere Antwortmöglichkeiten

Im Folgenden werden die /t/-Antworten gegenüber allen anderen analysiert. Aus Abbildung 21 wird ersichtlich, dass einzelne Konsonanten im EP bei allen Stimulustypen vertreten waren, während sie im BP ausschließlich bei /put/ vorkamen. Außerdem hatten sie einen größeren Anteil in der unbetonten als in der betonten Kondition. Die Prüfstatistik zeigte entsprechend einen signifikanten Einfluss der Sprechervarietät ($z=4.8$, $p<0.001$) und der Betonung ($z=2.4$, $p<0.05$) auf die Wahrnehmung von einzelnen Konsonanten.

Der C₁-Artikulationsort ergab insgesamt zwar keinen signifikanten Einfluss, die Post-Hoc-Tests zeigten jedoch, dass der

Artikulationsort einen signifikanten Einfluss auf die Antwortverteilung von /t/ bei den europäischen ($p < 0.001$), obgleich nicht bei den brasilianischen Daten hatte. Dies ist in Abbildung 21 daran zu erkennen, dass bilabiale Cluster einen deutlich größeren Anteil von einzelnen Konsonanten haben als velare C_1 . In der Produktion begann C_2 bei ersteren vor oder während des C_1 -Plateaus und somit zeigten sie eine engere Überlappung als velare Token. Die Wahrnehmung dieser Token scheint sich an ihrer Synchronisierung zu orientieren und der Produktionsrichtungseffekt (Chitoran et al. 2002) eine geringere Rolle zu spielen.

Die Betonung hatte einen Einfluss auf die bilabialen EP-Token ($p < 0.001$), aber weder auf die velaren EP-Token noch auf die velaren BP-Token. Die Hörervarietät zeigte sowohl einen signifikanten Einfluss auf ($z = 2.4$, $p < 0.05$) als auch eine signifikante Interaktion mit der Sprechervarietät ($\chi^2[5] = 6.6$, $p < 0.05$). Dies kommt darin zum Ausdruck, dass Hörer sich bei den BP-Stimuli unterschieden ($p < 0.05$), aber nicht bei den EP-Stimuli.

CVC vs. C(C)

Im Folgenden werden die Antworten mit Vokalen, d.h. / C_1et / und / C_1ut / behandelt, und diejenigen ohne Vokal, d.h. /t/ und / C_1t / gegenübergestellt. Abbildung 21 lässt erkennen, dass der Anteil der wahrgenommenen dreigliedrigen Elemente bei den BP-Daten sehr hoch (zwischen 78-97%), bei den EP-Daten hingegen mit der Ausnahme von <kut> sehr gering war (zwischen 10-30%). Die ist der Grund, weshalb die Sprechervarietät ($z = 20.0$, $p < 0.001$) einen hochsignifikanten Einfluss auf die Antwortverteilung hatte. Auch dieser Effekt war in den Produktionsdaten ersichtlich, da intendierte CVC-Sequenzen keine Überlappung im BP und eine starke Überlappung im EP aufwiesen (siehe 5.3 dieser Arbeit). Die CVC-Verteilung wurde außerdem noch vom Artikulationsort ($z = 3.4$, $p < 0.001$) und der Betonung ($z = 4.7$, $p < 0.001$) beeinflusst. Die Post-Hoc-Tests

belegen, dass der Einfluss des Artikulationsortes bei den EP-Daten signifikant ($p < 0.001$ in beiden Betonungsbedingungen) war, bei den BP-Daten aber lediglich in der betonten Kondition ($p < 0.05$). Die Betonung hatte einen signifikanten Einfluss nach bilabialem C_1 ($p < 0.001$ im BP und $p < 0.05$ im EP), aber nicht nach velarem C_1 . Die Hörervarietät zeigte ebenfalls einen signifikanten Einfluss ($z = 2.8$, $p < 0.01$) und die Tukey Post-Hoc-Tests ergaben, dass Hörer sich bei den BP- ($p < 0.05$), nicht jedoch bei den EP-Sprechdaten unterschieden.

6.5 Diskussion

Im Folgenden werden die Ergebnisse zu den Hypothesen in der Reihenfolge der Hypothesen erläutert.

6.5.1 Sprechervarietät: BP vs. EP

Die Sprechervarietät zeigte einen hochsignifikanten Einfluss auf die Identifikation des Stimulustyps Cluster vs. $\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$, wobei der Anteil der korrekten Identifikationen viel größer bei den BP- als bei den EP-Token ausfiel. Dies war vor allem aus zwei Gründen zu erwarten, die im Folgenden erörtert werden.

Im BP wurde bisher zwar eine gewisse Abschwächung der lexikalischen Unterschiede beschrieben, die sich in Konsonantenclustern mit einem intermediären Vokal äußern, allerdings ausschließlich bei Plosivclustern (z.B. Mateus & d'Andrade 2000, Bisol 1999). Daher wurde angenommen, dass die beschriebene Vokalepenthese die Identifikation der Typen nicht stark beeinträchtigen würde, was auch tatsächlich der Fall war. Im EP hingegen wurde die optionale Vokaltilgung bei zwei der vier möglichen unbetonten Vokale festgestellt und daher die Zusammenführung aller Stimulustypen (Cluster vs. $\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$) zu Konsonantenclustern erwartet. Dies traf teilweise zu, denn im EP ergab sich eine geringere Anzahl an korrekten Identifizierungen.

Dieses Ergebnis konnte in einem zweiten Schritt mit Koordinationsunterschieden in Verbindung gebracht werden: Wie in Kapitel 4 und 5 festgestellt wurde, waren Konsonanten sowohl in Clustern als auch in CVC-Sequenzen vergleichsweise weiter synchronisiert und überlappten sich weniger im BP als im EP. Bei der Perzeption war der Anteil der CVC-Antworten größer bei den BP- als bei den EP-Stimuli. Bei den EP-Stimuli war der Anteil der wahrgenommenen Cluster und der einzelnen Konsonanten immer größer als bei den BP-Daten. Dies kann wiederum mit der engen Konsonantenkoordinierung der EP-Token in Verbindung gebracht werden.

Auf diese Weise konnten wir – wie vorhergesagt – feststellen, dass in BP-Token mehr intermediäre Vokale und in EP-Token mehr einzelnen Konsonanten wahrgenommen wurden, wobei diese Tatsache vor allem mit Koordinationsunterschieden zusammenhing.

6.5.2 Stimulustyp: Cluster vs. $\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$

$\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$ konnten in den BP-Sprachdaten identifiziert werden, jedoch nicht in den EP-Sprachdaten, da diese Stimulustypen zum großen Teil ohne intermediären Vokal in der Perzeption wiedergegeben wurden. Dies war angesichts der in den letzten Jahrzehnten beschriebenen Vokalelision im EP (Cunha 2011a, Mateus & d'Andrade 2000, Silva 1998, 1997, Vigário 1998, Martins 1975) zu erwarten und kann durch Koordinationsunterschiede erklärt werden. Wie vorhergesagt, hatte die breite Koordinierung von $\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$ in den BP-Daten den Effekt, dass beide Vokale identifiziert werden konnten. Im EP hingegen war die Koordinierung von $\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$ enger als im BP. Die Überlappung der Konsonantengesten auch bei diesen Stimuli hatte für die Perzeption zur Folge, dass Hörer die lexikalisch intermediären Vokale teilweise nicht mehr wahrnehmen konnten. Diese Tendenz zeigte sich stärker bei $\langle C_1eC_2 \rangle$ als bei $\langle C_1uC_2 \rangle$. Die intergestische Koordination von Clustern vs. $\langle C_1eC_2 \rangle$ zeigte in der Produktion ebenso kleinere Koordinationsunterschiede als die von Clustern vs. $\langle C_1uC_2 \rangle$.

In den EP-Token wurden Cluster zu höheren Anteilen als bei den BP-Token korrekt identifiziert. Hier sind die Bezüge zur Produktion sehr deutlich, da Cluster im BP insgesamt breiter koordiniert wurden als im EP, was wiederum dazu geführt hat, dass sie manchmal mit CVC-Sequenzen verwechselt wurden. Die enge Koordinierung (bzw. starke Überlappung) der Cluster im EP hat außerdem dazu beigetragen, dass EP-Cluster teilweise als einzelne Konsonanten wahrgenommen wurden. Zudem hatten wir aufgrund der engeren Synchronisierung von [u] im BP und [i] im EP eine Interaktion zwischen Sprechervarietät und Stimulustyp erwartet: Cluster vs. <C₁uC₂> im BP und Cluster vs. <C₁eC₂> sollten im EP stärker neutralisiert werden. Dies war für das EP der Fall, nicht jedoch für das BP.

6.5.3 Artikulationsort: bilabialer vs. velarer C₁

Die Antwortverteilung wurde ebenfalls vom Artikulationsort bei den Plosiv-Token beeinflusst. Der Anteil einzelner Konsonanten überwog bei den intendierten /kt/-Token in beiden Varietäten. Die Gründe dafür waren die starke Überlappung der Konsonanten bei diesen Token und der sogenannten Produktionsrichtungseffekt (Chitoran et al. 2002). Dieser Effekt führt zu der Vorhersage, dass bei (Plosiv-)Clustern in der *back-to-front*-Richtung (/kt/) weniger Überlappung auftritt, weil sie perzeptiv schwerer wiederherzustellen sind als *front-to-back*-Cluster (/pt/). Dies war bei unseren Produktionsdaten nicht der Fall und daher hatte die starke Überlappung der /kt/-Token zur Folge, dass der zweite Plosiv den ersten maskierte oder verdeckte. Somit war der erste Verschluss für die Hörer nicht mehr wiederherstellbar und sie konnten nicht mehr zwei Konsonanten, sondern lediglich C₂ wahrnehmen.

Allerdings zeigten die Produktionsdaten, dass beide Konsonantengesten noch vorhanden sind, sich aber sehr stark überlappen – im Extremfall so stark, dass ihre Verläufe in manchen Token parallel sind. Die Perzeptionsdaten ergaben jedoch, dass Hörer nicht mehr in der Lage waren, beide Gesten zu rekonstruieren: Insofern deuten diese Ergebnisse auf eine

Unstimmigkeit zwischen Produktion und Perzeption hin. Wenn ein und derselbe Stimulus manchmal als Einzelkonsonant, manchmal als Cluster wahrgenommen wird, kann dieses Missverhältnis laut Ohala (1993) zu einem Lautwandel führen. Der wahrgenommene Token wird vom Hörer zunächst reanalysiert. Lautwandel kann dann dadurch entstehen, dass der Hörer in der Rolle als Sprecher sich für die wahrgenommene Form in der eigenen Produktion entscheidet, d.h. der Hörer nimmt eine bestimmte neue Form wahr, und indem er sie nicht durch die bekannte Form ersetzt, sondern sie möglicherweise als Sprecher weitergibt, treibt er einen Lautwandel voran. In diesem Sinne würde ein laufender Lautwandel entstehen, indem der einzelne Konsonant alternativ zu dem Cluster in der Sprechgemeinschaft produziert wird und beide Formen miteinander koexistieren würden, bis sich eine etabliert. In diesem Stadium wäre der Lautwandel vollzogen.

Solch ein aktuell stattfindender Lautwandel (Scheutz 2005:1704-1716 für engl. *sound change in progress*) im Sinne von Ohala (1993) würde erklären, weshalb die Konsonantentilgung in Plosivclustern in der brasilianischen Varietät weiter fortgeschritten ist: Die analysierten Produktionsdaten zeigten, dass BP eine breite Koordination von Konsonantenclustern (außer bei /kt/) aufweist, so dass auch zwei Konsonanten perzeptiv wiederhergestellt werden können. Bei /kt/ jedoch überlappen sich der velare und der apikale Konsonant zwar weniger als in EP, aber sehr stark im Vergleich zu allen anderen BP-Clustern. Da die starke Überlappung in dieser Varietät nicht erwartet wird, können die Hörer sie nicht kompensieren, d.h. sie brauchen einen größeren Abstand zwischen den Konstriktionsplateaus und den akustischen Verschlüssen, um die Plosive zu identifizieren. Da die Verschlüsse sehr nah aneinander erfolgten, schlossen die Hörer nicht mehr auf einen Cluster zurück und nahmen lediglich einen Konsonanten wahr (/t/).

Eine ähnliche Entwicklung beider Plosive könnte bei der breiten Synchronisierung der Plosive in dem bilabialen Cluster /pt/ im

BP erfolgt sein, allerdings in einer anderen Richtung; /pt/ wird im BP breiter synchronisiert und in unseren Perzeptionsdaten hatte das zur Folge, dass die Hörer aufgrund des größeren Abstands einen intermediären Vokal perzipieren. Das gleiche graduelle Lautwandelmuster zwischen Hörer und Sprecher könnte dazu geführt haben, dass ein Vokal von manchen Sprechern nach und nach eingefügt wurde. Und dies wäre der Grund, weshalb in manchen der lexikalischen Cluster Stimmhaftigkeit sowie wenige Formanten zwischen den Konsonanten zu finden sind.

6.5.4 Betonung: betonte vs. unbetonte Kondition

Mit der Betonungskondition wird der Einfluss der Betonung auf die korrekte Identifikation des Stimulustyps und auf die Anzahl der wahrgenommenen Elemente geprüft. Der Einfluss der Betonung auf die Identifikation der Stimulustypen war eher marginal. Lediglich bei der Interaktion zwischen Stimulustyp und Artikulationsort wurde das Missverhältnis zwischen Produktion und Perzeption von der Betonung intensiviert, indem es ausgeprägter in unbetonter Kondition auftrat. Dies führte zu einem größeren Anteil an einzelnen Konsonanten in unbetonter Kondition bei den Perzeptionsantworten.

6.5.5 Hörervarietät: BP vs. EP

Insgesamt hatte die Hörervarietät nur einen geringen Einfluss auf die korrekte Identifikation des Stimulustyps. Es wurden vor allem Unterschiede bei der Wahrnehmung der Plosivstimuli der BP-Sprecher festgestellt. Eine mögliche Erklärung wäre, dass der Unterschied zwischen Clustern und Sequenzen mit intermediärem Vokal vor allem an der Koordination liegt. Demnach würden Hörer beider Varietäten eine gewisse Variabilität zwischen eng und weit koordinierten Konsonanten zeigen. So würde vor allem die Koordination zwischen Clustern und CVC-Sequenzen innerhalb einer Varietät (generell enger für Cluster als für CVC) einen Einfluss auf die Wahrnehmung ausüben und weniger die Koordinationsunterschiede zwischen den Varietäten (insgesamt enger im EP als im BP).

6.6 Schlussfolgerung und Ausblick

Aus dem durchgeführten Perzeptionsexperiment wurde ersichtlich, dass die vom BP-Sprecher produzierten Stimulustypen insgesamt von den meisten der beiden Hörergruppen korrekt identifiziert wurden. Allerdings wurden lexikalische Cluster vergleichsweise seltener identifiziert als Sequenzen mit Vokal ($\langle C_1eC_2 \rangle$ und $\langle C_1uC_2 \rangle$). Bei den intendierten Clustern hatte der Artikulationsort einen deutlichen Einfluss auf die Anzahl der Elemente bei den Antwortverteilungen. Bei den EP-Daten stach die Variabilität zwischen einem und zwei Konsonanten bei der Wahrnehmung heraus, bei den BP-Daten wurden alle Antwortmöglichkeiten ausgewählt.

Der Grund für Letzteres könnte die enge Koordinierung bei /kt/ und die breite Koordinierung bei /pt/ im BP sein, die aus der unterschiedlichen Synchronisierung von /p/ und /k/ zum nachfolgenden /t/ entsteht: In der *front-to-back* Produktionsrichtung (/pt/) war die Überlappung beider Plosive kleiner als in der umgekehrten Richtung (/kt/, siehe 3.3.3). Der Produktionsrichtungseffekt (Chitoran et al. 2002) macht allerdings die umgekehrte Vorhersage: *front-to-back* Cluster überlappen mehr, da ihre Elemente leichter herzustellen sind, als Konsonanten in der *back-to-front* Produktionsrichtung. Bei der *back-to-front* Richtung fand mehr Überlappung statt, obwohl die Konsonanten in dieser Richtung schwerer wiederherzustellen sind.

Möglicherweise waren Hörer nicht in der Lage, diese Koordinationsvariabilität zwischen den Produktionsrichtungen zu kompensieren (siehe Mann & Repp 1980, Repp 1984 für Kompensations-effekte). Daher nahmen sie diese Sequenzen nicht als Cluster wahr, sondern /pt/ (weit koordiniert) überwiegend mit einem intermediären Vokal und /kt/ (eng koordiniert) teilweise als einzelnen Konsonant. Der zweite Konsonant in /pt/ tritt später auf als bei anderen Clustern (z.B. /pl/), was in der Sprachperzeption als kurzes vokalisches Element empfunden wird. Bei /kt/ wurden die Konsonanten so eng

miteinander koordiniert, dass sie sich stark überlappten, was dazu führte, dass die Hörer nicht mehr zwei, sondern lediglich einen Konsonanten wahrnahmen. Infolgedessen alternierte die Wahrnehmung von /pt/ aufgrund der breiteren Synchronisierung zwischen CC und CVC, während /kt/ aufgrund der engen Synchronisierung beider Konsonanten entweder als Cluster oder als Einzelkonsonant (C₂) wahrgenommen wurde. Diese Diskrepanzen zwischen Produktion und Perzeption sind typisch für einen stattfindenden Lautwandel im Sinne von Ohala (1993).

Bei den vom EP-Sprecher produzierten Token konnten die Cluster identifiziert werden und die Sequenzen mit intermediären Vokalen, vor allem <C₁eC₂>, wurden –wie erwartet – größtenteils als reine Konsonantenabfolgen wahrgenommen. Dies legt nahe, dass der lexikalische Unterschied zwischen Clustern und Sequenzen auf kontinuierliche Weise schwächer wurde, so dass Cluster vs. <C₁eC₂> in der Perzeption schwieriger zu unterscheiden ist als Cluster vs. <C₁uC₂>. Diese Unterschiede stimmen mit den Ergebnissen aus der Produktion überein, insofern als die Synchronisierung von Cluster vs. <C₁eC₂> weniger Unterschiede aufwies als Cluster vs. <C₁uC₂>. Inwiefern bei der Unterscheidung beider Vokale eine zusätzliche Zungenbewegung zu einem vokalischen Target in /u/ aber nicht in /e/ vorhanden war und ob diese in der Perzeption einen zusätzlichen Hinweis (*cue*) darstellt, wirft eine sehr interessante Frage für die weitere Forschung auf.

Da die Koordination bei Plosiv-Liquid-Token insgesamt nicht so eng war wie bei Plosiv-Token (vor allem im Fall der reinen Plosivcluster) wurden die Unterschiede in der Antwortverteilung der einzelnen Konsonanten mit der Koordination in Verbindung gebracht und nicht mit der Artikulationsart. Bezüglich der Unterscheidung Plosiv vs. Liquid in der C₂-Position konnten keine Aussagen getroffen werden, da die analysierten Cluster in unterschiedlichen Positionen vorkamen und sich zusätzlich ihre Koordinierung unterschied. Mit dem Ziel, den Einfluss der Artikulationsart bei der Wahrnehmung von eng koordinierten Konsonanten zu untersuchen, könnte man in einem neuen

Perzeptionsexperiment Plosive und Laterale in der C₂-Position gegenüberstellen. Dies würde möglicherweise Erkenntnisse zur Rolle der Artikulationsart liefern, da sowohl Plosive als auch Liquide in Konsonantenclustern Überlappung beider Konsonantengesten aufwiesen, was aber bei den zentralen Liquiden (/r/) nicht der Fall war (siehe Kapitel 4). Die Laterale wurden allerdings aus dem Perzeptionsexperiment ausgeschlossen, weil die beiden untersuchten Varietäten jeweils unterschiedliche Laterale aufweisen und die Hörer somit einen zusätzlichen Hinweis für die Identifikation der Token zur Verfügung gehabt hätten, was die Untersuchung unserer varietätenspezifischen Fragestellung gestört hätte.

Aufgrund der Produktionsunterschiede wurde vermutet, dass Muttersprachler der beiden Varietäten möglicherweise die Token im mentalen Lexikon im Sinne der Exemplar-Theorie (Johnson 1997, Pierrehumbert 2000) andersartig organisieren. Insgesamt deuten die Ergebnisse jedoch auf eine ähnliche Speicherung der Token bei EP- und BP-Hörern hin und die Unterschiede in der Perzeption wurden eher mit varietätenspezifischen Koordinationsunterschieden in Verbindung gebracht.

7 Conclusio

Die vorliegende Arbeit beschäftigte sich mit der Organisation von Konsonanten in reinen Clustern und CVC-Sequenzen in der Sprachproduktion in zwei portugiesischen Varietäten und ihrer Auswirkung auf die Perzeption der Sprache. Selbst wenn der Unterschied zwischen einer CV-Silbe und einem Konsonantencluster aus theoretischer Sicht (sowohl in der Sprachproduktion, siehe Mattingly 1981, als auch in mehreren phonologischen Ansätzen, siehe z.B. Clements & Keyser 1983, McCarthy 1981, 1986, Prince & Smolensky 2004, Selkirk 1986) eindeutig zu sein scheint, können in der gesprochenen Sprache Schwächungen oder Variationen dieser lexikalischen Unterschiede auftreten. Die Variabilität kann sowohl von den Clustern selbst ausgehen, die durch Einfügung eines Vokals oder Epenthese zu CV-Sequenzen werden, als auch in der umgekehrten Richtung erfolgen, indem lexikalische CV(C)-Sequenzen mittels Vokalschwächungen oder Tilgungen zu reinen Konsonantengruppen werden.

Solche Variabilität findet aktuell im portugiesischen Kontext statt. Dabei ist allerdings besonders rätselhaft, wieso beide Varietäten sehr unterschiedliche – in mancher Hinsicht scheinbar entgegengesetzte – Tendenzen aufweisen, die gegen eine Verbesserung des Systems im phonologischen Sinne, also die Entwicklung hin zur universell unmarkierten CV-Silbe (Vennemann 1988, Restle & Vennemann 2001), sprechen würden. Es sei denn, man würde annehmen, dass sich Systeme in mehrere Richtungen entwickeln können, womit die Tendenz zur CV-Silbe nicht für jedes System notwendigerweise erstrebenswert wäre (Clements & Keyser 1983, McCarthy 1981). Im portugiesischen Kontext werden im BP die Tendenzen zur Realisierung aller CV-Silben und zur Vermeidung aller Konsonantengruppen außer Plosiv-Liquid-Cluster durch Einfügung eines intermediären epenthetischen Vokals beschrieben (Bisol 1999, Mateus & d'Andrade 2000, Silveira & Seara 2009). Im EP hingegen tritt

häufig die Tilgung der unbetonten Vokale /i,u/ ein, so dass in der gesprochenen Sprache die Konsonantengruppen an Anzahl und Komplexität gewinnen. So werden beispielsweise im EP Gruppen mit bis zu sechs Konsonanten beschrieben, z.B. [dʃprvn] *desprevenido* ‘unvorbereitet’, [dʃprtʃt] *desprestigiar* ‘herabwürdigen’, siehe u.a. Magalhães 2009, Mateus & d’Andrade 2000: 43-44, Schwarzsinger 2006.

Aus dem Interesse an dieser spannenden Datenlage heraus beschäftigte sich die Arbeit mit der Frage, inwiefern sich innerhalb einer Varietät und zwischen den Varietäten die zeitliche und räumliche Organisation von Konsonanten in Clustern im Vergleich zu den entsprechenden CVC-Sequenzen mit hinterem und nicht-hinterem Vokal unterscheiden. Man vermutet, dass Hörer Cluster von CVC-Sequenzen unterscheiden können, da sie einen phonemischen Charakter besitzen und somit zur Bedeutung in der Kommunikation beitragen (z.B. *crer* ‘glauben’ vs. *querer* ‘wollen’). Allerdings könnte es sein, dass die Hörer solche Wortpaare aufgrund ihrer lautlichen Ähnlichkeit ausschließlich aus dem Kontext rekonstruieren. Daher wurden bei den durchgeführten Experimenten die Produktionskontexte mittels Einbettung der zu untersuchenden Cluster und CVC-Sequenzen in einen konstanten Trägersatz vereinheitlicht, so dass keinerlei zusätzliche kontextuelle Hinweise gegeben waren.

7.1 Cluster vs. CVC-Sequenzen

Die Ergebnisse zeigen eindeutige zeitliche Unterschiede zwischen Clustern im Vergleich zu den CVC-Sequenzen $\langle C_1eC_2 \rangle$ und $\langle C_1uC_2 \rangle$. Die Cluster wurden in beiden Varietäten enger koordiniert als beide CVC-Sequenzen. Bezüglich der CVC-Sequenzen überlappten sich die beiden Konsonanten stärker im EP bei $\langle C_1eC_2 \rangle$ als bei $\langle C_1uC_2 \rangle$. Daher waren $\langle C_1eC_2 \rangle$ -Sequenzen enger koordiniert als die Sequenz mit dem hinteren Vokal $\langle C_1uC_2 \rangle$. Dies betraf sowohl Plosiv-Liquid-Cluster und in geringerem Maße auch Plosiv-Token, bei welchen die Koordination von CVC-Sequenzen nicht immer eindeutig war. Bei den BP-Token wurde eine Interaktion zwischen Vokal und

Artikulationsort nachgewiesen, so dass $\langle C_1eC_2 \rangle$ mehr Überlappung nach /k/ und $\langle C_1uC_2 \rangle$ nach dem bilabialen /p/ zeigte.

Für die Perzeption hatte dies zur Folge, dass die präsentierten Stimuli aus dem BP korrekt identifiziert werden konnten, d.h. die Perzeptionsantworten stimmten mit dem von der Versuchsperson im Produktionsexperiment intendierten Stimulustyp überein. Bei den Perzeptionsantworten zu den EP-Daten konnten Cluster ebenso mit großen Anteilen korrekt identifiziert werden, allerdings zeigten beide CVC-Sequenzen einen deutlich geringeren Anteil an korrekten Antworten. Dieser Anteil fiel insbesondere bei $\langle C_1eC_2 \rangle$ sehr gering aus.

Aus dem Zusammenspiel zwischen Produktion und Perzeption ging hervor, dass beide Konsonanten im EP auch bei den CVC-Sequenzen so eng aneinander liegen, dass die Hörer nicht mehr in der Lage sind, den lexikalischen Vokal (vor allem /i/) wahrzunehmen. Diese Überlappung war so stark bei den Plosiv-Plosiv-Token, dass die Hörer bei Clustern großteils nur noch den zweiten Konsonant wahrnahmen. Bei den CVC-Sequenzen fiel die Überlappung vergleichsweise geringer aus, allerdings wurden selbst bei den lexikalischen CVC-Sequenzen einige Token als einzelne Konsonanten wahrgenommen. Dies zeigt sehr deutlich, wie eng die intergestische Koordination der Konsonantengesten im EP sein kann, so dass in manchen Token sogar zwei lexikalische Elemente in der Perzeption verloren gehen können. Dies könnte ein Hinweis darauf sein, dass sich Sprecher in ihrer alltäglichen Kommunikation zum Verständnis auf den Kontext stützen.

In der vorliegenden Arbeit wurde die Anzahl der untersuchten Konsonanten auf zwei beschränkt (unter Hinzufügung eines intermediären Vokals in den CVC-Sequenzen) und daher ist die Nicht-Wahrnehmung von zwei aus drei Elementen der mögliche Extremfall. Man erkennt jedoch bereits bei diesen Token mit wenigen Elementen die starke Tendenz zur Tilgung in der

Perzeption (d.h. die lexikalischen Elemente werden nicht wahrgenommen bzw. nicht wiedergegeben), welche aus der starken Überlappung in der Produktion resultiert (siehe auch den abschließenden Abschnitt in Kapitel 4).

Die physiologischen Daten zeigen eindeutig, dass beide Konsonanten von den EP-Sprechern zwar produziert werden, jedoch so eng miteinander koordiniert sind, dass sie zum Teil parallel verlaufen. Demnach verdeckt C_2 im akustischen Signal C_1 und die Hörer nahmen nur noch C_2 wahr. Somit zeigen die analysierten Daten eine ersichtliche Diskrepanz zwischen Produktion und Perzeption: Es werden zwar noch zwei Segmente produziert, sie konnten aber nicht mehr wahrgenommen werden. Diese Diskrepanz ist instabil in der Hinsicht, dass der Hörer die starke Überlappung kompensieren muss (siehe Mann & Repp 1980 für Kompensationseffekte in der Perzeption), indem er möglicherweise das im Sprachsignal fehlende Element aus dem Kontext rekonstruiert.

Dies ist sehr aufwändig und kann möglicherweise auf eine von zwei Möglichkeiten reduziert werden: In ihrer wechselnden Rolle als Sprecher und Hörer in der Kommunikation können Teilnehmer in ihrer Rolle als Sprecher a) entweder ihre Produktion den Hörern gegenüber anpassen und die Konsonanten so weit auseinander artikulieren, dass beide Konsonanten perceptiv wiederherstellbar sind, oder b) aufhören, den ersten Konsonanten zu produzieren. Die zweite Möglichkeit ist zwar ökonomisch, würde allerdings einen Lautwandelprozess in Gang setzen (Beddor et al. 2007⁵⁵), bei dem Cluster zu einzelnen Konsonanten reduziert würden, wie es ebenfalls in der historischen Entwicklung von Plosivclustern sowohl im Portugiesischen als auch in anderen Sprachen geschehen ist.

Die Alternanz zwischen einem und zwei Konsonanten im aktuellen portugiesischen Schriftbild in beiden Varietäten (<fato>

⁵⁵ Siehe Beddor et al. 2007 für die Anwendung von Methoden aus der Perzeption für die Erfassung von Sprachvariation und Sprachwandel.

vs. <facto> ‘Fakt’) sowie die neue Rechtschreibreform für die Regelung der Schreibweise von u.a. Plosivclustern deuten allerdings bereits an, in welche Richtung der Lautwandel der Plosivcluster im Portugiesischen geht. Die Neuregelung besagt, dass Doppelkonsonanten als Einzelkonsonant geschrieben werden, z.B. alt ‘facto’, neu ‘fato’.

In diesem Zusammenhang wurde im Laufe der Arbeit ersichtlich, dass Konsonanten in einem Cluster und in einer CVC-Sequenz grundsätzlich unterschiedliche Koordinierungen aufweisen, die jedoch in dem komplexen Zusammenspiel von Produktion und Perzeption geschwächt werden können. Die zwei analysierten Varietäten zeigten einerseits eine verhältnismäßig stabile Aufrechterhaltung dieser Unterscheidung im BP und andererseits ihre zunehmende Aufhebung vor allem im Plosivkontext im EP.

7.2 Varietätsspezifische Unterschiede

Die Analyse der zeitlichen Organisation beider Konsonantengesten deckte deutlich varietätsspezifische Unterschiede auf: Die Koordination der Konsonanten sowohl bei Plosiv-Liquid- als auch bei Plosiv-Plosiv-Token war ausnahmslos breiter im BP als im EP.

Bezüglich des räumlichen Verlaufs der Zungenspitze für die Produktion der drei apikalen Laute /l, r, t/ waren bei Plosiv-Liquid-Token die Unterschiede zwischen den Varietäten signifikant, indem die apikalen Gesten im BP deutlich schneller produziert wurden und einen größeren Umfang erreichten als im EP. Bei den Plosiv-Plosiv-Token unterschieden sich die /t/-Gesten zwischen den Varietäten nicht signifikant voneinander, aber es war dennoch die gleiche Tendenz zu schnelleren und umfangreicheren Gesten im BP als im EP gegeben.

Die deutlichen Unterschiede in der apikalen Geste bei den Plosiv-Liquid-Token können mit Produktionsunterschieden in Verbindung gebracht werden. Es wurde bereits beschrieben, dass in beiden Varietäten der zentrale Liquid als alveolarer Tap /r/

realisiert wird, allerdings kommt im BP eine Variation vor allem mit dem Trill /r/ vor (Silva 1996). Bei den Lateralen ist der Unterschied deutlicher, da in diesem Kontext (als zweites Elements eines Clusters) der Lateral ein klares /l/ im BP ist (Silva 1996, Pagan & Wertzner 2007), jedoch ein dunkles /ɫ/ im EP (Andrade 1999, Martins et al. 2011, Oliveira et al. 2011, Teixeira et al. 2011). Die Dauer der Zungenspitzenengeste (Sproat & Fujimura 1993: 307) und des Verschlusses ist länger bei dunklen als bei hellen Lateralen (s. Recasens & Farnetani 1994: 198 für den Vergleich von italienischen hellen und katalanischen dunklen Lateralen und Recasens & Espinosa 2005: 12 für den Vergleich von Valencianischem hellen und Mallorquinischem dunklen /ɫ/). Der vorgeschlagene Grund dafür ist, dass die Erhöhung des Zungendorsums für die hintere Artikulation die Zungenspitze in der vorderen Region verflacht (dental oder weiter anterior, vgl. Recasens 2009: 478).

Die Dauer der Schließgeste war zwar nicht Gegenstand der aktuellen Untersuchung, allerdings geht aus den Daten hervor, dass der Zungengestenumfang bei Lateralen im EP viel geringer ausfällt als im BP. Dies entsteht möglicherweise durch die eben beschriebene Verflachung im vorderen Mundraum.

Außerdem wurde ersichtlich, dass der Zungenspitzenumfang von zentralen und lateralen Liquiden einen ähnlichen Umfang aufweist. Diese Ähnlichkeit im Umfang und die Änderung der zeitlichen intergestischen Koordination beider Konsonanten könnte der Grund sein, weshalb im BP ein Rhotazismus von /l/ zu /r/ beschrieben wird (z. B. *bicireta* für *bicicleta* 'Fahrrad', vgl. Romano & Silva 2010). Bei der intragestischen Koordination des apikalen Plosivs deuten die nicht-signifikanten Unterschiede zwischen BP und EP auf keine besonderen Auffälligkeiten in der /t/-Produktion in den analysierten Varietäten hin.

Insgesamt kann festgehalten werden, dass die Unterscheide zwischen den Varietäten in der intragestischen Koordination mit varietätsspezifischen Eigenschaften zusammenhängen und sich beide Varietäten erheblich in der intergestischen Koordination

beider Konsonantengesten unterscheiden: Die Konsonanten lagen in allen Token weiter auseinander und überlappten sich weniger im BP als im EP. Im EP war die Koordination beider Konsonantengesten so eng, dass die lexikalischen Vokale innerhalb der CVC-Sequenzen stark gekürzt wurden und häufig so massiv von den Konsonanten überlappt wurden, dass sie in der Perzeption nicht mehr wahrnehmbar waren.

7.3 Artikulationsort und Produktionsrichtungseffekt

Der Artikulationsort hatte einen marginalen Einfluss auf die Konsonantenkoordination der Plosiv-Liquid-Token, zeigte jedoch einen erheblichen Einfluss im Fall der Plosiv-Plosiv-Token. Bei Letzteren war der Effekt des Artikulationsortes und der damit assoziierten Produktionsrichtung von zentraler Bedeutung für die Analyse. Dieser Unterschied in der Koordination kann mit einer größeren Überlappung der Konsonanten in medialen heterosilbischen Kontexten zusammenhängen oder mit einem größeren Einfluss der Produktionsrichtung auf Cluster, deren C₂ ein Plosiv ist (im Vergleich zu einem Liquid). Der Grund dafür ist die starke Maskierung anderer Konsonanten durch Plosive in der C₂ Position. Diese Unterschiede werden jedoch nicht näher betrachtet, da Plosiv-Liquid-Token und Plosiv-Plosiv-Token aufgrund ihrer unterschiedlichen Position im Wort und hinsichtlich ihres tauto- vs. heterosilbischen Charakters nicht vergleichbar sind.

Bezüglich des Produktionsrichtungseffekts bei den Plosiv-Plosiv-Token zeigten die Ergebnisse, dass sich entgegen den Erwartungen die Konsonantengesten bei Clustern eindeutig mehr bei *back-to-front* Clustern (/kt/) überlappten und ebenso signifikant kleinere intermediäre Abstände zwischen den Plateaus in dieser Richtung zeigten als bei *front-to-back* /pt/. Es geht somit eindeutig hervor, dass /kt/ trotz der laut Produktionsrichtungseffekt postulierten schlechteren perzeptiven Wiederherstellbarkeit erheblich enger koordiniert wurde als /pt/.

Trotz der starken Vorhersage der Produktion, dass Plosive in der *back-to-front* Produktionsrichtung aufgrund ihrer schlechteren Wiederherstellbarkeit weniger überlappen dürfen, um in der Perzeption wiederherstellbar zu bleiben, zeigten Plosive genau in dieser Richtung die stärkste Überlappung bei beiden Messverfahren. Die Übereinstimmung der Verfahren zeigt die Stabilität dieses Effekts im Portugiesischen, allerdings wirkt sich die starke Überlappung auf die Perzeption aus. Wie vorhergesagt, können beide Plosive in /kt/ nicht wiederhergestellt werden und infolgedessen werden sie von den Hörern auf /t/ reduziert. Auf diese Weise entstand der Einzelkonsonant /t/ in der Perzeption aus der starken Gestenüberlappung in der Produktion. Dies fand ebenso bereits früher in der Übertragung des Klassischen Lateins hin zum Portugiesischen (KL *septimum* > AltPT *seitimo* > Port. *séptimo* ‘siebte’, Williams 1961: 86) – womöglich aus dem gleichen Grund – statt.

Bei den CVC-Sequenzen gab es eine insbesondere im BP ausgeprägte Interaktion zwischen Produktionsrichtung und Stimulustyp. Daraus resultierte, dass die Plosive in <C₁et> mehr Überlappung nach velaren und in <C₁ut> nach bilabialen C₁ zeigten. In beiden Varietäten war die Überlappung beider Konsonanten so stark, dass /put/ im Hinblick auf das intergestische Timing in beiden Varietäten von dem lexikalischen Cluster /pt/ nicht zu unterscheiden war. Der Grund dafür war eine starke Überlappung beider Konsonanten im EP, so dass der lexikalische Vokal /u/ in <put> bedeckt wurde. Im BP war die Überlappung zwar größer in <put> als in <pet>, allerdings überlappten die Konsonanten bei den lexikalischen Clustern <pt> weniger als die zwei Plosive in /kt/. Das hatte zur Folge, dass im BP <put> in der Perzeption eindeutig korrekt zugeordnet werden konnte, allerdings waren bei <pt> über 55% der Antworten nicht korrekt, d.h. die Perzeptionsantworten stimmten nicht mit den intendierten Token überein. Da die Hörer vier Antwortmöglichkeiten zur Auswahl hatten, ist dieser geringe Anteil an korrekten Identifizierungen nicht zufällig entstanden. Kurzgefasst zeigen also die Ergebnisse aus dem Zusammenspiel zwischen

Produktion und Perzeption, dass im BP die Wahrscheinlichkeit (abgesehen vom Kontext), einen Plosivcluster /pt/ zu identifizieren, unter 50% liegt.

Im EP können Cluster eindeutig korrekt identifiziert werden, CVC-Sequenzen hingegen nicht. Im EP wurden die Konsonanten vor allem bei <put> so eng miteinander koordiniert, dass keine Unterschiede in der Perzeption zwischen dem intendierten Cluster <pt> und der CVC-Sequenz <put> ersichtlich waren. Auch zeigten unsere Experimente, dass *back-to-front* CVC-Folgen besser erkannt werden konnten als die Folgen in der umgekehrten Produktionsrichtung. Unter den bilabialen Token konnte zwischen intendierten Clustern und CVC-Sequenzen in der Wahrnehmung kaum unterschieden werden. Den vorliegenden Daten zufolge ging die Unterscheidung zwischen dem lexikalischen Cluster <pt> und der lexikalischen Sequenz <pet> in der Perzeption in beiden Varietäten verloren.

7.4 Betonung

Die Betonung hatte überwiegend bei den Plosiv-Plosiv-Token einen signifikanten Einfluss auf die Konsonantenkoordination. Der Einfluss der Betonung auf die Plosiv-Liquid-Token war eher marginal. Bei den Plosiv-Plosiv-Token zeichnet sich die Auswirkung der Betonung auf eine höhere Geschwindigkeit der Geste in betonter als in unbetonter Kondition aus sowie durch einen größeren Umfang der Gesten. Dies folgt aus der stärkeren Koartikulation der Segmente in unbetonter Position, die sich wiederum in der kürzeren Dauer der Konstriktion und geringerem Umfang der Geste widerspiegelt.

Im BP hatte die schnellere Geschwindigkeit in betonter als unbetonter Kondition keine Auswirkung auf den Umfang nach Stimulustyp. Betonte und unbetonte Token unterscheiden sich nicht sehr stark und möglicherweise ist dies der Hintergrund der sogenannten Silben-Isochronie im BP (Bisol 1999, Barbosa 2004), die durch eine Alternanz von betonten und unbetonten Silben gekennzeichnet ist.

Die Betonung hatte ebenso einen eindeutigen Einfluss auf die intergestische Koordination der Plosiv-Plosiv-Token, indem Token in unbetonter Kondition deutlich enger koordiniert wurden als betonte Token. Allerdings ging der Effekt der Betonung auf die Perzeption nicht deutlich hervor. In den aktuellen Daten ist nicht klar, ob Hörer in der Perzeption bei den unbetonten Token stärker die zusätzliche Koartikulation (s. dazu Kleber et al. 2012) bzw. stärkere Überlappung in der Produktion kompensieren.

7.5 Schlusswort

Das Ziel der vorliegenden Arbeit war es, zu untersuchen, inwiefern Konsonanten innerhalb eines Clusters und als Teil einer CVC-Sequenz unterschiedlich koordiniert werden und inwiefern Konsonantenkoordination im Rahmen der Artikulatorischen Phonologie (Browman & Goldstein 1986, 1988, ff.) die beschriebene Variation zwischen lexikalischen Clustern und lexikalischen CVC-Sequenzen erklären kann. Ausgehend vom Portugiesischen verfolgte die Arbeit zunächst den Ursprung der analysierten Formen zurück bis hin zum Klassischen Latein und skizzierte ihre Entwicklung im Sprachvergleich zu den nächstgelegenen romanischen Sprachen. Im Großrahmen der Arbeit wurden anhand von akustischen und physiologischen Daten folgende Fragen untersucht:

- (i) inwiefern zwei Konsonanten in der Sprachproduktion eine unterschiedliche Organisation aufweisen, in Abhängigkeit davon, ob sie unmittelbar aufeinander folgen oder mit einem intermediären lexikalischen Vokal vorkommen,
- (ii) wie sich Koordination in der Produktion auf die Sprachperzeption auswirkt, und schließlich
- (iii) wie Produktion und Perzeption die Entwicklung einer Sprache beeinflussen können.

Die Ergebnisse zeigten deutliche Unterschiede zwischen Clustern und CVC-Sequenzen, die robuster bei der intergestischen Koordination beider Konsonanten waren als bei der räumlichen Organisation der analysierten Zungenspitzenregionen. Wie aus den vorherigen Kapiteln hervorgeht, war die zeitliche Organisation der Konsonanten graduell und nach Sprachvarietät, Artikulationsort, Betonung und weiteren Faktoren anpassungsfähig. Gleichzeitig war sie zentral für die Übermittlung lexikalischer Unterschiede zwischen Clustern und CVC-Sequenzen (siehe dazu Browman & Goldstein 1990a, b). Diese doppelte Funktion der zeitlichen Koordination, die sowohl bedeutungsunterscheidende phonologische Information als auch kontextbezogene Zusatzinformationen wiedergibt, kann dazu führen, dass bedeutungsunterscheidende lexikalische Information von kontextbedingten Anpassungen der Koordination (wie z. B. das schnelle Sprechen) geschwächt wird und im Extremfall verloren geht. Dies war der Fall im EP, bei dem die Unterschiede zwischen lexikalischen Clustern und CVC-Sequenzen sehr schwach ausfielen.

Die Alternanz zwischen einem und zwei Plosiven in dem aktuellen portugiesischen Schriftbild konnte ebenso mit Koordinationsunterschieden in Verbindung gebracht werden und insbesondere auf die Diskrepanz zwischen Produktion und Perzeption hinweisen. In diesem Zusammenhang entstand die Vereinfachung von Clustern zu Einzelkonsonanten als Folge der starken Überlappung der Elemente und der resultierenden Nicht-Wiederherstellbarkeit in der Perzeption. Möglicherweise hat eine ähnliche Diskrepanz zwischen Produktion und Perzeption den gleichen Wandel bereits zu einem früheren Zeitpunkt der Sprachgeschichte ausgelöst, indem <kt> zu <t> vereinfacht wurde. Somit lässt sich abschließend festhalten, dass moderne empirische Methoden offenbar die historische Sprachwissenschaft mit neuen Hypothesen und Erklärungsvorschlägen bereichern können.

8 Literaturverzeichnis

Abaurre, M. B. und C. Galves 1998. Rhythmic differences between European and Brazilian Portuguese: An optimalist and minimalist approach. *D.E.L.T.A., Documentação de Estudos em Lingüística Teórica e Aplicada* 14, 377-403.

Abaurre, M. B. und F. Sândalo 2007. Acento secundário e duas variedades do português: Uma análise baseada na OT. In G. A. Araújo (Ed.). *O Acento do Português, Abordagens Fonológicas*, 145-168. São Paulo, Editora Parábula.

Albright, A. 1999. *The effect of context on epenthetic vowels in Lakota*. Los Angeles, UCLA.

Ali, A. N., M. Lahrouchi und M. Ingleby 2008. Vowel epenthesis, acoustics and phonology patterns in Moroccan Arabic. *Ninth Annual Conference of the International Speech Communication Association*. 1178-1181.

Allen, W. S. 1953. *Phonetics in Ancient India*. London, Oxford University Press.

Almeida, M. und J. Dorta 1993. Datos acústicos de las líquidas españolas. In C. Díaz Alayón (Ed.). *Homenaje a José Pérez Pidal*, 97-110. Tenerife, Litografía A. Romero.

Andrade, A. 1993. Estudo acústico das sequências oclusivas em Português Europeu. *Actas do IX Encontro da APL*, 1-16. Lisboa, APL.

Andrade, A. 1999. On /l/ velarization in European Portuguese. *Proceedings of the International Congress of Phonetic Sciences*. San Francisco.

d' Andrade, E. und M. C. Viana 1993a. Sinérese, diérese e estrutura silábica. *Actas do X Encontro da APL*, 224-233. Lisboa, APL.

- d' Andrade, E. und M. C. Viana 1993b. As sobradas da translineação. *Actas do Primeiro Encontro de Processamento da Língua Portuguesa Escrita e Falada*, 209-214. Lisboa, INESC, UNINIVA, CLUL.
- Baayen, R. H. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, Cambridge University Press.
- Barbosa, J. M. 1965. *Études de phonologique portugaise*. Lisboa, Junta de Investigações Científicas do Ultramar.
- Barbosa, P. A. 2000. Syllable-timing in Brazilian Portuguese: Uma crítica a Roy Major. *D.E.L.T.A., Documentação de Estudos em Linguística Teórica e Aplicada* 16 (2), 369-402.
- Barbosa, P. A. 2002. Explaining Brazilian Portuguese resistance to stress shift with a coupled-oscillator model of speech rhythm production. *Cadernos de Estudos Linguísticos* 43, 71-92.
- Barbosa, P. A. 2006. *Incurções em torno do ritmo*. Campinas, Pontes.
- Barbosa, P. A. und E. Albano 2004. Brazilian Portuguese. *Journal of the International Phonetic Association* 34, 227-232.
- Barbosa, P., P. Arantes und L. S. Silveira 2004. Unifying stress shift and secondary stress phenomena with a dynamical systems rhythm rule. *Proceedings of the Speech Prosody*, 49-52.
- Barry, M. 1992. Palatalization, assimilation and gestural weakening in connected speech. *Speech Communication* 11, 393-400.
- Bates, D. M., M. Maechler und B. Bolker 2012. lme4, Linear mixed-effects models using S4 classes. R package version 0.999999-0.
- Beckman, M. E., M. Swora, J. Rauschenberg und K. Jong 1990. Stress shift, stress clash, and polysyllabic shortening in a prosodically annotated discourse, 5-8. *Proceedings of the International Conference on Spoken Language Processing*.
- Beckman, M., E., K. De Jong, S. A. Jun und S. H. Lee 1992. The interaction of coarticulation and prosody in sound change. *Language and Speech* 35, 45-58.

- Beddor, P. S., A. Brasher und C. Narayan 2007. Applying perceptual methods to the study of phonetic variation and sound change. In M. J. Solé, P. S. Beddor und M. Ohala (Eds.). *Experimental Approaches to Phonology*, 127-143. Oxford, Oxford University Press.
- Bisol, L. 1990. Degeminação e elisão. *IX Congresso Internacional da Associação de Linguística e Filologia da América Latina*, Campinas, Brazil.
- Bisol, L. 1992. O acento e o pé métrico binário. *Cadernos de Estudos Linguísticos* 22, 69-80.
- Bisol, L. 1999. A sílaba e seus constituintes. In M. H. M. Neves (Ed.). *Gramática do Português Falado* 7, 701-741. Campinas, Humanitas/Unicamp.
- Bisol, L. 2000. O troqueu silábico no sistema fonológico: Um adendo ao artigo de Plínio Barbosa. *D.E.L.T.A., Documentação de Estudos em Linguística Teórica e Aplicada* 16, 403-413.
- Bisol, L., M. Garcia, O. Nazabal, M. Oñederra, A. Lopes, G. Rio-Torto, J. Veloso 2010. A simetria no sistema vocálico do Português Brasileiro. *Linguística - Revista de Estudos Linguísticos da Universidade do Porto* 5 (1), 41-52.
- Bisol, L. und C. Brescancini 2002. *Fonologia e variação, recortes do Português Brasileiro*. Porto Alegre, EDIPUCRS.
- Bisol, L. und C. Brescancini 2008. *Contemporary phonology in Brazil*. Cambridge, Cambridge Scholars Publishing.
- de Boer, B. 2000. Self-organization in vowel systems. *Journal of Phonetics* 28, 441-465.
- Bombien, L. 2011. *Segmental and prosodic aspects in the production of consonant clusters*. Ludwig-Maximilians-Universität München. Doctoral Dissertation.

- Bombien, L., C. Mooshammer, P. Hoole, B. Kühnert und J. Schneeberg 2006. An EPG study of initial /kl/ clusters in varying prosodic conditions in German. *Proceedings of the 7th International Seminar on Speech Production*, 457-460. Ubatuba.
- Bombien, L., C. Mooshammer, P. Hoole und B. Kühnert 2010. Prosodic and segmental effects on EPG contact patterns of word-initial German clusters. *Journal of Phonetics* 38, 388-403.
- Bombien, L., C. Mooshammer und P. Hoole 2013. Articulatory coordination in word-initial clusters of German. *Journal of Phonetics* 41, 546-561.
- Booij, G. 1995. *The Phonology of Dutch*. Oxford, Clarendon Press.
- Borgström, C. H. 1949. *Linguistic survey of the Gaelic dialects of Scotland, Vol. 1, The dialects of the Outer Hebrides*. Oslo, Norwegian Universities Press.
- Bosch, A. 1995. *A gestural analysis of epenthesis in Scottish Gaelic*. Kentucky, University of Kentucky. Doctoral Dissertation.
- Boyd-Bowman, P. 1952. La pérdida de vocales átonas en la altiplanicie Mexicana. *Nueva Revista de Filología Hispánica* 6 (6), 138-140.
- Boyd-Bowman, P. 1954. *From Latin to Romance in sound charts*. Washington D.C., George Town University Press.
- Bradley, T. 2001. A typology of rhotic duration contrast and neutralization. *Proceedings of NELS* 31, 79-98.
- Bradley, T. 2002. Gestural timing and derived environment effects in Norwegian clusters. *Proceedings of the West Coast Conference on Formal Linguistics* 21, 101-14.
- Bradley, T. 2004. Gestural timing and rhotic variation in Spanish codas. *Laboratory Approaches to Spanish Phonology*, 195-220.
- Bradley, T. 2006. Spanish complex onsets and the phonetics-phonology interface. *Optimality-Theoretic Studies in Spanish Phonology*, 15-38.

- Bradley, T. und B. Schmeiser 2003. On the phonetic reality of Spanish /r/ in complex onsets. In P. Kempchinsky und C. Piñeros (Eds.). *Theory, practice, and acquisition*, 1-20. Sommerville, MA, Cascadilla.
- Browman, C. P., L. Goldstein, J. Kelso, P. E. Rubin und E. Saltzman 1984. Articulatory synthesis from underlying dynamics. *Journal of the Acoustical Society of America* 75, 22-35.
- Browman, C. P. und L. Goldstein 1986. Towards an articulatory phonology. *Phonology yearbook* 3 (21), 9-252.
- Browman, C. P. und L. Goldstein 1988. Some notes on syllable structure in articulatory phonology. *Phonetica* 45, 140-155.
- Browman, C. P. und L. Goldstein 1989. Gestural structures and phonological patterns. *Haskins Laboratories Status Report on Speech Research* 92.
- Browman, C. P. und L. Goldstein 1990a. Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics* 18, 299-320.
- Browman, C. P. und L. Goldstein 1990b. 'Targetless' schwa: An articulatory analysis. *Papers in Laboratory Phonology* 2, 26-56.
- Browman, C. P. und L. Goldstein 1991. Gestural structures, distinctiveness, phonological processes, and historical change. *Modularity and the motor theory of speech perception* 1, 313-338.
- Browman, C. P. und L. Goldstein 1991. Tiers in articulatory phonology, with some implications for casual speech. *Papers in Laboratory Phonology* 1, 341-376.
- Browman, C. P. und L. Goldstein 1992. Articulatory phonology: An overview. *Phonetica* 49, 155-180.
- Browman, C. P. und L. Goldstein 1995. Dynamics and articulatory phonology. In T. van Gelder und R. F. Port (Eds.). *Mind as motion*, 175-193. Cambridge, MA, MIT Press.

- Browman, C. P. und L. Goldstein 2000. Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la Communication Parlée* 5, 25-34.
- Buchwald, A. B., B. Rapp und M. Stone 2007. Insertion of discrete phonological units: An articulatory and acoustic investigation of aphasic speech. *Language and Cognitive Processes* 22 (6), 910-948.
- Bybee, J. 2006. From usage to grammar, the mind's response to repetition. *Language* 82 (4), 711-718.
- Bybee, J. L. 2008. 5. Formal universals as emergent phenomena: The origins of structure preservation. *Linguistic Universals and Language Change* 1 (9), 108-123.
- Byrd, D. 1993. 54,000 American stops. *UCLA Working Papers in Phonetics* 83, 97-116.
- Byrd, D. 1994. *Articulatory timing in English consonant sequences*. Phonetics Laboratory, UCLA. Doctoral Dissertation.
- Byrd, D. 1996a. A phase window framework for articulatory timing. *Phonology* 13, 139-170.
- Byrd, D. 1996b. Influences on articulatory timing in consonant sequences. *Journal of Phonetics* 24, 209-244.
- Byrd, D. und C. Tan 1996. Saying consonant clusters quickly. *Journal of Phonetics* 24, 263-282.
- Byrd, D. und E. Saltzman 1998. Intra-gestural dynamics of multiple phrasal boundaries. *Journal of Phonetics* 26, 173-199.
- Byrd, D. und S. Choi 2010. At the juncture of prosody, phonology, and phonetics: The interaction of phrasal and syllable structure in shaping the timing of consonant gestures. *Laboratory Phonology* 10, 31-59.
- Câmara, J. 1970. *Estrutura da língua Portuguesa*. Petrópolis, Vozes.
- Carvalho, K. C. 2004. *Descrição fonético-acústica das vibrantes no Português e no Espanhol*. Assis, UNESP.

Castro, I. 1991. *Curso de história da língua Portuguesa*. Lisboa, Universidade Aberta.

Chitoran, I., L. Goldstein und D. Byrd 2002. Gestural overlap and recoverability: Articulatory evidence from Georgian. *Laboratory Phonology* 7, 419-448.

Chitoran, I., E. Marsico, F. Pellegrino und C. Coupé 2009. *Approaches to phonological complexity*. Berlin-New York, de Gruyter.

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

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

Clements, G. und S. Keyser 1983. *CV phonology: A generative theory of the syllable*. Cambridge, MA, MIT Press.

Collischonn, G. 2003. Epêntese vocálica no Português do Sul do Brasil: Variáveis extralinguísticas. *Revista Letras* 61, 285-297.

Costa, F. 2004. Intrinsic prosodic properties of stressed vowels in European Portuguese. *Proceedings of the 2nd International Conference on Speech Prosody*. Nara, Japan.

Costa, L. T. 2006. Ataque complexo no Português Brasileiro. *ReVEL – Revista Virtual de Estudos da Linguagem* 7, 1-15.

Cristófaros-Silva, T. 1999. *Fonética e fonologia do Português: Roteiro de estudos e guia de exercícios*. São Paulo, Editora Contexto.

Cristófaros-Silva, T. 2000. Sobre a quebra de encontros consonantais no Português Brasileiro. *Estudos Linguísticos* 29, 522-526.

Cristófaros-Silva, T. 2001. Difusão lexical: Estudo de casos do Português Brasileiro. In E. A. M. Mendes, P. M. Oliveira und V. Benn-Ibler (Eds.). *O Novo Milênio: Interfaces Linguísticas e Literárias*. 2092-2118. Belo Horizonte, Faculdade de Letras.

- Cristófaró-Silva, T. 2002. On Branching Onsets in Brazilian Portuguese. *Revista de Estudos da Linguagem* 10 (1), 91-107.
- Cruz-Ferreira, M. 1995. European Portuguese. *Journal of the International Phonetic Association* 25 (2), 90-94.
- Cunha, C. *in press.* captar-cap(i)tar, facto-fato: Variation und Wandel bei Plosivsequenzen im Portugiesischen. *Promptus*.
- Cunha, C. 2014. Neutralisierung lexikalischer Unterschiede im europäischen und brasilianischen Portugiesisch: Interaktion von Sprachproduktion und Perzeption. In T. Krefeld und E. Pustka (Eds.). *Perzeptive Linguistik: Phonetik, Semantik, Varietäten. (Zeitschrift für Dialektologie und Linguistik – Beihefte)*, 33-50. Franz Steiner.
- Cunha, C. 2012. Inter und intra-gestural coordination of initial stop-liquid consonant clusters in two varieties of Portuguese. *LabPhon13*. Stuttgart.
- Cunha, C. 2011. A elisão vocálica no Português Europeu. In M. Arden, C. Märzhäuser und B. Meisnitzer (Eds.) *Linguística do Português: Rumos e Pontes*, 151-166. München, Meidenbauer.
- Cunha, C. und J. Harrington 2011a. The gesture overlap in Portuguese lexical and post-lexical consonant clusters. *Proceedings of International Seminar on Speech Production*. Montréal, Canada.
- Cunha, C. und J. Harrington 2011b. The perception of /pt/ and /kt/ in European and Brazilian Portuguese. *Proceedings of 17th International Congress of Phonetic Sciences*. Hong Kong.
- Davidson, L. 2006. Schwa elision in fast speech: Segmental deletion or gestural overlap? *Phonetica* 63, 79-112.
- Davidson, L. und M. Stone 2004a. Epenthesis versus gestural mistiming in consonant cluster production: An ultrasound study. *Proceedings of the West Coast Conference on Formal Linguistics* 22, 165-178.

- Davidson, L., P. Smolensky und P. Jusczyk 2004b. The initial and final states, theoretical implications and experimental explorations of richness of the base. In R. Kager, W. Zonneveld und J. Pater (Eds.). *Fixing Priorities, Constraints in Phonological Acquisition*, 321-368. Cambridge, Cambridge University Press.
- Delforge, A. M. 2006. Gestural alignment constraints and unstressed vowel devoicing in Andean Spanish. *WECO*, 95-104.
- Delforge, A. M. 2008a. Unstressed vowel reduction in Andean Spanish. In L. Colantoni und J. Steele (Eds.) *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*, 107-124. Somerville, MA, Cascadilla.
- Delforge, A. M. 2008b. Gestural alignment constraints and unstressed vowel devoicing in Andean Spanish. In C. B. Chang and H. J. Haynie (Eds.). *Proceedings of the 26th West Coast Conference on Formal Linguistics*, 147-155. Somerville, MA, Cascadilla.
- Delforge, A. M. 2010. *The rise and fall of unstressed vowel reduction in the Spanish of Cusco, Peru: A sociophonetic study*. Davis, University of California. Doctoral Dissertation.
- Delgaard, P. 2008. *Introductory Statistics with R*. Heidelberg, Springer.
- Dicionário de Português-Latim 2007. 2. Aufl. Porto, Porto Editora.
- Dicionário de Latim-Português 2008. 3. Aufl. Porto, Porto Editora.
- Emiliano, A. 1990a. *Apologia do acordo ortográfico: Textos de intervenção em defesa da língua Portuguesa contra o acordo ortográfico de 1990*. Lisboa, Verbo.
- Emiliano, A. 1990b. *O fim da ortografia: Comentário razoado dos fundamentos técnicos do acordo ortográfico da língua Portuguesa*. Lisboa, Guimarães Editores.
- Emiliano, A. 2008 *Foi você que pediu Um Acordo Ortográfico?* Lisboa, Guimarães Editores.

- Escudero, P., P. Boersma, A. Schurt Rauber und R. A. H. Bion 2009. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *The Journal of the Acoustical Society of America* 126 (3), 1379-1393.
- Everett, C., Z. Miller, K. Nelson, V. Soare und J. Vinson 2011. Reduction of Brazilian Portuguese vowels in semantically predictable contexts. *Proceedings of the 17th International Congress on Phonetic Sciences*, 651-654. Hong Kong.
- Farnetani, E. und D. Recasens 1997. Coarticulation and connected speech processes. In J. Laver und W. Hardcastle (Eds.). *The Handbook of Phonetic Sciences*, 316-352. Oxford, Blackwell.
- Flemming, E. 2003. The relationship between coronal place and vowel backness. *Phonology* 20 (3), 335-373.
- Fokes, J. und Z. S. Bond 1993. The elusive/illusory syllable. *Phonetica* 50, 102-118.
- Fougeron C. 1999. Prosodically conditioned articulatory variations, a review. U.C.L.A. *Working Papers in Phonetics*, 97, 1-74.
- Fougeron C. und D. Steriade 1997. Does deletion of French schwa lead to neutralization of lexical distinctions? *Proceedings of Eurospeech97*, 2, 943-946. Rhodes.
- Fougeron, C. und P. Keating 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, 101, 3728-3740.
- Fowler, C. 2005. Parsing coarticulated speech in perception, effects of coarticulation resistance. *Journal of Phonetics* 33, 199-213.
- Freitas, E. 1999. *Sobre a quebra de encontros consonantais no Português Brasileiro*. Fale, UFMG. Doctoral Dissertation.
- Freitas, M. J. 1997. *Aquisição da estrutura silábica do Português Europeu*. Lisboa, FLUL. Doctoral Dissertation.

- Freitas, M. J. 2002. Estatutos de [j] na aquisição do Português Europeu. In I. Duarte, J. Barbosa, S. Matos und T. Hüsgen. *Encontro Comemorativo dos 25 Anos do Centro de Linguística da Universidade do Porto*, 99-110. Porto, CLUP.
- Freitas, M. J. 2003. The acquisition of Onset clusters in European Portuguese. *Probus* 15 (1), 27-46.
- Freitas, M. J. und A. L. Santos 2001. *Contar (histórias de) sílabas: Descrição e implicações para o ensino do Português como língua materna*. Lisboa, Edição Colibri.
- Frota, S., C. Galves und M. Vigário 2008. Ler a fonologia: Do Português clássico ao Português moderno. In M. Vigário (Ed.). *Actas do XXIII Encontro Nacional da Associação Portuguesa de Linguística*, 193-206. Lisboa, Edição Colibri.
- Frota, S., M. Vigário und F. Martins 2002. Language discrimination and rhythm classes: Evidence from Portuguese. *Proceedings of Speech Prosody 2002*, 315-318. Aix-en-Provence.
- Frota, S. und M. Vigário 2000. Aspectos de prosódia comparada, ritmo e entoação no PE e no PB. In R. V. Castro und P. Barbosa (Eds.). *Actas do XV Encontro Nacional da Associação Portuguesa de Linguística*, 533-555. Coimbra, APL.
- Frota, S. und M. Vigário 2001. On the correlates of rhythmic distinctions: The European/Brazilian Portuguese case. *Probus* 13, 247-273.
- Fujisaki, H. und T. Kawashima 1968. Some experiments on speech perception and a model for the perceptual mechanism. *Annual Report of the Engineering* 29, 67-73.
- Fujisaki, H. und T. Kawashima 1970. The role of auditory short-term memory in vowel perception. *Reports of the 6th International Congress on Acoustics*, 207-214.
- Gafos, A. 2002. A grammar of gestural coordination. *Natural Language and Linguistic Theory* 20, 269-337.

- Gafos, A., P. Hoole, D. Roon und C. Zeroual 2010. Variation in overlap and phonological grammar in Moroccan Arabic clusters. *Laboratory Phonology* 10, 657-698.
- Gafos, A. I., P. Hoole und C. Zeroual 2011. Preliminary study of Moroccan Arabic word-initial consonant clusters and syllabification using electromagnetic articulography. *Instrumental Studies in Arabic Phonetics* 29, 27-46.
- Garrapa, L. 2011. *Vowel elision in Florentine Italian*. Bern, Peter Lang.
- Geng, C., C. Mooshammer, H. Nam und P. Hoole 2010. Schwa deletion under varying prosodic conditions: Results of a pilot study. In P. Hoole, C. Mooshammer und M. Zygis (Eds.). *Between the regular and the particular in speech and language*, 145-169. Frankfurt am Main, Peter Lang.
- Gibbon, F., W. Hardcastle und K. Nicolaidis 1993. Temporal and spatial aspects of lingual coarticulation in /kl/sequences, a cross-linguistic investigation. *Language and Speech* 36, 261-277.
- Girelli, C. A. 1988. *Brazilian Portuguese syllable structure*. ETD Collection for University of Connecticut. Doctoral Dissertation.
- Glowacka, D. 2001. Unstressed vowel deletion and new consonant clusters in English. *Poznan Studies Contemporary Linguistics* 37, 71-94.
- Goldinger, S. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological review* 105 (2), 251-264.
- Goldstein, L., I. Chitoran und E. Selkirk 2007a. Syllable structure as coupled oscillator modes, evidence from Georgian vs. Tashlhiyt Berber. *Proceedings of the XVIth International Congress of Phonetic Sciences*, 241-244.
- Goldstein, L., M. Pouplier, L. Chen, E. Saltzman und D. Byrd 2007b. Dynamic action units slip in speech production errors. *Cognition* 103, 386-412.

- Gordon, A. 1980. Notas sobre la fonética del castellano en Bolivia. In A. Gordon und E. Rugg (Eds.). *Actas del Sexto Congreso Internacional de Hispanistas*, 349-352. Toronto, University of Toronto.
- Grosvald, M. 2009. Interspeaker variation in the extent and perception of long-distance vowel-to-vowel coarticulation. *Journal of Phonetics* 37, 17-188.
- Hall, N. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23 (3), 387-395.
- Hall, N. 2003. *Gestures and segments: Vowel intrusion as overlap*. Amherst, University of Massachusetts. Doctoral Dissertation.
- Hardcastle, W. J. 1985. Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication* 4, 247-263.
- Harrikari, H. 1999. Epenthesis, geminates, and the OCP in Finnish. *Nordic journal of linguistics* 22, 3-26.
- Harrington, J. 2010. *Phonetic analysis of speech corpora*. Chichester, Wiley.
- Hermes, A., M. Grice, D. Mücke und H. Niemann 2008. Articulatory indicators of syllable affiliation in word initial consonant clusters in Italian. *Rema* 151, 230-234.
- Hermes, A., M. Grice, D. Mücke und H. Niemann 2008. Articulatory indicators of syllable affiliation in word initial consonant clusters in Italian. *Proceedings of the 8th International Seminar on Speech Production*, 433-436. Strasbourg.
- Hermes, A., R. Ridouane, D. Mücke und M. Grice 2011. Kinematics of syllable structure in Tashlhiyt Berber: The case of vocalic and consonantal nuclei. *Proceedings of the 9th International Seminar on Speech Production*, 401-408. Montreal.

- Honorof, D. und C. P. Browman 1995. The centre or edge, how are consonant clusters organised with respect to the vowel? In K. Elenius und P. Branderud (Eds.) *Proceedings of the XIIIth International Congress of Phonetic Sciences*, 552-555. Stockholm.
- Hoole, P., L. Bombien, B. Kühnert und C. Mooshammer 2009. Intrinsic and prosodic effects on articulatory coordination in initial consonant clusters. In G. Fant, H. Fujisaki und J. Shen (Eds.). *Frontiers in Phonetics and Speech Science: Festschrift for Wu Zongji*, 275-286. Commercial Press, Beijing.
- Hoole, P. und A. Zierdt 2010. In B. Maassen, P. van Lieshout (Eds.). Five-dimensional articulography. *Speech Motor Control*, 331-349. Oxford, Oxford University Press.
- Hooper, J. 1978. Constraints on schwa-deletion in American English. In E. Fisiak (Ed.). *Recent Developments in Historical Linguistics*. The Hague, Mouton.
- Hothorn, T., F. Bretz, P. Westfall und R. M. Heiberger 2008. *multcomp: Simultaneous inference in general parametric models*. URL <http://CRAN.R-project.org>. R package version 1.0-0.
- Hundley, J. E. 1983. *Linguistic variation in Peruvian Spanish, unstressed vowel and /s/*. Ann Arbor, University Microfilms International. Doctoral Dissertation.
- Hundley, J. E. 1986. The effect of two phonological processes on syllable structure in Peruvian Spanish. *Hispania* 69, 665-668.
- Iskarous, K. und D. Kavitskaya 2010. The interaction between contrast, prosody, and coarticulation in structuring phonetic variability. *Journal of Phonetics* 38 (4), 625-639.
- Jancke, L., A. Bauer, P. Kaiser und K. T. Kalveram 1997. Timing and stiffness in speech motor control of stuttering and nonstuttering adults. *Journal of Fluency Disorders* 22 (4), 309-321.
- Jannedy, S. 1994. Rate effects on German unstressed syllables. *OSU Working Papers in Linguistics*.

- Jensen, F. 1999. *A comparative study of romance*. New York, Peter Lang.
- Jespersen, O. 1904. *Phonetische Grundfragen*. Leipzig, Berlin, Teubner.
- Johnson, K. 1997. The auditory/perceptual basis for speech segmentation. *Working Papers in Linguistics*, 101-113. Columbus, Ohio State University.
- Johnson, K. 2006. Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics* 34 (4), 485-499.
- Johnson, K. 2011. *Quantitative Methods in Linguistics*. John Wiley and Sons.
- Johnson, K., E. Strand und M. D'Imperio 1999. Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics* 27 (4), 359–384.
- Kabatek, J. 1996. *Die Sprecher als Linguisten: Für eine perzeptive Varietätenlinguistik*. Niemeyer, Tübingen.
- Kager, R. 1999. *Optimality theory*. Cambridge University Press.
- Kelso, J. und G. Schöner 1988. Self-organization of coordinative movement patterns. *Human Movement Science* 7 (1), 27–46.
- Kelso, J., E. Vatikiotis-Bateson, E. L. Saltzman und B. Kay 1985. A qualitative dynamic analysis of reiterant speech production, phase portraits, kinematics, and dynamic modelling. *The Journal of the Acoustical Society of America* 77 (1), 125-159.
- Kisler, T., F. Schiel und H. Sloetjes 2012. Signal processing via web services, the use case WebMAUS. *Proceedings Digital Humanities 2012*, 30-34. Hamburg.
- Kleber, F., J. Harrington, U. Reubold und J. Siddins 2012. Compensation for coarticulation in prosodically weak words. *Speech Prosody 2012*. Shanghai.

- Kohler, K. J. 2001. The investigation of connected speech processes: Theory, method, hypotheses and empirical data. *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel* 35, 1-32.
- Koike, D. und D. P. Macedo 1992. *Romance linguistics: The Portuguese context*. Greenwood Publishing Group.
- Krefeld, T. 1987. Romanische Vokalschwächung und rumänisches /i/. In G. A. Plangg und M. Iliescu (Eds.). *Akten der Theodor Gartner-Tagung* 14, 323-333. Innsbruck, Institut für Romanistik der Leopold-Franzens-Universität.
- Krefeld, T. 1999. *Wortgestalt und Vokalsystem in der Italoromania: Plädoyer für eine gestaltphonologische Rekonstruktion des romanischen Vokalismus*. Habilitationsschrift, Kiel.
- Krefeld, T. 2004. *Einführung in die Migrationslinguistik: Von der Germania Italiana in die Romania multipla*, Tübingen, Narr.
- Krefeld, T. und E. Pustka 2010. Für eine perzeptive Varietätenlinguistik. In T. Krefeld und E. Pustka. *Perzeptive Varietätenlinguistik: Spazi comunicativi - Kommunikative Räume*, 8-27. Frankfurt am Main, Peter Lang Publishing.
- Krötsch, M. 2004. La formation de groupes de consonnes au début des unités phoniques en français parlé. In T. Meisenburg und M. Selig. (Eds.). *Nouveaux départs en phonologie*, 218-233. Tübingen, Narr.
- Kühnert, B., P. Hoole, C. Mooshammer 2006. Gestural overlap and C-center in selected French consonant clusters. *Proceedings of the 7th International Seminar on Speech Production*, 327-334. UFMG Belo Horizonte.
- Kul, M. 2010. Towards a gradual scale of vowel reduction: A pilot study. *Poznań Studies in Contemporary Linguistics* 46 (4), 429-456.
- Ladefoged, P. und I. Maddieson 1996. *The sounds of the world's languages*. Cambridge, Blackwell.

Lawrence, M. A. 2011. "Package 'ez'". URL <http://cran.r-project.org/web/packages/ez/ez.pdf>

Lieberman, A. M., K. S. Harris, H. S. Hoffman und B. C. Griffith 1957. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54 (5), 358-372.

Lieberman, A. M., K. S. Harris, J. A. Kinney und H. Lane 1961. The discrimination of relative onset-time of the components of certain speech and non-speech patterns. *Journal of Experimental Psychology* 61 (5), 379-390.

Lieberman, A. M., F. S. Cooper, D. P. Shankweiler und M. Studdert-Kennedy 1967. Perception of the speech code. *Psychological Review* 74, 431-461.

Lieberman, A. M. und I. Mattingly 1985. The motor theory of speech perception. *Cognition* 21, 1-36.

Lindblom, B. 1990. Explaining phonetic variation: A sketch of the H and H theory. In W. Hardcastle und A. Marchal (Eds.). *Speech Production and Speech Modelling*, 403-439. Dordrecht, Springer.

Lindblom, B. 1996. Role of articulation in speech perception: Clues from production. *Journal of Acoustical Society of America* 99 (3), 1683-1692.

Lindblom, B., R. Diehl, S. H. Park und G. Salvi 2011. Sound systems are shaped by their users: The recombination of phonetic substance. In N. Clements und R. Ridouane (Eds.). *Where Do Phonological Features Come From*, 67-97. Amsterdam, John Benjamins.

Lipski, J. M. 1989. /S/-voicing in Ecuadorian Spanish: Patterns and principles of consonantal modification. *Lingua* 79 (1), 49-71.

Lipski, J. M. 1990. Aspects of Ecuadorian vowel reduction. *Hispanic Linguistics* 4 (1), 1-19.

Lipski, J. M. 1995. Literary Africanized' Spanish as a research tool, dating consonant reduction. *Romance Philology* 49, 130-167.

- Lope Blanch, J. 1963. En torno a las vocales caedizas del Español Mexicano. *Nueva Revista de Filología Hispánica* 17, 1-19.
- Lopes, M. A. und M. Santos 1995. Encontros consonantais tautossilábicos no Português Brasileiro. *Semana do NAPq*. Falé, UFMG.
- Lousada, M. L. und L. M. T Jesus 2006. Analysis of stop consonant production in European Portuguese. *ITRW on Experimental Linguistics*.
- Lousada, M. L., L. M. T. Jesus und A. Hall 2010. Temporal acoustic correlates of the voicing contrast in European Portuguese stops. *Journal of the International Phonetic Association* 40 (3), 261-275.
- Maddieson, I. 2003. The sounds of the Bantu languages. In D. Nurse und G. Philippson (Eds.). *The Bantu Languages*, 15-41. Oxon, Routledge.
- Madureira, E. 1997. Difusão lexical e variação fonológica: O fator semântico. *Revista de Estudos da Linguagem* 5, 5-22.
- Magalhães, P. 2009. *Silabificação de sequências de consonantes*. Porto, FLUP. Doctoral Dissertation.
- Mann, V. A. und B. H. Repp 1980. Influence of vocalic context on the perception of [f-s] distinction. *Perception und Psychophysics* 28, 213-228.
- Manuel, S., S. Shattuck-Hufnagel, S. Huffman, K. Stevens und S. Hunnicut 1992. Studies of vowel and consonant reduction. *Second International Conference on Spoken Language Processing*.
- Marin, S. und M. Pouplier 2008. Organization of complex onsets and codas in American English: Evidence for a competitive coupling model. *Proceedings of the 8th International Seminar of Speech Production*, 437-440.
- Marin, S. und M. Pouplier 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.

- Marslen-Wilson, W. D. 1987. Functional parallelism in spoken word-recognition. *Cognition* 25, 71-102.
- Martinet, A. 1960. *Éléments de linguistique générale*. Paris, Armand Colin.
- Martins, M. R. D. 1975. Vogais e consoantes do Português: Estatística do ocorrência, duração e intensidade. *Boletim de Filologia* 24, 1-11.
- Martins, P., I. Carbone, A. Pinto, A. Silva und A. Teixeira 2008. European Portuguese MRI Based Speech Production Studies. *Speech Communication* 50 (11), 925-952.
- Martins, P., C. Oliveira, C. Ferreira, A. Silva und A. Teixeira 2011. 3D MRI and semi-Automatic segmentation techniques applied to the study of European Portuguese lateral sounds. *Proceedings International Seminar on Speech Production*, 235-239.
- Martins, P. 2014. *Ressonância magnética em estudos de produção da fala*. Aveiro, Universidade de Aveiro. Doctoral Dissertation.
- Mateus, M. H. 2002. Variação e variedades, o caso do Português. In S. A. S. Grosse (Ed.). *Ex Oriente lux. Festschrift für Eberhard Gärtner zu seinem 60 Geburtstag*. Frankfurt am Main, Valentia.
- Mateus, M. H. und E. d' Andrade 1998. The syllable structure in European Portuguese. *D.E.L.T.A., Documentação de Estudos em Lingüística Teórica e Aplicada* 14 (1), 13-32.
- Mateus, M. H. und E. d' Andrade 2000. *The Phonology of Portuguese*. Oxford, Oxford University Press.
- Mateus, M. H. und M. R. D. Martins 1982. Contribuição para o estudo das vogais átonas [ə] e [u] no Português Europeu. *Biblos* 58, 111-125.
- Mateus, M. H., I. Falé und M. J. Freitas 2005. *Fonética e fonologia do Português*. Universidade Aberta.
- Matluk, J. 1952. La pronunciación del Español en el valle de México. *Nueva Revista de Filología Hispánica* 2, 109-120.

- Mattingly, I. G. 1981. Phonetic representation and speech synthesis by rule. In T. Myers, J. Laver und J. Anderson (Eds.). *The Cognitive Representation of Speech*, 415-420. Amsterdam, Elsevier.
- Mattos e Silva, R. V. 1991. *O Português arcaico fonologia*. Rio de Janeiro, Editora Contexto.
- Mattos e Silva, R. V. 1994. *O Português arcaico, morfologia e sintaxe*. Rio de Janeiro, Editora Contexto.
- McCarthy, J. J. und A. Prince 1993. Prosodic morphology I, Constraint interaction and satisfaction. *Linguistics Department Faculty Publication Series*. Amherst, New Brunswick, University of Massachusetts and Rutgers University.
- McCarthy, J. J. 1979. *Formal problems in Semitic phonology and morphology*. Cambridge, MA, MIT Press. Doctoral Dissertation.
- McCarthy, J. J. 1981. A prosodic theory of non-concatenative morphology. *Linguistic Inquiry* 12, 373-418.
- McCarthy, J. J. 1986. OCP effects, gemination and antigemination. *Linguistic Inquiry* 17, 207-263.
- McCarthy, J. J. 2003. *Optimality theory in phonology*. Wiley-Blackwell.
- Menendez-Pidal, R. 1926. *Orígenes del Español*. Madrid, Espasa-Calpe.
- Monaretto, V. N. O. 2009. Descrição da vibrante no Português do Sul do Brasil. In L. Bisol und G. Collischonn (Eds.). *Português no Sul do Brasil - Variação Fonológica*, 141-151. Porto Alegre, PUCRS.
- Moura, V. G. 2008. *Acordo ortográfico: A perspectiva do desastre*. Lisboa, Aletheia.
- Müller, D. 2011. *Developments of the lateral in Occitan dialects and their romance and cross-linguistic context*. Toulouse/Heidelberg, L'Université de Toulouse 2/Heidelberg. Doctoral Dissertation.
- Netto, W. F. 2001. *Introdução à fonologia da língua Portuguesa*. São Paulo, Hedra.

- Nicolaidis, K. 2001. An electropalatographic study of Greek spontaneous speech. *Journal of the International Phonetic Association* 31, 67-85.
- Nishida, G. 2009. *Análise acústica do tap em grupos no PB*. Curitiba, Universidade Federal do Paraná. Doctoral Dissertation.
- Nübling, D., A. Dammel, J. Duke und R. Szczepaniak 2006. *Historische Sprachwissenschaft des Deutschen: Eine Einführung in die Prinzipien des Sprachwandels*. Gunter Narr Verlag.
- Ó Siadhail, M. 1989. *Modern Irish, grammatical structure and dialectical variation*. Cambridge, Cambridge University Press.
- Oda, T. 2007. Further evidence on the underlying schwa of syllabic consonants in present-day English: Against the form without schwa. In O. Crouzet und J. P. Angoujard (Eds.). *Proceedings of JEL'2007 Schwa(s)*, 113-118. Nantes, University of Nantes Linguistics Department.
- Ohala, D. K. 1999. The influence of sonority on children's cluster reductions. *Journal of Communication Disorders* 32 (6), 397-422.
- Ohala, J. 1992. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. *Papers from the Parasession on the Syllable*, 319-338. Chicago, Chicago Linguistic Society.
- Ohala, J. 1993. The phonetics of sound change. In C. Jones (Ed.). *Historical Linguistics, Problems and Perspectives*, 237-278. London, Longman.
- Ohde, R. N. und D. J. Sharf 1977. Order effect of acoustic segments of VC and CV syllables on stop and vowel identification. *Journal of Speech and Hearing Research* 20 (3), 543-554.
- Öhman, S. 1967. Peripheral motor commands in labial articulation. *Speech Transmission Laboratory - Quarterly Progress and Status Report*, 30-63. Stockholm, Royal Institute of Technology.
- Oliveira, C. 2009. *Do grafema ao gesto: Contributos linguísticos para um sistema de síntese de base articulatória*. Aveiro, Universidade de Aveiro. Doctoral Dissertation.

- Oliveira, C. und A. Monteiro 2009. *Do grafema ao gesto: Contributos linguísticos para um sistema de síntese de base articulatória*. Aveiro, Universidade de Aveiro.
- Oliveira, C., P. Martins und A. Teixeira 2011. An articulatory and acoustic study of the European /l/. *Proceedings of the 17th International Congress of Phonetic Sciences*. Hong Kong.
- Oliveira, C., P. Martins, S. Silva und A. Teixeira 2012 An MRI study of the oral articulation of European Portuguese nasal vowels. *13th Annual Conference of the International Speech Communication Association*. Portland.
- Pagan, L. O. und H. F. Wertzner 2007. Análise acústica das consoantes líquidas do Português Brasileiro em crianças com e sem transtorno fonológico. *Revista da Sociedade Brasileira de Fonoaudiologia* 12 (2), 106-113.
- Pagan, L. O. 2003. *Estudo das líquidas laterais e vibrantes em crianças com distúrbio fonológico: Análise acústica e articulatória*. São Paulo, Faculdade de Filosofia, Letras e Ciências Humanas.
- Parkinson, S. 1983. Portuguese nasal vowels as phonological diphthongs. *Lingua* 61, 157-177.
- Pastore, R. E. 1987. Categorical perception: Some psychophysical models. In S. Harnad (Ed.). *Categorical Perception: The Groundwork of Cognition*, 29-52. Cambridge, Cambridge University Press.
- Pierrehumbert, J. 2000. Probabilistic phonology, discrimination and robustness. In R. Bod, J. Hay und S. Jannedy (Eds.). *Probabilistic Linguistics*, 177-228. Cambridge, MA, MIT Press.
- Pierrehumbert, J. 2001. Exemplar dynamics, word frequency, lenition and contrast. *Typological Studies in Language* 45, 137-158.
- Pierrehumbert, J. 2002. Word-specific phonetics. *Laboratory Phonology* 7, 101-139.
- Pierrehumbert, J. 2006. The next toolkit. *Journal of Phonetics* 34, 516-530.

- Pisoni, D. B. 1973. Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Attention, Perception and Psychophysics* 13 (2), 253-260.
- Pisoni, D. B. 1975. Auditory short-term memory and vowel perception. *Memory and Cognition* 3, 7-18.
- Port, R. und M. O'Dell 1985. Neutralization of syllable-final voicing in German. *Journal of Phonetics* 13, 455-471.
- Pouplier, M. und S. Benuš 2011. On the phonetic status of syllabic consonants: Evidence from Slovak. *Laboratory Phonology* 2, 243-273.
- Price, P. J. 1980. Sonority and syllabicity: Acoustic correlates of perception. *Phonetica* 37, 327-343.
- Prieto, V. P. 2004. *Fonètica i fonologia, els sons del català*. Barcelona, Editorial UOC.
- Prieto, P. und F. Torreira 2007. The segmental anchoring hypothesis revisited, syllable structure and speech rate effects on peak timing in Spanish. *Journal of Phonetics* 35, 473-500.
- Price, P. 1980. Sonority and syllabicity: Acoustic correlates of perception. *Phonetica* 37, 327-343.
- Prince, A. und P. Smolensky 2004. *Optimality theory, constraint interaction in generative grammar*. Oxford, Wiley.
- Proctor, M. 2009. *Gestural characterization of a phonological class, the liquids*. Yale, Yale University.
- Pustka, E. 2007. *Für eine perzeptive Varietätenlinguistik*. Tübingen, Narr.
- Pustka, E. 2011. *Einführung in die Phonetik und Phonologie des Französischen*. Schmidt, Berlin.
- Quillis, A. 1993. *Tratado de fonología y fonética Españolas*. Madrid, Editorial Gredos.
- Ramers, K.-H. und H. Vater 1995. *Einführung in die Phonologie*. Hürth-Efferen, Gabel.

- Ramos, A. P. 2000. Aquisição dos róticos em um dialeto do Português Brasileiro: Discussão de aspectos fonéticos e fonológicos e suas implicações para a clínica fonoaudiológica. *Revista da Sociedade Brasileira de Fonoaudiologia* 4 (6), 31-35.
- Recasens, D. 1996. *Fonètica descriptiva del català: Assaig de caracterització de la pronúncia del vocalisme i consonantisme del català al llarg del segle XX*. Barcelona, Institut d'Estudis Catalans.
- 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. 2006. Integrating coarticulation, assimilation and blending into a model of articulatory constraints. *Laboratory Phonology* 8, 611-634.
- Recasens, D. 2009. Articulatory and acoustic factors involved in the vocalization of dark /l/ and in /l/ elision in Romance. In F. Sánchez Miret (Ed.). *Romanística sin Complejos: Homenaje a Carmen Pensado*, 455-482. Frankfurt am Main, Peter Lang.
- Recasens, D. und E. Farnetani 1990. Articulatory and acoustic properties of different allophones of /l/ in American English, Catalan, and Italian. *Proceedings of the International Conference on Spoken Language Processing* 90, 961-964.
- Recasens, D. und E. Farnetani 1994. Spatiotemporal properties of different allophones of /l/, phonological implications In W. U. Dressler, M. Prinzhorn und J. R. Rennison (Eds.), *Proceedings of the 7th International Phonology Meeting*, 195-204, Turin.
- Recasens, D. und M. Pallarès 1999. A study of /j/ and /ɾ/ in the light of the ‚DAC‘ coarticulation model. *Journal of Phonetics* 27 (2), 143-169.
- Recasens, D. und A. Espinosa 2005. Articulatory, positional and coarticulatory characteristics for clear /l/ and dark /l/: Evidence from two Catalan dialects. *Journal of the International Phonetic Association* 1 (35), 1-25.

- Recasens, D., F. Sánchez Miret und K. J. Wireback 2010. *Experimental phonetics and sound change*. München, Lincom.
- Reich, U. 2004. Ritmo, saliencia prosódica y cliticación en Portugués y Español. In T. Meisenburg und M. Selig (Eds.). *Nouveaux départs en phonologie, les conceptions sub- et suprasegmentales*, 125-137. Tübingen, Narr.
- Reichel, U. 2012. PermA and Balloon, tools for string alignment and text processing, *Proceedings Interspeech*, 1874-1877. Portland.
- Repp, B. H. 1982. Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, 92 (1), 81-98.
- Repp, B. H. 1984. Categorical perception, issues, methods, findings. *Speech and Language: Advances in Basic Research and Practice*, 243-335.
- Restle, D. und T. Vennemann 2001. Silbenstruktur. In M. Haspelmath (Ed.). *Language Typology and Language Universals 2* (HSK 20.1), 1310-1336. Berlin, New York, de Gruyter.
- Reubold, U., J. Harrington, J. und F. Kleber 2010. Vocal aging effects on F0 and the first formant, a longitudinal analysis in adult speakers. *Speech Communication* 52, 638-651.
- Ridouane, R. 2008. Syllables without vowels, phonetic and phonological evidence from Tashlhiyt Berber. *Phonology* 25 (2), 321-359.
- Romano, V. und J. Silva 2010, Bicicleta ou bicicleta: Um estudo do rotacismo na fala paulista a partir de dados do atlas linguístico do Brasil. *Anais do 4. Celli*. Meringá, Universidade Estadual.
- Roon, K. D., A. I. Gafos, P. Hoole und C. Zeroul 2007. Influence of articulator and manner of stiffness. *Proceedings 16th International Congress of Phonetic Sciences*, 409-412. Saarbrücken.
- Saltzman, E. und J. A. Kelso 1987. Skilled actions: A task-dynamic approach. *Psychological Review* 94 (1), 84-92.

- Saltzman, E. L. und K. G. Munhall 1989. A dynamical approach to gestural patterning in speech production. *Ecological Psychology* 1 (4), 333-382.
- Saltzman, E., H. Nam, J. Krivokapic und L. Goldstein 2008. A task-dynamic toolkit for modelling the effects of prosodic structure on articulation. *Proceedings of the 4th International Conference on Speech Prosody*. Campinas.
- Sampson, R. 1999. *Nasal vowel evolution in Romance*. Oxford, Oxford University Press.
- Scheutz, H. 2005. Aktuell stattfindender Lautwandel / Sound change in progress. *Sociolinguistics. Soziolinguistik. An International Handbook of the Science of Language and Society. Handbücher zur Sprach- und Kommunikationswissenschaft*, 1704-1717. Berlin, New York, de Gruyter.
- Schiel, F. 1999. Automatic phonetic transcription of non-prompted speech, *Proceedings of the International Congress of Phonetic Sciences*, 607-610. San Francisco.
- Schmeiser, B. 2009a. A gestural account of Latin obstruent + lateral onset cluster realizations in Galego-Portuguese. *Actas do XXIV Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa, APL.
- Schmeiser, B. 2009b. An acoustic analysis of intrusive vowels in Guatemalan Spanish /rC/ clusters. *Proceedings of the 32nd Annual Penn Linguistics Colloquium* 15 (1), 451-485.
- Schuchardt, H. 1867. *Der Vokalismus des Vulgärlateins*. Leipzig, Teubner.
- Schwarzinger, T. 2006. Morphonotaktik bei Konsonantengruppen im Portugiesischen. *Wiener Linguistische Gazette* 73, 20-46.
- Selkirk, E. O. 1986. *Phonology and syntax: The relation between sound and structure*. Cambridge, MA, MIT Press.

- Serrano, J. 2006. En torno a las vocales vaedizas del Español Mexicano: Una aproximación sociolingüística. In P. Butragueño (Ed.). *Estudios de variación y cambio*, 37-59. México, El Colegio de México.
- Shaw, J. A. 2011. *The temporal organization of syllabic structure*. New York University. Doctoral Dissertation.
- Shaw, J. A., A. Gafos, P. Hoole und C. Zeroual 2009. Syllabification in Moroccan Arabic, evidence from patterns of temporal stability in articulation. *Phonology* 26, 187-215.
- Shaw, J. A und A. I. Gafos 2010. Quantitative Evaluation of Competing syllable parses. *Proceedings of the 11th Meeting of the ACL Special Interest Group on Computational Morphology and Phonology*, 54-62.
- Shaw, J. A., A. I. Gafos, P. Hoole und C. Zeroual 2011. Dynamic invariance in the phonetic expression of syllable structure: A case study of Moroccan Arabic consonant clusters. *Phonology* 28 (3), 455-490.
- Sievers, E. 1893. *Grundzüge der Phonetik*. Leipzig, Breitkopf und Härtel.
- Silva, A. 1996. *Para a descrição fonético-acústica das líquidas no Português Brasileiro: Dados de um informante paulistano*. Campinas, UNICAMP. Doctoral Dissertation.
- Silva, D. J. 1997. The variable deletion of unstressed vowels in Faialense Portuguese. *Language Variation and Change* 9 (3), 295-308.
- Silva, D. J. 1998. Vowel lenition in São Miguel Portuguese. *Hispania* 81, 166-178.
- Silveira, F. und I. Seara 2009. A vogal epentética em encontros consonantais heterosilábicos no Português Brasileiro: Um estudo experimental. *Revista do GEL* 2 (6), 9-35.
- Sproat, R. und O. Fujimura 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21, 291-311.

- Steriade, D. 1999. Alternatives to the syllabic interpretation of consonantal phonotactics. In O. Fujimura, B. Joseph und B. Palek (Eds.) *Proceedings of the 1998 Linguistics and Phonetics Conference*, 205-242. Prague, The Karolinum Press.
- Steriade, D. 2000. Paradigm uniformity and the phonetics-phonology boundary. *Papers in Laboratory Phonology* 5, 313-334.
- Szczepaniak, R. 2007. *Der phonologisch-typologische Wandel des Deutschen von einer Silben- zu einer Wortsprache*. Berlin, New York, de Gruyter.
- Tabain, M. 2000. Coarticulation in CV syllables: A comparison of locus equation and EPG data. *Journal of Phonetics* 28 (2), 137-159.
- Teixeira, A., P. Martins und C. Oliveira 2011. 3D MRI and semi-automatic segmentation techniques applied to the study of European Portuguese lateral sounds. *Proceedings International Seminar on Speech Production*, 243-247. Montreal.
- Teixeira, A. und F. Vaz 2001. European Portuguese nasal vowels: An EMMA study. *Seventh European Conference on Speech*, 1843-1846.
- Teruya, N. M. und H. F. Wertzner 2000. *Análise acústica das líquidas do Português Brasileiro*. São Paulo.
- Tesar, B. und P. Smolensky 2008. Learnability in optimality theory. In J. J. McCarthy (Ed.). *Optimality Theory in Phonology*, 118-140. Oxford, Blackwell.
- Torreira, F. und M. Ernestus 2011. Vowel elision in casual French: The case of vowel /e/ in the word c'était. *Journal of Phonetics* 39, 50-58.
- Trubetzkoy, N. 1939. *Grundzüge der Phonologie*. Prag. (Nachdruck (1989), Göttingen, Vandenhoeck und Ruprecht.
- Vasishth, S. und M. Broe 2010. *The foundations of statistics: A simulation-based approach*. Berlin, Heidelberg, Springer.
- Veloso, J. 2003. *Da influência do conhecimento ortográfico sobre o conhecimento fonológico*. Porto, FLUP. Doctoral Dissertation.

Veloso, J. 2006. Reavaliando o estatuto silábico das seqüências obstruinte + lateral em Português Europeu. *D.E.L.T.A., Revista de Estudos em Linguística Teórica e Aplicada* 22 (1), 127-158.

Veloso, J. 2007a. *Da influência do conhecimento ortográfico sobre o conhecimento fonológico*. München, Lincom.

Veloso, J. 2007b. Schwa in European Portuguese: The phonological status of [ɨ]. In O. Crouzet und J.-P. Angoujard (Eds.). *Actes des Proceedings of JEL*, 55-60. Nantes, Université de Nantes.

Vennemann, T. 1978. Universal Syllabic Phonology. *Theoretical Linguistics* 3, 175-215.

Vennemann, T. 1987. Muta cum liquida: Worttrennung und Syllabierung im Gotischen. Mit einem Anhang zur Worttrennung in der Pariser Handschrift der althochdeutschen Isidorübersetzung, *Zeitschrift für deutsches Altertum und deutsche Literatur* 116, 165-204.

Vennemann, T. 1988. *Preference laws for syllable structure and the explanation of sound change: With special reference to German, Germanic, Italian, and Latin*. Berlin-New York, de Gruyter.

Vennemann, T. 1989. Der Ursprung der Baiern in sprachwissenschaftlicher Sicht. In G. Sitzmann (Ed.). *Jahresberichte der Stiftung Aventinum* 3, 5-27. Abensberg.

Vennemann, T. 1993. Ein ubisches Lautgesetz. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 115, 367-399.

Vennemann, T. 1998. Etymology and phonotactics, Latin grandis vs. Basque handi 'big' and similar problems. *Journal of Indo-European Studies* 26, 345-390.

Viana, M., E. d' Andrade, L. Oliveira und I. Trancoso 1992. Uma questão de equilíbrio. *Actas do X Encontro da APL*. Lisboa, APL.

Vigário, M. 1998. Elisão da vogal não-recuada final e a palavra prosódica no Português Europeu. *XIII Encontro Nacional da Associação Portuguesa de Linguística*, 359-376. Lisboa, APL.

- Vigário, M. 2003. *The prosodic word in European Portuguese*. Berlin, New York, de Gruyter.
- Vigário, M. und I. Falé 1994. A sílaba no Português fundamental, uma descrição e algumas considerações de ordem teórica. *Actas do IX Encontro Nacional da APL*, 465-478.
- Vigário, M., S. Frota und M. J. Freitas 2003. From signal to grammar: Rhythm and the acquisition of syllable structure. *Proceedings of the 27th Annual Boston University Conference on Language Development*, 809-821.
- Warner, N., A. Jongman, A. Cutler und D. Mücke 2001. The phonological status of Dutch epenthetic schwa. *Phonology* 18 (3), 387-420.
- Wertzner, H. F., L. Pagan-Neves und M. Castro. 2007. Acoustic analysis and stimulability index for Brazilian Portuguese liquid sounds. *Revista CEFAC* 9 (3), 339-350.
- Wetzels, W. L. 1997. The lexical representation of nasality in Brazilian Portuguese. *Probus* 9, 203-232.
- Whitney, W. D. 1874. *Oriental and linguistic studies*. New York, Scribner, Armstrong, and Company.
- Williams, E. B. 1961. *Do Latim ao Português: Fonologia e morfologia históricas da Língua Portuguesa*. Rio de Janeiro, Tempo Brasileiro.
- Williams, E. B. 1968. *From Latin to Portuguese*. Philadelphia, University of Pennsylvania Press.
- Wireback, K. J. 1997. *The role of phonological structure in sound change from Latin to Spanish and Portuguese*. New York, Peter Lang.
- Zierdt, A. 2007. Ema and the crux of calibration. *XVIIth International Congress of Phonetic Sciences*, 593-596.