

Conceição Cunha (München)

captar – cap(i)tar, facto – fato:

Variation und Wandel bei Plosivsequenzen im Portugiesischen

The present study is concerned with the production and perception of the stop consonant clusters /pt/ and /kt/ as well as CVC-Sequences /pVt/ and /kVt/ in European and Brazilian Portuguese (EP and BP, respectively). European and Brazilian Portuguese have the same syllable structure (Bisol 1999: 731, Mateus/d'Andrade 2000: 39), but are postlexically affected by contrary phenomena. After the occurrence of phonological processes such as vowel reduction and deletion in European clusters and vowel epenthesis in Brazilian consonant clusters, the difference between lexical consonant clusters and CVC-sequences would be diminished in both varieties, so that EP would realize both as CC, while BP would realize both as CVC. In order to test whether clusters and CVC-sequences can be distinguished in production and perception, we discuss a physiological experiment and a perceptual study with participants of both varieties. The results show less overlap in BP than in EP. The reason for which is seen in the perception of intervocalic epenthetic elements even in lexical clusters in BP and more consonant clusters in EP.

1. Einleitung

Plosive werden mit einem vollständigen Verschluss gebildet und können benachbarte Laute beeinflussen. Dieser Effekt fällt bei Vokalen am geringsten aus, da Vokale ohne eine starke Konstriktion produziert werden (Öhman 1967: 47) und ihr Verlauf gleichzeitig mit der Bewegung für den Plosiv beginnen kann, ohne dass dabei die Wahrnehmung beider Segmente beeinträchtigt wird. Aus diesem Grund ist bei der aufeinanderfolgenden Produktion eines Konsonanten und eines Vokals (CV) die parallele Übertragung maximal (Mattingly 1981: 417). Bei Konsonantensequenzen oder Clustern (CC) ist dies nicht der Fall, da Konsonant-

en andere Konsonanten unterschiedlich stark beeinflussen bzw. maskieren können. Plosivlaute maskieren andere Konsonanten am stärksten (id.), so dass ein Verschlusslaut das vorangehende Element komplett verdecken kann. In einem reinen Konsonantencluster /C₁C₂/ kommen Plosive daher relativ selten vor und reine Plosivcluster sind in den Sprachen der Welt vor allem in wortinitialer und finaler Position extrem selten (Ladefoged/Maddieson 1996: 128). Diachronisch verhalten sich reine Plosivcluster selbst in wortmedialer Position relativ instabil.

1.1 Variabilität bei Plosivclustern in Diachronie und Synchronie

Plosivcluster aus dem Klassischen Latein wurden im Italienischen zu Geminaten (KL FACTUM > it. *fatto* 'Fakt'; KL SCRIPTUM > it. *scritto* 'geschrieben') und im Portugiesischen wurde bei den Erbwörtern¹ der erste Plosiv entweder vokalisiert oder getilgt, bei später aufgenommenen gehobenen Wörtern – den so genannten entlehnten Kultismen oder Latinismen – wurde ebenfalls der erste Plosiv getilgt. Die Übernahme und Eingliederung der intermediären Cluster in das portugiesische Lexikon fand erst für gelehrte Wörter statt, die keine mündliche Übertragung durchmachten, sondern direkt aus dem Klassischen Latein ab dem 16. Jahrhundert entlehnt wurden. Bei der Übertragung von Lexemen mit /pt/-Clustern wurden die folgenden vier Möglichkeiten beschrieben (Castro 1991: 104-126, Williams 1968: 86):

¹ Um die 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*), 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 (Jensen 1999: 46-47, Williams 1968: 28-29). Dies ist daran erkennbar, dass die gehobenen Wörter abweichende marginale Lösungen für die lautliche Entwicklung aufweisen, die wenig über das Sprachsystem aussagen.

- (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 *APTARE* > port. *atar* ‘binden, schnüren’ und KL *SCRIPTUM* > 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* > aport. *auto* > port. *apto* ‘fähig’ und KL *AD-OPTARE* > aport. *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* > aport. und port. *conceito* ‘Begriff’ und KL *SEPTIMUM* > aport. *seitimo* > port. *séptimo* ‘siebte’).
- (iv) In gehobenen Wörtern wurde /pt/ beibehalten (KL *CAPTARE* > port. *captar* ‘fangen’)

Das Cluster /kt/ wurde zu zwei verschiedenen hohen Vokalen vereinfacht (Williams 1968: 95):

- (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* > aport. *fruito*², 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 festzu-

² Im Galicischen ist dies auch heute noch die Form des Lexems.

stellen (KL DOCTOREM > port. *doutor* ‘Doktor’ und KL CONTRACTUM > aport. *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. *actolato* ‘Akt’ und KL FACTUM > port. *factofato* ‘Fakt’).

Auch synchronisch weisen Plosivcluster 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 (cf. Mateus/d’ Andrade 2000: 89). 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* > *estresse* (BP)/*stress* (EP) und eng. *scanner* > [is'kaner] *escaner* (BP) und ['skaner] *scaner* (EP).

1.2 Postlexikalische Plosivcluster in EP

Im EP wurde ebenfalls die gegensätzliche Tendenz festgestellt: Die unbetonten hohen Vokale (/u, i/) können elidiert werden, so dass sowohl Sequenzen mit intermediärem /i/ (geschrieben $\langle C_1eC_2 \rangle$), 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 [tə'ptar] und [kõ'ptar]) realisiert werden können. Empirische Belege für diese Vokaltilgung findet man bereits in den 1970er Jahren: Martins (1975: 6) führte die erste empirische Studie zur Vokaltilgung durch. Die Ergebnisse zeigen, dass von den 394 erwarteten Vokalen 49 nicht realisiert wurden, was insgesamt einer Tilgung von 11% entspricht. Die nasalen Vokale [i, 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 [i], gefolgt von [a]. Eine spätere Perzeptionsstudie zeigte, dass die Teilnehmer *prece* vs. *perce* identifizieren konnten, aber für die velaren Stimuli (*crer* vs. *querer*) wurde als Antwort meist 'eine Silbe' ausgewählt 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. Dabei zeigte /i/ den größeren Anteil von Tilgungen sowohl in finaler als auch in nicht-finaler Position, gefolgt von /u/. Der Vokal /v/ wurde hingegen in den meisten Fällen realisiert (Silva 1997: 303). Eine weitere Studie zum EP von Porto zeigte ein klar geteiltes Bild: Einerseits werden [i, u] in über 80% der phonologisch erwarteten Positionen getilgt, während [i, v, a, ε, ɔ] eine Tilgung von unter 10% aufweisen (Cunha 2011a: 146).

1.3 Stand der Forschung und Fragestellung

Im BP werden lexikalische Cluster und CVC-Sequenzen mit einem intermediären Vokal beschrieben, während im EP – mit Ausnahme vereinzelter älterer Sprecher (Veloso 2007: 58) – beide aufgrund der häufigen Tilgung der lexikalisch unbet-

onten Vokalen /i, u/ als Konsonantengruppe auftreten. Dies hat zur Folge, dass trotz der angenommenen unterschiedlichen Koordinierung von Clustern und CVC-Sequenzen diese innerhalb einer Varietät beinahe homophon sind und die Unterschiede zwischen lexikalischen Kategorien innerhalb einer Varietät kleiner ausfallen als die Variation zwischen den beiden Varietäten.

Das Auftreten intermediärer Vokale bei Konsonantenclustern wurde auch in anderen Sprachen beschrieben, z.B. *Ingalaterra* statt *Inglaterra* 'England', *egelesia* statt *eglesia* 'Kirche', *coronica* statt *cronica* 'Chronik' im Spanischen (cf. Menendez-Pidal 1926: 127) oder [blo] vs. [belo] für *blow* im Englischen (cf. Ohala 1992: 409, Browman/Goldstein 1990: 101-102, Price 1980: 331-139). Diese Realisierungen wurden bislang mit phonologischen Prozessen wie Vokalepenthese, Konsonanten- und Vokaltilgung beschrieben. Da Plosivcluster aufgrund ihrer Sonorität nicht in ein Cluster integriert werden können, sie ebenso wenig zum Auslaut der letzten Silbe gehören, weil aufgrund von phonotaktischen 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 2000: 228, Bisol 1999: 731). Bei der Vokaltilgung bleiben die Silben erhalten, da der Vokal jederzeit hinzugefügt werden kann und somit in der Repräsentation des Wortes erhalten bleibt (cf. Mateus/d'Andrade 2000: 96-101).

Eine weitere mögliche Erklärung für die letzteren Aussprachevarianten ist, dass die längere Dauer des Sonoranten nach dem bilabialen Plosiv in *blow* bei den Hörern ein anderes perceptives Konstrukt induziert, das besser mit dem akustischen Signal übereinstimmt (Ohala 1992: 409). Mit anderen Worten, Hörer interpretieren den verlängerten Sonorant als Abfolge von Schwa und Sonorant ([əɫ]). Dies kann eintreten, weil z.B. VC-Transitionen weniger wichtig für die Wahrnehmung der einzelnen Segmente sind als CV-Transitionen (Ohde/Sharf 1977: 550) oder weil Schwa im Englischen ohne ein vokalisches Target (Browman/Goldstein 1990: 101-102) 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 2011: 550), beschreibbar im Rahmen der Artikulatorischen Phonologie (Browman/Goldstein 1986: 9-252, 1988: 149-151, 2000: 27-31). 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.

Frühere Studien über Konsonantenkoordination von englischen Clustern und CVC-Sequenzen wiesen darauf hin, dass Probanden bei der Überlappung beider Konsonantengesten eher ein Cluster und bei Nicht-Überlappung eher eine CVC-Sequenz wahrnahmen (Browman/Goldstein 1990: 101-102). Bei portugiesischen Clustern und CVC-Sequenzen mit intermediären /e, u/ ($\langle C_1eC_2 \rangle$ vs. $\langle C_1uC_2 \rangle$) wurden in Cunha (2012: 108-144, in *Begutachtung*) kleine, aber konsistente Unterschiede bei der Koordinierung von Plosiven und Liquiden festgestellt, so dass jeder Stimulustyp auch in Abhängigkeit weiterer Faktoren in der Produktion unterschieden werden konnte. Die Frage zu beantworten, ob dieser auch bei Plosivtokens der Fall sein wird, ist das vorrangige Ziel dieser Untersuchung.

Das zweite Ziel ist es, am Beispiel derselben Varietäten EP und BP zu überprüfen, ob lexikalische Cluster und lexikalische CVC-Sequenzen in Bezug auf möglicherweise neutralisierte Formen in der Perzeption identifiziert werden können. In diesem Zusammenhang wird auch der Frage nachgegangen, inwiefern dabei der Zusammenhang zwischen intendierten³ und wahrgenommenen Token durch Koordinationsunterschiede erklärt werden kann. Sollte sich die Alternanz zwischen lexikalischen Clustern und Clustern mit einem intermediären epenthe-

³ Der Begriff 'intendiert' bezieht sich in dieser Arbeit nicht auf kognitive Prozesse beim Sprechen, sondern bezeichnet einfach die zugrundeliegende Form, die die Probanden laut Stimuluspräsentation produzieren sollten.

tischen Vokal (CəC) sowie lexikalischen CVC-Sequenzen und deren reduzierten Formen (CC) vor allem durch die zeitliche Koordinierung ihrer Konsonanten unterscheiden, stellt sich die Frage, wie stabil diese Unterscheidung vor allem mit Hinblick auf Sprachentwicklung und Wandel ist.

So untersucht die hier vorliegende Studie konkret, inwiefern sich die Koordination von medialen, heterosilbischen, lexikalischen Plosivclustern vs. <C₁et> vs. <C₁ut> (C = /p,k/) in den portugiesischen Varietäten (EP vs. BP) unterscheidet und inwiefern die große Variabilität auf die unterschiedliche Koordination der einzelnen Konsonanten miteinander zurückzuführen ist. Dabei wird die Rolle des Artikulationsorts des ersten Konsonanten (labial vs. velar) und der Betonung (betont vs. unbetont) miteinbezogen. Der Einfluss des Artikulationsorts wird mit dem sogenannten Produktionsrichtungs-Effekt (Chitoran et al. 2002: 436-444) assoziiert. Der Produktionsrichtungseffekt besagt, dass sich zwei Plosive in der *front-to-back*-Richtung stärker überlappen können, als in der umgekehrten Produktionsrichtung (*back-to-front*). Der Grund dafür ist die leichtere Wiederherstellbarkeit der Segmente in der *front-to-back*-Richtung im Vergleich zur *back-to-front*-Richtung, möglicherweise aufgrund der visuellen Unterstützung des ersten Plosivs bei der Produktion von C₁ im vorderen Bereich des Mundraums. Folglich verringert sich die Gefahr, dass C₁ aufgrund seiner nicht-parallelen Übertragbarkeit (Mattingly 1981: 417) von C₂ verdeckt wird und damit seine Verschlusslösung akustisch/bei der Perzeption verloren geht, bei einer Produktion im vorderen Mundbereich. Aufgrund der zusätzlichen visuellen Information könnten Hörer diesen vorderen Konsonanten (womöglich unbewusst) wahrnehmen. Infolgedessen ist es unwahrscheinlicher, dass C₁ bei einer ähnlichen Überlappung der Gesten im vorderen als im hinteren Mundbereich verloren geht, bei dem dieser visuelle Hinweis fehlt. Da sich dieser Effekt auf Plosiv-Token besonders stark auswirken könnte, wird ebenfalls mehr Überlappung bei den *front-to-back* Token als bei den *back-to-front* Token erwartet.

Zuletzt wird der Einfluss von Betonung untersucht. Da die untersuchte Vokaltilgung ausschließlich in unbetonter Position auftritt, wurde die Betonungskondition für CVC-Sequenzen mithilfe von C₂ definiert. Dabei werden

also Sequenzen mit C_2 in betonter (z.B. *computar*, ‘rechnen’) und in unbetonter Silbe (z.B. *computador*, ‘Rechner’) verglichen. Der Zielvokal musste 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: 49, Byrd 1993: 112, Iskarous/Kavitskaya 2010: 637, Tabain 2000: 148), was mit der steigenden Überlappung der Konsonantengesten in diesem Kontext erklärt wurde (z.B. Browman/Goldstein 1991: 321). Infolgedessen wird ebenso mehr Überlappung bei den unbetonten als bei betonten Token vorhergesagt.

2. Produktionsexperiment

2.1 Methode

Für die Beantwortung dieser Fragestellungen wurden akustische und artikulatorische Daten mithilfe des 5D-Artikulatographen AG500 der Firma Carstens Medizinelektronik GmbH am Institut für Phonetik und Sprachverarbeitung in München erhoben (cf. Cunha 2012: 55 für eine detaillierte Beschreibung des gesamten Experiments und der Vorteile / Beschränkungen dieser Methode).

2.1.1 Sprachmaterial

Zur Erstellung unserer Stimuli wurden die zu analysierenden Cluster und CVC-Sequenzen in Zielwörter integriert, die in Tabelle 1 aufgelistet werden. Sie wurden in folgenden Trägersatz *Ele leu ___ bem* (‘Er las ___ richtig’) eingebettet, randomisiert und von einem Computerbildschirm abgelesen.

Cluster	compact <u>ado</u> compact <u>ador</u>	‘kompakt’ ‘Verdichter’	cap <u>tado</u> cap <u>tador</u>	‘empfangen’ ‘Empfänger’
<C₁et>	que <u>tado</u> ¹ que <u>tadado</u>	‘wie doof’ ‘wie doof’	tap <u>etado</u> tap <u>etador</u>	‘mit Teppich’ ‘Teppichverleger’
<C₁ut>	cut <u>ano</u> cut <u>anoso</u>	‘Eigename’ ‘E., Genitiv’	comput <u>ado</u> comput <u>ador</u>	‘ausgerechnet’ ‘Rechner’

Tabelle 1: Liste der aufgenommenen Plosiv-Token mit den Zielclustern und Sequenzen (hier *kursiv* dargestellt) und den betonten Silben (als unterstrichene Silben markiert).

Die fünf EP-Sprecher produzierten jeweils acht Wiederholungen von jedem Satz und die vier BP-Sprecher jeweils neun 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 (12 Wörter x 5 Sprecher x 8 Wiederholungen) und für das BP insgesamt 432 Stimuli (12 Wörter x 4 Sprecher x 9 Wiederholungen).

2.1.2 Auswertung

Die statistische Auswertung wurde mit 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. /t/) und Betonung (betonte vs. unbetonte Kondition). Die abhängige Variable war die wie folgt definierte relative Überlappung:

$$\text{ÜBERLAPPUNG} = 1 - (C_2.\text{START} - C_1.\text{TARGET}) / (C_1.\text{LÖSUNG} - C_1.\text{TARGET})$$

2.2 Ergebnisse

Abbildung 1 veranschaulicht die Synchronisierungsmittelwerte von fünf EP-Sprechern (links) und vier BP-Sprechern (rechts). Dabei geht es um die Überlappung von C₂ in Bezug auf C₁, d.h. es wurde untersucht, ob C₂ vor, während

oder nach dem C_1 -Plateau begann. Diese Messwerte entsprechen einer starken Überlappung bzw. einer engen Koordinierung, einer normalen Überlappung und schließlich keiner Überlappung bzw. einer weiten Koordinierung. In der oberen Hälfte befinden sich die betonten, in der unteren die unbetonten Token.

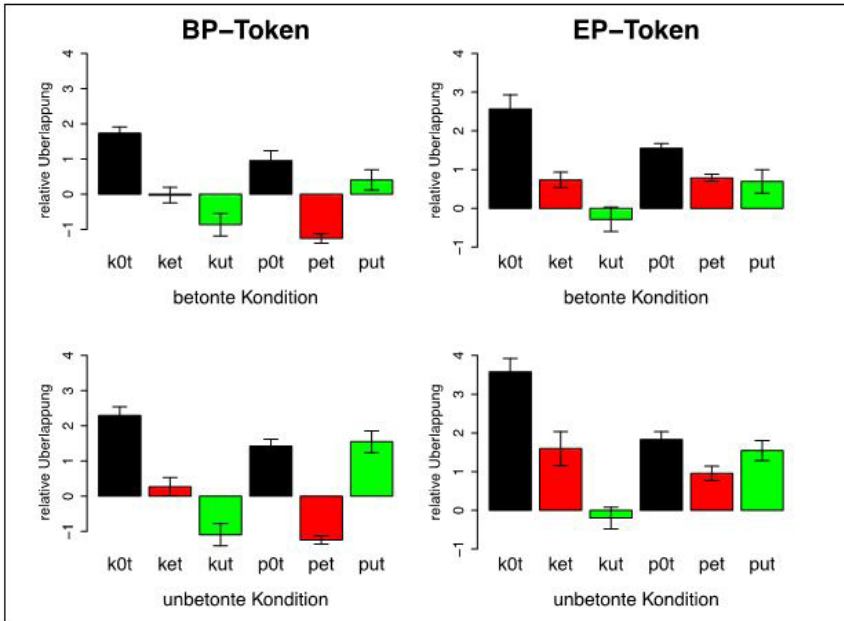


Abbildung 1: Übersicht der Synchronisierung der Gesten nach Varietäten und Betonungsbedingungen getrennt (Mittelwerte über fünf EP und vier BP Sprecher mit Standardabweichungen)

Betrachtet man die Gesamtergebnisse, so besitzen die Sprechervarietäten, die eine größere Überlappung bei EP als bei BP zeigen ($\chi^2[1] = 13.1, p < 0.01$), und die Stimulustypen, die eine größere Überlappung bei Clustern als bei den anderen Sequenzen zeigen ($\chi^2[2] = 226.1, p < 0.001$), signifikante Haupteffekte auf die Synchronisierung beider Plosive. Diese werden allerdings von der signifikanten Interaktion ($\chi^2[2] = 20.9, p < 0.01$) zwischen beiden Faktoren eingeschränkt. In der Koordination der Plosive wurden die gleichen Tendenzen in beiden Varietäten ersichtlich: Plosivcluster zeigen starke Überlappung und eine

enge Koordinierung der Konsonanten. Allerdings waren die Gesten ausnahmslos stärker überlappt in EP als in BP.

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. Die Plosive in $\langle C_1uC_2 \rangle$ wurden eng koordiniert nach dem bilabialen und weit auseinander nach dem velaren Plosiv. Der Faktor Artikulationsort an sich hatte keinen Einfluss auf die Fragestellung. Der Stimulustyp hatte ebenso eine hochsignifikante Auswirkung auf die Plosivkoordination im BP, 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 beiden Varietäten in unbetonten Token früher innerhalb von C_1 als in betonten in beiden Varietäten.

Die zeitliche Koordination der Cluster wurde ebenfalls stark von C_1 und dem damit assoziierten Produktionsrichtungseffekt bestimmt, allerdings nicht in der erwarteten Richtung: Die Konsonantengesten überlappten sich unerwartet eindeutig stärker bei *back-to-front*- Clustern (/kt/) und zeigten zudem signifikant kleinere intermediäre Abstände zwischen den Plateaus als bei *front-to-back* (/pt/). Die Sprechervarietät zeigte sich als signifikanter Haupteffekt bei beiden Verfahren, bei denen die Plosivgesten im EP ausnahmslos geringer überlappten als im BP. Die Übereinstimmung der Ergebnisse beider Verfahren bekräftigt die Stabilität dieses Effekts. Angesichts der beschriebenen Vokaltilgungen im EP (Martins 1975: 6, Silva 1997: 296-306, 1998: 170-174, Cunha 2011a: 145) sowie der zahlreichen Studien zu Rhythmus-Unterschieden in den beiden Varietäten (Abaurre/Sandalo 2007: 148-152, Barbosa 2006: 175-211, Frota/Vigário, 2007: 285-291) waren die varietätsspezifischen Unterschiede nicht unerwartet.

Die Betonung beeinflusste auf signifikante Weise die Synchronisierung beider Konsonanten miteinander, bedingt durch die Tatsache, dass unbetonte Token größere positive und kleinere negative Mittelwerte aufwiesen, so dass die Überlappung ausnahmslos stärker bei unbetonten als bei betonten Token war. Dies ist möglicherweise eine Folge der deutlich stärkeren Hypo-Artikulation der Segmente (Lindblom 1990: 420-435) bzw. der damit assoziierten stärkeren Koartikulation der Segmente in nicht prominenten und vorhersagbaren Kontexten, wie es bei unbetonten Silben oft der Fall ist.

2.3 Diskussion der Ergebnisse

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. Die Vorhersage basiert auf zahlreichen Arbeiten zu Konsonantenclustern in einer Vielzahl von Sprachen (Bombien 2011: 116-178, Byrd 1993: 105-116, Hardcastle 1985: 254-258, Marin/Pouplier 2010: 392-403, Recasens/Espinosa 2005: 13-22 u.v.a.) und konkret auf Plosivclustern (Chitoran et al. 2002, Gafos et al. 2010: 667-678, Shaw et al. 2011: 467-479). Die erneute Bestätigung dieser Ergebnisse bekräftigt, wie robust die Organisation von Konsonanten in Clustern tatsächlich ist.

Es wurden außerdem erhebliche Koordinationsunterschiede zwischen beiden Varietäten festgestellt, da BP ein weiteres intergestisches Timing aufwies als EP. Dieser Effekt war über die Gesamtdaten betrachtet sehr robust und trat in beiden Messverfahren auf. Die starke Überlappung der Konsonanten im EP könnte der Grund für die häufigen Vokaltilgungen (Cunha 2011a: 145, Silva 1997: 296-306, 1998: 170-174, Vigário 1998: 359-374) in dieser Varietät sein. Wenn Konsonanten einen großen Druck auf die benachbarten Segmente ausüben, können wenig prominente Segmente graduell kürzer und im Extremfall getilgt werden. Die zeitliche Koordination wurde ebenfalls stark vom C_1 -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 aus der Untersuchung eindeutig hervor, dass /kt/trotz der laut Produktionsrichtungseffekt postulierten schlechteren perceptiven Wiederherstellbarkeit 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 rückt der Artikulationsort des Plosivs weiter nach hinten bei hinteren Vokalen (Flemming 2003: 348). Außerdem ergab sich in beiden Varietäten eine starke Überlappung beider Konsonanten in /put/, 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 Vorhersage der Sprachperzeption für die Übertragung von Plosivclustern besagt, dass sich Plosive nicht stark überlappen können, ohne die Wiederherstellbarkeit beider Segmente für den Hörer in Gefahr zu bringen. Angesichts der erheblichen zeitlichen Überlappung beider Konsonantengesten in der Produktion wird für die Perzeption erwartet, dass beide Konsonanten zu großen Teilen vom Hörer nicht mehr wiederherstellbar sind und daher als einzelne Konsonanten wahrgenommen werden.

3. Perzeptionsstudie

Das erste Ziel des Perzeptionsexperiments 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 <C₁et> vs. lexikalische <C₁ut> zu identifizieren. Hierbei fließen bisherige Forschungserkenntnisse in die zwei folgenden Hypothesen mit ein: Entweder werden kleine Unterschiede in der Koordination der Konsonanten von den Hörern nicht wahrgenommen und sie nehmen daher den Unterschied zwischen Cluster vs. <C₁et> vs. <C₁ut>

nicht wahr (Abschwächung des lexikalischen Unterschieds) oder bestimmte phonetische Details ermöglichen es den Hörern, den Stimulustyp zu identifizieren (unvollständige Neutralisierung, Port/O'Dell 1985: 463). 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.

Das zweite Ziel des Experiments war es, zu testen, ob Hörer zu jedem präsentierten Stimulustyp (Cluster vs. $\langle C_1et \rangle$ vs. $\langle C_1ut \rangle$) 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:

- (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).

3.1 Methode

3.1.1 Stimuli

Anstelle von synthetischen graduellen Stimuli wurden hier natürliche Stimuli von jeweils einem EP- und einem BP-Sprecher verwendet. Bei den Aufnahmen wurden die Zielwörter in folgende Trägersätze eingebaut, bei denen der Satz-

akzent variiert wurde. Beim ersten Satz wurde das Zielwort in akzentuierter (*O Pedro leu ___ bem*, ‘Peter las ___ richtig’) und beim zweiten in deakzentuierter Position (*O PEDRO leu ___ bem*, ‘PETER las ___ richtig’) produziert. Jede Satzfolge wurde viermal pro Sprecher in randomisierter Reihenfolge wiederholt. 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.

3.1.2 Perzeptionsstimuli

Für das Perzeptionsexperiment wurden die akustischen Stimuli aus der zweiten Wiederholung von zwei Probanden (einmal EP, einmal BP) mit MAUS (Kisler et al. 2012, Reichel 2012) vorsegmentiert und anschließend in EMU (Harrington 2010: 1-110) manuell korrigiert. Die Zielsilben wurden zwischen dem vorangehenden Vokal-Offset und dem nächsten Vokal-Onset extrahiert. Eine Auflistung der exakten Stimuli folgt in Tabelle 2:

	/k/	/p/
Plosiv-Token	/kt/	/pt/
	/ket/	/pet/
	/kut/	/put/

Tabelle 2: Liste der Produktionsstimuli für das Perzeptionsexperiment

Diese Token wurden mit weiteren Stimuli bei einer größeren Studie getestet (insgesamt fünf Blöcke: K-Plosiv, P-Plosiv, K-Liquid, P-Liquid und Minimalpaare, wovon wir hier die ersten zwei auswerten). Die Analyse berücksichtigt 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 Lautsprechers 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 (/t/) sowie drittens und viertens die entsprechenden CVC-Folgen mit jeweils /e/oder /u/in der V-Position, d.h. die Probanden hatten <Ct>, <t>, <Cet>, <Cut> (wobei $C = [p, k]$) zur Auswahl. Diese Stimuli wurden von 40 EP-(13m/27w) und 37 BP-Probanden (23m/13w) beurteilt. Alle Teilnehmer waren entweder Studenten der Universität Campinas oder der Universität Porto. Sie waren im Alter zwischen 23 und 36 Jahren und stammten aus den gleichen Städten wie beim Produktionsexperiment (Porto in Portugal und Campinas 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.

3.1.3 Statistische Auswertung

Die Antworten wurden mit einem *Generalized Linear Mixed Model* (GLMM) in R ausgewertet. Die festen Faktoren waren Sprechervarietät und Stimulustyp, Artikulationsort von C_1 , Betonung und Hörervarietät.

3.2 Ergebnisse

Die Perzeptionsantworten wurden zu den Produktionstypen für die getrennten Artikulationsarten in Abhängigkeit von Sprechervarietät, Artikulationsort von C_1 , Betonungskondition sowie Hörervarietät ausgewertet.

3.2.1 Plosivcluster

Bei den entsprechenden zwei Blöcken des Perzeptionsexperiments mussten die Teilnehmer entscheiden, ob sie den jeweiligen Stimulus als Einzelkonsonanten (<t>), Cluster (< C_1 t>) oder CVC-Sequenz mit /e/oder /u/(< C_1 et>, < C_1 ut>) wahrnahmen. Das Gesamtbild zeigt, dass alle vier Möglichkeiten ausgewählt wurden:

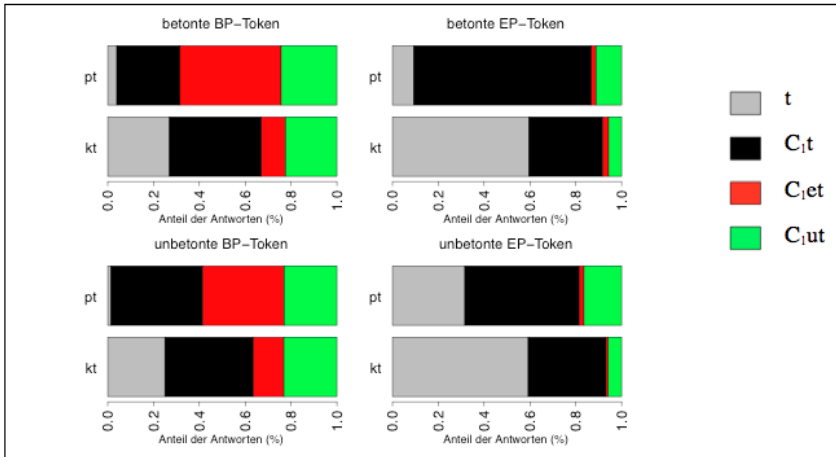


Abbildung 2: Anteilige Antworten zu den intendierten reinen Plosivclustern.

Die Ergebnisse zeigen deutlich mehr Wahrnehmungen von Einzelkonsonanten bei den EP-Stimuli als bei den BP-Stimuli. 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.

Insgesamt zeigt die Verteilung der /t/-Antworten 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 wurden im Vergleich zu bilabialen mehr einzelne Konsonanten wahrgenommen.

Die Betonungskondition hatte keine Auswirkung, aber es gab eine hochsignifikante Interaktion ($\chi^2[4] = 32.0$, $p < 0.001$), wobei die Betonung bilabiale Cluster im EP ($p < 0.001$) beeinflusste, jedoch nicht velare Cluster. Bei velaren Clustern begann C_2 bereits vor dem C_1 -Plateau und somit fand eine größere Überlappung beider Konsonantengesten statt bzw. eine engere Koordinierung beider Konsonanten, was die Verdeckung der ersten Verschlusslösung zur Folge hatte.

Dies ist möglicherweise der Grund dafür, weshalb Hörer lediglich C_2 wahrnehmen. Außerdem stellten Chitoran et al. (2002: 436-444) fest, dass Plosive in der *back-to-front*-Produktionsrichtung für den Hörer schwerer erkennbar sind als in der umgekehrten Richtung. Im BP begann C_2 entweder während oder nach dem C_1 -Plateau, weshalb BP insgesamt eine signifikant weitere Koordinierung als EP aufwies.

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

Bezüglich der Antworten mit Vokalen (/C₁et/, /C₁ut/) im Vergleich zu denen ohne Vokal (/t/und /C₁t/) zeigten die Ergebnisse 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). Dies ist in Abbildung 2 ersichtlich: Der Anteil der dreigliedrigen Antworten (/C₁et/und /C₁ut/) ist größer in den BP- als in den EP-Daten und bei bilabialen als bei velaren C₁. 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 (cf. Chitoran et al. 2002: 436-444) 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 unterschiedli-

che 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

Obwohl beide Versuchspersonen dieselben Token als Sprechvorlage hatten, zeigt sich in Abbildung 5, 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.

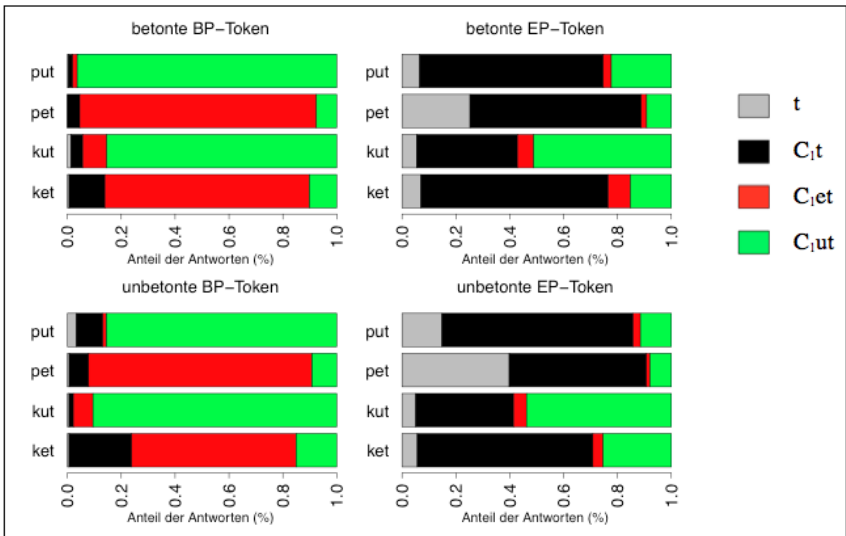


Abbildung 3: Anteilige Perzeptions-Antworten zu den intendierten CVC-Sequenzen mit Plosiven.

Aus Abbildung 3 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 C_1 -Artikula-

tionsort 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 3 daran zu erkennen, dass bilabiale Cluster einen deutlich größeren Anteil von wahrgenommenen 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: 436-444, Kühnert et al. 2006: 332) 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 ($z = 2.4$, $p < 0.05$) als auch eine signifikante Interaktion mit der Sprechervarietät ($\chi^2[5] = 6.6$, $p < 0.05$). Dies kommt darin zum Ausdruck, dass Hörergruppen sich bei den BP-Stimuli unterschieden ($p < 0.05$), aber nicht bei den EP-Stimuli.

Bezüglich der lexikalischen Vokale lässt Abbildung 3 erkennen, dass der Anteil der wahrgenommenen CVC 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%). Dies erklärt den hochsignifikanten Einfluss der Sprechervarietät ($z = 20.0$, $p < 0.001$). Auch dieser Effekt war in den Produktionsdaten ersichtlich, da intendierte CVC-Sequenzen keine Überlappung im BP und eine starke Überlappung im EP aufwiesen. 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 ($p < 0.001$ in BP und $p < 0.05$ in 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 Hörergruppen sich lediglich bei den BP-Daten ($p < 0.05$) unterschieden.

3.3 Diskussion

Die Sprechervarietät zeigte einen hochsignifikanten Einfluss auf die Identifikation des Stimulustyps Cluster vs. <C₁et> vs. <C₁ut>, 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: 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. Bisol 1999: 731-732, Mateus/d'Andrade 2000: 45). 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. <C₁et> vs. <C₁ut>) 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: Konsonanten waren 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 findet man – wie vorhergesagt – mehr intermediäre Vokale in BP-Token und mehr einzelne Konsonanten in EP-Token. <C₁et> vs. <C₁ut> 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 kann durch Koordinationsunterschiede erklärt werden sowie möglicherweise durch die in den letzten Jahrzehnten beschriebene Vokalelision im EP (Cunha 2011a: 133-148, Mateus/d'Andrade 2000: 42-45, Silva 1997: 296-306, 1998: 170-174, Vigário 1998: 359-374, Martins 1975: 1-11).

Lexikalische Cluster wurden allerdings in den EP-Token zu höheren Anteilen als in den BP-Token korrekt identifiziert. Hier sind die Bezüge zur Produktion ebenfalls deutlich, da Cluster im BP insgesamt breiter koordiniert wurden als im EP und somit CVC-Sequenzen ähnlicher sind, was wiederum dazu geführt hat, dass sie manchmal mit Letzteren verwechselt wurden. Die enge Koordinierung (bzw. extreme 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] in BP und [i] in EP eine Interaktion zwischen Sprechervarietät und Stimulustyp erwartet: Cluster vs. <C₁ut> in BP und Cluster vs. <C₁et> sollten in EP stärker neutralisiert werden. Dies war für das EP der Fall, nicht jedoch für das BP.

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 der bereits erwähnte Produktionsrichtungseffekt (Chitoran et al. 2002: 436-444). 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. Dieses Missverhältnis zwischen Produktion und Perzeption kann laut Ohala (1993) zu Lautwandel führen. Er kann dadurch entstehen, dass der Hörer in der Rolle des Sprechers 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.

Solch ein aktuell stattfindender Lautwandel im Sinne von Ohala (id.) 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 perceptiv wiederhergestellt werden können. Bei /kt/ jedoch überlappen sich der velare und der apikale Konsonant zwar weniger als im EP, aber immer noch deutlich stärker als 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. die Rezipienten brauchen einen größeren Abstand zwischen den Konstriktionplateaus und den akustischen Verschlüssen, um die Plosive zu identifizieren. Da die Verschlüsse sehr nah aufeinanderfolgten, schlossen die Hörer nicht mehr auf ein Cluster und nahmen lediglich einen Konsonanten wahr (/t/).

Eine ähnliche Divergenz 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 perzipierten. Das gleiche graduelle Lautwandelmuster zwischen Hörer und Sprecher könnte dazu geführt haben, dass ein Vokal von manchen Sprechern nach und nach eingefügt wurde. Dies wäre der Grund, weshalb in manchen Stimmhaftigkeit sowie wenige Formanten zwischen den Konsonanten zu finden sind. Mit der Betonungskondition wird der Einfluss der Betonung auf die korrekte Identifikation des Stimulustyps und auf die Anzahl der wahrgenommenen Elemente geprüft. 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 bei den Perzeptionsantworten zu einem größeren Anteil an einzelnen Konsonanten in unbetonter Position. Insgesamt

hatte die Hörervarietät nur einen geringen Einfluss auf die korrekte Identifikation des Stimulustyps trotz große Unterschiede bei der Produktion.

4. Schlussfolgerung und Ausblick

Aus den durchgeführten Experimenten 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 korrekt erkannt als Sequenzen mit Vokal (<C₁et> und <C₁ut>), was mit der Realisierung von epenthetischen Vokalen in diesem Kontext in Verbindung gebracht wurde. 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/ in 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/), was die Vorhersage des Produktionsrichtungseffekts (Chitoran et al. 2002: 436-444) widersprach. Daher waren Hörer nicht in der Lage, für diese Variabilität zu kompensieren (siehe Mann/Repp 1980 für Kompensationseffekte) und nahmen weit koordinierte /pt/-Sequenzen überwiegend mit einem intermediären Vokal war und eng koordinierte /kt/-Sequenzen teilweise als einzelnen Konsonant. 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. Solche 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 <C₁et>, wurden – wie erwartet – größtenteils als reine Konsonantenabfolge wahrgenommen. Dies zeigt, dass der lexikalische Unterschied zwischen Clustern und Sequenzen auf kontinuierliche Weise in EP schwächer wurde, so dass Cluster vs. <C₁et> in der Perzeption schwieriger zu unterscheiden sind als Cluster vs. <C₁ut>. Diese Unterschiede stimmen mit den Ergebnissen aus der Produktion überein, insofern als die Synchronisierung von Cluster vs. <C₁et> weniger Unterschiede aufwies als Cluster vs. <C₁ut>. Aufgrund der Produktionsunterschiede wurde vermutet, dass Muttersprachler der beiden Varietäten möglicherweise die Token im mentalen Lexikon im Sinne der Exemplar-Theorie (Johnson 1997: 101-113, Pierrehumbert 2003: 189-211) 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.

Die Alternanz zwischen einem und zwei Plosiven im aktuellen portugiesischen Schriftbild konnte ebenso mit Koordinationsunterschieden in Verbindung gebracht werden und insbesondere auf die Diskrepanz zwischen Produktion und Perzeption hinweisen. Im Zusammenhang mit dieser Diskrepanz 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 mit neuen Hypothesen und Erklärungsvorschlägen die historische Sprachwissenschaft augenscheinlich bereichern können.

Bibliographie

- Abaurre, Maria Bernardete/Sândalo, Filomena. 2007. «Acento secundário e duas variedades do português: uma análise baseada na OT». In: Araújo, Gabriel Antunes (Hg.): *O acento em português: abordagens fonológicas*. São Paulo: Editora Parábula, 145-168.
- Barbosa, Plínio. 2006. *Incursões em torno do ritmo*. Campinas: Pontes.
- Beckman, Mary et al. 1992. «The interaction of coarticulation and prosody in sound change». In: *Language and Speech* 35, 45-58.
- Bisol, Leda. 1999. «A sílaba e seus constituintes». In: *Gramática do Português Falado* 7, 701-741.
- Bombien, Lasse. 2011. *Segmental and prosodic aspects in the production of consonant clusters*. München: Universitätsbibliothek.
- Browman, Catherine/Goldstein, Louis. 1986. «Towards an articulatory phonology». In: *Phonology yearbook* 3/21, 219-252.
- , 1988. «Some notes on syllable structure in articulatory phonology». In: *Phonetica* 45/2-4, 140-155.
- , 1990. «'Targetless' Schwa: An Articulatory Analysis». In: *Haskins Laboratories Status Report on Speech Research*. Bd. SR-101/102.
- , 1991. «Gestural structures: Distinctiveness, phonological processes, and historical change». In: *Modularity and the motor theory of speech perception*, 313-338.
- , 2000. «Competing constraints on intergestural coordination and self-organization of phonological structures». In: *Bulletin de la Communication Parlée* 5, 25-34.
- Byrd, Dany. 1993. «54,000 American Stops». In: *UCLA Working Papers in Phonetics* 83, 97-116.
- Castro, Ivo. 1991. *Curso de História da Língua Portuguesa*. Lisboa: Universidade Aberta.
- Chitoran, Ioana et al. 2002. «Gestural Overlap and Recoverability: Articulatory Evidence from Georgian». In: *Laboratory Phonology*, 419-448.

- Cunha, Conceição. 2011a. «A elisão vocálica no Português Europeu». In: Arden, Mathias et al. (Hg.) In: *Linguística do português. Rumos e pontes*. München: Meidenbauer, 133-148.
- , 2011b. «Die Perzeption von lexikalischen Plosivclustern und CVC Sequenzen im Portugiesischen». *JournalLIPP* 1, 71-88.
- , 2012. *Die Organisation von Konsonantenclustern und CVC-Sequenzen in zwei portugiesischen Varietäten*. Ludwig-Maximilians-Universität München: Noch unveröffentl. Dissertationsschrift.
- Cunha, Conceição/Harrington, Jonathan. 2011. «The perception of /pt/and /kt/ in European and Brazilian Portuguese». In: *Proc. of 17th. International Congress of Phonetic Sciences*, 548-552.
- Flemming, Edward 2003. «The Relationship Between Coronal Place and Vowel Backness». In: *Phonology* 20/3, 335-373.
- Frota, Sónia/Vigário, Marina. 2007. «Intonational Phrasing in two varieties of European Portuguese». In: Gussenhoven, Carlos/Riad, Tomas (Hg.): *Tones and Tunes. Vol. I*. Berlin: Mouton de Gruyter, 265-291.
- Gafos, Adamantios. 2002. «A grammar of gestural coordination». In: *Natural Language and Linguistic Theory* 20, 269-337.
- Gafos, Adamantios et al. 2010. «Variation in overlap and phonological grammar in Moroccan Arabic clusters». In: *Laboratory Phonology* 10, 657-698.
- Hardcastle, William. 1985. «Some phonetic and syntactic constraints on lingual coarticulation during /kl/sequences». In: *Speech Communication* 4/1-3, 247-263.
- Harrington, Jonathan. 2010. *Phonetic Analysis of Speech Corpora*. Malden: Wiley-Blackwell.
- Iskarous, Khalil/Kavitskaya, Darya. 2010. «The interaction between contrast, prosody, and coarticulation in structuring phonetic variability». In: *Journal of phonetics* 38/4, 625-639.
- Jensen, Frede. 1999. *A comparative study of romance*. New York: Peter Lang Publishing.

- Johnson, Keith. 1997. «The auditory/perceptual basis for speech segmentation». In: *OSU Working Papers in Linguistics* 50, 101-113.
- Kisler, Thomas et al. 2012: «Signal processing via web services: the use case WebMAUS». In: *Proceedings Digital Humanities*, 30-34.
- Koike, Dile/Macedo, Donaldo. 1992. *Romance Linguistics: The Portuguese Context*. Greenwood Publishing Group.
- Kühnert, Barbara et al. 2006. «Gestural overlap and C-center in selected French consonant clusters». In: *Proceedings 7th International Seminar on Speech Production, UFMG Belo Horizonte*, 327-334.
- Ladefoged, Peter/Maddieson, Ian. 1996. *The sounds of the world's languages*. Cambridge: Blackwell.
- Lindblom, Björn. 1990. «Explaining Phonetic Variation: a sketch of the H & H theory». In: Hardcastle, William/Marchal, Alain (Hg.): *Speech Production and Speech Modeling*. Dordrecht: Kluwer, 403-439.
- Mann, Virginia/Repp, Bruno. 1980. «Influence of Vocalic Context on the Perception of [f-s] Distinction: I. Temporal factors». In: *Perception & Psychophysics* 28, 213-228.
- Martins, Maria Raquel Delgado. 1975. «Vogais e Consoantes do Português: Estatística do Ocorrência, Duração e Intensidade». In: *Boletim de Filologia* 24, 1-11.
- Marin, Stefania/Poupplier, Marianne. 2010. «Temporal Organization of Complex Onsets and codas in American English: Testing the predictions of a gestural coupling model». In: *Motor Control* 14 (3), 380-407.
- Menendez-Pidal, Ramón. 1926. *Orígenes del Español*. Madrid: Espasa-Calpe.
- Mateus, Maria Helena. 2002. «Variação e Variedades: o caso do Português» In: Grosse, Sybille/Schönberger, Axel (Hg.): *Ex Oriente lux. Festschrift für Eberhard Gärtner zu seinem 60. Geburtstag*. Frankfurt am Main: Valentia, 287-296.
- Mateus, Maria Helena/d' Andrade, Eugénio. 2000. *The Phonology of Portuguese*. Oxford: Oxford University Press.

- Mateus, Maria Helena/Martins, Maria Raquel Delgado. 1982. «Contribuição para o Estudo das Vogais Átonas [ə] e [u] no Português Europeu». In: *Biblos* 58, 111-125.
- Mattingly, Ignatius. 1981. «Phonetic Representation and Speech Synthesis by Rule». In: Anderson, John et al. (Hg.): *The cognitive representation of speech*. Amsterdam: North-Holland, 415-420.
- Ohala, Diane. 1999. «The Influence of Sonority on Children's Cluster Reductions». In: *Journal of Communication disorders* 32 (6): 397-422.
- Ohala, John. 1992. «Alternatives to the Sonority Hierarchy for Explaining Segmental Sequential Constraints». In: Deaton, Karen et al. (Hg.): *Papers from the Parasession on the Syllable in Phonetics & Phonology*. Chicago: Chicago Linguistic Society, 319-338.
- , 1993. «The Phonetics of Sound Change». In: Jones, Charles (Hg.): *Historical Linguistics: Problems and Perspectives*. London: Longman, 237-278.
- Ohde, Ralph/Sharf, Donald. 1977. «Order Effect of Acoustic Segments of VC and CV Syllables on Stop and Vowel Identification». In: *Journal of Speech and Hearing Research* 20/3, 543-554.
- Öhman, Sven. 1967. «Peripheral Motor Commands in Labial Articulation». In: *Quarterly Progress and Status Report* 4 (Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden), 30-63.
- Pierrehumbert, Janet. 2003. «Probabilistic Phonology: Discrimination and Robustness». In: Bod, Rens et al. (Hg.): *Probabilistic Linguistics*. Cambridge: MIT Press, 177-228.
- Port, Robert/O'Dell, Michael. 1985. «Neutralization of Syllable-Final Voicing In German». In: *Journal of phonetics* 13, 455-471.
- Price, Patti Jo. 1980. «Sonority and Syllabicity – Acoustic Correlates of Perception». In: *Phonetica* 37, 327-343.
- Recasens, Daniel/Espinosa, Aina. 2005. «Articulatory, positional and coarticulatory characteristics for clear /l/ and dark /l/: evidence from two Catalan dialects». In: *Journal of the International Phonetic Association* 35/1, 1-25.

- Reichel, Uwe. 2012. «PermA and Balloon: Tools for string alignment and text processing». In: *Proceedings Interspeech*, 1874-1877.
- Shaw, Jason et al. 2011. «Dynamic Invariance in the Phonetic Expression of Syllable Structure: a Case Study of Moroccan Arabic Consonant Clusters». In: *Phonology* 28/3, 455-490.
- Silva, David James 1997. «The Variable Deletion of Unstressed Vowels in Faialense Portuguese». In: *Language Variation and Change* 9/3, 295-308.
- 1998. «Vowel Lenition in São Miguel Portuguese». In: *Hispania* 81, 166-178.
- Tabain, Marija. 2000. «Coarticulation in CV syllables: a Comparison of Locus Equation and EPG data». In: *Journal of phonetics* 28/2, 137-159
- Veloso, João. 2007. «Schwa in European Portuguese: The Phonological Status of [i]». In: Crouzet, Olivier/Angoujard, Jean-Pierre (Hg.): *Actes des Proceedings of JEL'2007*. Nantes: Université de Nantes, 55-60.
- Vigário, Marina. 1998. «Elisão da Vogal Não-Recuada Final e a Palavra Prosódica no Português Europeu». In: Marquilhas, Rita/Mota, Maria Antonia (Hg.): *XIII Encontro Nacional da Associação Portuguesa de Linguística II*. Lisboa: APL, 359-376
- Williams, Edwin Bucher. 1968. *From Latin to Portuguese*. Philadelphia: University of Pennsylvania Press.