
Zusammenhänge zwischen
synchroner Variation und diachronem
Lautwandel im Sibilantsystem des
Polnischen

Véronique Bukmaier



München 2016

Zusammenhänge zwischen
synchroner Variation und diachronem
Lautwandel im Sibilantsystem des
Polnischen

Véronique Bukmaier

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

vorgelegt von
Véronique Bukmaier
aus Augsburg

München, den 13.10.2016

Erstgutachter: Prof. Dr. Jonathan Harrington

Zweitgutachter: Prof. Dr. Philip Hoole

Tag der mündlichen Prüfung: 31.01.2017

Inhaltsverzeichnis

Zusammenfassung	xv
1 Einleitung	1
1.1 Synchrone Variation in Produktion und Perzeption	1
1.2 Synchrone Variation und Kontrastneutralisierungen	5
1.3 Forschungsstand zur /s ʂ/ Neutralisierung	7
1.4 Die stimmlosen Sibilanten /s ʂ ʐ/ des Polnischen	11
1.4.1 Konsonanten und Vokale des Standard-Polnischen	11
1.4.2 Regionale Varietäten des Polnischen	16
1.4.3 Artikulation der stimmlosen Sibilanten /s ʂ ʐ/	19
1.4.4 Akustische Eigenschaften der Sibilanten /s ʂ ʐ/	23
1.5 Lautwandel	26
1.5.1 Aufbau der Arbeit	33
2 Artikulatorische und akustische Eigenschaften der Polnischen Sibilanten (EMA)	35
2.1 Zusammenfassung	35
2.2 Einleitung	36
2.3 Experimentaufbau	39
2.3.1 Datenerhebung und Sprecher	39
2.3.2 Sprachmaterial und Experimentaufbau	40
2.3.3 Analyse der physiologischen Daten	42
2.3.4 Analyse der akustischen Daten	44
2.4 Ergebnisse	46
2.4.1 Ergebnisse I: Friktion – Physiologische Analyse	46
2.4.2 Ergebnisse II: Friktion – Akustische Analyse	48

2.4.3	Ergebnisse III: Koartikulatorische Einflüsse auf den Vokal – Physiologische Analyse	51
2.4.4	Ergebnisse IIII: Koartikulatorische Einflüsse auf den Vokal – Akustische Analyse	53
2.5	Diskussion	55
3	Perzeption des Dreier-Kontrasts /s ʂ ɕ/ im Polnischen in akzentuiertem und deakzentuiertem Kontext	61
3.1	Zusammenfassung	61
3.2	Einleitung	62
3.3	Experimentaufbau	71
3.3.1	Erstellung der Stimuli	71
3.3.2	Datenerhebung und Hörer	77
3.3.3	Experimentdurchführung	78
3.3.4	Kategoriale Wahrnehmung	79
3.3.5	Datenanalyse	80
3.4	Ergebnisse	81
3.4.1	Ergebnisse I: Rolle der Formanttransitionen für den /s ɕ/-Kontrast	81
3.4.2	Ergebnisse II: In-/Stabilität des /s ʂ/-Kontrasts	85
3.5	Diskussion	91
4	Variation im Spracherwerb: Sibilantproduktionen von Kindern und Erwachsenen	103
4.1	Zusammenfassung	103
4.2	Einleitung	104
4.3	Parameter zur Analyse von Sibilanten bei Kindern	114
4.4	Experimentaufbau	120
4.4.1	Datenerhebung und Sprecher	120
4.4.2	Sprachmaterial und Experimentaufbau	120
4.4.3	Datenanalyse	124
4.5	Ergebnisse	126
4.6	Diskussion	139
5	Zusammenfassung und Diskussion	149
5.1	Artikulation der Sibilanten /s ʂ ɕ/	151

5.2	Perzeption der Polnischen Sibilanten /s ʂ ʐ/ in verschiedenen prosodischen Kontexten	153
5.3	Sibilantproduktionen im Spracherwerb und in erwachsener Sprache	155
5.4	Die Sibilanten /s ʂ ʐ/ und der Faktor Geschlecht	157
5.5	Fazit und Ausblick	158
A	Anhang zu Kapitel 2	167
A.1	Multitaper-Methode	167
B	Anhang zu Kapitel 3	169
B.1	Hintergrundinformationen zu den Teilnehmern des Perzeptionsexperiments	169
B.2	Hörerspezifische Perzeptionskurven	170
C	Anhang zu Kapitel 4	177
C.1	Produktionsdaten separat für die männlichen und weiblichen L1-Polnisch sprechenden kindlichen Versuchsteilnehmer	177
D	Publikationsliste	181
	Literaturverzeichnis	182
	Danksagung	201

Abbildungsverzeichnis

1.1	Konsonantsystem des Polnischen	12
1.2	Lautwandel im Polnischen Sibilantinventar	14
1.3	Vokalsystem des Polnischen	15
1.4	Regionale Einteilung der Polnischen Dialekte	18
1.5	MRT Bilder während der Produktion der Polnischen Sibilanten /s ʂ ʐ/	22
2.1	Position der Zungensensoren	40
2.2	Schematische Repräsentation der Position der Orientierungspunkte der artikulatorischen Daten	43
2.3	Lobanov-normalisierte vertikale und horizontale TT Bewegungsverläufe separat für jeden der drei Sibilanten	47
2.4	Lobanov-normalisierte TT Orientierung und TT Position separat für jeden Sprecher und jeden Sibilanten	48
2.5	Ensemble-gemittelte Frikativspektren zum zeitlichen Mittelpunkt des Frikativs	49
2.6	Erster und zweiter mel-skaliertes DCT-Koeffizient separat für die drei Sibilanten /s ʂ ʐ/	50
2.7	Bewegungsverlauf der vertikalen und horizontalen Zungenrückenposition separat für jeden Sibilanten und Sprecher	52
2.8	Gemittelte und zeitnormalisierte Trajektorien im F2 (y-Achse) × F3 (x-Achse) Raum	53
2.9	Formanten F1 und F2 zum zeitlichen Vokalmittelpunkt separat für /ʐ/ und /s ʂ/	54
3.1	Stimuli Auswahl /s-ʂ/ ambiger Vokal	75
3.2	Stimuli Auswahl /s-ʐ/ ambiger Vokal	75
3.3	Stimuli Auswahl /s-ʐ/ manipulierter Vokal	75

3.4	Exemplarische Darstellung der F_0 -Manipulationen für die beiden prosodischen Kontexte	77
3.5	Schematische Darstellung psychometrischer Kurven	80
3.6	Psychometrische Kurven der beiden $/s\varepsilon-\zeta\varepsilon/$ Kontinua im akzentuierten und deakzentuierten Kontext	82
3.7	Hörerspezifische Neigungen der beiden $/s\varepsilon-\zeta\varepsilon/$ Kontinua im akzentuierten und deakzentuierten Kontext	83
3.8	Psychometrische Kurven der Kontinua $/s\varepsilon-\xi\varepsilon/$ und $/s\varepsilon-\zeta\varepsilon/$ im akzentuierten und deakzentuierten Kontext	85
3.9	Hörerspezifische Neigungen der Kontinua $/s\varepsilon-\xi\varepsilon/$ und $/s\varepsilon-\zeta\varepsilon/$ im akzentuierten und deakzentuierten Kontext	87
3.10	Hörerspezifische Neigungen der Kontinua $/s\varepsilon-\xi\varepsilon/$ und $/s\varepsilon-\zeta\varepsilon/$ im akzentuierten und deakzentuierten Kontext getrennt für die männlichen und weiblichen Hörer	89
3.11	Differenzen zwischen den Neigungswerten des akzentuierten und deakzentuierten Kontexts getrennt für die Kontinua $/s\varepsilon-\xi\varepsilon/$ und $/s\varepsilon-\zeta\varepsilon/$ und separat für die männliche und weibliche Versuchspersonengruppe	90
3.12	Hypothetische Verteilung und Verschiebung der Sibilantkategorien $/s \xi/$ im Sinne der Exemplartheorie	100
4.1	DCT's zum zeitlichen Mittelpunkt der Sibilanten $/s \xi \zeta/$ für L1-Polnisch sprechende Kinder	117
4.2	DCT's zum zeitlichen Mittelpunkt der Sibilanten $/s \xi \zeta/$ für L1-Polnisch sprechende Erwachsene	117
4.3	Gemittelte und zeitnormalisierte F2-Transitionen vom Onset bis zum Mittelpunkt des den Sibilanten $/s \xi \zeta/$ nachfolgenden Vokals für L1-Polnisch sprechende Kinder und Erwachsene	119
4.4	Bildschirmfoto der Zielwortpräsentation im Verlauf der Experimentdurchführung für die Bildbenennung der Kinder und die Leseaufgabe der Erwachsenen	124
4.5	Logarithmisierte Verhältnisse der Euklidischen Distanzen im $C_1 \times C_2$ Raum getrennt für die drei Sibilanten und separat für Kinder und Erwachsene	127
4.6	Lobanov-normalisierte F2-Werte zum zeitlichen Vokalonsset getrennt für die kindlichen, die männlichen und die weiblichen erwachsenen Versuchspersonen	130

4.7	Lobanov-normalisierte F3-Werte zum zeitlichen Vokalonsset getrennt für die drei Sibilanten und separat für die kindlichen, die männlichen und die weiblichen erwachsenen Versuchspersonen	132
4.8	Differenzen der d_{sib} -Werte zwischen den drei Sibilanten getrennt für die kindlichen, die männlichen und die weiblichen erwachsenen Sprecher	133
4.9	Differenzen der F2-Werte zum zeitlichen Onset des post-sibilantischen Vokals / ϵ / für die drei dem Vokal vorangehenden Sibilanten / $s \ \zeta \ \epsilon$ / separat für die Kinder, die männlichen und weiblichen Erwachsenen	135
4.10	Differenzen in den F3-Werten zum Onset des post-sibilantischen Vokals / ϵ / für die drei Sibilanten / $s \ \zeta \ \epsilon$ / separat für die Kinder, die männlichen und die weiblichen Erwachsenen	137
A.1	Ensemble-gemittelte Spektren berechnet anhand der Methode in Abb. 2.5 (original) und anhand der Multitaper-Methode (taper) separat für alle drei Sibilanten	167
B.1	Hörerspezifische psychometrische Kurven der Proportion an NICHT-/s/-Antworten als Funktion der Stimulus Nummer für das / $s\epsilon$ - $\zeta\epsilon$ / Kontinuum mit ambigem Vokal und in akzentuierter Bedingung	170
B.2	Hörerspezifische psychometrische Kurven der Proportion an NICHT-/s/-Antworten als Funktion der Stimulus Nummer für das / $s\epsilon$ - $\zeta\epsilon$ / Kontinuum mit manipuliertem Vokal und in akzentuierter Bedingung	171
B.3	Hörerspezifische psychometrische Kurven der Proportion an NICHT-/s/-Antworten als Funktion der Stimulus Nummer für das / $s\epsilon$ - $\zeta\epsilon$ / Kontinuum mit ambigem Vokal und in akzentuierter Bedingung	172
B.4	Hörerspezifische psychometrische Kurven der Proportion an NICHT-/s/-Antworten als Funktion der Stimulus Nummer für das / $s\epsilon$ - $\zeta\epsilon$ / Kontinuum mit ambigem Vokal und in deakzentuierter Bedingung	173
B.5	Hörerspezifische psychometrische Kurven der Proportion an NICHT-/s/-Antworten als Funktion der Stimulus Nummer für das / $s\epsilon$ - $\zeta\epsilon$ / Kontinuum mit manipuliertem Vokal und in deakzentuierter Bedingung	174
B.6	Hörerspezifische psychometrische Kurven der Proportion an NICHT-/s/-Antworten als Funktion der Stimulus Nummer für das / $s\epsilon$ - $\zeta\epsilon$ / Kontinuum mit ambigem Vokal und in deakzentuierter Bedingung	175

C.1	Logarithmisierte Verhältnisse der Euklidischen Distanzen zu den Mittelwerten von Dental /s/ und Retroflex /ʂ/ im $C_1 \times C_2$ -Raum für die drei Sibilanten /ç/ (links), /s/ (mittig) und /ʂ/ (rechts), separat für die männlichen und weiblichen kindlichen Versuchspersonen	177
C.2	Lobanov-normalisierte F2-Werte zum zeitlichen Vokalonsset getrennt für die drei Sibilanten /s ʂ ç/ und separat für die männlichen und weiblichen kindlichen Versuchspersonen	178
C.3	Lobanov-normalisierte F3-Werte zum zeitlichen Vokalonsset getrennt für die drei Sibilanten /s ʂ ç/ und separat für die männlichen und weiblichen kindlichen Versuchspersonen	179

Tabellenverzeichnis

2.1	EMA-Experiment: Korpus	41
2.2	Gesamtverteilung CV-Sequenzen	42
2.3	Gaußsche Klassifikation der drei Sibilanten /s ʃ ʒ/ im $C_1 \times C_2$ -Raum	51
2.4	Gaußsche Klassifikation der drei Sibilanten /s ʃ ʒ/ basierend auf F2 und F3-Werten der nachfolgenden Vokale	55
4.1	Produktionsexperiment: Korpus der kindlichen Probanden	121
4.2	Produktionsexperiment: Korpus der erwachsenen Probanden	122
4.3	Verteilung der Sibilantproduktionen in den zwei Teilnehmergruppen	123
4.4	Gemittelte d_{sib} -Werte separat für die drei Sibilanten /s ʃ ʒ/ und getrennt für die drei Sprechergruppen	129
B.1	Hintergrundinformationen zu allen 13 Versuchsteilnehmern, die ursprünglich am Perceptionsexperiment teilgenommen haben	169

Zusammenfassung

In dieser Dissertation wurden die Sibilanten /s ʂ ʐ/ des Standard-Polnischen untersucht. Genauer gesagt standen im Fokus dieser Arbeit mögliche Zusammenhänge zwischen synchroner Variation und der potentiellen diachronen Instabilität dieses Kontrasts. Für einige regionale Varietäten des Polnischen, sowie auch für die drei, den polnischen Lauten ähnlichen, Sibilanten des Mandarin-Chinesischen, wurde eine Neutralisierung des Dreier-Kontrasts hin zu einem Zweier-Kontrast beschrieben. Dabei wird für beide Sprachen davon ausgegangen, dass der Kontrast zwischen dem retroflexen Frikativ /ʂ/ und dem dentalen Laut /s/ von der Neutralisierung betroffen ist, sodass neben dem alveolopalatalen Frikativ /ʐ/ nur der dentale Laut bestehen bleibt. Während für die Kontrastneutralisierung im Mandarin-Chinesischen Studien zeigten, dass interne und extralinguistische Faktoren den Grad der Neutralisierung der Sibilanten /s ʂ/ beeinflussen, ist für das polnische Sibilantensystem bisher nicht bekannt, welche Wirkung interne und externe Faktoren, auf die Stabilität des Standard-Polnischen Dreier-Kontrasts haben. Diese Arbeit war ein erster Versuch den Einfluss verschiedener interner und externer Faktoren auf die Stabilität der drei Sibilanten in der Produktion und Perzeption der Standardsprache des Polnischen zu testen, um mögliche Erklärungen für eine größere Wahrscheinlichkeit eines /s ʂ/ Zusammenfalls zum Vorteil des dentalen Sibilanten zu liefern.

In einer physiologischen Studie wurde die relative synchrone Stabilität der Sibilanten, produziert von L1-Polnisch sprechenden Erwachsenen, getestet. Hierfür wurden akustische und Zungenbewegungsdaten von Sibilantproduktionen in Logatomen und in polnischen Echtwörtern in zwei verschiedenen Sprechgeschwindigkeiten erhoben. Die Analyse der artikulatorischen Daten brachte eine nahezu komplette Separierung der drei Sibilanten hervor. Jedoch zeigte die Analyse der akustischen Daten keine so deutliche Trennung der untersuchten Laute auf. Während die Informationen im Friktionsgeräusch der Sibilanten den dentalen Laut /s/ durch seinen hohen Frequenzschwerpunkt von den beiden anderen Frikativen trennt, werden für /ʂ ʐ/ Ähnlichkeiten im Friktionsanteil der Sibilanten gefunden.

Zu Beginn des nachfolgenden Vokals finden sich Informationen, welche den alveolopalatalen Laut klar von den Sibilanten /s ʂ/ trennen, wobei letztere in ihren F2-Transitionswerten stark miteinander überlappen. Für die Kategorisierung des dentalen und des alveolopalatalen Sibilanten reicht damit jeweils ein Parameter – Friktionsgeräusch für /s/ und F2-Wert für /ɕ/ – aus, während für den retroflexen Sibilanten die Merkmale im Sibilanten selbst, wie auch die Information im Onset des nachfolgenden Vokals zur Kategorisierung benötigt werden.

In einem Perzeptionsexperiment wurde der Einfluss des prosodischen Kontexts auf die Kategorisierung eines Kontinuums zwischen dem scheinbar instabilen /s ʂ/-Kontrast und einem Kontinuum zwischen dem vermutlich stabilen /s ɕ/-Kontrast getestet. Die Hypothese hierzu lautete, dass eine weniger kategoriale Perzeption des scheinbar instabilen /s ʂ/-Kontrasts in deakzentuierter Position, im Vergleich zum stabilen /s ɕ/-Kontrast, gefunden werden kann. Die Auswertung der Daten deutete, sowohl einen Effekt des prosodischen Kontexts als auch einen Einfluss des Geschlechts der Hörer auf die Wahrnehmung der Kontinua an.

Eine Untersuchung der Produktion der Sibilanten von kindlichen und erwachsenen L1-Polnisch sprechenden Versuchspersonen brachte Hinweise dafür hervor, dass die F2 und F3-Werte der Sibilanten bei den kindlichen und den männlichen erwachsenen Sprechern nicht zur Kategorisierung des /s ʂ/-Kontrasts beitragen. Genauer gesagt konnte für diese beiden Sprechergruppen keine Unterscheidung in den besagten Merkmalen gefunden werden. Dem gegenüber stehen die weiblichen Probanden, welche alle drei sibilantischen Kategorien in den spektralen Merkmalen des Friktionsgeräusches, den F2 und in den F3-Werten zu Beginn des nachfolgenden Vokals voneinander trennten.

Diese Dissertation liefert in allen experimentellen Kapiteln Hinweise auf eine Instabilität des retroflexen Sibilanten und Erklärungsansätze für die größere Wahrscheinlichkeit einer Neutralisierung des /s ʂ/-Kontrasts.

Kapitel 1

Einleitung

1.1 Synchrone Variation in Produktion und Perzeption

Sprache ist variabel und anpassungsfähig. Es ist uns möglich zu sprechen, während wir anderen Aktivitäten nachgehen. Dabei ist wichtig, dass wir unsere Sprache an die verschiedensten Situationen anpassen können, sodass das Verständnis des Gesagten erhalten bleibt.

Verschiedene Faktoren, die dazu führen, dass sprachliche Komponenten variieren, sind mannigfaltig in der alltäglichen Kommunikation zu finden. Eine Äußerung, welche von ein und demselben Sprecher mehrmals produziert wurde, variiert mitunter aufgrund ebendieser Faktoren von Wiederholung zu Wiederholung in den feinen phonetischen Details des Sprachsignals. Ein Beispiel für einen solchen Faktor, welcher die spektralen Eigenschaften im Sprachsignal verändern kann, ist die emotionale Lage des Sprechers. Murray und Arnott (1993) zeigten im Vergleich von neutraler Sprache und Sprache in verschiedenen Emotionen, dass lautliche Einheiten je nach untersuchter Emotion variieren. Frikative zeigten in ihrer Studie Veränderungen in der Verteilung der Schwerpunktfrequenz, d.h. während für die Emotion „glücklich“ mehr Energie in den höheren Frequenzen zu finden war, wurde für die Emotion „traurig“ die Energie des spektralen Schwerpunkts in die niedrigeren Frequenzen gesenkt. Auch das Alter und das Geschlecht des Sprechers beeinflusst die lautlichen Komponenten. Reubold, Harrington und Kleber (2010) fanden in ihrer longitudinalen Studie, dass für ihre weibliche Versuchsperson mit steigendem Alter sowohl die Grundfrequenz als auch die Werte des ersten Formanten niedriger wurden. Für Kinder zeigten Vorperian und Kent (2007), dass durch das Wachstum des Vokaltrakts die Formantwerte der kindlichen Vokale mit steigendem Alter meist niedriger werden. Betrachtet man Variationen zwischen

den Sprachproduktionen von weiblichen und männlichen Sprechern können Veränderungen in der lautlichen Struktur teilweise den anatomischen und biologischen Gegebenheiten und zum Teil jedoch auch sozio-phonetischen Faktoren zugeschrieben werden. Der oftmals kleinere Vokaltrakt bei Frauen kann beispielsweise höhere Grundfrequenzen zur Folge haben (Hillenbrand, Getty, Clark & Wheeler, 1995). Doch nicht nur solche physiologischen Merkmale haben einen Effekt auf die Akustik der Sprache, denn Sprecher manipulieren ihren Vokaltrakt, um ihre soziale Identität und Zugehörigkeit zu markieren (Stuart-Smith, 2007). Schon Labov (1990) zeigte in seinen sozio-phonetischen Studien, dass Frauen häufig die Standardsprachlichen Ausspracheformen nutzen, während die männlichen Mitglieder der Sprachgemeinschaft eine Tendenz zum Gebrauch von regional- und dialektal-gefärbter Sprache aufweisen. Ebenfalls lässt die Umgebung und der Gesprächspartner die Sprachproduktion variieren (Kent & Read, 1992; Myers, 2002). Der Sprecher passt die Klarheit beziehungsweise das Maß an Reduktionen innerhalb einer Äußerung an die jeweilige Situation und deren Anforderungen an die Gesprächsteilnehmer an (Lindblom, 1990). Hyperartikulierte Sprache, welche als klare und kontrastreiche Sprache beschrieben werden kann, findet sich meist in formellen Gesprächssituationen und in Umständen, in welchen das Verständnis des Gesagten im Vordergrund steht. Die Hypoartikulation kommt in weniger formellen sprachlichen Umgebungen vor, sowie in Äußerungsteilen, welche als schon bekannt vorausgesetzt werden können. Die hypoartikulierte Sprache ist meist durch Kontrastreduktionen, Elisionen und durch weniger periphere Realisierungen von Lautkontrasten geprägt (Mitterer & Ernestus, 2006).

Ein weiterer Faktor, welcher die Variation von Sprachproduktionen beeinflusst, ist der lautliche Kontext, in welchem sprachliche Segmente vorkommen (Yu, Abrego-Collier, Phillips, Pillion & Chen, 2015). Heinz und Stevens (1961) und Repp (1986) zeigten, wie die englischen Sibilanten /s ʃ/ von den umliegenden Vokalen beeinflusst werden können. Folgte dem Sibilanten der hohe, hintere, gerundete Vokal /u/ wurde der Frequenzschwerpunkt gesenkt, während für die hohen, vorderen, ungerundeten Vokale (/i e/) der umgekehrte Effekt eintrat, nämlich ein Anheben des Frequenzschwerpunkts des sibilantischen Friktionsgeräusches. Ruch (2013) zeigte außerdem, dass die Länge der *voice onset time*¹ (kurz

¹Für die Unterscheidung von stimmhaften, stimmlosen und stimmlos-aspirierten Plosiven kann die *voice onset time* (kurz VOT) eingesetzt werden. Die VOT bezeichnet die Zeit zwischen oraler Verschlusslösung und dem Einsetzen der Stimmlippenschwingung. Ist die VOT negativ, d.h. Schwingen die Stimmlippen schon vor der Verschlusslösung, ist das ein Hinweis auf einen stimmhaften Laut. Ist die VOT positiv, d.h. fangen die Stimmlippen erst nach der Verschlusslösung zu schwingen an, so kann der Laut als stimmlos bezeichnet werden. Eine positive VOT im Bereich zwischen 0 ms bis ungefähr 30 ms beschreibt einen stimmlosen nicht-aspirierten Plosiven. Eine positive VOT von mehr als 30 ms zeichnet meist stimmlos

VOT) in post-aspirierten /st/-Sequenzen mit dem Vokalkontext variierte: kürzere VOT im Kontext eines nachfolgenden hohen, vorderen Vokals /i/, längere VOT im Kontext des Vokals /u/ im Falle der älteren spanischen Versuchspersonengruppe. Die Akzentuierung hat ebenfalls einen Einfluss auf die Realisierung der VOT in Plosiven des Amerikanischen Englisch und zwar insofern, als dass in akzentuierter Position eine Verstärkung der gemessenen Parameter – VOT, F_0 und Dauer des Verschlusses – für stimmhafte und stimmlose Plosive gefunden wurde (Cole, Kim, Choi & Hasegawa-Johnson, 2007). In deJong (1995) wurde dieser Effekt der Verstärkung von Segmenten in akzentuierter Bedingung für die Artikulation der Sprachgesten gezeigt. Das Ergebnis brachte größere Ausmaße der Sprechbewegungen und der Gestendauer von Segmenten in akzentuierten Silben hervor. In der Akustik konnten deJong (1998) diesen Einfluss der Betonung und Akzentuierung ebenfalls nachweisen. In betonten Silben konnten mehr spektrale Energie und längere Dauern für die jeweiligen Segmente gefunden werden. Neben dem Effekt der Akzentuierung auf koartikulatorische Prozesse berichtete Recasens (2015) auch von einer Beeinflussung der VCV-Koartikulation (V = Vokal, C = Konsonant) durch verschiedene Sprechgeschwindigkeiten, d.h. unbetonte Laute und in schneller Sprechgeschwindigkeit produzierte Segmente zeigten ein größeres Maß an Koartikulation. Für die lexikalische Frequenz², d.h. die Worthäufigkeit, und die Nachbarschaftsdichte³ wurde ebenfalls ein Einfluss auf die phonetische Variation festgestellt. Je höher die lexikalische Frequenz eines Wortes, desto häufiger sind Reduzierungen zu finden (Bybee & Hopper, 2001). Ist die Nachbarschaftsdichte eines Wortes hoch, treten eher kanonische, weniger reduzierte Ausspracheformen auf, während für Äußerungen mit einer niedrigen Nachbarschaftsdichte vermehrt Reduktionen der Ausspracheform vorkommen (Munson, 2007).

Die oben aufgeführten Beispiele der verschiedenen Arten der Variation können grob in zwei Kategorien unterteilt werden:

Interne Variation fasst alle die sprachlichen Veränderungen zusammen, welche beispielsweise durch den lautlichen Kontext und die prosodischen Strukturen hervorgebracht werden.

aspirierte Plosive aus (Lisker & Abramson, 1964).

²Lexikalische Frequenz bezieht sich darauf, wie häufig ein bestimmtes Wort in einer bestimmten Sprache vorkommt (Pierrehumbert, 2002).

³Nachbarschaftsdichte bezieht sich auf die Anzahl an phonologisch ähnlichen Worten im Lexikon. Zur Messung der Nachbarschaftsdichte wird häufig die Anzahl an Worten festgehalten, welche durch das Zufügen, Löschen und das Ersetzen eines Lautes in einem bestimmten Wort entstehen (Luce & Pisoni, 1998).

Externe (und extra-linguistische) Variation beinhaltet Veränderungen der Sprachproduktion zwischen verschiedenen Sprechern. Sprecherspezifisches situationsabhängiges Verhalten gegenüber verschiedenen Gesprächspartnern, Emotionen, sowie regionale und soziale Gegebenheiten prägen diese Art der Variation.

Wichtig ist bei all der vielen sprachlichen Variation, dass es trotzdem möglich ist die Kommunikation und die sprachliche Verständigung innerhalb der Sprachgemeinschaft aufrecht zu erhalten. Vor diesem Hintergrund suchte Lindblom (1996) nach Merkmalen im Sprachsignal, welche über Kontexte hinweg unverändert bleiben und damit die Perzipierbarkeit des Gesagten für den Hörer aufrechterhalten. Studien, die die Verbindung von Koartikulation in der Produktion und deren Perzeption im Hörer untersuchten, stellten einen engen Zusammenhang zwischen den beiden Modalitäten fest (Delattre, Liberman & Cooper, 1962; Hughes & Halle, 1956). Mann und Repp (1980) fanden das Aufrechterhalten der Verständigung durch die Kompetenz der Hörer gegeben, für verschiedene Arten von Variationen zu kompensieren. In ihrer Studie konnte gezeigt werden, dass die Kategoriengrenze in einem synthetischen Kontinuum zwischen den Sibilanten /ʃ s/ je nach nachfolgendem Vokal variierte. Folgte der sibilantischen Sequenz der hintere, gerundete Vokal /u/, so wurde die perzeptive Kategoriengrenze zugunsten des alveolaren Sibilanten verschoben, d.h. die Hörer gaben, verglichen mit der Perzeptionskurve, die für das /ʃa-sa/ Kontinuum gefunden wurde, mehr /s/-Antworten. Diese Kompensation für Koartikulation ist darauf zurückzuführen, dass die Hörer um den koartikulatorischen Effekt des nachfolgenden Vokals auf das Friktionsgeräusch der Sibilanten wissen und die niedrigen Frequenzen der Lippenrundung des Vokals /u/ zuschreiben. Doch nicht nur für interne Variation kann kompensiert werden, wie einige Perzeptionsstudien zeigten. Regionale und soziale Variation in der Sprache wird von den Hörern wahrgenommen und beeinflusst die Art und Weise, wie das Sprachsignal von ebendiesen interpretiert wird (Evans & Iverson, 2004; Hay, Warren & Drager, 2006; Sumner & Samuel, 2009). Die Studien zur Sprachwahrnehmung und zur Kompensation für Variation weisen allesamt auf ein Merkmal der Produktion und Perzeption von Sprache hin, nämlich dass die Variation von Sprachlauten nicht kategorial abläuft, sondern dass die kontextuelle Veränderung von Segmenten in den feinen phonetischen Details zu sehen ist, welche vom Hörer wahrgenommen und im Zuge der Verarbeitung des Sprachsignals berücksichtigt werden.

1.2 Synchrone Variation und Kontrastneutralisierungen

Im vorangegangenen Abschnitt wurde generell gezeigt, welche Faktoren es gibt, die die Variabilität von Sprache steigern können. In manchen Fällen von synchroner Variation – sowohl hinsichtlich interner als auch externer Variation – können Lautkontraste unter bestimmten Bedingungen reduziert oder sogar komplett neutralisiert werden. Yu (2011, S. 291) definiert phonologische Kontraste und Reduzierungen wie folgt:

„Two sounds are phonological contrastive if they are in opposition with each other, that is, if they are capable of signaling a meaning difference between two lexical items in a particular language. [...] When a phonological opposition is suspended, neutralization or merger obtains.“

Bei Neutralisierungen wird davon ausgegangen, dass zwei unterschiedliche zugrundeliegende Formen bei der Oberflächenrealisierung zusammenfallen. Jedoch können diese Oberflächenrealisierungen laut Yu (2011) in drei unterschiedliche Formen des Endzustands unterteilt werden:

Struktur-erhaltende Reduktionen sind Fälle von Neutralisierungen, bei welchen zwei Laute nach einer Neutralisierung durch eine phonetische Form dargestellt werden, welche einem der zwei ursprünglichen Laute entspricht. Ein Beispiel für eine Struktur-erhaltende Reduktion ist die Neutralisierung des Unterschieds zwischen dem einfachen velaren Laut /k/ und dem labialisierten velaren Gegenstück /k^w/ vor dem Vokal /ɔ/ in einigen Varietäten des Kantonesischen.

Struktur-bildende Reduktionen sind gegeben, wenn nach einer Kontrastneutralisierung eine phonetische Form besteht, welche zwischen den zwei neutralisierten Phonemen liegt, wie es zum Beispiel für das Flapping im Englischen der Fall ist. Hier werden die Laute /t d/ in intervokalischer Position, wobei dem koronalen Laut eine unbetonte Silbe folgen muss, zu einem Flap /ɾ/ neutralisiert.

Freie Variation bezeichnet eine phonetische Form, welche zwischen zwei Varianten frei wechselt. Eine solche freie Variation besteht für die Laute /n/ und /l/ bei einem Großteil von Sprechern des Kantonesischen-Chinesisch.

(für Details zu den drei Formen der Reduktionen und den aufgeführten Beispielen vgl. Yu, 2011, S. 291ff.)

Ein weiteres Beispiel für eine Struktur-erhaltende Reduzierung eines lautlichen Kontrasts zwischen zwei Lautpaaren ist die Neutralisierung des Artikulationsorts des Alveolaren /s/ und des postalveolaren Sibilanten /ʃ/ in regionalen Varietäten des Deutschen (Bukmaier, Harrington & Kleber, 2014), im Europäischen Portugiesisch (Mateus & d’Andrade, 2000), sowie in sCV-Sequenzen (s = alveolarer Sibilant, C = Konsonant, V = Vokal) in einigen Varietäten des Englischen (Baker, Archangeli & Mielke, 2011). Für diese Neutralisierungen, die alle durch den Wegfall des alveolaren und den Erhalt des postalveolaren Sibilanten definiert sind, ist gezeigt worden, dass das Maß an Reduzierung dieses Kontrasts vom lautlichen Kontext, der Position in der Silbe, von der dialektalen Varietät und vom Alter der Sprecher abhängig ist. In der schwäbischen Varietät des Deutschen wird dieser Kontrast in silbenfinaler, präkonsonantaler Position neutralisiert, wobei eben nur der postalveolare Sibilant in diesen Bedingungen bestehen bleibt. In ihrer Studie zeigten Bukmaier, Harrington und Kleber (2014), dass nicht nur die Position in der Silbe, der lautliche Kontext und die dialektale Varietät des Standarddeutschen einen Einfluss auf diese Kontrastneutralisierung hat, sondern auch das Alter der Sprecher/Hörer. Während die jungen Dialekt-Sprecher in der Produktion und der Perzeption den Mustern der Standardsprachlichen Versuchspersonen ähnelten, zeigten die älteren Sprecher noch eine etwas stärkere Neigung hin zur /s ʃ/ Neutralisierung. Das Portugiesische zeigt einen weniger eingeschränkten Fall der /s ʃ/ Neutralisierung. Der alveolare Frikativ wird dabei grundsätzlich in wortfinaler Position als postalveolarer Frikativ realisiert (vgl. Mateus, 2000). Das Maß an Rückverlagerung des alveolaren Sibilanten (besser bekannt unter dem Begriff der *s-retraction*) in den englischen /sC sCr/-Klustern wird vom Geschlecht der Sprecher beeinflusst (M. Stevens & Harrington, 2016). Der Frequenzschwerpunkt des alveolaren Sibilanten zeigt für die männlichen Versuchspersonen niedrigere Werte als für die weiblichen Sprecher. Weiterhin wurde für die *s-retraction* des alveolaren Sibilanten in einigen Varietäten des Englischen berichtet, dass der spektrale Frequenzschwerpunkt des Sibilanten in /str spr skr st sp sk/-Klustern so stark gesenkt wird, dass ein Einfluss auf die Perzeption beobachtet werden kann. Hörer ordnen die präkonsonantalen, alveolaren Sibilanten, im Vergleich zu prävokalischen, alveolaren Frikativen der postalveolaren Kategorie zu (für mehr Details zur Produktion und Perzeption der *s-retraction* vgl. M. Stevens und Harrington, 2016).

Für all diese /s ʃ/ Neutralisierungen ist, wie oben schon erwähnt, bekannt, dass es der Kontrast zwischen diesem einen bestimmten Lautpaar /s ʃ/ ist, welcher reduziert wird. Es gibt jedoch auch Neutralisierungen, welche eine ganze Lautserie betreffen, wie es zum Beispiel

im Zusammenfall der retroflexen Laute /tʂ tʂʰ ʂ z/ mit ihren dentalen Gegenstücken /ts tsʰ s z/ in einigen regionalen, Polnischen und Mandarin-Chinesischen Varietäten der Fall ist.

1.3 Forschungsstand zur /s ʂ/ Neutralisierung

Wie im vorangegangenen Abschnitt erwähnt, ist für einige dialektale Varietäten des Polnischen und des Mandarin-Chinesischen eine Neutralisierung bekannt, welche die Serie der retroflexen Lautklasse betrifft. Dabei fällt durch diese Neutralisierung der Standardsprachliche Dreier-Kontrast zwischen dem dentalen Frikativ /s/, dem retroflexen /ʂ/ und dem alveolopalatalen Laut /ç/ im Sibilantensystem dieser Sprachen zu einem Zweier-Kontrast zwischen /s ç/ zusammen. Für beide Sprachen wird angenommen, dass die Neutralisierung des sibilantischen Kontrasts durch einen Merger⁴ der Kategorien /s ʂ/ passiert (Duanmu, 2006; Gussmann, 2007; Nowak, 2006a).

Während der Prozess der /s ʂ/ Lautneutralisierung im Polnischen kaum untersucht wurde, zeigten einige Studien zur Reduzierung des Dental-Retroflex Kontrasts im Mandarin-Chinesischen, dass der Zusammenfall dieses Kontrasts durch linguistische, sowie durch extra-linguistische Faktoren beeinflusst wird (Beckman, Li, Kong & Edwards, 2014; Chang, 2012; Y. Li, 2009; Shih & Kong, 2011). Die bisher bekannten Studien zur Reduzierung des Dreier-Kontrasts im Mandarin-Chinesischen brachten hervor, dass der Grad der Neutralisierung des Dental-Retroflex Kontrasts in der Sprachproduktion und in der Perzeption beispielsweise vom prosodischen Kontext, von der lautlichen Umgebung, vom Geschlecht der Sprecher und der dialektalen Herkunft der Sprecher und Hörer beeinflusst wird. Chang (2011) berichtete in einem Vergleich der Produktion des retroflexen Sibilanten bei Sprechern des Taiwan-Mandarin und des Beijing-Mandarin, dass für die Beijing-Mandarin Sprecher eine stärkere Retroflexion gefunden werden kann, was sich akustisch durch niedrigere Schwerpunktfrequenzen in den retroflexen Lauten bemerkbar machte. Für die Sprecher der dialektalen Varietät, also für das Taiwan-Mandarin, wurden in den Sibilantproduktionen höhere Schwerpunktfrequenzen festgestellt, was auf eine weniger ausgeprägte Retroflexion von /ʂ/ und damit auf ein geringeres Maß an Unterscheidung vom dentalen

⁴In Yu (2011) wird angemerkt, dass sowohl der Begriff Neutralisierung als auch der Begriff Merger, den Prozess des Aufhebens eines phonologischen Kontrasts beschreibt. Häufig wird der Begriff Merger hinsichtlich eines diachronen Zusammenfalls zweier Kategorien verwendet, während die Neutralisierung synchrone Kontrastreduktionen beschreibt (für mehr Details vgl. Yu, 2011).

Laut /s/ hinwies. Weiterhin wurde in dieser Studie gefunden, dass die Neutralisierung des /s ʃ/-Kontrasts nicht durch einen kategorialen Sprung von einer retroflexen hin zu einer alveolaren Realisierung definiert ist, sondern eher durch einen graduellen Übergang von einer starken Realisierung des retroflexen Sibilanten der Standardsprachlichen Sprecher des Mandarin-Chinesischen bis hin zur vollständigen Reduzierung der Retroflexion mit vielen, dazwischenliegenden Stufen hinsichtlich der Retroflexion für die Sprecher der dialektalen Varietät – Taiwan-Mandarin.

Die Ergebnisse von Studien zu weiteren Einflussfaktoren auf die Neutralisierung sind weniger eindeutig. Hinsichtlich des Vokalkontexts berichtete Jeng (2006) von einem kleineren Unterschied zwischen dem dentalen und dem retroflexen Sibilanten im Kontext des hohen, gerundeten Vokals /u/, während Y. Li (2009) ein größeres Maß an akustischer Annäherung von /s ʃ/ grundsätzlich im Kontext von hohen Vokalen beschreibt. Der prosodische Kontext brachte ebenfalls kein klares Muster hinsichtlich der /s ʃ/ Neutralisierung hervor. Chuang und Fon (2010) zeigten, dass die Verstärkung des Kontrasts in akzentuierter Position und die Reduzierung des Kontrasts in prosodisch schwacher Position sprecher- und geschlechtsspezifisch realisiert wurde. Während alle weiblichen Sprecher den Dental-Retroflex Kontrast sowohl in der Akzentuierung als auch in der Deakzentuierung aufrechterhielten, konnte für den Großteil der männlichen Versuchspersonen eine Neutralisierung der Sibilanten gefunden werden. Bezogen auf die Frage, ob in der Akzentuierung der Kontrast verstärkt wird, während der Kontrast zwischen /s ʃ/ in der Deakzentuierung reduziert wird, waren die Ergebnisse jedoch weniger konsistent. Die Verstärkung der Sibilanten in der akzentuierten Position wurde von den Versuchspersonen auf unterschiedliche Arten realisiert. Für einige Versuchspersonen konnte eine Verstärkung beider Sibilanten in akzentuierter Bedingung gefunden werden, was bedeutet, dass der retroflexe Sibilant in die niedrigeren Frequenzen geschoben wurde, während der dentale Frikativ mit höheren Frequenzen realisiert wurde. Eine weitere Strategie zur Verstärkung des Kontrasts war die alleinige Anhebung der Frequenzen im dentalen Sibilanten, während der retroflexe Laut über die beiden prosodischen Kontexte hinweg unverändert blieb. Wie oben schon erwähnt, konnte für einige männliche Versuchspersonen ein Zusammenfall des Lautkontrasts in beiden prosodischen Kontexten, d.h. sowohl in akzentuierter als auch in deakzentuierter Position, berichtet werden. Für diese Sprecher galt, dass in der akzentuierten Bedingung der dentale und der retroflexe Sibilant mit höherem Frequenzschwerpunkt realisiert wurden als in der deakzentuierten Position.

Des weiteren wird für die /s ʃ/ Neutralisierung im Mandarin-Chinesischen berichtet, dass

die Sprachgemeinschaft um diesen Prozess weiß. In einigen Situationen führt dieses Wissen zu vermehrter Realisierung des retroflexen Sibilanten. Faktoren, welche den Grad der Retroflexion beeinflussen, sind laut Chung (2006), neben den schon aufgeführten kontextuellen, dialektalen und biologischen Gegebenheiten, das Alter, der Sprachstil und das Bildungsniveau der jeweiligen Sprecher. In informellen Situationen, in der älteren Generation und in der Sprache der niedrigeren Bildungsschichten sind wenige retroflexe und vermehrte dentale Sibilantproduktionen zu finden. Demgegenüber steigt der Gebrauch der retroflexen Sibilanten in formellen Situationen, in jüngeren Generationen und in Sprechern mit höherem Bildungsniveau an. In manchen formellen Diskursen finden sich sogar hyperkorrekte Formen, d.h. eine Ersetzung von zugrundeliegenden dentalen Lauten durch den retroflexen Frikativ. Solche Situationen kommen meist auf, wenn der Sprecher den Fokus auf die Formalität der Situation, die Hervorhebung eines wichtigen Aspekts oder die Auflösung von Zweideutigkeiten legt (Chung, 2006).

An diesen Ausführungen lässt sich erkennen, dass in den Varietäten des Mandarin-Chinesischen, welche von der Neutralisierung betroffen sind, von den Sprechern trotzdem auch die Standardsprachliche Form realisiert werden kann. Daraus kann wiederum geschlossen werden, dass auch für diese neutralisierenden Varietäten der retroflexe Laut der zugrundeliegenden Repräsentation entspricht, welche dann je nach Situation, tatsächlich als retroflexer, oder eben als dentaler Sibilant realisiert wird. Dieses Wissen und die Akzeptanz der Ersetzung des retroflexen Sibilanten durch den dentalen Laut zeigt auch die Studie von Chuang und Fon (2011). In ihrem Perzeptionsexperiment wurden den Hörern Wörter mit den Sibilanten /s ʃ/ präsentiert, wobei in einer Bedingung des Experiments die ursprünglichen Sibilanten im Wort vorkamen und in der zweiten Bedingung der retroflexe Sibilant durch den dentalen ausgetauscht wurde, und umgekehrt. Für die Wörter der Bedingung, die die Neutralisierung nachahmte, d.h. Wörter mit ursprünglich initialem Laut /ʃ/, welcher jedoch durch den dentalen Frikativ ersetzt wurde, waren die Reaktionszeiten der Worterkennung gleich den Reaktionszeiten für die Bedingung, in welcher der ursprüngliche retroflexe Laut wortinitial erhalten wurde.

Auch in der Perzeption werden beide Sibilantkategorien von Standardsprachlichen und von den dialektalen Hörern wahrgenommen. Chang (2012) zeigte in einem Diskriminationsexperiment, dass beide Sibilantkategorien sowohl von Mandarin-Chinesischen Hörern als auch von der Taiwanesischen Hörergruppe, wobei letztere Gruppe die Laute in der Produktion

neutralisiert, perzipiert werden können. In einem ABX-Diskriminationsexperiment⁵ wurden den Versuchspersonen Stimuli aus einem Kontinuum zwischen dem retroflexen und dem dentalen Sibilanten in zwei Vokalkontexten – /a u/ – präsentiert. Das Ergebnis zeigte, dass das Kontinuum zwischen dem retroflexen und dem dentalen Sibilanten von beiden Hörergruppen kategorial wahrgenommen werden konnte, wobei in beiden Vokalkontexten für die Standardsprachlichen Hörer niedrigere Schwerpunktfrequenzen nötig waren, um /ʂ/-Antworten zu erhalten. Im Vergleich dazu ordneten die Taiwan-Mandarin Hörer höhere Frequenzen der retroflexen Kategorie zu. Eine weitere Perzeptionsstudie zur Wahrnehmung von Kontrasten zwischen verschiedenen Frikativen im Mandarin-Chinesischen zeigte dennoch, dass die kleinste perzeptive Distanz für Taiwan-Mandarin Hörer zwischen dem retroflexen und dem dentalen Sibilanten besteht (Shih & Kong, 2011).

In der hier vorliegenden Studie soll der Kontrast zwischen den Sibilanten des Polnischen näher betrachtet werden. Zum einen, da für diese Kontrastreduktion noch keine experimentellen Studien bekannt sind, und zum zweiten um herauszufinden, welche Faktoren – neben dem der dialektalen Varietät – die Wahrscheinlichkeit für einen /s ʂ/ Zusammenfall erhöhen. Hierfür werden artikulatorische Daten der Standardpolnischen Sibilanten /s ʂ ʐ/ in der physiologischen Studie in Kapitel 2 erhoben, sodass geprüft werden kann, ob in der Artikulation ein Hinweis auf die scheinbare Instabilität des retroflexen Sibilanten zu finden ist. Im darauffolgenden Kapitel 3 soll in einem Perzeptionsexperiment der Einfluss von Akzentuierung und Deakzentuierung auf die Wahrnehmung der Sibilanten getestet werden, um zu sehen ob der prosodische Kontext größere Auswirkungen auf den scheinbar instabilen /s ʂ/-Kontrast hat, verglichen mit dem vermeintlich stabileren Kontrast zwischen /s ʐ/. Kapitel 4 befasst sich mit dem Zusammenhang von Spracherwerb und der Produktion der Sibilanten. Genauer wird in Kapitel 4 ein Vergleich von kindlichen und erwachsenen Produktionsdaten angestellt, um zu prüfen, ob in den akustischen Daten der kindlichen Sibilantproduktionen Merkmale beobachtet werden können, welche die Wahrscheinlichkeit eines Zusammenfalls der Laute /s ʂ/ erhöhen. Des Weiteren wird in der Perzeptionsstudie in Kapitel 3, sowie in der letzten experimentellen Studie (Kapitel 4) der Wirkung des Geschlechts der Hörer/Sprecher auf die Kategorisierung und die Produktion des /s ʂ/ Kontrasts nachgegangen.

⁵Diskriminationsexperimente wurden von Liberman, Harris, Hoffman und Griffith (1957) zur Untersuchung der kategorialen Wahrnehmung von Sprachlauten genutzt. Bei einem ABX-Diskriminationsexperiment hören Versuchspersonen direkt nacheinander drei Stimuli, wobei A und B unterschiedlich sind und der dritte Stimulus mit einem der beiden vorangegangenen übereinstimmt. Die Aufgabe der Versuchspersonen ist es zu bestimmen, ob der dritte Stimulus mit A oder mit B übereinstimmt.

1.4 Die stimmlosen Sibilanten /s ʃ ʂ/ des Polnischen

Das Polnische wird, neben dem Sorbischen, dem Tschechischen und dem Slowakischen, zu den West-Slawischen Sprachen gezählt (Janda, 2005). Die drei großen slawischen Sprachgruppen – Süd-Slawisch, Ost-Slawisch und West-Slawisch – entstammen dem Proto-Slawischen, welches sich als dialektale Varietät aus der indoeuropäischen Sprachgruppe heraus entwickelte (Sussex & Cubberley, 2006). Bis zur politischen Teilung Polens in den Jahren 1772/1773 konnte das Land als eine Art europäische Großmacht angesehen werden, deren politischer, kultureller und vor allem auch linguistischer Einfluss über die Landesgrenzen hinweg reichte (Sussex & Cubberley, 2006). Mit der oben genannten Aufteilung Polens zwischen Russland, Preußen und Österreich und den damit einhergehenden Restriktionen hinsichtlich des polnischen Sprachgebrauchs im öffentlichen Leben, kam eine Gefährdung der Weiterentwicklung der Polnischen Sprache auf. Erst nach dem Ende des ersten Weltkriegs war Polen als eigenständiger Staat wiederhergestellt, sodass ebenfalls der Gebrauch der Polnischen Sprache den Weg zurück in das öffentliche Leben der Sprachgemeinschaft fand. Trotz der Besetzung und Besiedlung einiger polnischer Gebiete im Verlauf des Zweiten Weltkriegs durch die Deutschen, kann die heutige Sprach- und Dialektsituation in Polen, laut Sussex und Cubberley (2006), als relativ traditionell und einheitlich angesehen werden. In den folgenden Abschnitten wird ein kurzer Einblick in das Lautsystem der Standard-Polnischen Sprache und deren dialektale Varietäten gegeben.

1.4.1 Konsonanten und Vokale des Standard-Polnischen

Das Phoneminventar des Standard-Polnischen umfasst ein umfangreiches Konsonantsystem (vgl. Abb. 1.1). Ein markantes Merkmal der polnischen Sprache ist die Palatalisierung einiger Konsonanten, welche durch einen nachfolgenden vorderen, hohen Vokal /i/ ausgelöst wird. Ein Beispiel ist der initiale Sibilant im polnischen Wort „sinus“ (Sinus), welcher durch den nachfolgenden Vokal /i/ palatalisiert wird, sodass die Realisierung [sʲinus] hervorgebracht wird (Gussmann, 2007). Einige palatalisierte Konsonanten, die nicht in Abb. 1.1 aufgeführt sind, kommen ausschließlich in Lehnwörtern vor, zu diesen Konsonanten gehören: /tʲ dʲ sʲ zʲ ʃʲ ʂʲ tsʲ dzʲ tʃʲ dʒʲ rʲ/ (für weitere Ausführungen und Beispiele zur Palatalisierung im Polnischen vgl. Gussmann, 2007).

Der Unterschied zwischen stimmhaften (Symbole rechts in den Spalten in Abb. 1.1) und stimmlosen (Symbole links in den Spalten in Abb. 1.1) Lauten ist mit der Realisierung des Stimmhaftigkeitskontrasts im Französischen, im Spanischen und im Russischen zu verglei-

	Labial	Palatalized labial	Dental	Alveolar	Alveo-palatal	Palatal	Velar
Stop	p b	p ^j b ^j	t d			c ɟ	k g
Affricate			ts dz	tʂ dz _ɹ	tɕ dʑ		
Fricative	f v	f ^j v ^j	s z	ʂ z _ɹ	ɕ ʑ	ç	x
Nasal	m	m ^j	n		ɲ		
Trill				r			
Lateral				l			
Glide	w					j	

Abbildung 1.1: Konsonantsystem im Standard-Polnischen. In der linken Hälfte der jeweiligen Spalten sind die stimmlosen, auf der jeweils rechten Seite die stimmhaften Varianten der Konsonanten aufgeführt. (Abbildung entnommen aus Nowak, 2006b)

chen (Malisz & Żygis, 2015), d.h. die stimmhaften Obstruenten werden mit einer Stimmlippenschwingung über die Verschlussphase hinweg produziert, während für ihre stimmlosen Gegenstücke das Fehlen der Stimmlippenschwingung in der Verschlussphase und eine kurze VOT bestimmend sind. Die stimmlosen Obstruenten unterscheiden sich damit von den stimmlosen Lauten des Englischen und des Deutschen durch das nur sehr geringe Maß an Aspiration (Malisz & Żygis, 2015). Der Kontrast zwischen stimmhaften und stimmlosen Obstruenten wird in silbenfinaler Position im Polnischen, wie im Deutschen und im Niederländischen, zum Vorteil der stimmlosen Variante neutralisiert. Ein Beispiel aus Gussmann (2007) illustriert die Auslautverhärtung im Polnischen: die zugrundeliegende Form für das Polnische Wort „chleb“ (Brot) ist /xlɛb/, doch die tatsächliche Realisierung zeigt wortfinal

die stimmlose Variante des Plosivs auf, nämlich [xlɛp].

Ein interessanter Aspekt des polnischen Konsonantsystems ist der Dreier-Kontrast in der Klasse der Frikative und Affrikaten. Maddieson (1984) berichtet, dass ungefähr 83%, der von ihm untersuchten Sprachen (n=317) einen sibilantischen Laut aufweisen, welcher dental oder alveolar artikuliert wird. Für Sprachen, welche zwei Sibilanten aufweisen, wird neben dem dentalen/alveolaren Laut meist noch der postalveolare Frikativ /ʃ/ gefunden. Der eher seltene Fall sind Sprachen, die drei Sibilanten kontrastiert, wie eben das Standard-Polnische, das Schwedische und das Mandarin-Chinesische. Maddieson (1984) gibt an, dass im Dreier-Kontrast der Sibilanten häufig ein dentaler, ein palatoalveolarer und ein alveolopalataler Sibilant gefunden werden kann. Für die koronalen Laute des Polnischen sind jeweils stimmhafte und stimmlose Varianten der dentalen Serie /s z ts dz/, der retroflexen Klasse /ʂ ʐ tʂ dʐ/ und der alveolopalatalen Laute /ɕ ʑ tɕ dʑ/ beschrieben. Laut Żygiś und Padgett (2010) ist der palatalisierte palato-alveolare Sibilant [ʃʲ] ein Allophon des retroflexen Phonems /ʂ/. Die [ʃʲ] Variante des Phonems wird hervorgerufen, wenn /ʂ/ vor den Lauten [i j] vorkommt. In Nowak's (2006b) Darstellung sind in der Spalte der alveolaren Frikative die Symbole /ʂ z/ aufgeführt, welche einer retroflexen Realisierung der Sibilanten entsprechen. Die Frage, ob dieser Sibilant tatsächlich retroflex realisiert wird, oder eher palato-alveolar, ist in der Literatur kontrovers diskutiert. Demgegenüber stehen Autoren, wie Gussmann (2007), der diese postalveolaren Laute mit dem IPA Symbol /ʃ ʒ/ beschreibt. Neuere Studien von Hamann (2002a) und Padgett und Żygiś (2003) bringen artikulatorische, akustische und perzeptive Gründe hervor, warum dieser Laut eher als flacher retroflexer Laut beschrieben werden sollte. Die retroflexen Sibilanten entstanden laut Padgett und Żygiś (2003) im Zuge eines Lautwandels im Standard-Polnischen im 16. Jahrhundert. Im Laufe dieses Wandels wurden die palatalisierten palato-alveolaren Laute durch eine Depalatalisierung zu ihrer jetzigen retroflexen Variante verändert. Diesem Lautwandel ging eine weitere Veränderung im Polnischen Phoneminventar voraus, welche die palatalisierten dentalen Sibilanten, um das 13. Jahrhundert herum, hin zu der heute bekannten alveolopalatalen Ausspracheform wandelte (vgl. Abb. 1.2). Für eine solche Verteilung spricht die perzeptive Distanz der Sibilanten zueinander, welche, laut Żygiś und Padgett (2010), durch die heute bekannten Exemplare – dental /s/, retroflex /ʂ/ und alveolopalatal /ɕ/ – der Sibilantserie optimiert zu sein scheint. Genauer gesagt argumentieren Żygiś und Padgett (2010), dass in Sprachen, welche einen Kontrast zwischen zwei postalveolaren Lauten aufweisen, der Kontrast zwischen /ʂ ɕ/ einer akustischen und perzeptiven Aufteilung entspricht, welche von Sprechern und Hörern präferiert wird, wäh-

rend das für einen Kontrast zwischen /ɕ ʃ/ scheinbar nicht der Fall ist. Da nach dem ersten Lautwandelprozess im 13. Jahrhundert der alveolopalatale Sibilant aufkam, wurde der palatoalveolare Laut in der zweiten Stufe des Lautwandels zum retroflexen Frikativ weiterentwickelt, sodass eine bessere Verteilung in der postalveolaren Klasse der Frikative hergestellt war (Żygis & Hamann, 2003; Żygis & Padgett, 2010). Weiterhin kann aus einer artikulatorischen Perspektive für die retroflexe Natur des postalveolaren Lautes argumentiert werden. Hamann (2002a) beschrieb die Artikulation der retroflexen Laute anhand von Röntgenaufnahmen. Sie berichtete, dass der Zungenrücken bei der Lautproduktion flach und rückverlagert ist, sowie dass die Zungenspitze den aktiven Artikulator darstellt, was den beschriebenen Merkmalen von retroflexen Lauten in Ladefoged und Maddieson (1996) und Keating (1991) entspricht. Weitere Details zur Artikulation und zur Akustik der sibilantischen Lautklasse des Polnischen folgen in den Abschnitten 1.4.2 und 1.4.3.

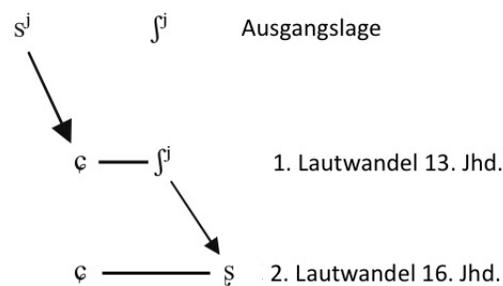


Abbildung 1.2: Lautwandel im Polnischen Sibilantinventar. (Abbildung entnommen aus Żygis und Padgett, 2010)

Neben der Vielzahl an Konsonanten umfasst das polnische Vokalsystem sechs Oral-, und zwei Nasal-Vokale (vgl. Abb. 1.3).

Die oralen Vokale /i ɛ a ɔ u/ haben kein in der Gespanntheit kontrastierendes Gegenüber, wie es etwa im Deutschen der Fall ist.

Die Nasalvokale werden als eine Art Diphthong beschrieben, da sie in den meisten Fällen vor einem nasalierten labio-velaren Approximanten [w̃] vorkommen (Gussmann, 2007). Der

Vokal /i/ hat zum einen eine rein orthographische Funktion, d.h. er zeigt die Palatalisierung des vorgehenden Konsonanten an. Ein Beispiel hierfür ist das Wort „siano“ /ɕano/ (Heu), in welchem der Buchstabe <i> keinen vokalischen Wert hat. Zum anderen steht der Buchstabe <i> für den hohen, vorderen Vokal /i/, wie im Wort „sini“ /ɕini/ (pl. Pfahl).

i	i	u
ɛ ɛ̃		ɔ ɔ̃
	a	

Abbildung 1.3: Vokale im Standard-Polnischen. (Abbildung entnommen aus Nowak, 2006b)

Die Konsonanten und Vokale des Polnischen verbinden sich in verschieden Onset und Reim Kombinationen zu Silben. Im wortinitialen Onset sind Kluster mit bis zu vier Konsonanten erlaubt, wortmedial und in der wortfinalen Koda sogar bis zu fünf Konsonanten (Rubach & Booij, 1990; Zydorowicz, 2010). Der Wortakzent im Polnischen fällt, mit nur einigen wenigen Ausnahmen, grundsätzlich auf die vorletzte Silbe (Gussmann, 2007; Sawicka, 2001).

1.4.2 Regionale Varietäten des Polnischen

Die Dialektlandschaft des Polnischen lässt sich traditionell in vier große Dialektregionen unterteilen, welche wiederum in viele kleine regionale Varietäten aufgegliedert sind (Rothstein, 2005; Sussex & Cubberley, 2006).

Großpolnischer Dialekt in der Region der Städte Posen und Gnesen im Westen Polens (vgl. braune Region „The Greater Poland dialect“ in Abb. 1.4). Diese Varietät beeinflusste die Form der literarischen Sprache in Polen und weist damit weniger regionale und eher Standardsprachliche Merkmale auf.

Kleinpolnischer Dialekt in der Region der Stadt Krakau im Süd-Osten Polens (vgl. graue Region „The Lesser Poland dialect“ in Abb. 1.4), welcher neben dem Großpolnischen Dialekt ebenfalls ein Einflussfaktor auf die moderne Polnische Sprache war.

Masowischer Dialekt östlich und nord-östlich der Hauptstadt Warschau (vgl. orange Region „The Mazovia dialect“ in Abb. 1.4).

Schlesischer Dialekt im Süd-Westen in der Region um die Stadt Kattowitz (vgl. grüne Region „The Silesia dialect“ in Abb. 1.4).

Ein fünfter Dialekt, dessen Status nicht eindeutig geklärt ist, ist das Kaschubische im Norden Polens (vgl. graue Region „The Kashubian language“ in Abb. 1.4). Diese Varietät unterscheidet sich stark von der polnischen Standardsprache und auch von den anderen polnischen dialektalen Varietäten, was die Frage nach sich zieht, ob das Kaschubische nicht als eigenständige literarische Sprache angesehen werden kann (Rubach, 2005; Sussex & Cubberley, 2006). Des weiteren zählen zu den polnischen regionalen Varietäten die sogenannten *new mixed dialects* (vgl. gelbe Region „New mixed dialects“ in Abb. 1.4), welche im Norden, Westen und Nord-Osten Polens vorkommen (Milobog & Garrett, 2011). Die *new mixed dialects* sind das Resultat der politischen Situation in Poland nach dem Zweiten Weltkrieg, in welcher bestimmte Teile der polnischen Bevölkerung durch die kommunistische Regierung in eben die Region der neuen Dialekte umgesiedelt wurde. Diese Umsiedlung und die daraus resultierende Mischung verschiedener polnischer regionaler Varietäten, brachte diese neuen Varietäten hervor (Milobog & Garrett, 2011).

Traditionell werden die Dialekte zunächst grob hinsichtlich des Vorkommens des Phänomens *Mazurzenie* untergliedert, welches den Zusammenfall der dentalen /s z ts dz/ und

der retroflexen /ʂ z, tʂ dz/ Sibilantklasse beschreibt (Sussex & Cubberley, 2006). Dieser Zusammenfall kommt in der masowischen, der kleinpolnischen und im nördlichen Teil der schlesischen Varietät vor. Dieses Phänomen fand bisher ausschließlich in rein deskriptiven Erläuterungen zur Polnischen Sprache eine kurze Erwähnung, der Autorin dieser Arbeit sind jedoch keine näheren Erläuterungen und experimentellen Studien zum Prozess *Ma-zurzenie* bekannt.

Weiterhin divergieren die polnischen Dialekte hinsichtlich der regressiven Assimilation der Stimmhaftigkeit über die Wortgrenzen hinweg. Für die Dialektregionen Groß- und Kleinpolen und Schlesien werden wortfinale stimmlose Konsonanten stimmhaft realisiert, wenn ihnen ein Wort mit einem initialen Vokal oder Sonorant folgt. In der masowischen Dialektregion ist diese Assimilation der Stimmhaftigkeit nicht zu finden.

Neben diesen beiden weitgreifenden Dialektmerkmalen finden sich in den einzelnen Varietäten weitere spezifischere dialektale Unterschiede, wie beispielsweise die Denasalisierung der Nasalvokale in der kleinpolnischen Region, oder das Ausbleiben der Velarisierung von /n/ hin zu [ŋ] vor velaren Lauten in der masowischen Varietät (Rothstein, 2005; Sussex & Cubberley, 2006), auf die in dieser vorliegenden Arbeit nicht weiter eingegangen wird (für weitere Details zu den einzelnen dialektalen Merkmalen im Polnischen vgl. Sussex & Cubberley, 2006).



Abbildung 1.4: Regionale Einteilung der Polnischen Dialekte. (Abbildung entnommen aus Milobog & Garrett, 2011)

Die Literatur, die in der vorliegenden Studie, zum heutigen regionalen Sprachgebrauch im Polnischen genannt werden kann, ist eher spärlich. Gründe dafür sind zum einen die fehlenden dialektologischen und soziolinguistischen Studien zum Polnischen (Sussex & Cubberley, 2006), zum anderen ist die vorhandene Literatur zu dieser Thematik oftmals ausschließlich polnischsprachig.

In einer Studie, welche sich hauptsächlich mit dem Unterrichten der polnischen Sprache in Polen befasste, wurde jedoch angemerkt, dass in der heutigen Sprachsituation die vielfältigen regionalen Varietäten des Standard-Polnischen der Standardsprachlichen Form weichen und, dass gut erhaltene dialektale Sprachformen hauptsächlich bei älteren Sprechern und in ländlichen Gegenden gefunden werden können (Berlinska, 1998). Auch in Milobog und Garrett's (2011) Studie, welche die Haltung zweier regionaler Sprechergruppen gegenüber den verschiedenen dialektalen Varietäten untersuchte, wiesen die Ergebnisse darauf hin, dass die heutigen regionalen Varietäten in ganz Polen eher der Standardsprache und we-

niger den dialektalen Varietäten entsprechen.

1.4.3 Artikulation der stimmlosen Sibilanten /s ʃ ʒ/

Sibilanten zählen zur Lautklasse der Frikative. Die Artikulation dieser Lautklasse fordert einen hohen Grad an Präzision, denn bei ihrer Produktion bilden die Zunge und der Gaumen eine Verengung im Vokaltrakt, welche dazu führt, dass die aus der Lunge strömende Luft verwirbelt wird. Jedoch darf die Verengung nicht zu stark sein, d.h. es darf kein Verschluss gebildet werden, da dieser Verschluss die ausströmende Luft blockieren würde. Wichtig ist dennoch, dass die Verengung klein genug ist, um Turbulenzen des Luftstroms entstehen zu lassen (Iskarous, Shadle & Proctor, 2011). Für die Leitung des Luftstroms spielt bei Sibilanten weiterhin nicht nur der vordere Zungenteil – Zungenspitze oder Zungenblatt – eine Rolle, sondern auch der Zungenrücken, welcher eine Rille bildet, um den ausströmenden Luftstrom zielgerichtet hin zur Verengung zu leiten. Diesem ersten Hindernis – der Verengung zwischen Gaumen und Zunge – folgt bei Sibilanten ein weiteres Hindernis. So bezeichnen Ladefoged und Maddieson (1996) Sibilanten auch als *obstacle fricatives*, da bei ihrer Artikulation zusätzlich zu den Luftverwirbelungen, die durch die Verengung zwischen Zunge und Gaumen entstehen, weitere Turbulenzen hervorgerufen werden, indem der ausströmende Luftstrom auf die Schneidezähne trifft. Hierbei entsteht das für Sibilanten typische hochfrequente Zischen, welches ihnen auch den Namen *Zischlaute* einbringt. Ebendiese präzise und oft auch als statisch angesehene Artikulation dieser Lautklasse (Bladon & Nolan, 1977), sowie die Beteiligung des Dorsums bei der Artikulation von Sibilanten, um den ausströmenden Luftstrom zu leiten (Stone & Lundberg, 1996), führte dazu, dass Sibilanten, im Vergleich zu Plosiven oder Nasalen als eher koartikulationsresistent angesehen werden (Recasens & Espinosa, 2009). Die koartikulatorische Resistenz ist als der Grad an kontextabhängiger, artikulatorischer Variation von Vokalen und Konsonanten definiert. Es wird angenommen, dass das Maß an koartikulatorischer Resistenz mit steigenden artikulatorischen Einschränkungen zunimmt (Recasens & Espinosa, 2009; Recasens, Pallarès & Fontdevila, 1997). Bei der Klassifikation der Lautklassen in weniger beziehungsweise in stärker resistent für Koartikulation werden die Abstufungen je nach beteiligtem Artikulator gemacht: alveolare Laute, welche mit der Zungenspitze produziert werden, weisen eine weniger starke Resistenz für Koartikulation auf, während beispielsweise alveopalatale Laute durch die Beteiligung des Zungendorsums bei deren Artikulation eine starke Resistenz gegenüber koartikulatorischen Effekten aufzeigen sollten. Recasens und Espinosa (2009) beschreiben auch, dass die Artikulationsart der Laute ein Einfluss-

faktor auf den Grad der koartikulatorischen Resistenz ist, sodass eben Sibilanten mit ihren artikulatorischen Einschränkungen, welche notwendig sind, um das Friktionsgeräusch zu erhalten, vermutlich stark resistent gegenüber Koartikulation sind.

Betrachtet man den Vokaltrakt aus der seitlichen, d.h. aus der sagittalen, Perspektive so können die Sibilanten anhand ihres Konstriktionsorts gegliedert werden. Angefangen bei der am nächsten an den Lippen liegenden Produktion, nämlich bei dentalen Sibilanten. Der dentale Frikativ weist eine Konstriktion zwischen der Zunge und den oberen Schneidezähnen auf, was die *front cavity*, d.h. den Teil des Vokaltrakts, der zwischen der Verengung und den Lippen liegt, sehr kurz werden lässt (vgl. Abb. 1.5). Dem dentalen Sibilanten folgt der alveolare Laut /s/, welcher durch eine Verengung zwischen Zunge und Alveolen⁶ produziert wird. Sprach- und sprecherspezifisch ist die Wahl zwischen Zungenspitze und Zungenblatt als artikulierendes Organ bei der Artikulation der dentalen und alveolaren Sibilanten zulässig (Ladefoged & Maddieson, 1996). Der Zungenrücken liegt für den dentalen und alveolaren Sibilanten meist tiefer mit hochgeklappten Zungenrändern, sodass eine Rille entlang der anterioren-posterioren-Achse der Zunge entsteht. Diese Rille hilft dabei den ausströmenden Luftstrom in Richtung Konstriktion zu leiten. Generell kann der anteriore Sibilant /s/ je nach sprachspezifischen und individuellen anatomischen Gegebenheiten auch dental realisiert werden (vgl. Keating, 1991; Dart, 1991), wie es zum Beispiel für den anterioren Sibilanten im Polnischen (Gussmann, 2007) und im Mandarin-Chinesischen (Ladefoged & Maddieson, 1996) beschrieben wurde.

Der Konstriktionsort der Sibilanten, die in die Klasse eines eher posterioren Artikulationsorts fallen, ist hinter den Alveolen und am Anfang des weichen Gaumens zu finden. Unter diesem weiter hinten gelegenen, postalveolaren Artikulationsort können der retroflexe, der palato-alveolare sowie der alveolopalatale Sibilant zusammengefasst werden. Für die Unterscheidung der postalveolaren Sibilantkategorie ist neben der Lage der Konstriktion die Zungenform von Bedeutung, da diese weitere Veränderungen des Resonanzraumes hervorruft und damit ebenso wie der Artikulationsort zu den charakteristischen akustischen und perceptiven Merkmalen dieser Lautklasse beiträgt (Shadle, 1991). Der palato-alveolare Sibilant /ʃ/, wie in deutsch „Busch“ /buʃ/ ist durch eine Annäherung des Zungenblattes an die Stelle direkt hinter dem Zahndamm definiert, wobei sich hier der Zungenrücken sprachspezifisch ebenfalls in Richtung Gaumen wölben kann (Keating, 1991; Ladefoged & Maddieson, 1996). In einigen Sprachen der Welt ist der palato-alveolare Sibilant zusätzlich

⁶Die Alveolen entsprechen der Region im Vokaltrakt, die direkt hinter den oberen Schneidezähnen und vor dem weichen Gaumen liegt. Diese Region wird auch Zahndamm genannt.

durch gerundete Lippen gekennzeichnet, wie es zum Beispiel im Deutschen, Englischen und Französischen der Fall ist, während im Vergleich hierzu im Russischen gewöhnlich keine Lippenrundung bei der Produktion des Sibilanten /ʃ/ gefunden wird (Ladefoged & Maddieson, 1996). In Gussmann (2007) und Spencer (1986) wird zwar zur Darstellung der polnischen, retroflexen Sibilanten das IPA Symbol /ʃ/ verwendet, welches für den palato-alveolaren Laut steht, jedoch werden von Hamann (2002a, 2002b) einige Argumente für den retroflexen Status des polnischen postalveolaren Lauts geliefert. Der retroflexe Sibilant variiert in seiner Artikulation stärker als die vorangegangenen Laute, wobei die verschiedenen Arten der Artikulation mit der jeweils untersuchten Sprache variieren. Genauer gesagt gibt es im Ausmaß der Retroflexion eine graduelle Abstufung von starken/„echten“ retroflexen Sibilanten, wie sie zum Beispiel im Sibilantinventar der Sprache Toda gefunden werden können und bei deren Artikulation sich die Zungenspitze nach oben und hinten rollt und sich damit dem weichen Gaumen für eine Engebildung nähert, bis hin zu schwächeren retroflexen Sibilanten, die eher mit einer flachen Retroflexion beschrieben werden, wie es für die Sibilanten im Polnischen der Fall ist. Ladefoged und Maddieson (1996) beschreiben die Konstriktion der retroflexen Laute des Polnischen als eher rückverlagert verglichen mit dem anterioren Laut /s/, wobei sich im Falle des retroflexen Lautes die Oberfläche der Zungenspitze dem Konstriktionsort nähert (vgl. Abb. 1.5). Die Zunge ist für die weniger ausgeprägten retroflexen Laute des Polnischen eher flach, was diese Lautklasse ebenfalls von den etwas gewölbten palato-alveolaren Lauten des Englischen unterscheidet. Während für den dentalen/alveolaren Laut die Zunge die Zähne berührt, entsteht für den retroflexen Laut durch die rückverlagerte Zunge eine Art „Höhle“ unter der Zunge – auch bekannt als *sublingual cavity*. Hamann (2002a) beschreibt die polnischen Sibilanten ebenfalls als retroflexe Laute. Ihre Argumente für die Retroflexion der Laute ist kompatibel mit denen von Ladefoged und Maddieson (1996): zum einen beschreibt sie die Velarisierung der retroflexen Laute, d.h. die Rückverlagerung des kompletten Zungenkörpers, um ein Anheben der Zungenspitze zu ermöglichen, zum anderen kommt durch das nach hinten Ziehen und Dehnen der Muskeln im Zuge der Rückverlagerung das zweite Merkmal der Retroflexion auf, nämlich der flache Zungenrücken.

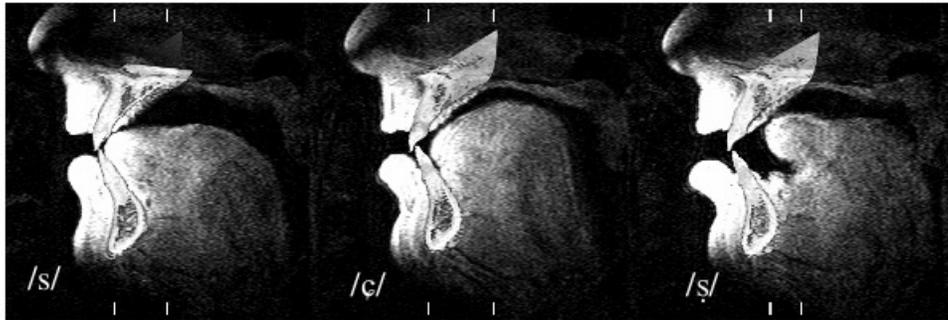


Abbildung 1.5: MRT Bilder während der Produktion der Polnischen Sibilanten: dental /s/ links, retroflex /ʂ/ rechts, alveopalatal /ɕ/ mittig. (Abbildung entnommen aus Toda, Maeda & Honda, 2010)

Der alveopalatale Frikativ /ɕ/ ist durch seine Artikulation mit dem Zungenblatt definiert, welches sich in Richtung des Zahndamms bewegt (Ladefoged & Maddieson, 1996). Bei der Artikulation des alveopalatalen Frikatives ist in den meisten Fällen sowohl der vordere Teil der Zunge als auch das Dorsum zum Gaumen hin gewölbt, sodass die Zunge relativ hoch im Vokaltrakt liegt und mit dem Gaumen einen langen flachen Kanal bildet (vgl. Abb. 1.5). Der Konstriktionsort dieses Sibilanten im Polnischen ist zwar weiter hinten als es für /s/ der Fall ist, er muss aber nicht so weit rückverlagert sein, wie es für den retroflexen Sibilanten gefunden wird. An der Beschreibung oben und in Abb. 1.5 kann man erkennen, dass sich der dentale und der retroflexe Laut im Polnischen hauptsächlich hinsichtlich der Lage der Konstriktion im Vokaltrakt in der anterior-posterior Dimension unterscheiden. Der alveopalatale Laut /ɕ/ grenzt sich dagegen vorwiegend in der Zungenform von den beiden anderen Sibilanten ab. Die artikulatorischen Merkmale führen

zu unterschiedlichen akustischen Charakteristika der drei polnischen Sibilanten, welche im nachfolgenden Abschnitt 1.4.4 näher erläutert werden sollen.

1.4.4 Akustische Eigenschaften der Sibilanten /s ʂ ʐ/

Im vorangegangenen Abschnitt (1.4.3) wurden die artikulatorischen Unterschiede des sibilantischen Dreier-Kontrasts im Polnischen beschrieben. Ein wichtiges physiologisches Merkmal ist die Lage der Konstriktion in der anterior-posterior Dimension des Vokaltrakts, welche Unterschiede in der Verteilung der Energie im Spektrum des Friktionsgeräusches mit sich bringt (Hughes & Halle, 1956; Shadle, 1991). Genauer gesagt beeinflusst die Lage der Konstriktion die spektralen Frequenzschwerpunkte der jeweiligen Sibilanten. Diese spektralen Unterschiede zwischen Sibilanten sind vor allem auf die Resonanzen des Teils des Vokaltraktes zurückzuführen, der zwischen der Verengung und den Lippen liegt, also auf die *front cavity*. Je weiter vorne der Sibilant produziert wird und je kleiner der Raum zwischen der Verengung und den Lippen, desto höher liegt der Großteil der spektralen Energie. Bezogen auf die Sibilanten des Polnischen bedeutet dies, dass der dentale Frikativ eine Energieansammlung in den hohen Frequenzbereichen meist bei über 6 kHz aufweist (Nowak, 2006a; Żygiś & Hamann, 2003). Der retroflexe Laut hingegen mit seinem zurückverlagerten Konstriktionsort und der zusätzlichen Lippenrundung hat eine längere *front cavity*, was die spektrale Energie dieses Lautes in einen Frequenzbereich unter 4 kHz sinken lässt. Der Konstriktionsort für den alveolopalatalen Sibilanten ist ebenfalls weiter hinten im Vokaltrakt, was laut Nowak (2006a) für die akustische Ähnlichkeit der Laute /ʂ ʐ/ spricht, jedoch berichten Flemming (2002) und Żygiś und Hamann (2003) für diesen Laut von einem Frequenzschwerpunkt, der zwischen dem des dentalen und dem des retroflexen Sibilanten im Bereich zwischen 4 und 6 kHz liegt.

Kudela (1968) zeigte, dass sich die polnischen Sibilanten hinsichtlich ihres spektralen Gipfels unterscheiden lassen. Für die postalveolaren Sibilanten /ʂ ʐ/ wurde gefunden, dass für beide Laute prominente spektrale Gipfel im Frequenzbereich zwischen dem dritten und vierten Formanten vorkommen. Jedoch wurde für den retroflexen Sibilant ein weiterer hoher Gipfel im Bereich des zweiten Formanten gemessen, wodurch dieser vom alveolopalatalen Frikativ unterschieden werden konnte. Żygiś und Padgett (2010) merken in ihrer Studie zu den polnischen Frikativen jedoch an, dass die Analyse spektraler Gipfel in Studien zu Sibilanten einige Schwierigkeiten aufbringt, da in einem Sibilantspektrum mehrere Gipfel vorkommen können und die Präsenz dieser Gipfel von Token zu Token variieren kann.

Ein Vorschlag für eine stabilere Methode zur akustischen Unterscheidung von Sibilanten, ist die Analyse der spektralen Momente (Forrest, Weismer, Milenkovic & Dougall, 1988; Jongman, Wayland & Wong, 2000). In der Analyse der spektralen Momente wird das Spektrum als eine Wahrscheinlichkeitsverteilung betrachtet, sodass die statistischen Momente berechnet werden können (Forrest et al., 1988). Mittels der ersten vier spektralen Momente – Mittelwert, Varianz, Schiefe und Kurtosis – besteht die Möglichkeit die Form des Spektrums von Sibilanten, zu beschreiben. Das erste spektrale Moment, auch bekannt als *center of gravity* (kurz CoG), zeigt den Frequenzschwerpunkt von Sibilanten an. Anhand dieses Parameters konnten die drei polnischen Sibilanten in Żygis und Hamann (2003) voneinander unterschieden werden und zwar insofern, als dass der dentale Sibilant den höchsten CoG und der retroflexe Laut den niedrigsten CoG-Wert aufwies, während der spektrale Schwerpunkt des alveolopalatalen Frikativs zwischen den beiden ersteren lag. Für das zweite Moment, die Varianz, gilt, dass es die Abweichung des Spektrums vom Frequenz-Mittelwert angibt. Es scheint weniger gut geeignet, um Sibilanten voneinander zu unterscheiden. Das zweite Moment ist ein Parameter, durch welchen sich eher Sibilanten von nicht-Sibilanten unterscheiden lassen (Jongman et al., 2000). Das dritte spektrale Moment, die Schiefe, kann negative und positive Werte haben, da für die Berechnung die Energien oberhalb und unterhalb des Frequenz-Mittelwerts voneinander subtrahiert werden. Das dritte spektrale Moment ist ein weiterer Parameter, welcher die Unterscheidung der Sibilanten ermöglicht. Für den englischen Frikativ /ʃ/ wurden positive Werte des dritten spektralen Moments gefunden, welche auf eine Energiekonzentration unterhalb des Mittelwertes hinwiesen, während für den alveolaren Sibilanten negative Werte gezeigt wurden, die für eine Energiekonzentration oberhalb des Mittelwerts sprachen (Forrest et al., 1988). Die Kurtosis – das vierte spektrale Moment – gibt an, inwieweit das Spektrum von einer Normalverteilung abweicht. Jongman et al. (2000) zeigten für diesen Parameter, der die Unterschiede in den Frequenzgipfeln im Spektrum wiedergibt, die Möglichkeit Sibilanten von nicht-sibilantischen Lauten zu unterscheiden.

Eine weitere Methode, welche von Watson und Harrington (1999) zur Kategorisierung von Sprachlauten vorgeschlagen wurde, ist die *Discrete Cosinus Transformation* (kurz DCT). Die DCT zerlegt das Signal in Cosinuskurven. Werden diese Cosinuskurven wieder zusammengefasst, wird das ursprüngliche Signal rekonstruiert (Harrington, 2010). Die Amplituden der einzelnen Cosinuskurven geben die DCT-Koeffizienten wieder. Durch die verschiedenen Koeffizienten ist es möglich die Form verschiedener Sprachsignale zu be-

schreiben. Die DCT-Koeffizienten [0, 1, 2, 3] geben den Mittelwert, die Steigung, die Krümmung und die Amplitude der höheren Frequenzen eines Spektrums wieder. Janney, Weirich und Helmeke (2015) zeigten in ihrer Studie zur Akustik der Sibilanten /ç ʃ/ in einer nicht-Standardsprachlichen Varietät des Deutschen eine gute Trennung der Laute anhand der DCT-Koeffizienten [1, 2, 3]. Für das Polnische wurde eine Unterscheidung der Sibilanten mittels der DCT-Koeffizienten [1, 2] gefunden (Guzik & Harrington, 2007). In Bukmaier, Harrington, Reubold und Kleber (2014) konnten die drei polnischen Sibilanten ausschließlich mit Hilfe des zweiten DCT-Koeffizienten unterschieden werden. Genauer gesagt wurde in ihrer Studie das logarithmisierte Verhältnis der Euklidischen Distanzen der Sibilanten zu den Zentroiden /s ʂ/ im DCT [2]-Raum berechnet. Es wurde festgestellt, dass der dentale Sibilant /s/ sehr gut von /ʃ ʂ/ unterschieden werden kann, während die beiden letzteren ähnlichere Werte zeigten (für weitere Details zum methodologischen Vorgehen siehe Bukmaier et al., 2014, sowie Kapitel 2 und 4 dieser Arbeit).

Während der Konstriktionsort bei Sibilanten charakteristische Eigenschaften in Hinblick auf den spektralen Frequenzschwerpunkt hervorbringt, werden durch die unterschiedlichen Zungenformen die Transitionen hin zu den umliegenden Vokalen beeinflusst. Im Abschnitt zur Artikulation der Sibilanten wurde aufgezeigt, dass sich die Sibilanten dahingehend in ihrer Zungenform unterscheiden, als dass der dentale und der retroflexe Sibilant eher eine flache Zungenform aufweisen, während der alveolopalatale Frikative mit einem gewölbten Zungendorsum artikuliert wird (vgl. Abb. 1.5). Diese Zungenformen gehen einher mit unterschiedlichen Einflüssen auf die umliegenden Laute. Das Anheben des Zungendorsums in Richtung Gaumen verkürzt für den alveolopalatalen Laut /ç/ den Bereich hinter der Konstriktion – auch bekannt als *back-cavity* –, was wiederum einen erhöhten zweiten Formanten zur Folge hat, da der zweite Formant invers mit der Länge der *back-cavity* korreliert ist (Nitttrouer, Studdert-Kennedy & McGowan, 1989). Die Sibilanten, die mit einem flacheren Zungenrücken artikuliert werden, zeigen hingegen niedrige und im Falle des dentalen und des retroflexen Sibilanten des Polnischen sehr ähnliche Transitionen des zweiten Formanten. K. Stevens, Li, Lee und Keyser (2004) fanden in ihrer Studie zu den Sibilanten des Mandarin-Chinesischen unterschiedliche F2-Werte zum zeitlichen Onset des Vokals, welcher den jeweiligen Sibilanten folgte. Für das Polnische beobachtete Nowak (2006a), dass die Transitionen des zweiten Formanten von Hörern bei der perceptiven Kategorisierung der polnischen Sibilanten genutzt werden, wobei Bukmaier, Harrington, Reubold und Kleber (2014) und Nowak (2006a) zeigen konnten, dass dies gerade für den alveolopalatalen

Sibilanten ein wichtiges Merkmal ist, welches /ɕ/ von /s ʂ/ unterscheidet.

Des Weiteren bilden für den retroflexen Laut die Werte des dritten Formanten ein charakteristisches Merkmal. Durch die angehobene und nach hinten gerollte Zungenspitze entsteht eine Höhle unter der Zunge – die *sublingual cavity* –, welche die Transitionswerte des dritten Formanten vom retroflexen Frikativ hinein in die benachbarten Vokale sinken lässt (Hamann, 2002a; Ladefoged & Maddieson, 1996).

In den meisten bisher bekannten Studien zu dem Dreier-Kontrast der Sibilanten /s ʂ ɕ/ wurden ausschließlich die spektralen Merkmale im Friktionsanteil der Laute und die Transitionen des zweiten Formanten zum Onset des Vokals, welcher den Sibilanten nachfolgte, untersucht. In der vorliegenden Studie werden die Transitionen des dritten Formanten ebenfalls in die Analyse der polnischen Sibilanten eingebunden, um festzustellen, ob in den spektralen Eigenschaften im Friktionsgeräusch, den F2 und den F3-Transitionen Charakteristika gefunden werden können, welche die relative Instabilität des Dreier-Kontrasts in einigen polnischen Varietäten erklären können.

1.5 Lautwandel

Wie in den Abschnitten 1.1 und 1.2 anhand einiger Beispiele aufgezeigt, ist Sprache flexibel und wandelbar. Es existieren Variationen aufgrund vieler verschiedener Faktoren. Sprachliche Variationen und Veränderungen finden sich sowohl über Jahrhunderte und damit über verschiedene Sprachstufen hinweg als auch in einem einzelnen Sprecher zwischen zwei Äußerungen. Einige bekannte Lautwandelprozesse kommen nicht in einer einzelnen Sprache vor, sondern sind in mehreren nicht miteinander verwandten Sprachen zu verschiedenen Zeitpunkten zu beobachten. Hinsichtlich solcher universeller Sprachwandel fragte Ohala (1971), ob in diesen Lautwandelprozessen universelle Tendenzen des menschlichen Lautsystems zu finden sind und, ob diese Tendenzen in den artikulatorischen und perzeptiven Mechanismen begründet liegen, die die Sprecher und Hörer unterschiedlichster Herkunft gemeinsam haben und die auf der ganzen Welt zum Zweck der Kommunikation genutzt werden. Eine ähnliche Herangehensweise ist in Garrett und Johnson's (2012) Artikel *Phonetic biases in sound changes* zu erkennen, denn sie argumentieren wie folgt: Während man manche Lautwandelmuster sehr häufig und immer wieder beobachten kann, sind andere mögliche Formen von Wandel in den Sprachen der Welt bisher nicht zu finden. Das lässt sie schließen, dass es bestimmte Tendenzen – oder eben bestimmte *biases* – gibt, die gewisse Lautwandelmuster wahrscheinlich beziehungsweise andere wiederum eher unwahr-

scheinlich machen.

Die heutigen vielfältigen Erkenntnisse hinsichtlich Variation und Wandel verdanken wir zu einem Großteil der Arbeit der historischen Sprachwandelforschung des 18. Jahrhunderts. Schon die frühe Linguistik befasste sich mit Lautwandel. Dabei wurde in den anfänglichen linguistischen Untersuchungen komparativ gearbeitet, d.h. ausgehend von sehr frühen Sprachstufen, welche hauptsächlich in schriftlichen Quellen zu finden sind, wurden Vergleiche mit anderen Sprachen und den dort vorkommenden Sprachstufen, sowie mit nachfolgenden Sprachstufen angestellt und diese beschrieben. Dies geschah zum einen, um gemeinsame Ursprachen zu rekonstruieren und zum anderen, um die Entwicklung der Sprachen der Welt aufzuzeigen. Durch diese schrittweisen Vergleiche mit immer jüngeren Sprachstufen konnten anhand von unterschiedlichen Wortformen und Schreibweisen Lautwandelprozesse erkannt und nachvollzogen werden. Jedoch konnten in diesen frühen Beobachtungen nur Start- und Endzustand der verschiedenen Lautwandel betrachtet und beschrieben werden. Das Wissen zum „Wie genau?“, und vorallem zum „Warum?“ hinter diesen Lautwandelprozessen blieb bei den anfänglichen Betrachtungsweisen zum Großteil aus. Eine Analogie zur Chemie, die von Ohala (1971) sehr passend beschrieben wird, zeigt den Wandel in der Beschreibung von Variationen und von Veränderungen mit der Weiterentwicklung der Forschung. In der frühen Beschreibung von chemischen Prozessen wurden grundlegende Vorgänge, wie zum Beispiel *Eisen + Luft = Rost*, entdeckt und beschrieben. Mit dem Fortschritt der Forschung konnten die Beschreibungen von verschiedenen Prozessen präzisiert werden, sodass die genaue Zusammensetzung der beteiligten Komponenten aufgeführt werden konnte. Man kam zu einer Darstellung der chemischen Reaktion, die bis heute bekannt ist: $4\text{Fe} + 3\text{O}_2 \rightarrow 2\text{Fe}_2\text{O}_3$

Auch wenn in der ersten Darstellung (*Eisen + Luft = Rost*) der zu beobachtende Prozess vollkommen richtig niedergeschrieben wird, bleiben dem Beobachter tieferliegende Details und Strukturen in einer solchen Veränderung verborgen. In der zweiten Art und Weise den chemischen Prozess zu beschreiben, werden wesentlich mehr Informationen und Details zu dieser chemischen Reaktion preisgegeben.

Ähnlich verhält es sich auch mit der komparativen Methode der historischen Linguistik. In der Lautwandelforschung sind viele verschiedene Lautwandelprozesse bekannt, dennoch fehlt eine Begründung für sich wandelnde Sprachen und die genaue Motivation von Lautwandelprozessen. In der modernen Lautwandelforschung stehen die Fragen „Wie?“ und „Warum?“ hinsichtlich sprachlicher Veränderungen im Fokus. Die schon von Weinreich,

Labov und Herzog (1968) hierzu aufgeführten fünf Punkte sind sowohl für die heutige Erforschung von Sprachwandel als auch für die Bildung von Lautwandeltheorien, nach wie vor ein geeigneter Leitfaden:

The constraints problem: Welche Sprachwandelprozesse sind möglich? Wo liegen Einschränkungen hinsichtlich Sprachwandel?

The transition problem: Wie weiten sich sprachliche Innovationen innerhalb der Sprachgemeinschaft aus?

The embedding problem: Wie sieht die Sprachstruktur in Hinblick auf den sprachlichen Wandel aus? Inwiefern hat die Veränderung einen Einfluss auf die soziale Struktur der Sprachgemeinschaft?

The evaluation problem: Wie beeinflusst die Wahrnehmung und die Evaluierung der Variation den Lautwandel?

The actuation problem: Wie hängen synchrone Variation und Lautwandel zusammen? Welche Rolle spielen dabei interne und externe Faktoren?

Um neue Erkenntnisse zu den oben aufgeführten Fragen zu bekommen, fokussieren neuere Methoden hinsichtlich der Lautwandelforschung in ihren Untersuchungen aktuell stattfindende Lautwandelprozesse. Seit den sozio-phonetischen Studien von Labov (1963) ist bekannt, dass es in der Lautwandelforschung nicht grundsätzlich notwendig ist, in die Vergangenheit zu blicken, um Sprach- und Lautwandel zu finden und zu beschreiben. Lautwandel und dessen Fortschritt können im Hier und Jetzt analysiert werden. Zwei Vorgehensweisen machen das möglich: 1. das sogenannte *apparent time design* und 2. das *real time design*. Ersteres, ist auf die Beobachtungen von Lautwandel im Fortschritt von Labov (1963) zurückzuführen. In seinen Studien wurde die Sprache von verschiedenen nebeneinander lebenden Generationen phonetisch analysiert und miteinander verglichen, um den Zusammenhang von linguistischen und extralinguistischen Faktoren und Lautwandel zu untersuchen. Die *real time* Methode hingegen untersucht einzelne Individuen und deren Sprache über einen längeren Zeitraum hinweg (Harrington, 2007). Mit diesen Untersuchungsmethoden wird auch die Annahme der Sprachforschung gestützt, dass Lautwandel

nicht kategorial ablaufen müssen, sondern dass es die feinen phonetischen Details im akustischen Signal und in der Perzeption sind, welche Erklärungen für das, von Generation zu Generation, schrittweise Fortschreiten eines Wandels liefern.

Die Frage nach den Gründen für Lautwandel fokussiert die verschiedenen Umstände, welche Lautwandel bedingen können. Es existieren sprachliche Wandel, welche durch die interne linguistische Struktur und andere, welche durch extra-linguistische Umstände hervorgerufen werden. Lautwandel durch interne linguistische Strukturen sind, laut Hickey (2012), durch die *properties of language*, d.h. durch die Eigenschaften des linguistischen Systems an sich, zu erklären. Unter solche Lautwandel fallen beispielsweise nicht-phonemische Wandel, wie die *vowel shifts*, und phonemische Veränderungen, wie *splits* und *mergers* (Torgersen & Kerswill, 2004). Die Unterscheidung zwischen phonemischen und nicht-phonemischen Sprachwandel besteht darin, dass phonemische Wandel die Anzahl an Phonemen in einem Lautsystem verändern, während nicht-phonemische, oder auch allophonische, Wandel die Anzahl der Phoneme einer Sprache nicht beeinflussen (für konkrete Beispiele zu phonemischen und nicht-phonemischen Wandel siehe Campbell, 2004).

Extra-linguistische Gegebenheiten fassen die Umstände zusammen, bei welchen das Verhalten des Sprechers im Vordergrund steht (Hickey, 2012). Beispiele für extra-linguistisch-motivierte Lautwandel sind Sprachkontakt und soziolinguistische Faktoren, wie die Gruppenzugehörigkeit der verschiedenen Sprecher und/oder das Prestige von verschiedenen Sprechergruppen.

Die Trennung von internen und extra-linguistischen Faktoren in einem Lautwandel ist jedoch nicht immer deutlich. Beim genaueren Hinsehen, können Lautwandel aufgedeckt werden, bei welchen Spuren von einem Zusammenspiel von internen und externen Faktoren erkannt werden können (Hickey, 2012; Torgersen & Kerswill, 2004).

Die Fragen nach dem „Warum?“ und dem „Wie?“ in Zusammenhang mit Sprachwandel umfassen Informationen hinsichtlich des Sprachsystems und bezüglich der Sprachgemeinschaft und ihren kulturellen und sozialen Veränderungen. In phonetischen Lautwandel-Theorien, welche sich mit der sprachinternen Struktur und Lautwandel befassen, steht die Frage nach den Bedingungen für Lautwandel, also die Frage nach dem „Warum?“, im Mittelpunkt. Genauer gesagt steht die Frage im Raum, ob es Laute gibt, welche in Bezug auf ihre Produktion und/oder ihre Perzeption von Grund auf instabil sind. Harrington, Kleber, Reubold und Stevens (2016) argumentieren, dass eine solche Veranlagung für einen sprach-

lichen Wandel in der typologischen Verteilung von Lauten in den Sprachen der Welt zu sehen sein sollte. Ein Beispiel, welches ihr Argument darlegt, ist das eher seltene Vorkommen von silbeninitialen /kn gn/-Sequenzen in den Sprachen der Welt und die diachrone Neigung hin zur Tilgung der Plosive /k g/ in eben solchen Sequenzen (für weitere Ausführungen und Beispiele vgl. Harrington, Kleber, Reubold & Stevens, 2016).

In phonetischen Modellen zu Lautwandel wird eine starke Verbindung von synchroner Variation in den Sprachen der Welt und diachronem Wandel angenommen. Ein bekanntes Beispiel zum Verhältnis von synchroner Variation und historischem Lautwandel wird von Beddor (2009) aufgezeigt. In ihrer Studie wird der Prozess der Nasalierung von Vokalen in VN Sequenzen (V=Vokal, N=Nasal) zu \tilde{V} (\tilde{V} = nasaliertes Vokal) erklärt. Beddor (2009) beschreibt den Vorgang, durch welchen Wörter, wie „manus“ (Hand) des Lateinischen im heutigen Französisch mit einem Nasalvokal produziert werden, nämlich als /mɛ̃/ („main“, Hand), als eine Art Misinterpretation der Nasalierung. Es wird argumentiert, dass in einem solchen Lautwandelprozess das gesenkte Velum in der Perzeption, dem Vokal und nicht dem koartikulatorischen Einfluss des Nasals auf den Vokal zugeschrieben wird (Krakow, 1999). Diese Idee eines Zusammenhangs zwischen Misinterpretationen des Gesagten, genauer gesagt die fehlerhafte Interpretation der Koartikulation im Sprachsignal, durch den Hörer und Lautwandel ist auf das Lautwandelmodell von Ohala (1981, 1993, 2012) zurückzuführen. In seinem Modell wird angenommen, dass die Bedingungen für Lautwandel dann gegeben sind, wenn die Koartikulation in der Sprachproduktion vom Hörer in der Perzeption nicht eindeutig interpretiert wird. In den Abschnitten 1.1 und 1.2 wurde aufgezeigt wie sich Sprachlaute gegenseitig beeinflussen können. Für die Koartikulation, d.h. für den gegenseitigen Einfluss von Sprachlauten im kontinuierlichen Sprachfluss, wird gewöhnlich seitens des Hörers normalisiert. In Abschnitt 1.1 wurde die Koartikulation und die perzeptive Kompensation für die Koartikulation, anhand der Perzeptionsstudie von Mann und Repp (1980) zu den englischen Sibilanten /s ʃ/ gezeigt. Hier wurde gefunden, dass in der Perzeption die Kategorisierung des Kontinuums zwischen den Sibilanten /ʃ s/ in unterschiedlichen Vokalkontexten variierte. Folgte den Sibilanten der Vokal /u/, verglichen mit demselben Kontinuum im /a/-Kontext, gaben die Hörer mehr /s/-Antworten. Dieses Verhalten wurde darauf zurückgeführt, dass die Versuchspersonen für den koartikulatorischen Einfluss der Lippenrundung kompensieren, d.h. sie rechnen einen größeren Anteil der niedrigen Frequenzen der Lippenrundung im Vokal an und nicht dem Sibilanten /ʃ/. Diese Kompensation für Koartikulation kann aber bei unerfahrenen Hörern fehlerhaft sein, was wiederum zu Lautwandelprozessen, wie der Nasalierung von Vokalen im Beispiel von

Beddor (2009), führen kann.

Während das Modell von Ohala (1981, 2012) zum Großteil dem Hörer die Verantwortung in einem Lautwandel überträgt, gibt es ebenfalls Modelle, welche die Rolle des Sprechers im Zusammenhang von synchroner Variation und Lautwandel in den Vordergrund stellen. Zu dieser Kategorie zählen beispielsweise die Modelle von Browman und Goldstein (1991) und von Lindblom, Guion, Hura, Moon und Willerman (1995). Diese beiden unterscheiden sich nicht nur dadurch von Ohala's Modell, dass sie dem Sprecher einen größeren Einfluss im Lautwandel beimessen, sondern auch in dem Punkt, dass sie davon ausgehen, dass Lautwandel willentlich passiert. Im Gegensatz dazu argumentiert Ohala (2012) dafür, dass Lautwandel nicht geplant ist. Die Misinterpretation des Gesagten resultiert aus Einschränkungen der Perzeption, d.h. aus fehlerhaften Verhalten des Hörers und kann somit, laut Ohala (1993), per definitionem nicht geplant sein kann.

Die Idee hinter dem Modell von Browman und Goldstein (1991) liegt in Aspekten der *Articulatory Phonology*, in der angenommen wird, dass die Sprachproduktion aus vielen verschiedenen einzelnen Gesten zusammengesetzt ist. Diese Gesten beeinflussen die Akustik des Signals zu unterschiedlichen Zeitpunkten in unterschiedlichem Ausmaß. Der Grund für gesteigerte Variabilität in der Sprache wird hier in der informellen Sprache gesehen, für welche vermehrt Reduzierungen und/oder Überlappungen in Gesten vorkommen (Browman & Goldstein, 1991). Zusammen mit den Reduzierungen und Überlappungen können Gesten zum einen eine Neuordnung und zum anderen eine fehlerhafte perzeptive Analyse erfahren und dadurch Lautwandel bedingen.

Ähnlich dazu fokussieren Lindblom (1990) und Lindblom et al. (1995) den Sprecher sowie die verschiedenen Sprachstile, entlang des Kontinuums zwischen hypo- und hyperartikulierter Sprache, in ihren Überlegungen zu Lautwandel. Der Sprecher passt das Maß der sprachlichen Information an die Bedürfnisse des Hörers in verschiedenen Kommunikationssituationen an. Dieser Aspekt spricht für den teleologischen Charakter dieser Variation, denn der Sprecher ist im Stande willentlich die Breite an Informationen in seinen Äußerungen zu steuern. Vonseiten des Hörers wird in einer normalen Kommunikationssituation der Fokus darauf gelegt *Was* gesagt wird. Doch ändert sich dieser Fokus des Hörers dahingehend, dass das *Wie* des Gesagten zum Hauptaugenmerk wird, so sind in diesem Modell die Bedingungen für einen Lautwandel gegeben. Im *Wie*-Modus ist es möglich, dass neue Aussprachevarianten, den schon bestehenden Varianten im mentalen Lexikon, hinzugefügt werden. Für Lindblom et al. (1995) bieten bestimmte prosodische Kontexte Raum für gesteigerte Variabilität und Lautwandel. Begründet liegt dies in den Reduzierungen, d.h.

der Hypoartikulation, von Sprachlauten in prosodisch schwachen Kontexten. Beckman, deJong, Jun und Lee (1992) zeigten für einige diachrone Wandel, wie zum Beispiel die diachrone Epenthese von Plosiven in Nasal-Frikativ Sequenzen, eine Verbindung von prosodisch schwachen Kontexten, Hypoartikulation und dem jeweiligen Lautwandel. Zu einem ähnlichen Ergebnis kommen Harrington, Kleber und Reubold (2013) in ihrer Studie zur antizipatorischen transkonsonantischen Koartikulation von Vokalen. In ihrer Studie wird die Möglichkeit aufgezeigt, die verringerte Kompensation für Koartikulation, d.h. die Bedingung für Lautwandel in Ohala's (1993) Modell, mit prosodisch schwachen Kontexten in Verbindung zu bringen (für mehr Details vgl. Harrington et al., 2013).

Die Frage nach dem „Wie?“, d.h. danach wie sich Sprachwandel in der Sprachgemeinschaft ausbreitet, ist ein komplexes Thema und wird häufig in soziolinguistischen und sozio-phonetischen Studien untersucht. Dabei werden oftmals die extra-linguistischen Faktoren *Geschlecht*, *Alter*, *Identität* und *Prestige* näher betrachtet. Labov (2001) argumentiert, dass das Geschlecht bei der Verbreitung von Lautwandel insofern eine Rolle spielen kann, als dass weibliche Sprecher/Hörer eher dazu neigen innovative Sprachformen zu übernehmen. Weiterhin wird das kindliche Sprachverhalten im Laufe des Spracherwerbs vermehrt in die Untersuchung von Lautwandelprozessen eingebunden, wobei Labov (2001, 2007) für eine schrittweise Steigerung der sprachlichen Innovation in Richtung des Lautwandelmusters durch die jüngeren Generationen argumentiert.

In exemplar-basierten Modellen zur Sprachproduktion und Perzeption (Pierrehumbert, 2001, 2002) wird die Erfahrung des Sprechers/Hörers und dessen linguistisches Wissen mit Bedingungen für Lautwandel in Zusammenhang gebracht. In diesem Modell wird davon ausgegangen, dass vom Hörer wahrgenommene lautliche Exemplare im Lexikon gespeichert werden. Da davon ausgegangen werden kann, dass die lautlichen Exemplare, sowohl durch interne als auch durch externe Faktoren beeinflusst werden, ist das Exemplar-Modell insofern besonders, als dass es die Bedingungen für Lautwandel, wie sie beispielsweise in Ohala's Modell beschrieben werden, und die eher soziolinguistische Frage nach der Ausbreitung von Lautwandelformen, die in vielen Studien von Labov untersucht wird, in Beziehung setzt.

1.5.1 Aufbau der Arbeit

Die vorliegende Studie untersucht die synchrone Variation der drei stimmlosen Sibilanten des Polnischen /s ʂ ʐ/ und deren diachronen Wandel hin zu einem Zweier-Kontrast, welcher in einigen Varietäten des Polnischen und im Mandarin-Chinesischen gefunden werden kann. Die Dissertation beschäftigt sich dabei mit zwei zentralen Fragen. Die erste beinhaltet die möglichen Faktoren, die die Stabilität dieses Lautkontrasts beeinflussen können. In der modernen Lautwandelforschung wurde gezeigt, dass diachrone Wandel von Sprache durch interne (z.B. Koartikulation), sowie durch externe Faktoren (z.B. Sprachkontakt) hervorgerufen werden können. Doch welche der vielen Einflussmöglichkeiten auf die Sprachproduktion und die Perzeption können als Prädispositionen für den Wandel in den polnischen Sibilanten von einem Dreier- hin zu einem Zweier-Kontrast angesehen werden? Die zweite Frage ist, welche Gründe es für die Neutralisierung von /s ʂ/ zum Vorteil des dentalen Sibilanten /s/ gibt. Genauer gesagt ist das Anliegen hier herauszufinden, warum ausgerechnet der retroflexen Sibilant im Zuge der Neutralisierung wegfällt und warum gerade das /s ʂ/-Sibilantenpaar von der Neutralisierung in den dialektalen Varietäten des Polnischen und im Mandarin-Chinesischen betroffen ist.

Um diesen verschiedenen Aspekten auf den Grund zu gehen, wurden in Kapitel 1.1 verschiedene Beispiele von synchroner Variation in der Sprachproduktion und der Perzeption aufgeführt. Darauf folgte in Abschnitt 1.2 die Verbindung von internen und externen Faktoren, welche die synchrone Variation bedingen, und daraus resultierenden Kontrastneutralisierungen. In Abschnitt 1.3 wurden Studien aufgeführt, welche die gesteigerte Variabilität des sibilantischen Dreier-Kontrasts hinsichtlich interner und externer Faktoren prüften. Im Anschluss wurde genau auf die drei Sibilanten des Polnischen eingegangen (Kapitel 1.4), wobei zunächst in Abschnitten 1.4.1 und 1.4.2 grob das Lautsystem des Polnischen, sowie die regionale Verteilung der polnischen dialektalen Varietäten beschrieben wurde. Die Abschnitte 1.4.3 und 1.4.4 betrachteten artikulatorische Charakteristika der drei Sibilanten und deren akustische Konsequenzen. In Abschnitt 1.5 wurden bekannte Lautwandeltheorien beschrieben, vor allem lag hier der Fokus auf Erkenntnissen darüber wie synchrone Variation, die hervorgerufen durch verschiedenste Faktoren mannigfach in der Sprache zu finden ist, zu diachronem Wandel führen kann.

In Kapitel 2 werden Ergebnisse einer physiologischen Studie vorgestellt, welche anhand von artikulatorischen und akustischen Aufnahmen von neun L1-Polnisch sprechenden er-

wachsenen Versuchspersonen, die Variation in der Produktion der polnischen Sibilanten untersucht. Interessant ist hier vor allem, das Ergebnis, welches Überlegungen hinsichtlich des nicht-linearen Verhältnisses zwischen der Artikulation der Sibilanten und dem daraus resultierenden Effekt auf die Akustik der drei Laute zu lässt. In einem weiteren Experiment, auf welches in Kapitel 3 eingegangen wird, soll herausgefunden werden, inwieweit die Instabilität des Dreier-Kontrasts in der Perzeption zu sehen ist und ob die Variabilität des Kontrasts unter verschiedenen prosodischen Bedingungen (Deakzentuierung vs. Akzentuierung des Zielwortes) verstärkt werden kann. Dieses Experiment lässt Spekulationen hinsichtlich eines Einflusses des prosodischen Kontexts auf die Kategorisierung der Sibilanten zu, da für das Kontinuum zwischen den potentiell instabilen Sibilanten /s ʂ/ die Kategoriengrenze etwas stärker beeinflusst wurde als für das scheinbar stabile Kontinuum zwischen /s ʧ/. Im letzten experimentellen Teil dieser Arbeit (Kapitel 4) wird von einem Experiment berichtet, welches untersucht, ob es die variabelere Sprachproduktion während des Spracherwerbs ist, die zur Instabilität des Kontrasts beiträgt und, ob möglicherweise in den Sibilantproduktionen von Kindern Hinweise auf eine Neutralisierung des /s ʂ/-Kontrasts gefunden werden können. Hierfür werden Sprachproduktionsdaten von Kindern und Erwachsenen miteinander verglichen. Die Ergebnisse zeigen sowohl altersbedingte (Kinder vs. Erwachsene) als auch geschlechtsspezifische (männliche vs. weibliche Sprecher) Unterschiede im Maß der Kontrasterhaltung in den untersuchten Parametern im Friktionsanteil der Sibilanten selbst und in den F2/F3-Werten zum zeitlichen Onset des Vokals, welcher den Sibilanten folgt. Im letzten Kapitel (Kapitel 5) der vorliegenden Dissertation werden die Ergebnisse aus den drei vorgestellten Experimenten in Zusammenhang gebracht und vor den Erkenntnissen aktueller Lautwandeltheorien diskutiert.

Kapitel 2

Artikulatorische und akustische Eigenschaften der Polnischen Sibilanten (EMA)¹

2.1 Zusammenfassung

Die vorliegende Studie beschäftigt sich mit der relativen synchronen Stabilität der drei im Polnischen kontrastiven Sibilanten /s ʂ ʐ/. Es wurden akustische und Zungenbewegungsdaten während der Produktion von symmetrischen Echt- und Nicht-Wörtern (letztere mit einer CVCV-Struktur) in drei Vokalkontexten von neun Polnisch (als Erstsprache) sprechenden Sprechern erhoben. Ein Gaußsches Modell wurde herangezogen, um die Sibilanten sowohl anhand der spektralen Information im Friktionsgeräusch als auch anhand der Formantfrequenzen zu Beginn des Vokals zu klassifizieren. Die physiologische Analyse zeigte eine nahezu komplette Separierung zwischen /s ʂ ʐ/ bezogen auf die ausgewählten Parameter – Position und Orientierung – der Zungenspitze und des Zungendorsums. Die akustische Analyse zeigte, dass durch die Energieansammlung in höheren Frequenzregionen /s/ hinsichtlich des Friktionsgeräusches von den beiden anderen Sibilantkategorien unterschieden werden kann. Zu Beginn des Vokals konnten in den akustischen Daten für /ʐ/, welcher einen palatalisierenden Einfluss auf den nachfolgenden Vokal ausübt, prägnante Informationen gefunden werden. Für die Identifikation von /s ʐ/ reichten somit jeweils entweder das Friktionsgeräusch oder der Vokalonset aus, während jedoch für die Trennung

¹Eine Version dieses Kapitels wurde im *Journal of the International Phonetic Association* veröffentlicht (Bukmaier & Harrington, 2016).

des retroflexen Sibilanten von /s ɕ/ alle beide Parameter notwendig waren. Die Befunde dieses Kapitels liefern Hinweise auf mögliche Gründe für die größere synchrone Instabilität von /ɕ/ sowie für den relativ späten Erwerb des retroflexen Sibilanten während des Spracherwerbs: Die größere synchrone Instabilität von /ɕ/ könnte damit von der hohen artikulatorischen Komplexität zusammen mit deren vergleichsweise niedrigen akustischen Auswirkungen stammen. Die Daten lassen weiterhin vermuten, dass der relativ späte Erwerb von /ɕ/ während des kindlichen Spracherwerbs in der schwachen Information im Vokal begründet liegen könnte, welche notwendig ist, um /ɕ/ von /s/ zu unterscheiden.

2.2 Einleitung

In den letzten 30 Jahren berichteten viele Studien über die Akustik (Cheon & Anderson, 2008; Evers, Reetz & Lahiri, 1998; Jongman, Wayland & Wong, 1998; Maniwa, Jongman & Wade, 2009; Nowak, 2006a), Perzeption (Cheon & Anderson, 2008; F. Li, Munson, Edwards, Yoneyama & Hall, 2011; McGuire, 2007) und die Artikulation von Sibilanten. Die meisten dieser Studien untersuchten jedoch den Zweier-Kontrast zwischen dem alveolaren Sibilanten /s/ und dem post-alveolaren Laut /ʃ/. Die vorliegende Studie beschäftigt sich mit dem vergleichsweise seltenen Kontrast zwischen drei unterschiedlichen Artikulationsorten von Sibilanten im Polnischen. Standard-Polnisch ist, neben dem Schwedischen und Mandarin-Chinesischen, eine der wenigen Sprachen, welche drei Artikulationsorte bei Sibilanten unterscheiden, für die stimmlosen Varianten ist die Distinktion wie folgt (Gussmann, 2007): dental /s/ (z.B. sali /sali/, „Zimmer“ (gen.)), retroflex /ɕ/ (z.B. szali /ɕali/, „Waage“ (gen.)), und alveolopalatal /ɕ/ (e.g. siali /ɕali/, „angesät“ (Adj.)).

In den letzten Jahren wurden diese drei Sibilanten physiologisch für das Polnische in Toda et al. (2010), für Mandarin-Chinesisch in Proctor, Lu, Zhu, Goldstein und Narayanan (2012) und für beide Sprachen von Hu (2008) analysiert. Diese Studien zeigten, dass sich die drei Sibilanten artikulatorisch nicht nur in der Zungenposition unterscheiden, sondern auch in ihrer Zungenform. Die Frikative zeigen auch Unterschiede in zwei weiteren Zungenmerkmalen. Erstens ist für /ɕ s/ die vertikale Orientierung der Zungenspitze meist nach oben gerichtet, wohingegen für /ɕ/ die Zungenspitze nach unten orientiert ist. Zweitens, während die Zungenspitze in /ɕ/ tendenziell stärker rückverlagert ist als in /s/, wird gezeigt, dass der Grad der Rückverlagerung im Polnischen und Mandarin-Chinesisch geringer ausfällt als in indischen Sprachen (Hamann, 2002a, 2002b; Hu, 2008; Toda et al., 2010):

aus diesem Grund ähneln sich im Polnischen /s ʂ/ bezogen auf ihre Zungenform stärker als in indischen Sprachen (für Details hierzu vgl. Ladefoged und Maddieson, 1996, S. 154).

Die physiologische Analyse von vier L1-Polnisch Sprechern in Bukmaier, Harrington, Reubold und Kleber (2014) brachte Hinweise auf eine größere Variabilität in /ʂ/ als in den beiden anderen Sibilanten. In langsam gesprochener Sprache konnten /s ʂ/ gut anhand der unterschiedlichen Zungenspitzenorientierung voneinander unterschieden werden, d.h. /ʂ/ wurde hier sub-laminal produziert, sodass die Unterseite der Zungenspitze Kontakt zum Artikulationsort hatte. In schneller Sprechgeschwindigkeit hingegen, verschwand die Verschiedenheit der beiden Sibilanten in Bezug auf ihre Zungenspitzenorientierung, sodass /ʂ/ dem dentalen Sibilanten durch eine eher supra-laminale und einer weniger sub-apikale Produktion ähnelte. Hu's (2008) physiologische Analyse des Mandarin-Chinesischen weist ebenfalls auf eine stärkere artikulatorische Variabilität in /ʂ/ verglichen mit den beiden anderen Sibilanten hin.

Was die Akustik der Sibilanten anbetrifft, haben viele Studien der letzten 50 Jahre gezeigt, dass der Unterschied im Artikulationsort zwischen den englischen Lauten /s ʃ/ zum Großteil auf spektralen Eigenschaften des Friktionsgeräusches basiert (Evers et al., 1998; Jongman et al., 1998; Shadle & Mair, 1996; Whalen, 1991): genauer gesagt, verschiebt das kürzere der Schallquelle/dem Konstriktionsort vorgelagerte Ansatzrohr im Sibilant /s/, die Energie im Spektrum in Richtung höherer Frequenzen, sodass es akustisch und perceptiv (Fujisaki & Kunisaki, 1978; Mann & Repp, 1980) von /ʃ/ durch einen höheren spektralen Frequenzschwerpunkt (besser bekannt als *center of gravity*; kurz COG) (Forrest et al., 1988) unterschieden werden kann. Für den Dreier-Kontrast im Polnischen ist es möglich /s/ durch die spektralen Eigenschaften des Friktionsgeräusches von den beiden anderen Sibilanten zu trennen. Einige Studien (Jassem, 1995; Nowak, 2006a; Żygis & Hamann, 2003) konnten jedoch zeigen, dass der Frequenzschwerpunkt im Friktionsgeräusch nicht ausreicht, um /ʂ ʐ/ zu separieren.

Die Frage, ob Formanttransitionen zur akustischen und perceptiven Unterscheidung des Artikulationsorts bei Frikativen beitragen, ist noch nicht vollständig geklärt. Die ersten Studien von Harris (1958) und Heinz und Stevens (1961), die sich mit diesem Thema beschäftigten, zeigten, dass im Englischen Formanttransitionen zur perceptiven Unterscheidung von Sibilanten nicht notwendig sind, wohl aber für die Separierung der Nicht-Sibilanten

/f θ/. Trotz der Hinweise aus akustischen Studien, dass Vokaltransitionen weit in den Frikativ hineinreichen (Soli, 1985), deuteten spätere Studien darauf hin, dass sowohl bei Sibilanten als auch bei nicht-sibilantischen Frikativen Vokaltransitionen perzeptiv nur eine kleine Rolle für die Wahrnehmung von Artikulationsortkontrasten spielen (Jongman, 1989; LaRiviere, Winitz & Herriman, 1975). Die meisten dieser Studien basierten jedoch auf Sprachen, welche nur zwei Sibilanten kontrastieren. Im Gegensatz dazu steht eine neuere sprachensübergreifende Studie von Wagner, Ernestus und Cutler (2006), die zeigte, dass der Wirkungsgrad von Formanttransitionen als Hinweis bei der Kontrastwahrnehmung sprachabhängig ist: genauer gesagt wurde gezeigt, dass zum Beispiel Hörer von Sprachen wie dem Polnischen, welche zwischen den Frikativen /ʂ ʐ/ kontrastieren, die wiederum in Bezug auf das Friktionsgeräusch großteils undifferenziert sind, stärker auf Informationen angewiesen sind, die sie aus Formanttransitionen beziehen können. Diese Erkenntnisse gehen einher mit Ergebnissen aus Nowak (2006a), welcher hervorbrachte, dass L1-Polnisch Hörer polnische Sibilanten sowohl allein anhand der isolierten Friktionsanteile als auch in VCV-Sequenzen, bei welchen die Transitionen in den nachfolgenden Vokal entfernt wurden, identifizieren konnten. Nowak's (2006a) Ergebnisse zeigten, dass Frikative zuverlässig anhand des Geräuschteils identifiziert werden können, wobei für die Unterscheidung von /ʂ ʐ/ in VCV-Sequenzen Formanttransitionen unbedingt notwendig sind. In Einklang damit zeigten Toda et al. (2010) inwiefern die unterschiedlichen Zungenformen für /ʂ ʐ/ zur Unterscheidung der beiden Sibilanten bezogen auf ihre Formanttransitionen beitragen können.

Studien zum kindlichen Erwerb von polnischen Sibilanten haben gezeigt, dass /ʂ/ relativ spät und vor allem auch nach den beiden anderen Sibilanten erworben wird (Łukaszewicz, 2006, 2007). Die artikulatorische Variabilität/Instabilität von /ʂ/ wie auch die Erkenntnisse aus dem Erstspracherwerb, hängen möglicherweise mit dem diachronen Wandel des Dreier-Kontrasts /s ʂ ʐ/ hin zu einem Zweier-Kontrast als Ergebnis eines Zusammenfalls von /s ʂ/ zusammen, welcher schon sowohl in der Min Varietät im Mandarin (Chuang & Fon, 2010; Duanmu, 2006) als auch in einigen Polnischen Dialekten (Żygis, Pape & Czaplicki, 2012) gefunden werden kann.

Einer der Hauptbeweggründe für die vorliegende Studie ist es die synchrone Basis für den diachronen Zusammenfall des /s ʂ/-Kontrasts in Richtung /s/ zu untersuchen. Die spezifischeren Ziele sind es sowohl die physiologischen als auch die akustischen Eigenschaften der drei Frikative zu analysieren, um herauszufinden, ob die Identifikation von /ʂ/ im

Vergleich zu den beiden anderen Sibilanten benachteiligt ist. Hierfür wurden sowohl artikulatorische (elektromagnetische Artikulografie; kurz EMA) als auch akustische Aufnahmen von neun L1-Polnisch sprechenden Erwachsenen während der Äußerung von Sibilanten gemacht, um feststellen zu können inwieweit sich die Sibilanten physiologisch und auch in Bezug auf ihre Friktionsgeräusche und Formanttransitionen unterscheiden.

2.3 Experimentaufbau

2.3.1 Datenerhebung und Sprecher

Akustische, ebenso wie Sprechbewegungsdaten wurden mit der elektromagnetischen Artikulografie (EMA) in einer Aufnahmekabine am IPS in München erhoben (AG501, Carstens Medizinelektronik). Durch eine solche Aufnahme ist es möglich, Messungen der horizontalen, vertikalen und lateralen Position der Artikulatoren zu erhalten. Für die EMA-Aufnahmen wurden zwei Sensoren² auf der Zunge angebracht (Fig. 2.1): einer an der Zungenmittellinie 1 cm hinter der Zungenspitze (kurz TT; hergeleitet von Engl. *tongue tip*) und ein weiterer auf einer Höhe mit den hinteren Backenzähnen auf dem Zungenrücken (kurz TB; hergeleitet von Engl. *tongue back*). Außerdem wurden zwei Sensoren auf der Ober- und Unterlippe, d.h. auf der Haut direkt über und unter den Lippen, angebracht. Um Kopfbewegungen aus den Bewegungsdaten der Zunge heraus korrigieren zu können, wurden vier weitere Sensoren am Oberkiefer (Gewebe direkt über den Zähnen), auf dem Nasenrücken, sowie am rechten und am linken Mastoid (hinter den Ohren liegender Teil des Schläfenbeins) befestigt. Für diese Studie wurden die Daten der beiden Zungensensoren ausgewertet. Das akustische Sprachsignal wurde synchron zu den physiologischen Daten mit einem Sennheiser ME66 (supercardioid) Mikrophon, welches im Abstand von einem Meter zur Versuchsperson positioniert wurde, mit eingeschaltetem Bass Roll-off Filter (-6dB bei 200 Hz) aufgenommen. Die Audiodaten wurden mit einem National Instruments Compact DAQ *multichannel data acquisition front-end* erhoben, welches durch einen USB-Anschluss mit einem Notebook verbunden war. Im Anschluss an die Aufnahme wurde bei der Nachverarbeitung der Daten die Synchronisierung von Audio- und Bewegungssignal durchgeführt (für mehr Details zur Nachverarbeitung von akustischen und artikulatorischen Daten siehe

²Die Anzahl der Sensoren wurde auf nur zwei Sensoren gesetzt, da für zukünftige Untersuchungen der Sibilanten, physiologische Aufnahmen mit kindlichen Sprechern geplant sind. Aufgrund der kleineren Zunge bei Kindern, war geplant nur zwei Sensoren auf deren Zunge zu befestigen. Damit die erwachsenen und die kindlichen Daten vergleichbar gehalten werden, wurde in dieser ersten Studie mit den erwachsenen Sprechern also vorab entschieden, die Anzahl der Sensoren auf zwei Zungensensoren zu beschränken.

Hoole und Zierdt, 2010).



Abbildung 2.1: Platzierung der zwei Zungensensoren.

Die Versuchspersonen in diesem Experiment waren neun L1-Polnisch sprechende Erwachsene im Alter zwischen 19 und 28 Jahren (4 männlich und 5 weiblich). Der Geburts- und Schulort von vier der Versuchspersonen befindet sich in Dialektregionen, in welchen alle drei Sibilanten kontrastiert werden (jeweils zwei aus Schlesien und Großpolen). Die verbleibenden fünf Versuchspersonen stammten aus und lebten die meiste Zeit ihres Lebens (d.h. gingen dort zur Schule) in Dialektregionen, in welchen der dental/retroflex Kontrast neutralisiert wird (jeweils zwei aus Masowien und Kleinpolen, sowie eine aus Kaschubien). Diese Versuchspersonen wurden trotzdem in die Analyse einbezogen, da sie von einem L1-Polnisch Sprecher mit linguistischem Hintergrund als Sprecher des Standard-Polnischen ohne wahrnehmbare regionale Färbung eingestuft wurden. Zum Zeitpunkt der Aufnahme hatte keine der Versuchspersonen länger als zwei Jahre außerhalb Polens gelebt.

2.3.2 Sprachmaterial und Experimentaufbau

Die Versuchspersonen produzierten sowohl symmetrische CVCV (z.B. sese /sɛsɛ/) Logatome (wobei C=/s ɕ ʑ/ und V=/a ɛ ɔ/) als auch zweisilbige polnische Echt-Wörter (Tab. 2.1) mit initialen CV-Sequenzen (wobei gilt C=/s ɕ ʑ/ und V=/a ɛ ɔ/). Alle Zielwörter waren in den Trägersatz „Ania woła [ZIELWORT] aktualnie“ (wörtlich „Ania ruft [ZIELWORT] momentan“) eingebettet, wobei das Zielwort mit einem nuklearen Pitch-Akzent produziert wurde. Die Versuchsteilnehmer lasen die Sätze laut von einem Computerbildschirm ab, auf welchem die einzelnen Sätze nacheinander und in randomisierter Reihenfolge erschienen. Bei Versprechern oder Produktionen von falschen Betonungsmustern wurden die Versuchs-

personen aufgefordert, den Satz zu wiederholen. Die Aufnahmesitzung bestand aus zehn Blöcken, die zwischen langsamer und schneller Sprechgeschwindigkeit alternierten. Vor der eigentlichen Aufnahme wurden die Versuchspersonen in einem Vortest gebeten, Beispiele aus dem Sprachmaterial in selbstgewählter schneller und langsamer Sprechgeschwindigkeit zu lesen. Dieser Vortest diente dazu die individuellen Sprechgeschwindigkeiten festzulegen und die jeweilige Aufnahmezeiten im eigentlichen Experiment anzupassen. Auf dem Bildschirm war eine Fortschrittsanzeige zu sehen, welche mit der jeweils gewünschten Sprechgeschwindigkeit verbunden war (die wiederum für jeden Sprecher und jede Geschwindigkeitsbedingung basierend auf der gemittelten Dauer des Vortest definiert wurde), und damit den zeitlichen Rahmen jedes einzelnen Tokens vorgab.

	Wort1	Wort2	Wort3	Logat.
/sa/	sara (/sara/, fem. Vorname)	sama (/sama/, „alleine“)	sawa (/sava/, m. Vorname)	sasa
/sɛ/	serce (/sertsɛ/, „Herz“)	-	-	sese
/sɔ/	sowa (/sɔva/, „Eule“)	-	-	soso
/ʂa/	szary (/ʂari/, „grau“)	szafa (/ʂafa/, „Regal“)	-	szasza
/ʂɛ/	szereg (/ʂɛrɛk/, „Linie“)	-	-	szesze
/ʂɔ/	szorstka (/ʂɔrstka/, „rau“)	-	-	szoszo
/ɕa/	siatka (/ɕatka/, „Netz“)	siano (/ɕano/, „Heu“)	-	siasia
/ɕɛ/	siekacz (/ɕɛkaɕ/, „Schneidezähne“)	-	-	siesie
/ɕɔ/	siodło (/ɕɔdɔ/, „Sattel“)	-	-	siosio

Tabelle 2.1: Verteilung der initialen CV-Sequenzen in den Echt-Wörtern (Wort 1-3) und den Logatomen (Logat.), welche als Zielwörter verwendet wurden.

Jedes der 22 Zielwörter, die sich aus neun Logatomen und 13 Echt-Wörtern zusammensetzen (vgl. Tab. 2.1), wurde in jeder Sprechgeschwindigkeit zehn Mal in randomisierter Reihenfolge wiederholt. Einige der wortinitialen CV-Sequenzen, wie zum Beispiel /sa/-, /ɕa/- und /ʂa/, kommen häufiger im Anlaut von zweisilbigen polnischen Echt-Wörtern vor, sodass für das Korpus dieser Studie mehr (nahe-)Minimalpaare mit diesen CV-Sequenzen

gefunden und in die Zielwortliste aufgenommen wurden (vgl. Tab. 2.1, Reihe 1: /sara/, /sama/, /sava/; Reihe 4: /sari/, /safa/; Reihe 7: /catka/, /canɔ/). Aufgrund ebendieser ungleichen Verteilung von CV-Sequenzen beinhaltet das Sprachmaterial dieser Studie zwischen zwei (/ca/- und /sa/-Anlaute) und drei (/sa/-Anlaute) Zielwörter für die gängigeren Sequenzen, während andere, seltenere Sequenzen mit nur einem Zielwort vertreten sind (z.B. /sɛ sɔ sɛ sɔ ɛɛ ɔɔ/). Tab. 2.1 zeigt die komplette Verteilung der CV-Sequenzen dieser Studie.

Das Experiment bestand aus 3960 (22 Zielwörter × 10 Wiederholungen pro Sprechgeschwindigkeit × 9 Sprecher) Sätzen, von denen 3895 für diese Studie analysiert werden konnten. Der Datenverlust von 65 Tokens ist auf technische Probleme sowohl während der Aufnahme als auch bei der Nachverarbeitung der Daten zurückzuführen. Die Gesamtanzahl der analysierten initialen Sibilant-Vokal Kombinationen aus den Echt- und den Nicht-Wörtern ist in Tabelle 2.2 zu sehen.

	/a/	/ɛ/	/ɔ/
/s/	708	354	358
/s̺/	530	348	354
/s̻/	534	355	354

Tabelle 2.2: Gesamtanzahl von Echt-Wörtern und Nicht-Wörtern getrennt für alle Sibilant-Vokal Kombinationen.

2.3.3 Analyse der physiologischen Daten

Die Nachverarbeitung der physiologischen Rohdaten wurde halb-automatisch in Matlab durchgeführt (Version MathWorks MATLAB R2012a) und beinhaltete die Rotation der Daten, sodass diese parallel zur Okklusionsebene³ lagen (Hoole & Zierdt, 2010).

Die artikulatorische Annotation der drei Sibilanten wurde anhand der vertikalen Bewegung der Zungenspitze (TT) und der tangentialen TT Geschwindigkeit festgesetzt. Die physiologischen Labels beinhalteten sieben Orientierungspunkte (Abb. 2.2). Ein kompletter CVC-Bewegungszyklus wurde unterteilt in eine CV-Öffnungsphase, einen Nukleus beziehungsweise eine Zielphase und eine VC-Verschlussphase. Der Onset und Offset der Öffnungs- und Verschlussbewegung wurde durch einen 20% Schwellenwert im Signal der

³Die Okklusionsebene, auch Kauebene genannt, beschreibt die räumliche Lage, auf der sich die Zähne des Ober- und Unterkiefers beim Zusammenbeißen der Zähne treffen.

Tangentialgeschwindigkeit markiert (Hoole, Bombien, Kühnert & Mooshammer, 2009; Hoole & Mooshammer, 2002). Das Intervall zwischen CV-Offset und VC-Onset wurde als Vokalnukleus definiert.

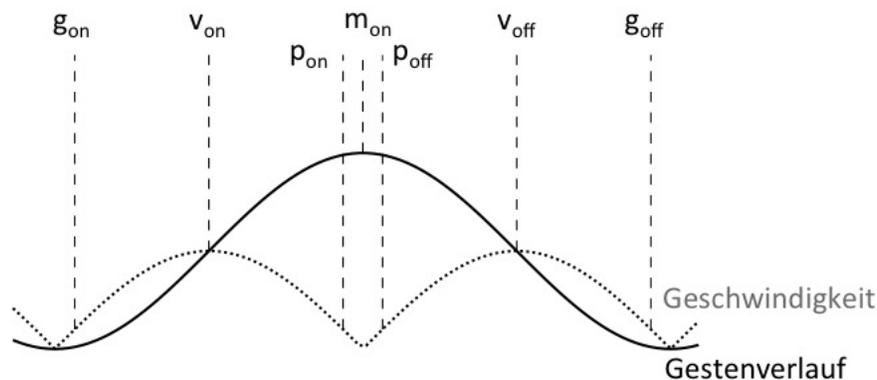


Abbildung 2.2: Schematische Repräsentation der Position der Orientierungspunkte: Gesten-Onset (g_{on}), maximale Geschwindigkeit beim Gesten-Onset (v_{on}), Onset des Verschlussplateaus (p_{on}), Maximum der Konstriktion (m_{on}), Offset des Verschlussplateaus (p_{off}), maximale Geschwindigkeit beim Gesten-Offset (v_{off}), und Gesten-Offset (g_{off}). (Abbildung entnommen aus Bombien, 2011)

Im EMA System werden die Informationen zu den Sensoren zum einen anhand der Kartesischen Koordinaten x , y und z und zum anderen durch zwei Winkelkoordinaten (*Azimuth* und *Elevation*) bereitgestellt. Diese Winkelkoordinaten zeigen in einem kugelförmigen Koordinatensystem die Richtung, in welche der jeweilige Sensor zeigt, an. Während *Azimuth* die Winkelposition in der xy -Ebene, d.h. also in der axialen Ebene, bezeichnet, gibt die *Elevation* die Winkelposition parallel zur xy -Ebene wieder (für mehr Details zur den Freiheitsgraden der Sensoren vgl. Hoole und Zierdt, 2010). Anhand der Orientierungspunkte in Abb. 2.2 wurden die Daten der vertikalen und horizontalen Positionsdaten (y und x) der Zungenspitze (TT) und des Zungendorsums (TB) extrahiert. Weiterhin wurde die Zungenspitzenorientierung (*Elevation*) analysiert, da diese Informationen zum retroflexen

Sibilanten liefert, in welchem bekanntermaßen die Zungenspitze nach hinten-oben gedreht wird (Bukmaier, Harrington, Reubold & Kleber, 2014; Hamann, 2002a, 2002b; Ladefoged, 2001; Toda et al., 2010).

2.3.4 Analyse der akustischen Daten

Die synchronisierten akustischen Daten wurden mit 25,600 Hz digitalisiert, und mit der Methode des *forced alignment* automatisch segmentiert, gelabelt und wenn nötig manuell nachkorrigiert (Munich Automatic Segmentation tool; Schiel, 2004). Es wurden sowohl Spektren (256 Punkte diskrete Fouriertransformation mit einer 40 Hz Frequenzauflösung, 5 ms Blackman Fenster, und einer Fensterverschiebung von 5 ms) als auch Formantfrequenzen berechnet (F1-F4; Pre-emphase von -0.8, 20 ms Blackman Fenster mit einer Fensterverschiebung 5 ms). Bei der Berechnung der Formantfrequenzen in EMU/R (Harrington, 2010) wurde die nominale Frequenz des ersten Formanten je nach Geschlecht des Sprechers festgelegt: für die männlichen Sprecher auf 500 Hz und für die weiblichen Versuchspersonen auf 600 Hz. Für die männlichen Sprecher wird der nominale F1 Wert auf 500 Hz gesetzt, da das der geschätzte Wert des ersten Formanten einer einfachen, einseitig geöffneten Röhre der Länge 17.5 cm ist, welche ein gutes Modell für den Vokal Schwa eines erwachsenen, männlichen Sprechers darstellt. In Anbetracht des meist kürzeren Vokaltraktes (kürzer als 17.5 cm) bei weiblichen Sprechern, ist der erste Formant eines Schwa-Vokals etwas höher verglichen mit den F1 Werten von männlichen Sprechern, sodass für die Formantberechnung der weiblichen Sprecher der nominale F1 Wert auf 600 Hz angehoben wird. Durch das Anheben des nominalen F1 Wertes werden in der Linearen Prädiktion – ein Algorithmus, welcher für die Formantberechnung verwendet wird – die LP-Koeffizienten reduziert, was zur Folge hat, dass für einen bestimmten Frequenzbereich weniger Formantwerte gefunden werden, was wiederum für einen kürzeren Vokaltrakt spezifisch ist (für weitere Details zur Formantberechnung in EMU/R vgl. Harrington, 2010).

Für die akustische Analyse des Friktionsgeräusches wurden die Spektren zum zeitlichen Mittelpunkt zwischen dem akustischen Onset und Offset jedes einzelnen Sibilanten extrahiert. Nachdem die spektralen Daten von der Hertz- in die mel-Skala konvertiert wurden, reduzierte man anhand der diskreten Kosinustransformation diese spektralen Daten zu einer Reihe von Koeffizienten. Für ein N-Punkte mel-skaliertes Spektrum, $x(n)$, welches sich in der Frequenz von $n = 0$ bis $N-1$ Punkten über eine Frequenzbandbreite zwischen

500-3500 mel (414-10 313 Hz) erstreckt, wurde der m te DCT-Koeffizient C_m ($m = 0, 1, 2$) durch (2.1) berechnet.

$$C_m = \frac{2k_m}{N} \sum_{n=0}^{N-1} x(n) \cos\left(\frac{2n + 1m\pi}{2N}\right) \quad (2.1)$$

Diese drei Koeffizienten C_m ($m = 0, 1, 2$) verschlüsseln jeweils den Mittelwert, die Steigung und die Krümmung des Signals, auf welches die DCT Transformation angewendet wurde (Harrington, 2010). Nachdem frühere Studien gezeigt haben, dass Sibilanten in ihrem Friktionsgeräusch optimal durch C_1 und C_2 (d.h. durch die Steigung und die Krümmung des Spektrums) voneinander unterschieden werden können, werden alle weiteren Quantifizierungen anhand dieser Koeffizienten durchgeführt.

Sowohl die artikulatorischen Daten als auch die Formantdaten wurden durch die Standardnormalisierung (Lobanov, 1971) sprechernormalisiert. Genauer gesagt, gesetzt den Fall, dass $\chi_{P.i.T}$ ein Rohwert eines artikulatorischen oder Formant-Parameters P eines Sprechers i zum Zeitpunkt T ist, ist der dazugehörige normalisierte Wert durch (2.2) gegeben:

$$\chi_{P.i.T} = \frac{\chi_{P.i.T} - \chi_{P.i.m}}{\chi_{P.i.s}} \quad (2.2)$$

dabei sind $\chi_{P.i.m}$ und $\chi_{P.i.s}$ der Mittelwert und die Standardabweichung, welche über alle Erhebungen desselben Parameters und desselben Sprechers berechnet wurden. Für die Normalisierung der artikulatorischen Daten des Friktionsgeräusches wurden ($\chi_{P.i.m}$ und $\chi_{P.i.s}$) über alle Datenerhebungen zwischen dem akustischen Onset und Offset des Frikativs berechnet; für die Normalisierung der Daten der Formantfrequenzen wurden ($\chi_{P.i.m}$ und $\chi_{P.i.s}$) über alle Datenerhebungen zwischen dem akustischen Onset und Offset des Vokals berechnet.

Um den Grad der Unterscheidung zwischen den Artikulationsorten der drei Sibilanten festzulegen, wurde eine Gaußsche Klassifikation der akustischen (spektralen, Formant-) Daten durchgeführt. Die Klassifikation basierte auf der quadratischen Diskriminantenanalyse (Srivastava, Jermyn & Joshi, 2007), bei welcher es eine Trainings- und eine Testphase gab. Während der Trainingsphase wurde jede Frikativklasse, bestehend aus einer bestimmten Anzahl an Beobachtungen in einem zweidimensionalen akustischen Raum als eine bivariate Gaußsche Verteilung modelliert; in der Testphase, wurden die Beobachtungen, basierend

auf der größten A-posteriori-Wahrscheinlichkeit als eine der drei Frikativkategorien klassifiziert. Die Verbindung zwischen der Test- und der Trainingsphase wurde durch die *leave-one-out* Validierung durchgeführt. Dabei wurden der Reihe nach die Daten eines jeden der 9 Sprecher klassifiziert, nachdem sie auf die Daten der 8 anderen Sprecher trainiert wurden. Für das Friktionsgeräusch waren die zwei Parameter, *C1* und *C2*, wie oben schon erwähnt, entnommen zum zeitlichen Mittelpunkt zwischen dem akustischen On- und Offset des Frikativs; für den Vokal waren die zwei Parameter F2 und F3 zum Zeitpunkt des akustischen Vokalonsets. Für die Vokalklassifikationen wurde das Trainieren und Testen anhand der *leave-one-out* Methode zusätzlich in jedem der drei /a ε ɔ/ Vokalkontexte durchgeführt. Grundsätzlich wurden alle Klassifikationen, wie sie oben beschrieben sind, separat für die zwei Kontexte – schnelle und langsame Sprechgeschwindigkeit – ausgeführt (d.h. vier Klassifikationen: zwei (langsam/schnell) basierend auf *C1* und *C2* sowie zwei (langsam/schnell) auf F2 und F3 zum Zeitpunkt des akustischen Vokalonsets).

2.4 Ergebnisse

Die nachfolgenden Ergebnisse werden für das Friktionsgeräusch (2.4.1) und für die Vokaltransitionen (2.4.3) getrennt voneinander dargestellt. In beiden Fällen war das Ziel herauszufinden, in welchem Maße die drei Frikative in Bezug auf ihren Artikulationsort unterschieden werden können und inwieweit die beiden verwendeten Merkmalszusammenstellungen ergänzende Informationen zu diesem Zweck liefern.

2.4.1 Ergebnisse I: Friktion – Physiologische Analyse

Die gemittelten Zungenspitzendaten in Abb. 2.3 zeigen für jeden der neun Sprecher eine klare Trennung der drei Frikative. Für die meisten Versuchspersonen haben /s ʂ/ die jeweils am weitesten nach vorne und zurück verlagerte Position, während sich /ɕ/ bezogen auf die anterior-posterior-Dimension zwischen den beiden anderen Sibilanten befindet. Weiterhin ist die Zungenspitze für /s/ generell weiter gesenkt im Vergleich zu den beiden anderen Frikativen; /ʂ/ tendiert dazu die höchste Zungenspitzenposition zu erreichen, möglicherweise da sich die Zungenspitze von einer anfangs nach hinten-oben rotierten Position entrollt.

Anschließende Analysen zeigten, dass verschiedene Kombinationen der beiden physiologischen Parameter eine klare Trennung der drei Frikative bezogen auf ihren Artikulationsort liefern. Am geeignetsten war jedoch die Kombination der horizontalen Position der Zungenspitze und ihre vertikale Orientierung (Abb. 2.4). Zur Erinnerung: letztere ent-

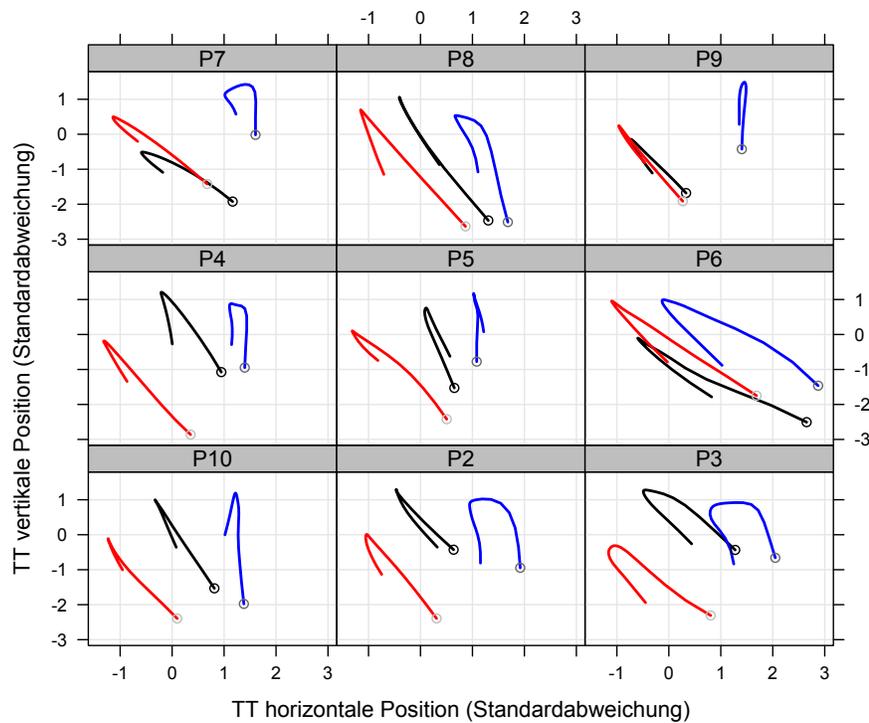


Abbildung 2.3: Lobanov-normalisierte vertikale (y-Achse) und horizontale (x-Achse) TT Bewegungsverläufe zwischen den beiden Geschwindigkeitsmaxima, gemittelt über die drei Vokalkontexte, getrennt für den dentalen (rot), retroflexen (blau) und den alveolopalatalen (schwarz) Frikativ (der Kringel markiert den Anfangspunkt des Bewegungsverlaufs).

hält Informationen zu den Winkelkoordinaten des Sensors parallel zur links-rechts Achse. Nachdem erwartet werden kann, dass die Zungenspitze in $/\zeta/$, jedoch nicht in $/\xi/$, nach hinten-oben gedreht ist, sollte der Zungenspitzensensor, der direkt hinter der Zungenspitze angebracht ist, für $/\xi/$ also entlang der Achse rotieren, welche senkrecht zur sagittalen Ebene liegt – beziehungsweise zumindest in höherem Maße als es für $/\zeta/$ der Fall ist. Wie in Abb. 2.4 zu sehen, ist dies für 8/9 Sprechern der Fall, für welche die Zunge im $/\xi/$ stärker rotierte als für $/\zeta/$: besonders auffällig ist, dass zwei der Sprecher (P5 und P8) $/\xi/$ und $/\zeta/$ nur durch dieses Merkmal unterscheiden, während die beiden Laute bezogen auf ihre horizontale TT Position undifferenziert bleiben. Weiterhin wird in Abb. 2.4 gezeigt, dass mit Ausnahme von Sprecher P6 anhand dieser beider Dimensionen alle drei Frikative für die restlichen Sprecher nahezu komplett unterschieden werden können. Hieraus wird gefolgert, dass $/\xi/$ $/\zeta/$ hinsichtlich der Zungenspitzenstellung voneinander getrennt werden.

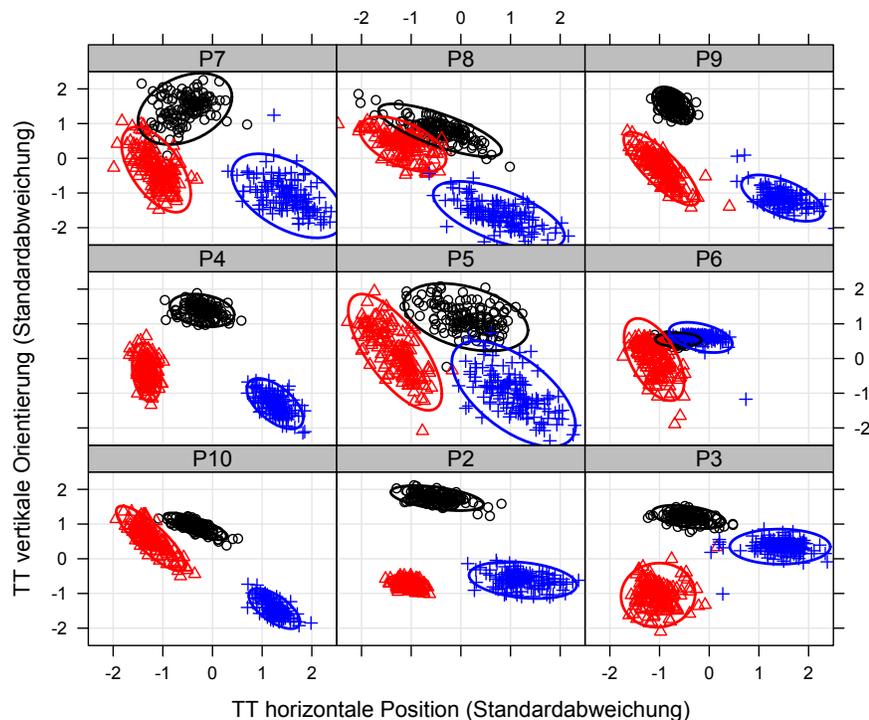


Abbildung 2.4: Lobanov-normalisierte TT Orientierung (y-Achse) und horizontale TT Position (x-Achse) zum Zeitpunkt der maximalen Konstriktion (m_{on} in Abb. 2.2) separat für jeden Sprecher und getrennt für /s/ (rotes Dreieck), /ʂ/ (blaues Kreuz) und /ɕ/ (schwarzer Kringel). Die Konfidenz-Ellipsen sind 2.47 Standardabweichungen and beinhalten 95% der Datenpunkte.

2.4.2 Ergebnisse II: Friktion – Akustische Analyse

Nachfolgend wird untersucht, inwieweit die klare physiologische Unterscheidung der drei Frikative und ihre akustischen Merkmale übereinstimmen. Die Ensemble-gemittelten Spektren in Abb. 2.5 zeigen, dass /s/ durch die Energie in höheren Frequenzen von den beiden anderen Frikativen getrennt werden kann, während sich die Ensemble-gemittelten Spektren von /ʂ ɕ/ relativ ähnlich sind⁴. Es wurden verschiedene Kombinationen von spektralen

⁴Eine Anmerkung eines anonymen Reviewers war es, dass für Ensemble-gemittelte Spektren die Bedingungen der Multitaper-Methode gegeben sein sollten (Jesus & Shadle, 2002; Shadle, 2006, 2010). Um dies zu testen, berechneten wir die Spektren mittels der Multitaper-Methode in Matlab (Percival & Walden, 1993). Wir verwendeten die Standard-Einstellungen mit einem Wert von vier für das Produkt der Zeit-Bandbreite; weiterhin wurde dabei Thomson's (1982) adaptive nicht-lineare Methode angewendet, um die einzelnen spektralen Schätzwerte zu kombinieren. Wie in Abb. A.1 in Anhang A.1 zu sehen, erhielten wir in unserem Vorgehen (Abb. 2.5) und in der Multitaper-Methode fast identische Ergebnisse. Unsere Ergebnisse sind daher konsistent mit denen aus Reidy (2015), der zeigte, dass die Multitaper-Methode für die

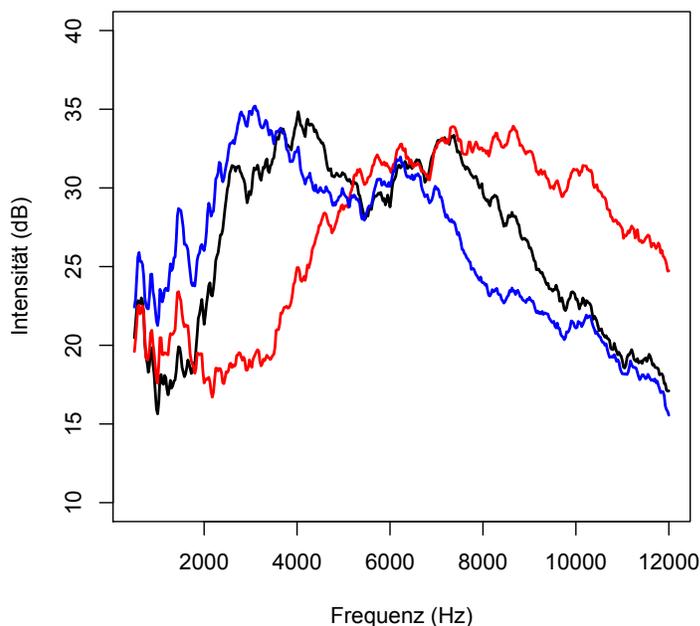


Abbildung 2.5: Ensemble-gemittelte Frikativspektren zum zeitlichen Mittelpunkt der Frikative getrennt für die dentalen (rot), retroflexen (blau) und alveolopalatalen (schwarz) Sibilanten. Gemittelt wurde über alle Frikative, die von allen Sprechern produziert wurden.

Parametern, darunter auch spektrale Momente, zum zeitlichen Mittelpunkt der Frikative getestet (Forrest et al., 1988). Es stellte sich heraus, dass die zwei Parameter, welche die Artikulationsorte der Sibilanten erfolgreich voneinander trennen konnten, proportional zur linearen Steigung (C_1) und Krümmung (C_2) liegen. Abgeleitet aus der diskreten Kosinustransformation wurden diese, wie in 2.3.1 beschrieben, berechnet, nachdem die Frequenzachse in die mel-Skala konvertiert wurde. Für C_1 gilt: wenn eine Regressionslinie durch die drei Spektren gezogen würde, dann würde /s/, wie in Abb. 2.5 angedeutet, durch eine steigende im Gegensatz zu einer fallenden Steigung von den beiden anderen Lauten unterschieden werden können. Im Fall von C_2 gilt, je ähnlicher das Ensemble-gemittelte Spektrum der Krümmung einer Parabel, desto höher die Werte von C_2 . Für das Ensemble-gemittelte /s/-Spektrum zeigt sich eine klare parabolische Form in Abb. 2.5, diese Beobach-

Schätzung von Gipfeln und Tälern von Vorteil ist, dass ihre Anwendung jedoch bei der Parametrisierung der Daten in spektralen Momenten (oder in den von uns berechneten DCT-Koeffizienten), die auf der Summe der Amplituden-Schätzungen basieren, einen nur sehr geringen Effekt hat.

tung zusammen mit der generell höheren Amplitude über den mittleren Frequenzbereich für /ɕ/ verglichen mit /ʂ/, bereitet die Basis zur Unterscheidung der Sibilanten in diesem Parameter.

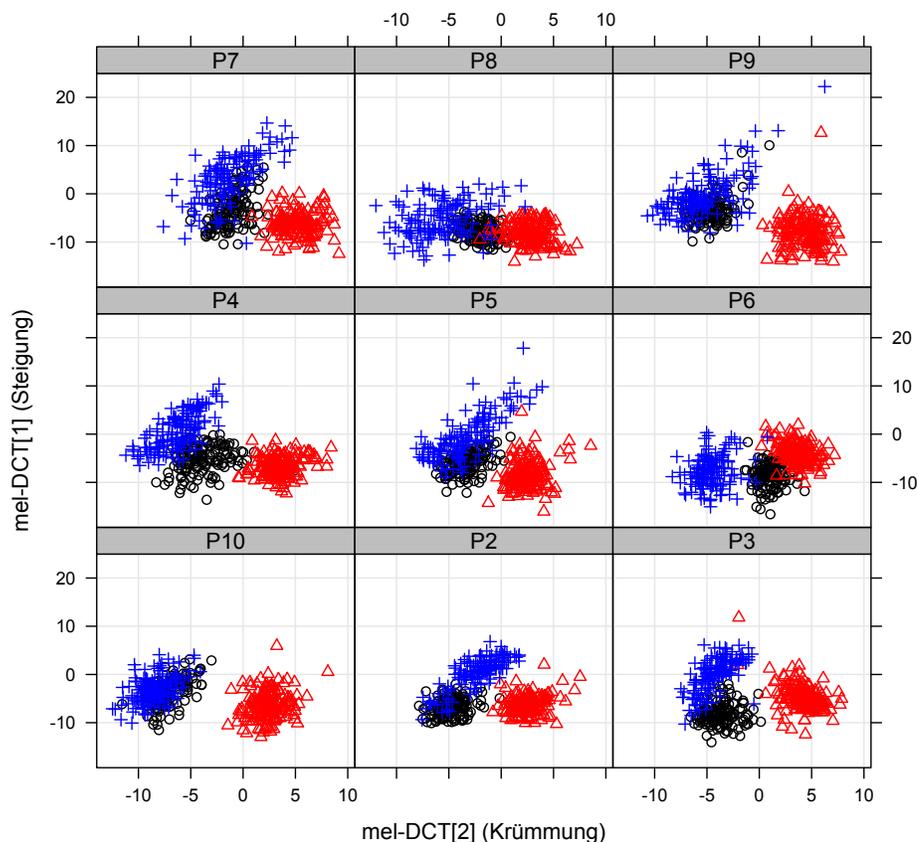


Abbildung 2.6: Erster (Steigung) und zweiter (Krümmung) mel-skalierten DCT-Koeffizient separat für den dentalen (rotes Dreieck), den retroflexen (blaues Kreuz) und den alveolopalatalen (schwarzer Kringel) Sibilanten gemittelt über alle Vokalkontexte sowie über beide Sprechgeschwindigkeiten.

Für die meisten Sprecher zeigt Abb. 2.6 eine Überlappung von /ʂ ɕ/ im $C_1 \times C_2$ Raum, während /s/ klar von den anderen beiden Sibilanten separiert wird (ausgenommen Sprecher P6). Die Daten in Abb. 2.6 sind konsistent mit der Klassifikation (für Details zur Klassifikation siehe Abschnitt 2.3.1), die in der langsamen Sprechgeschwindigkeit (Tab. 2.3) 96% korrekte Klassifikationen für /s/ im Gegensatz zu 77% und 63% jeweils für /ʂ ɕ/ zeigte. Tab. 2.3 zeigt einen höheren Grad an Verwechslung von /ʂ ɕ/ in der langsamen Sprechgeschwindigkeit, mit einer 25%-igen Falschklassifizierung der /ɕ/ Sibilanten als /ʂ/ und einer

23%-igen Missklassifizierung von /ʃ/ als /ç/. Weiter zeigt Tab. 2.3, dass die Auswertung der Klassifikation in der schnellen Sprechgeschwindigkeit ein ähnliches Muster hervorbringt.

Es wurde der Einfluss von Artikulationsort und von Sprechgeschwindigkeit auf die Klassifikationsrate getestet. Außerdem wurde getestet, ob es einen Einfluss hat, ob der Sibilant in einem Echt-Wort oder in einem Logatom vorkommt. Hierfür, wurde ein gemischtes Modell mit der binären Antwort korrekt oder inkorrekt klassifizierter Sibilant als abhängige Variable, sowie mit Artikulationsort (drei Stufen /s ʃ ç/), Worttyp (zwei Stufen: Echt-Wort/Logatom), und Sprechgeschwindigkeit (zwei Stufen: langsam/schnell) als *fixed factors*, angewendet; Sprecher (9 Stufen) und Wort (22 Stufen: die verschiedenen Echt-Wörter und Logatome in Tab. 2.1) wurden als Randomfaktoren in das Modell eingebunden. Desweiteren wurden alle Interaktionsmöglichkeiten zwischen den unabhängigen Faktoren in dieses Modell eingebunden. Der Einfluss von Worttyp und Sprechrate wurde durch zwei Modelle bewertet: eines in welchem alle Faktoren, wie oben beschrieben, eingebunden waren; sowie ein Modell für welches die Faktoren Worttyp und Sprechrate ausgeschlossen wurden. Ein Vergleich dieser beider Modelle (eines mit und eines ohne Worttyp und Sprechrate), zeigte keine signifikanten Unterschiede: weder Worttyp noch Sprechgeschwindigkeit haben einen signifikanten Einfluss auf die Klassifikationsrate. Wie vermutet, wurde die Klassifikationsrate jedoch durch den Artikulationsort der Sibilanten beeinflusst ($\chi^2 = 67.2$, $p < .001$).

	/ç/	/s/	/ʃ/
/ç/	.63 (.61)	.12 (.11)	.25 (.28)
/s/	.03 (.04)	.96 (.95)	.01 (.01)
/ʃ/	.23 (.23)	.00 (.00)	.77 (.77)

Tabelle 2.3: Ergebnisse für die korrekte Klassifikation der drei Sibilanten /s ʃ ç/ anhand Gaußschem Trainieren/ Testen im $C_1 \times C_2$ -Raum basierend auf der *leave-one-out* Methode, d.h. $\forall 9$ Sprecher ($k = 1, 2 \dots 9$), teste auf Sprecher k , trainiere auf alle anderen 8 Sprecher. Die Ergebnisse zeigen die Klassifikationen für die langsame Sprechgeschwindigkeit und für die schnelle Sprechrate in Klammern.

2.4.3 Ergebnisse III: Koartikulatorische Einflüsse auf den Vokal – Physiologische Analyse

Im vorangegangenen Abschnitt konnte gezeigt werden, dass die klare Trennung der Sibilanten bezogen auf die Zungenstellung nicht mit der akustischen Analyse des Friktions-

geräusches, welche eine wesentliche Verwechslung zwischen /s̺ ɕ/ zeigte, einhergeht. Im nachfolgenden Abschnitt soll eine ähnliche Analyse auf den Onset der Vokaltransitionen angewendet werden.

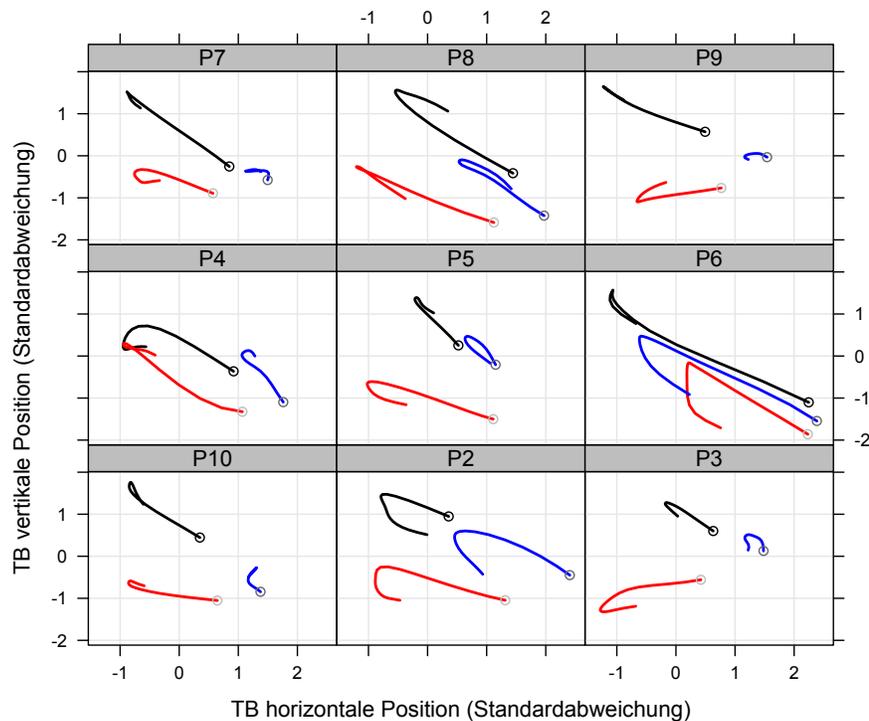


Abbildung 2.7: Für jeden Sprecher einzeln, Bewegungsverläufe für die vertikale und horizontale TB (Zungenrücken) Position zwischen dem akustischen Vokaloffset und -onset, welche dem dentalen (rot), dem retroflexen (blau) und dem alveolopalatalen (schwarz) Sibilanten folgten.

Ausgenommen Sprecher P4 (für welchen die TB Bewegungsverläufe für den dentalen und alveolopalatalen Sibilanten ähnlich sind), zeigt Abb. 2.7 für die restlichen Sprecher, dass die vertikale TB Position in den Vokalen, welche /ɕ/ folgen, höher ist (erkennbar an den höheren vertikalen TB Werten). Im Gegensatz dazu, sind die vertikalen TB Positionen in den Vokalen, die /s̺/ nachfolgen, niedriger (erkennbar an den niedrigeren TB Werten). Diese Ergebnisse lassen vermuten, dass /ɕ/ den nachfolgenden Vokal koartikulatorisch stark beeinflusst. Abb. 2.7 zeigt außerdem, dass (mit der Ausnahme von Sprecher P6) /s̺/ im Vergleich zu /s/ eine eher rückverlagerte Zungenposition aufweist. Demnach beinhaltet das Zungendorsum am Vokaloffset, sowie auch tiefer im Vokal, Informationen

zur Unterscheidung der drei Sibilanten.

2.4.4 Ergebnisse III: Koartikulatorische Einflüsse auf den Vokal – Akustische Analyse

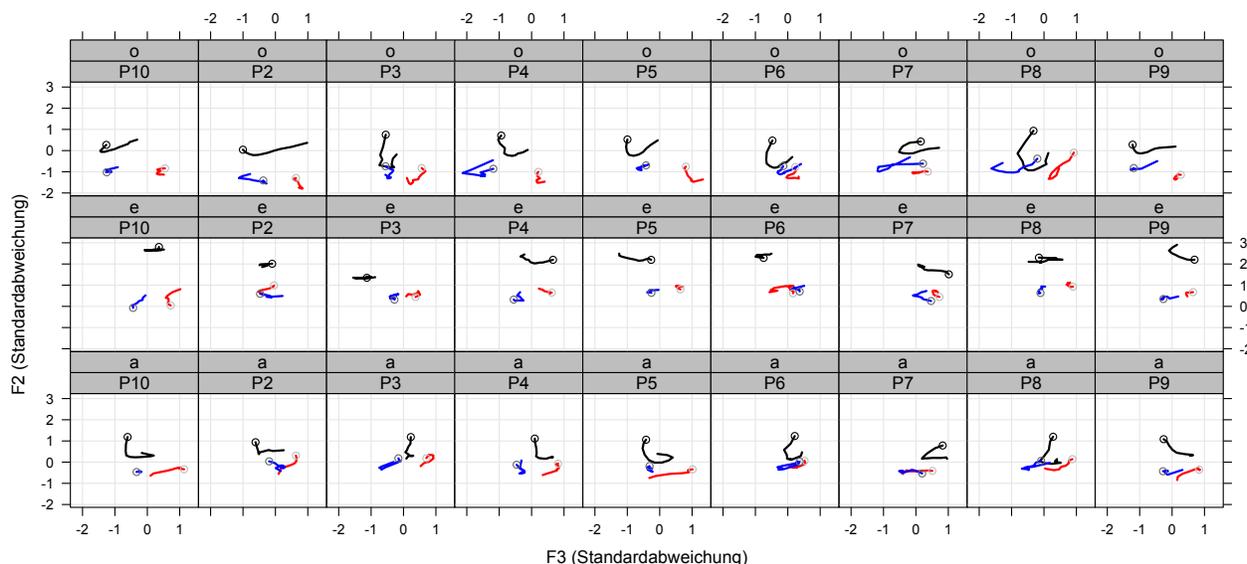


Abbildung 2.8: Gemittelte und zeitnormalisierte Trajektorien im F2 (y-Achse) × F3 (x-Achse) Raum zwischen dem akustischen Vokalonsset und seinem zeitlichen Mittelpunkt separat für jeden Vokal, der den drei Sibilanten folgt: dental (rot), retroflex (blau) und alveolopalatal (schwarz). Das berechnen der Mittelwerte wurde nach der Lobanov-Normalisierung der Sprecherdaten durchgeführt.

Für alle Sprecher wurden höhere F2-Transitionswerte für Vokale gefunden, die /ç/ nachfolgten. Dies geht einher mit der Beobachtung aus der physiologischen Analyse in Abb. 2.7. Obwohl /s ç/ in ihren F2-Transitionen überlappen, können sie zu einem gewissen Grad durch die niedrigeren F3-Werte von /ç/ unterschieden werden.

Abb. 2.9 veranschaulicht den starken koartikulatorischen Einfluss von /ç/ auf den nachfolgenden Vokal, welcher merklich für alle Vokale F2 anhebt und F1 im /ε/ senkt. Dahin gehend liefern diese Daten Beweise für die Palatalisierung des Vokals im /ç/ Kontext.

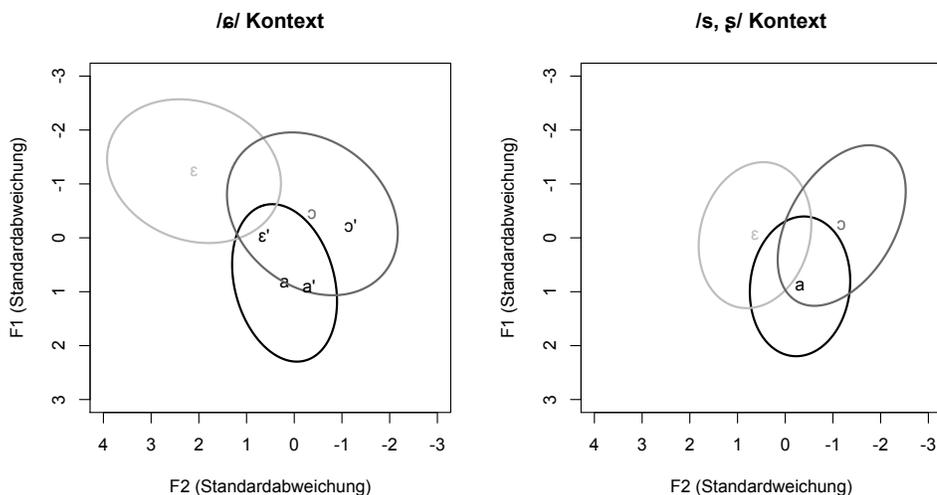


Abbildung 2.9: F1 (y-Achse) und F2 (x-Achse) extrahiert zum zeitlichen Mittelpunkt des Vokals für /ɛ/ (linkes Feld) und für /s ʂ/ (rechtes Feld) gefolgt von /ε/ (grau), /a/ (schwarz) und /ɔ/ (dunkelgrau). Die schwarzen Vokalsymbole (a' ε' ɔ') im linken Feld sind aus den Ellipsen im rechten Feld extrahierte Zentroide, d.h. die Formantzentroide für die Vokale, die /s ʂ/ folgen. Die Ellipsen sind 2.47 Standardabweichungen und beinhalten mindestens 95% der Tokens.

Die Ergebnisse der *leave-one-out* Klassifikation basierend auf einem Modell, das aus zwei Parametern besteht – F2 und F3 Werte zum zeitlichen Vokalonset –, zeigen in der langsamen Sprechrate eine hohe Klassifikationsrate von 91% für /ɛ/, während zwischen den beiden anderen Sibilanten in diesem Parameter eine relativ hohe Verwechslung gefunden wird (Tab. 2.4). Auch wenn die Identifikationsraten von /s ʂ/ in der langsamen Sprechgeschwindigkeit (jeweils 82%, 71%) über der Zufallswahrscheinlichkeit liegen (33%), wurde trotzdem eine deutliche Verwechslung zwischen den beiden gefunden (26% /ʂ/ missklassifiziert als /s/ and 13% /s/ missklassifiziert als /ʂ/). Ähnliche Klassifikationsraten sind in Tab. 2.4 für die schnelle Sprechgeschwindigkeit zu finden. Ein gemischtes Modell mit der binären Antwort korrekte/inkorrekte Klassifikation basierend auf ebendiesen Klassifikationen der Parameter F2 und F3 Onset, sowie mit den fixen und Random-Faktoren, wie sie schon im vorangehenden Abschnitt (2.4.1) erläutert wurden, zeigte weder für Sprechrate noch für Worttyp signifikante Effekte. Damit blieben die Klassifikationen abermals von Sprechrate (langsam oder schnell) und Worttyp (Echt-Wort oder Logatom) unberührt. Wie vorherzusehen war, wurden die Klassifikationsraten vom Artikulationsort des Konsonanten beeinflusst ($\chi^2 = 23.6, p \leq .001$).

	/ɕ/	/s/	/ʂ/
/ɕ/	.91 (.88)	.04 (.06)	.05 (.06)
/s/	.05 (.03)	.82 (.83)	.13 (.14)
/ʂ/	.04 (.05)	.26 (.24)	.71 (.71)

Tabelle 2.4: Ergebnisse für die korrekte Klassifikation der drei Sibilanten /s ɕ ʂ/ anhand Gaußschem Trainieren/Testen bezogen auf die F2 und F3-Werte der drei Vokale mit der *leave-one-out* Methode, d.h. \forall 9 Sprecher ($k = 1, 2 \dots 9$), teste für Sprecher k , trainiere auf alle anderen 8 Sprecher. Die Ergebnisse zeigen Klassifikationen der langsamen Sprechgeschwindigkeit, sowie der schnellen Sprechrate in Klammern.

2.5 Diskussion

Das Hauptziel dieser Studie war es, Aufschluss über die akustischen und artikulatorischen Charakteristika der drei polnischen Sibilanten /s ɕ ʂ/ zu geben und zu testen, ob die höhere phonetische Instabilität von /ʂ/ eine mögliche Quelle für die Reduktion des Dreier-Kontrasts zu einem Zweier-Kontrast sein kann, welche schon in einigen polnischen Varietäten sowie im Mandarin-Chinesisch beobachtet werden konnte. Angefangen wird hier mit dem Grad der Trennung der drei Frikative hinsichtlich des Friktionsgeräusches und den Vokaltransitionen.

Berichten früherer Studien zufolge können Hörer /s ɕ ʂ/ allein anhand des Friktionsgeräusches gut voneinander unterscheiden (Nowak, 2006a). Die hier vorliegenden physiologischen Daten von neun Sprechern zeigen relativ eindeutig, dass /s ɕ ʂ/ durch ihre Zungenposition und -form zu unterscheiden sind. Insbesondere zeigte /s/ (wie auch schon vermutet) einen vorverlagerten Konstriktionsort der Zungenspitze, während /ʂ/ eine Konstriktion vorweist, welche am weitesten rückverlagert ist: dieses Ergebnis geht mit physiologischen Studien zu Frikativen im Mandarin-Chinesisch von Hu (2008) und Proctor et al. (2012) einher, die ebenfalls eine stärker rückgelagerte Zungenspitzenposition für /ʂ/ im Vergleich zu /s ɕ/ fanden. Die Rückverlagerung der Zungenspitze in der vorliegenden Studie ist dadurch zu erklären, dass die Zungenspitze (wie es für retroflexen Konsonanten üblich ist) nach hinten in Richtung des harten Gaumens gerollt ist. Diese Form ist auch das Hauptmerkmal, das den retroflexen Sibilanten von /ɕ/ unterscheidet; das bedeutet, /ʂ ɕ/ unterscheiden sich in der Rotation der Zungenspitze um die Achse, welche parallel zur sagittalen Ebene ist. Allein diese beiden Parametrisierungen der Zungenspitze (horizontale Position, Rotation) sind ausreichend für eine nahezu komplette Trennung von /s ɕ ʂ/, sodass die gute

perzeptive Unterscheidung dieser Frikative anhand des Friktionsgeräusches, wie in Nowak (2006a) berichtet, mit ihren deutlichen physiologischen Unterschieden in der Zungenspitzenposition und auch -orientierung, wie sie in der vorliegenden Studie gefunden wurden, zusammenhängen könnte.

Die hier durchgeführte akustische Analyse des Geräuschanteils der Frikative war insofern konsistent mit der von Nowak (2006a) und anderen (z.B. Jassem, 1995) als dass sie eine klare Trennung von /s/ von den beiden anderen Kategorien zeigt, die auf dessen Energiekonzentration in höheren Frequenzen basiert. Halle und Stevens (1997) zufolge lassen theoretische Überlegungen zu Vokaltraktmodellierungen vermuten, dass die Energie, die in der Region gefunden wird, die typischerweise mit F2-Frequenzen assoziiert ist, für /ʂ/ niedriger ist als für /ʑ/; ersichtlich wird dies auch anhand der Tatsache, dass /ʑ/ eine engere palatale Konstriktion aufweist, die die Resonanzen unterdrückt, die in dem Teil des Vokalraums entstehen, der hinter der Konstriktion ist und damit zum Anstieg der Energie führt, die in der spektralen Region nahe F2 liegt. Prinzipiell sollte ein solcher Unterschied dazu führen, dass die spektrale Energie der niedrigeren Frequenzwerte für /ʂ/ höher gewichtet wird als für /ʑ/. Genau das geht aus den Ensemble-gemittelten Spektren in Abb. 2.5 hervor, die für /ʂ/ einen spektralen Gipfel in der Umgebung von 2 kHz (d.h. in der Region von F2) aufzeigen, der jedoch für /ʑ/ nicht gefunden wird. Diese beobachteten Unterschiede gehen mit den Ergebnissen von F. Li, Edwards und Beckman (2007) einher, die berichteten, dass die Energie in dieser F2-Region im Geräuschspektrum der Mandarin-Chinesischen Sibilanten, /ʂ ʑ/ klar voneinander unterscheidet. Über diese Unterschiede hinaus, und konsistent mit Nowak (2006a), zeigt die vorliegende Studie sehr ähnliche spektrale Formen für /ʂ ʑ/: das heißt, /ʂ ʑ/ unterscheiden sich generell darin, dass für den retroflexen Laute eine dem alveolopalatalen Sibilanten ähnliche spektrale Gestalt in niedrigeren Frequenzen zu sehen ist. Passend dazu berichteten Żygiś und Hamann (2003), dass für eine weibliche Sprecherin /ʂ/ durch den niedrigeren Frequenzschwerpunkt im Geräuschanteil von /ʑ/ getrennt wird, während dies für ihren männlichen Sprecher nicht der Fall ist. Im Klassifikationstest in dieser Arbeit, in dem die Kategorisierung der drei Frikative basierend auf einer DCT-Parametrisierung trainiert und getestet wurde, wird gezeigt, dass selbst bei Klassifikationsraten oberhalb der Zufallswahrscheinlichkeit, nichtsdestoweniger 25% von /ʂ ʑ/ durcheinandergebracht wurden. Dieses Ergebnis lässt vermuten, dass trotz der klaren physiologischen Unterscheidung die Akustik der Frikative allein nicht genug Information für ihre Trennung in eher informeller, spontaner Sprache liefert.

Zahlreiche Studien in den letzten 50 Jahren haben gezeigt, dass Formanttransitionen ergänzende Informationen zur Unterscheidung von Artikulationsorten bei Frikativen liefern. Allen voran wurde das für die nicht-Sibilanten /f θ/ im Englischen bewiesen (Harris, 1958). Andere Studien brachten hervor, dass Formanttransitionen in den nachfolgenden Vokal ebenfalls für die Unterscheidung des /s ʃ/ Kontrasts von Bedeutung sind (Delattre et al., 1962; Gordon, Barthmaier & Sands, 2002; F. Li et al., 2007; Lisker, 2001; Soli, 1985; Wagner et al., 2006; Whalen, 1991, für eine umfassende Zusammenfassung). Mit ihrer Analyse von Frikativen in Shona waren Bladon, Clark und Mickey (1987) unter den ersten, die vermuteten, dass Formanttransitionen vor allem in Sprachen mit Sibilanten, die in drei verschiedenen Artikulationsorten kontrastieren, ausschlaggebend sind. In den Ergebnissen der vorliegenden Studie wird gezeigt, dass die Transition des zweiten Formanten ganz besonders für die Identifikation von /ɕ/ von Nutzen ist. Wie auch schon Nowak (2006a) und Sawicka (2001) fanden, konnte in der vorliegenden Studie beobachtet werden, dass /ɕ/ einen starken koartikulatorischen Einfluss auf den angrenzenden Vokal ausübt, der zur Palatalisierung von letzterem führt: genauer gesagt ist in den physiologischen Daten im /ɕ/ Kontext zum Zeitpunkt des Vokalonssets eine gehobene Position des Zungendorsums zu sehen, welche auch weiter in den Vokal hineinreicht. Dieses Muster findet sich für alle Sprecher und alle Vokalkontexte (Abb. 2.4) und wird, ebenfalls für alle Sprecher und Vokalkontexte, über die erste Hälfte des Vokals hinweg von einem erhöhten F2-Wert begleitet. Dieses Ergebnis ist konsistent mit einer perzeptiven Studie von Lisker (2001), der zeigte, dass englische Hörer in der Lage sind, /ʃ ɕ/ allein durch die akustischen Informationen im Vokal recht zuverlässig voneinander zu unterscheiden. Eine weitere neue Erkenntnis der vorliegenden Studie ist, dass hier nicht nur F2, sondern auch F3 für ebendiese Unterscheidung von Nutzen sein kann. Schon in der akustischen Theorie zur Sprachproduktion wird vermutet, dass retroflexe Konsonanten mit gesenktem F3 verknüpft sind (Fant, 1960), was durch die vorliegende Studie bestätigt werden kann, die zeigt, dass F3 bei /ʃ/ niedriger ist als für die beiden anderen Frikative. Ein niedriger F3 für /ʃ/ in Toda wurde weiterhin in der akustischen Analyse von Gordon et al. (2002) berichtet. Die Ergebnisse der hier vorgestellten Klassifikationen, basierend auf einem zweidimensionalen F2/F3 Raum zu Beginn des Vokals, brachten hervor, dass die Klassifikationsrate über der Zufallswahrscheinlichkeit liegt, und dass fast 90% der /ɕ/ Frikative durch ebendiese Information im Vokal identifiziert werden können. Obgleich die Klassifikationsraten für /ʃ/ mit über 70% ebenfalls hoch sind, geht aus den Daten auch eine deutliche Verwechslung zwischen /s ʃ/ hervor, d.h. 12% /s/ werden als /ʃ/ missklassifiziert und 26% /ʃ/ als /s/. Diese Verwechslungsrate

wird jedoch vermutlich in der Kombination mit der Information aus dem Friktionsgeräusch aufgelöst, da letztere eine 95%-ige Separierung zwischen /s/ und den zwei anderen Frikativkategorien ermöglichte. Ganzheitlich gesehen ist die Haupteckenerkenntnis der Studie, dass das Friktionsgeräusch ausreichend Informationen beinhaltet, um /s/ von /ʃ ʒ/ zu unterscheiden, sowie Vokaltransitionen genügen, um /ɛ/ von /s ʃ/ zu differenzieren; daraus folgt wiederum, dass die erfolgreiche Identifikation von /ʃ/ anhand akustischer Informationen sowohl vom Friktionsgeräusch (zur Trennung von /s/) als auch von Informationen im Vokal (zur Unterscheidung von /ɛ/) abhängig ist. In der hier vorliegenden Arbeit wird gezeigt, dass /s/ von den beiden anderen Frikativkategorien durch Informationen im Geräuschanteil allein (durch die Energiekonzentration des Friktionsgeräusches in den oberen Bereichen des Spektrums) und /ɛ/ ausschließlich durch Informationen im Vokal (durch hohe F2-Werte zu Beginn des Vokals) von den beiden anderen Frikativen unterschieden werden können. Im Gegensatz dazu sind für die Identifikation von /ʃ/ beide Merkmale notwendig, um den Laut von den beiden anderen Frikativkategorien zu unterscheiden: Informationen aus der Friktion (wobei die Energiekonzentration zur Unterscheidung von /s/ in den niedrigeren Bereichen des Spektrums zu finden sein muss) und Informationen aus dem Vokal (wobei F2 und F3 niedrig sein müssen, um die Unterscheidung von /ɛ/ zu ermöglichen).

Außerdem zeigt die Studie, entgegen einer früheren Analyse in Bukmaier, Harrington, Reubold und Kleber (2014), dass die Sprechgeschwindigkeit weder einen Einfluss auf das Friktionsgeräusch noch auf die Vokaltransitionen hat. Dies war jedoch nicht auf mangelnde Variation der Sprechgeschwindigkeit der Sprecher zurückzuführen: für jeden einzelnen Sprecher war in der schnellen Sprechgeschwindigkeit, im Vergleich zu langsamen Sprechgeschwindigkeit, sowohl das Friktionsgeräusch als auch die Vokaltransitionen kürzer. Im Moment gibt es für die divergierenden Ergebnisse zwischen der aktuellen und der Studie von Bukmaier, Harrington, Reubold und Kleber (2014) keine Erklärung, jedoch wird vorläufig angenommen, dass die Auswirkung von Sprechgeschwindigkeit vermutlich kein synchroner Einflussfaktor hinsichtlich des diachronen Verlustes des Sibilanten /ʃ/ darstellt.

Abschließend wird die Möglichkeit erläutert, dass der retroflexive Konsonant sowohl von einem synchronen als auch von einem diachronen Standpunkt aus betrachtet der instabilste der drei Frikativkategorien ist. Die Instabilität des retroflexen Lautes wurde schon unabhängig von der vorliegenden Studie von Duanmu (2006) und Nowak (2006a) angeführt. Beide Studien deuten auf die Möglichkeit eines Zusammenbruches des Dreier-Kontrasts in

Varietäten des Mandarin-Chinesischen und des Polnischen aufgrund einer Neutralisierung der /s ʂ/-Kategorien, hin. Damit Hand in Hand gehen Analysen zur Varietät des Taiwan-Mandarin, die den Dreier-Kontrast nicht aufrechterhält, da der retroflexe Sibilant unter dem Einfluss der Min-Varietät (in welcher retroflexe Laute fehlen) häufig durch den dentalen Frikativ ersetzt wird (Chuang & Fon, 2010). Weiter wurde gezeigt, dass in prosodisch prominentem Kontext gewöhnlich weniger beide, sondern eher nur einer der beiden /s ʂ/ verstärkt wird, wobei die Verstärkung in den meisten Fällen /s/ betrifft. Schon Ladefoged und Bhaskararao (1983) weisen darauf hin, dass es die Komplexität der Gesten während der Produktion von retroflexen Konsonanten ist, die sowohl diese Art von diachronem Wandel als auch das typologisch seltene Vorkommen von retroflexen Lauten erklärt, die nur in Sprachen mit einem großen Inventar an koronalen Lauten zu finden sind (sprich, es ist keine Sprache bekannt, die als einzigen koronalen Laut einen Retroflex aufweist). Hier kann hinzugefügt werden, dass es möglicherweise nicht nur die artikulatorische Komplexität, sondern viel mehr das nicht-lineare Verhältnis zwischen Artikulation und Akustik ist, welches die Instabilität von /ʂ/ begünstigt: genauer gesagt zeigen die retroflexen Konsonanten in der aktuellen Arbeit bezogen auf ihre Zungenposition zwar eine eindeutige Unterscheidung von den beiden anderen Kategorien, jedoch ergibt sich in den Klassifikationen andererseits auch eine hohe Wahrscheinlichkeit an Verwechslung von /ʂ/ sowohl mit /ç/ im Friktionsgeräusch als auch mit /s/ im Vokal. Daher ist /ʂ/ möglicherweise mit Lindblom's (1998) Modell in Verbindung zu bringen, in welchem dieser Sibilant ein Beispiel für eine aufwändige und komplexe Artikulation darstellen könnte, die trotzallem jedoch nur sehr wenige Auswirkungen auf dessen akustische und perzeptive Hervorhebung, von den anderen Frikativkategorien, mit denen der Laut kontrastiert, hat.

Studien und Analysen zum kindlichen Spracherwerb liefern ebenfalls Hinweise in Bezug auf die Instabilität von /ʂ/. Für das Polnische wird angedeutet, dass /ʂ/ erst nach /s ç/, und/oder dass der Kontrast zwischen dem dentalen und dem retroflexen Artikulationsort sehr spät erworben wird (Łukaszewicz, 2006). Weiterhin zeigen Nittrouer und Studdert-Kennedy (1987) sowie Nittrouer (1992, 2002), dass Kinder im Vergleich mit Erwachsenen stärker auf dynamische und weniger auf statische Informationen bei der phonetischen Kategorisierung angewiesen sind: beispielsweise, nutzen Kinder stärker die Vokaltransitionen als das Friktionsgeräusch bei der Frikativkategorisierung. Im Laufe der Entwicklung und mit vorschreitendem Alter, verändert sich dieses Verhältnis schrittweise dahingehend, dass sich die Heranwachsenden Informationen im Friktionsgeräusch zunutze machen. Bezogen

auf die vorliegenden polnischen Daten lässt ein solches Modell vermuten, dass die /s ʂ/ Unterscheidung am ehesten für Verwechslungen anfällig ist. Zum einen, da, wie in der vorliegenden Studie gezeigt, unzureichend Informationen zur Trennung von /s ʂ ʐ/ im Vokal zu finden sind als auch, da Kinder, Nitttrouer's Modell zufolge, weniger fähig wären die kritischen Merkmale zur Trennung von /ʂ/ und den beiden anderen Frikativen aus den eher statischen Geräuschanteilen der Frikative zu extrahieren. Weiterführende empirische Analysen der Frikative sind notwendig, um zu testen, ob die diachrone Instabilität von /ʂ/ ihren Ursprung in der höheren Verwechslungswahrscheinlichkeit von /s ʂ/ in der Produktion und Perzeption von Kindern findet, wie es die Ergebnisse der vorliegenden Studie vermuten lassen.

Kapitel 3

Perzeption des Dreier-Kontrasts /s ʂ ʐ/ im Polnischen in akzentuiertem und deakzentuiertem Kontext

3.1 Zusammenfassung

In dieser Studie wurde die perzeptive Stabilität des /s ʂ ʐ/ Kontrasts im Standard-Polnischen untersucht. Dieser Kontrast basiert auf zwei Hauptmerkmalen: (1) auf den spektralen Eigenschaften des Frikativs, die /s/ von /ʂ ʐ/ separieren und (2) auf den Transitionen des zweiten Formanten in den nachfolgenden Vokal hinein, welche die Distinktion zwischen /ʐ/ und /s ʂ/ erhalten. Einige Studien weisen darauf hin, dass dieser eher seltene Dreier-Kontrast sowohl in einigen polnischen Dialekten als auch in einigen regionalen Varietäten des Mandarin-Chinesischen zu einem Zweier-Kontrast neutralisiert wird. Von dieser Neutralisierung ist dabei der Dental-Retroflex Kontrast betroffen, wobei der Kontrast, welcher erhalten bleibt, der zwischen /s/ und /ʐ/ ist. Bisher befassten sich nur relativ wenige Studien mit dieser Lautneutralisierung, sodass noch nicht geklärt ist, was genau diesen Merger ausgelöst hat und welche linguistischen und extralinguistischen Faktoren die Wahrscheinlichkeit für den Zusammenfall dieses phonemischen Kontrasts erhöhen. Um Hinweise hinsichtlich dieser Fragen zu erhalten, wurde in der vorliegenden Studie ein Perzeptionsexperiment mit männlichen und weiblichen L1-Polnisch sprechenden Versuchspersonen durchgeführt, wobei die Probanden Stimuli aus den synthetisch erzeugten Kontinua zwischen /s-ʂ/ und /s-ʐ/ identifizieren und kategorisieren sollten. Die Stimuli wurden durch die Manipulation des Friktionsgeräusches – von hohem Frequenzschwerpunkt in 11 Schritten

zu einem niedrigen Frequenzschwerpunkt – und im Falle des /s-ɕ/ Kontinuums durch die zusätzliche Manipulation des Vokals /ɛ/, welcher den Sibilanten nachfolgte, erzeugt. Weiterhin wurden die CVCV-Sequenzen, welche die jeweiligen Stimuli in initialer und medialer Position enthielten in zwei prosodischen Kontexten präsentiert, nämlich als deakzentuiertes und akzentuiertes Zielwort. Ausgehend von Studien, die zeigten, dass in hypoartikulierter Sprache Lautkontraste reduziert werden können, verglichen mit der Hyperartikulation, für welche man eher größere phonemische Kontraste gefunden hat, war die Erwartung für die vorliegende Studie, dass die Hörer im deakzentuierten Kontext größere Schwierigkeiten haben sollten, den Kontrast zwischen /s ʂ/ zu kategorisieren. Die Perception des vermeintlich stabileren Kontrasts zwischen /s ɕ/ sollte hingegen weniger von der Deakzentuierung beeinflusst werden. Zwar konnte in den Neigungswerten der Perzeptionskurven eher gegenteilige Ergebnisse gezeigt werden, nämlich für beide prosodischen Kontexte steilere Kurven für das scheinbar instabile Kontinuum, während die psychometrischen Kurven für das vermeintlich stabilere /s-ɕ/ Kontinuum im Vergleich dazu grundsätzlich flacher waren. Dennoch deuteten die Ergebnisse einen etwas größeren Einfluss des prosodischen Kontexts auf die Kategoriengrenze des potentiell instabilen Kontinuums zwischen dem dentalen und dem retroflexen Laut an: die Kategoriengrenze wurde hier in deakzentuierter Position hin zu vermehrten /s/-Antworten verschoben. Das scheinbar stabilere Kontinuum zwischen /s-ɕ/ hingegen, zeigte eine weniger große Verschiebung der Perzeptionskurve im Vergleich von akzentuierter und deakzentuierter Position. Dieser Fund deutet darauf hin, dass die Wahrscheinlichkeit für eine Neutralisierung des /s ʂ/-Kontrasts mitunter in der von Ohala (1993) beschriebenen Theorie zu Lautwandelprozessen Erklärungsansätze finden könnte. Diese sieht den Hörer und seine mögliche fehlerhafte Verarbeitung von koartikulatorischen Mustern als Faktor an, welcher die Bedingung für einen Lautwandelprozess entstehen lassen könnte. Weiterhin lassen sich die Beobachtungen von steileren Perzeptionskurven bei gleichzeitig stärkerem Einfluss des prosodischen Kontexts auf die Kategoriengrenze für das /s-ʂ/ Kontinuum mit exemplar-basierten Theorien (Johnson, 1997; Pierrehumbert, 2001, 2002) erklären.

3.2 Einleitung

Synchrone Variationen in der Sprache kommen aufgrund vieler verschiedener Faktoren vor. Sprachliche Variationen finden sich zwischen verschiedenen Sprechern (Peterson & Barney, 1952), wie auch innerhalb eines Sprechers (Labov, 1968), wobei beide sprachlichen Mod-

litäten – Produktion und Perzeption – betroffen sind. Des weiteren können akustische Informationen im Sprachsignal aufgrund von unterschiedlichen Lautumgebungen (Lieberman, Cooper, Shankweiler und Studdert-Kennedy 1967; Port und O’Dell 1985), infolge unterschiedlicher prosodischer Bedingungen (Peterson & Lehiste, 1960), durch Unterschiede in der Sprechgeschwindigkeit (J. Miller, 1981) und auch durch verschiedene Sprachstile (Moon & Lindblom, 1994) variieren, um nur einen Teil der mannigfachen Variationen, die in der Sprache zu finden sind, zu nennen. Die Verarbeitung der vielen Variationen durch den Hörer ist zwingend notwendig, um die Kommunikation in der Sprachgemeinschaft zwischen Sprecher und Hörer zu garantieren, was die linguistische Forschung dazu brachte, das direkte Verhältnis von Produktion und Perzeption genauer zu untersuchen (Fowler & Galantucci, 2005). Fowler und Galantucci (2005) argumentieren für die enge Verbindung zwischen Produktion und Perzeption, da der Sprecher zum Zweck der Kommunikation Sprachlaute produziert, welche der Hörer in einer erfolgreichen Kommunikation wahrnehmen und verstehen soll. Es wird davon ausgegangen, dass sprachliche Gesten einen Bestandteil der sprachlichen Kompetenz, der Sprachproduktion und zeitgleich auch der Sprachperzeption darstellen. Diese Verknüpfung lässt Fowler und Galantucci (2005), in Anlehnung an Lieberman und Mattingly’s (1989) Motor Theorie, weiter argumentieren, dass der Hörer sprachliche Gesten wahrnimmt.

Auch aktuelle Modelle zu diachronem Lautwandel, die sich mit der Verbindung von synchroner Variation und Lautwandel beschäftigen, sind bestrebt sowohl die Sprachproduktion als auch die Perzeption in den Erklärungen von Lautwandelprozessen unterzubringen und eine Verbindung zwischen diesen beiden Modalitäten hinsichtlich Sprachwandelprozessen herzustellen. Ein Beispiel für das Verbinden von Sprachproduktion und Perzeption in der Erklärung von vergangenen und aktuellen Lautwandelprozessen ist Ohala’s (1981, 1993) Modell zur synchronen Variation und diachronem Wandel. Für Ohala (1981, 1993) liefert die fehlerhafte Aufgliederung/Verarbeitung der Koartikulation durch den Hörer einen Erklärungsansatz, wie ein möglicher Schritt von synchroner Variation hin zu diachronem Wandel zustandekommen kann. Doch betrachtet man die Fülle an synchronen Variationen, die in der sprachlichen Kommunikation zu finden ist, so ist auch zu sehen, dass doch nicht jede, ja sogar relativ wenige sprachliche Variationen in einem diachronen Lautwandel enden. Ein Grund für die Seltenheit von Lautwandelprozessen, welche durch synchrone Variation herbeigeführt werden, liegt darin, dass der Hörer im Stande ist für Koartikulationen und andere phonetische Prozesse perzeptiv zu normalisieren beziehungsweise, wie

man auch sagt, zu kompensieren. Mann und Repp (1980) zeigten in ihrer Studie zur Perception von Sibilanten, dass Hörer für den Vokalkontext kompensieren. Im Laufe ihres Perceptionsexperiments wurde den Hörern ein synthetisches /ʃ-s/ Kontinuum präsentiert, wobei in einem Fall der nachfolgende Vokal /a/ war, während in der zweiten Bedingung dem Friktionsgeräusch der Vokal /u/ folgte. Diese Variation im Vokalkontext hatte zur Folge, dass die Kategoriengrenzen für das Sibilant-Kontinuum im /u/ Kontext in Richtung der niedrigeren Frequenzen verschoben wurden: die Versuchspersonen klassifizierten mehr Stimuli als /s/, da sie die niedrigen Frequenzen der Lippenrundung im Vokal /u/ zuschrieben und nicht den schrittweise sinkenden Frequenzen im Friktionsgeräusch. Doch gibt es Fälle, in welchen diese Art von Normalisierung für Koartikulation durch den Hörer nicht gegeben beziehungsweise fehlerhaft ist. Diese fehlerhafte Kompensation/Normalisierung für Koartikulation ist eine Möglichkeit, diachrone Lautwandel, wie beispielsweise die Tonogenese (Hombert, Ohala & Ewan, 1979), oder die Frontierung von hinteren Vokalen in der Umgebung von apikalen Konsonanten (Michailovsky, 1975, für das Tibetische; Kleber, Harrington & Reubold, 2012, für das Britische) zu erklären. Ohala (1992) findet weiterhin die Vorbedingungen für Lautwandel dann gegeben, wenn der Hörer unaufmerksam oder unerfahren hinsichtlich der Interpretation des Sprachsignals ist, oder aber wenn die Perception durch die Umgebung (z.B. durch Hintergrundgeräusche) gestört wird. Diese instabilen Gegebenheiten in der Sprachverarbeitungen können dann dazu führen, dass Laute miteinander verwechselt werden. G. Miller und Nicely (1955) zeigten die Möglichkeiten von Verwechslungen in der Perception von Frikativen des Englischen aufgrund von akustischen Ambiguitäten, welche zu perceptiven Ähnlichkeiten von Lauten führen. So wurden in ihrer Studie vor allem die englischen dentalen und labio-dentalen Frikative /θ f/ miteinander verwechselt, was G. Miller und Nicely (1955) als Erklärungsansatz ansahen, warum bei Kindern, welche als unerfahrene Hörer eingeordnet werden können, häufig die Aussprache „Keif“ /ki:f/ für den englischen Eigennamen „Keith“ /ki:θ/ auftritt. Jedoch ist auch hier anzumerken, dass der Hörer nicht grundsätzlich ein ambiges Signal falsch interpretiert, d.h. es gibt lediglich eine gewisse Wahrscheinlichkeit, dass Verwechslungen auftreten. (Ohala, 1992, S. 10) formuliert dies wie folgt:

„[...] an ambiguous acoustic signal will not always be misjudged. Rather, listeners may do so with a certain probability, always less than 1.0. Faced with an unclear signal, they guess the speaker's intention; sometimes they guess right and there is no deviation from what was spoken but sometimes they get it wrong and only then is there an error in perception.“

Es stellt sich hinsichtlich Lautwandel und synchroner Variation daher grundsätzlich die Frage, welche Gegebenheiten die Ambiguitäten im akustischen Signal erhöhen und zu Misinterpretationen des Gesagten durch den Hörer führen können? Beckman et al. (1992), wie auch Cole und Hualde (2013) befassten sich mit den prosodischen Bedingungen, mit welchen Reduktionen in der Sprachproduktion einhergehen können. In diesen Studien wurde gefunden, dass Segmente in prosodisch schwachen Kontexten in der Dauer reduziert werden können, was weiter zu einem *target undershoot* führen kann, was bedeutet, dass das artikulatorische und/oder akustische Ziel eines Sprechers für ein geplantes Segment in unbetonten Silben nicht erreicht wird. Cole und Hualde (2013) zeigten am Beispiel der Vokalalternationen im Spanischen, dass Vokale in weniger prominenter Position in ihrer Quantität und Qualität reduziert werden können, während in betonten Silben ein größeres Maß an phonologischen Kontrasten gefunden werden kann. Genauer gesagt werden dabei offene Vokale in betonten Silben zu Diphthongen erweitert, wobei man in unbetonten Positionen eher reduzierte Varianten der Vokale findet, welche aufgrund des Zusammenfalls von offenen und geschlossenen Vokalen aufkommen. Doch diese Reduktionen haben nicht nur Auswirkungen auf die Produktion, wie Lin, Beddor und Coetzee (2014) in ihrer Ultraschall-Studie zu /l/-Artikulationen im Englischen fanden. Spezifischer war das Ergebnis ihrer Untersuchung, dass eine geringe Reduktion der artikulatorischen Geste der Zungenspitze, die Akustik des Segments in einem solchen Maße beeinflussen kann, dass Unterschiede in der Perzeption und der Interpretation dieser reduzierten Segmente entstehen können. Resultate dieser Art sind mit Stevens' Quantal Theorie (1972, 1989) der Sprachperzeption in Einklang zu bringen. Diese besagt, dass im Vokaltrakt für die Artikulation bestimmter Laute Regionen bestehen, in welchen eine artikulatorische Veränderung kaum akustische Auswirkungen hat. Jedoch gibt es auch Regionen, in denen schon kleine artikulatorische Unterschiede einen großen Effekt auf die Akustik und damit auf die Perzeption des jeweiligen Lautes haben.

Lindblom et al. (1995) fassten Alternationen zwischen klarer und reduzierter Sprache, d.h. zwischen Hyper- und Hypoartikulation, in einem Modell zu Lautwandel zusammen, in welchem beschrieben wird, wie durch die Änderung der Produktion von hyperartikulierter zu hypoartikulierter Sprache eine Verbindung zu Lautwandelprozessen geschlagen werden kann. Dabei wird argumentiert, dass die Wahrscheinlichkeit von Lautwandelprozessen und von phonetischen Neuerungen durch das höhere Maß an synchroner Variation in den Teilen des Signals steigt, welche von den Gesprächsteilnehmern als schon bekannt beziehungswei-

se als vorhersagbar angesehen werden können. Harrington et al. (2013) zeigten, wie die Kompensation für Koartikulation, d.h. die perzeptive Normalisierung für die Koartikulation in der Sprachproduktion, in prosodisch schwachen Äußerungsteilen nicht mehr in dem Maße greift, wie es in prosodisch starken Komponenten der Fall ist, was dann weiter eine mögliche Bedingung für Lautwandel schaffen kann. Harrington und Kollegen (2013, 2015) stellen dadurch eine Möglichkeit dar, wie Ohala's (1993) und Lindblom's (1995) Modelle zu synchroner Variation und Lautwandel miteinander verbunden werden könnten.

Im Vergleich der artikulatorischen und akustischen Daten in Kapitel 2 konnten Hinweise darauf gefunden werden, dass im Polnischen der retroflexe Sibilant im Vergleich zu den beiden Sibilanten /s ʐ/ eine gewisse Instabilität aufweist. Dabei wurde die Variabilität/Instabilität des retroflexen Frikativs /ʂ/ dadurch begründet, dass für eine korrekte akustische Kategorisierung von /ʂ/ eine Abhängigkeit von zwei akustischen Merkmalen besteht: die spektralen Merkmale im retroflexen Sibilanten bringen die Trennung von /s/ hervor, während die F2-Transitionen, welche vom Sibilanten in den nachfolgenden Vokal hinein verlaufen, eine Unterscheidung von /ʐ/ gewähren. Ein zweiter Aspekt, welcher in der artikulatorischen Studie in Kapitel 2 aufkam, und welcher eine mögliche Erklärung für die größere Instabilität von /ʂ/ liefern könnte, ist das nicht-lineare Verhältnis zwischen Artikulation und Akustik des retroflexen Sibilanten. Genauer gesagt wurde in Kapitel 2 gefunden, dass man zwar eine gute Trennung der Sibilanten in den physiologischen Daten erkennen kann, jedoch hatte die relativ komplexe artikulatorische Geste des retroflexen Sibilanten, welcher mit einer angehobenen und rückverlagerten Zungenspitze artikuliert wird, nur sehr geringe akustische Auswirkungen. Der Sibilant /ʂ/ zeigte eine Überlappung mit /s/ in den F2-Transitionen in den nachfolgenden Vokal hinein, während Ähnlichkeiten zwischen /ʂ ʐ/ in den spektralen Eigenschaften im Sibilanten selbst gefunden werden konnten.

Wie schon in der Einleitung dieser Arbeit (1.4.3) und in Kapitel 2 beschrieben, ist es gerade der retroflexe Sibilant, welcher sowohl in einigen dialektalen Varietäten des Standard-Polnischen als auch im Mandarin-Chinesischen (einer Sprache, welche ein Sibilantinventar aufweist, welches dem des Standard-Polnischen sehr ähnlich ist), neutralisiert wird. Bei diesen Neutralisierungen fällt der Dental-Retroflex Kontrast zusammen, wobei letzten Endes neben dem alveopalatalen Sibilanten nur der dentale Frikativ erhalten bleibt (Duanmu, 2006; Nowak, 2006a). Für das Standard-Polnische wurde diese konkrete Neutralisierung bisher, nach bestem Wissen der Autorin, nur von Bukmaier, Harrington,

Reubold und Kleber (2014) und in der Studie in Kapitel 2 dieser Arbeit untersucht. Zu der Kontrastneutralisierung im Mandarin-Chinesischen existieren hingegen einige Studien, welche den Zusammenfall der Sibilanten /s ʃ/ untersuchten, wobei sowohl Produktionsdaten (Chang, 2011, 2012; Chang & Shih, 2015; F. Li, 2008) als auch Perzeptionsdaten (Chang, Shih & Allen, 2013; Chiu, 2009; Chuang & Fon, 2010, 2011; Shih & Kong, 2011) in Bezug auf den /s ʃ/-Merger analysiert wurden. In diesen Studien wurde gefunden, dass die Neutralisierung des Kontrasts sowohl von der Sprechgeschwindigkeit, dem Vokalkontext und der Akzentuierung als auch vom Geschlecht der Sprecher/Hörer abhängig sein kann. Chuang und Fon (2010) untersuchten in ihrer Studie zur Produktion der Sibilanten des Mandarin-Chinesischen, inwiefern sich die Produktion der Sibilanten /s ʃ/ von den prosodischen Gegebenheiten – Akzentuierung vs. Deakzentuierung – beeinflussen lässt. Die männlichen und weiblichen Versuchspersonen zeigten zum einen unterschiedliche geschlechtsspezifische Muster hinsichtlich der Kontrasterhaltung der beiden Frikative, zum anderen konnten ebenfalls geschlechterübergreifende Unterschiede in der Realisierung der Frikative in de- beziehungsweise in akzentuiertem Kontext gefunden werden. Drei der insgesamt vier männlichen Probanden tendierten zu einer Neutralisierung des Dental-Retroflex Kontrasts, während im Gegensatz dazu für alle vier weiblichen Versuchsteilnehmer keine Neutralisierung des /s ʃ/-Kontrasts gefunden werden konnte. Dieses Ergebnis kann mit Labov's (1990) sozio-phonetischer Studie zum Zusammenspiel von Geschlecht und Sprachgebrauch in Verbindung gebracht werden. Chuang und Fon (2010) zeigten eine Asymmetrie im Sprachgebrauch hinsichtlich nicht-/Standardsprachlicher Varianten der weiblichen und männlichen Probanden. Die weiblichen Sprecher nutzten eher die Standardsprachliche Aussprachevariante – Erhalt des Kontrasts –, während die männlichen Sprecher eher die dialektale und nicht-Standardsprachliche Form, in welcher der Kontrast neutralisiert wird, zeigten.

Bezogen auf die beiden prosodischen Kontexte – Akzentuierung und Deakzentuierung – konnten keine so klaren Muster für die Sibilantproduktionen gefunden werden. Gemessen anhand des ersten spektralen Moments wurden für einige der Sprecher kein Effekt von Betonung auf den Kontrast zwischen den beiden Sibilanten beobachtet. Genauer gesagt wurden in diesen Fällen zum einen überhaupt keine Unterschiede im ersten spektralen Moment (M1) hinsichtlich der beiden prosodischen Kontexte gefunden, zum anderen wurden beide Laute in der akzentuierten Bedingung in die gleiche Richtung verstärkt, nämlich meist mit höheren M1-Werten, sodass die akustische Distanz zwischen den beiden Kategorien gleich groß blieb. Für die Sprecher, welche den Kontrast zwischen den beiden Sibilanten in

deakzentuiertem Kontext neutralisierten, wurde die Neutralisierung ebenfalls in der Akzentuierung gefunden, wobei diese neutralisierten Sibilanten in der akzentuierten Bedingung generell mit erhöhten Frequenzschwerpunkten produziert wurden. Zwei der insgesamt acht Probanden zeigten eine tatsächliche Verstärkung des Kontrasts in akzentuiertem Kontext, in dem ihre Realisierung der Sibilanten in akzentuierter Position jeweils in eine extremere Richtung geschoben wurde: höhere Frequenzen für den dentalen und niedrigere Frequenzen für den retroflexen Frikativ. Auch Chang und Shih (2015) fanden keine konsistenten Unterschiede hinsichtlich der prosodischen Position, in welcher die Sibilanten produziert wurden. Jedoch wurden Einflüsse des Vokalkontexts auf den Kontrast der beiden Sibilanten /s ʂ/ gefunden. Im /i/-Kontext wurde ein größerer Kontrast zwischen den Sibilanten berichtet als in den Vokalkontexten /a u/. Diese Ergebnisse hinsichtlich des Einflusses des Vokalkontexts können mit den Ergebnissen aus der Studie von Jeng (2006) ergänzt werden, in welcher gefunden wurde, dass der /s ʂ/-Kontrast im /u/-Kontext am geringsten ausfällt.

Die Erkenntnisse in Bezug auf den Vokalkontext und die damit verbundene Distinktion der Sibilanten führt weiter zu der Frage nach der Relevanz der Vokaltransitionen für die Trennung der drei Laute /s ʂ ʐ/. In einer Perzeptionsstudie von Lee-Kim (2015), welche sich mit dem /ʂ ʐ/-Kontrast im Polnischen beschäftigte, und dabei vor allem mit der Gewichtung der Formanttransitionen für diesen Kontrast, wurde gefunden, dass die Informationen im nachfolgenden Vokal für den alveopalatalen Sibilanten von großer Bedeutung sind. Dabei wurde vorhergesagt, dass für den Laut /ʐ/ im /u/-Kontext der größte perzeptive Effekt hinsichtlich der Transitionen zu finden sein sollte, da in diesem Fall große Formantverläufe erwartet werden können. Im /ε/-Kontext liefern die schon hohen Werte des zweiten Formanten währenddessen eher wenig Informationen in Hinblick auf den Alveopalatal-Retroflex Sibilantkontrast. Um diese Vermutungen zu testen, wurden in Lee-Kim's (2015) Studie in den Stimuli, welche aus CV-Sequenzen (C=/ʂ ʐ/; V=/ε a u/) bestanden, die Informationen der Formanttransitionen schrittweise reduziert bis zu Beginn des Vokals keine Transitionen mehr vorhanden waren. Die Ergebnisse zeigten, dass für ein ambiges Friktionsgeräusch im Anlaut der CV-Sequenzen mit nachfolgenden Vokalsequenzen, die Reduktion der Formantinformaton im /ε/-Kontext relativ wenig Einfluss auf die Beurteilung der Stimuli hatte, d.h. im /ε/-Kontext gaben die Versuchspersonen fast ausschließlich /ʐ/-Antworten. Sogar mit einem Friktionsgeräusch, das dem retroflexen Sibilanten zuzuordnen war, wurden im /ε/-Kontext auch ohne Formantinformaton noch immer /ʐ/-Antworten gegeben. Hieraus wurde von Lee-Kim (2015) geschlossen, dass

in Vokalkontexten, welche schon kompatibel mit einem bestimmten Laut sind, wie es für den /ɛ/-Kontext und den alveolopalatalen Sibilanten der Fall ist, der Fokus auf die Frikationsgeräusche steigen muss, um eine Identifikation der anderen Sibilanten, und in ihrer Studie speziell des retroflexen Lautes, zu garantieren. Diese Ergebnisse sind mit den Daten in Kapitel 2 dieser Arbeit kompatibel, in welchen ein starker palatalisierender Effekt des Sibilanten /ɕ/ auf den nachfolgenden Vokal, beobachtet werden konnte.

Ausgehend von den bisherigen Studien sollen in dieser hier vorliegenden Arbeit für das Standard-Polnische drei Forschungsfragen untersucht werden: zum einen steht der Einfluss der Vokaltransitionen auf die Identifikation des alveolopalatalen Sibilanten im Fokus. Dabei soll getestet werden, welche Rolle die Informationen in den Formanttransitionen in einem Kontinuum, für welches die spektralen Eigenschaften im Sibilanten von /s/ (hoher spektraler Frequenzschwerpunkt) hin zu einem /ɕ/ (niedriger spektraler Frequenzschwerpunkt) geändert wurden, für die Kategorisierung des Dental-Alveolopalatal Kontrasts spielen. Die spezifische Frage ist dabei, ob die Informationen im sibilantischen Anteil ausreichen, um das scheinbar stabile Lautpaar /s ɕ/ zu kontrastieren, oder ob erst die Transitionen in den Vokal hinein eine Identifikation der Sibilanten möglich machen.

Die zweite und die dritte Frage dieser Studie beschäftigte sich mit der Lautneutralisierung von /s ʂ/ im Standard-Polnischen. Dabei steht zunächst im Vordergrund herauszufinden, ob in der Perzeption von L1-Polnisch sprechenden Versuchspersonen Hinweise auf eine größere Instabilität des Dental-Retroflex Kontrasts gefunden werden können, welche zeigen, dass es eben dieser Kontrast und nicht etwa der /ɕ ʑ/-Kontrast ist, welcher Tendenzen zu einer größeren Variabilität/Instabilität zeigt. In diesem Zusammenhang sollen auch geschlechtsspezifische Unterschiede in der Perzeption der Sibilanten untersucht werden, vor allem vor dem Hintergrund von Chuang und Fon's (2010) Studie zur Produktion dieser Laute und deren Ergebnis, dass eher die männlichen Probanden eine Neutralisierung des /s ʂ/-Kontrasts im Mandarin-Chinesischen zeigten, während die weiblichen Sprecher den Kontrast beibehielten.

Die dritte Frage, die sich hinsichtlich der /s ʂ/ Kontrastneutralisierung stellt, ist, ob in hypoartikuliertem Kontext und damit kompatibel mit Lindblom's (1990) Modell zu Hyper- und Hypoartikulation und Lautwandel, mögliche Bedingungen gefunden werden können, die Erklärungen für eine erhöhte Wahrscheinlichkeit einer /s ʂ/ Neutralisierung aufzeigen. Genauer gesagt soll hier geklärt werden, ob in deakzentuierter Position die Wahrnehmung

des /s ɣ/-Kontrasts, im Vergleich zur vermeintlich relativ stabilen Perzeption in akzentuiertem Kontext, stärker variiert. Dabei wird erwartet, dass im Falle einer perceptiven Instabilität des /s ɣ/-Kontrasts, der Effekt der Deakzentuierung für dieses Lautpaar größer ist als für den vermutlich stabileren Kontrast zwischen /s ɛ/.

Zusammengefasst lauten die Hypothesen für diese Studie wie folgt:

H1 (Rolle der Formanttransitionen für den /s ɛ/-Kontrast):

Der Kontrast zwischen den Lauten /s ɛ/ kann kategorial wahrgenommen werden, wenn den Versuchspersonen nur Informationen im sibilantischen Anteil zur Verfügung stehen. Sind zusätzlich Informationen in den Formanttransitionen vorhanden verbessert sich die Wahrnehmung des Kontrasts.

H2 (Instabilität des /s ɣ/-Kontrasts):

- a) Das /s-ɣ/ Kontinuum wird in deakzentuierter Position nicht in gleichem Maße kategorisiert, wie es in der akzentuierten Bedingung der Fall ist, was sich durch flachere Neigungen der Perzeptionskurven bemerkbar macht.
- b) Das /s-ɛ/ Kontinuum zeigt keinen so großen Unterschied in der Kategorisierung zwischen deakzentuierter und akzentuierter Position, wie es für das /s-ɣ/ Kontinuum der Fall ist. Für das Kontinuum zwischen /s-ɛ/ werden, im Vergleich zum /s-ɣ/ Kontinuum, steilere Perzeptionskurven erwartet.
- c) Männliche Hörer zeigen flachere Perzeptionskurven für das /s-ɣ/ Kontinuum als die weiblichen Hörer.

Um Erklärungsansätze für die oben genannten Fragen zu erhalten, wurden Perzeptionsdaten von Standard-Polnisch sprechenden weiblichen und männlichen erwachsenen Hörern elizitiert und nicht von einem der polnischen Dialekte, welcher von der /s ɣ/ Neutralisierung betroffen ist. Diese Vorgehensweise lässt Rückschlüsse auf Ohala's (1981, 2012) Idee zu, dass die Voraussetzungen für einen Lautwandel grundsätzlich in der Produktion einer Sprache gegeben sind, wobei ebendiese Tendenzen in der Produktion, wiederum die Perzeption beeinflussen und zu einem dauerhaften Lautwandel führen können (für weitere Argumente hinsichtlich dieser Vorgehensweise siehe Stevens & Harrington, 2016).

3.3 Experimentaufbau

3.3.1 Erstellung der Stimuli

Für dieses Perzeptionsexperiment bildeten drei verschiedene Kontinua zwischen den polnischen Sibilanten, die aus Sprachaufnahmen einer weiblichen L1-Polnisch sprechenden Versuchsperson extrahiert wurden, die Grundlage. Die Tokens, die zur Erstellung der Kontinua herangezogen wurden, wurden aus den Daten der artikulatorischen Aufnahmen in Kapitel 2 ausgesucht. Aus den ausgewählten Logatomen *sese* /sɛsɛ/, *szesze* /ʂɛʂɛ/, sowie aus dem Echt-Wort *siekacz* /ɕɛkaɕ/ (Dt. pl. „Schneidezahn“), die im Laufe der langsamen Sprechbedingung der artikulatorischen Sprachaufnahmen in Kapitel 2 elizitiert wurden, wurden anschließend für das Morphen die initialen CV-Sequenzen (C=/s ʂ ɕ/; V=/ɛ/) herausgeschnitten (für Details zum Experimentdesign der artikulatorischen Sprachaufnahme und der darin elizitierten langsamen und schnellen Sprechgeschwindigkeit vgl. Kapitel 2). Die Auswahl der oben genannten CV-Sequenzen zur Erstellung der Stimuli erfolgte aufgrund ohrenphonetischer Beurteilung der Klarheit der Aussprache des jeweiligen Sibilanten durch eine L1-Polnisch sprechende Hilfskraft mit linguistischen Kenntnissen. Weiterhin wurde eine akustische Analyse der Sibilanten (erstes spektrales Moment zum zeitlichen Mittelpunkt des ausgewählten Sibilanten) und des F2-Wertes zum zeitlichen Onsets des nachfolgenden Vokals durchgeführt, um sicherzugehen, dass die ausgewählten Sibilanten eindeutige Tokens der jeweiligen Sibilantkategorie darstellen. Dabei wurde darauf geachtet, dass die Sibilanten eine, für die jeweilige Sibilantkategorie, typische Verteilung – bezogen auf das erste spektrale Moment und die F2-Werte im nachfolgenden Vokalonset – zeigten. Für die /sɛ/-Sequenz wurde zum zeitlichen Mittelpunkt im Sibilanten ein erstes spektrales Moment (kurz M1) von 7294.2 Hz und zum Vokalonset ein F2-Wert von 1707.1 Hz gemessen. Für die Sequenz /ʂɛ/ wurde zum zeitlichen Mittelpunkt des Sibilanten ein M1-Wert von 4851.2 Hz und zum Vokalonset ein F2-Wert von 1601.1 Hz gemessen. Die Sequenz /ɕɛ/ wies zum zeitlichen Mittelpunkt des Frikativs einen M1-Wert von 4757.7 Hz und zu Beginn des Vokals einen F2-Wert von 2445.8 Hz auf.

Aus diesen drei CV-Sequenzen wurden mit dem Programm TANDEM-STRAIGHT (Kawahara, Takahashi, Morise & Banno, 2009) die drei nachfolgenden 22-Schritt-Kontinua erzeugt.

- /sɛ–ʂɛ/ mit ambigem nachfolgendem Vokal /ɛ/
- /sɛ–ʐɛ/ mit ambigem nachfolgendem Vokal /ɛ/
- /sɛ–ʐɛ/ mit Manipulation der Formantwerte im nachfolgenden Vokal /ɛ/

Für das erste Kontinuum wurde die /sɛ/ mit der /ʂɛ/-Sequenz gemorpht, während für das Morphen des zweiten und des dritten Kontinuums die Sequenzen /sɛ/ und /ʐɛ/ die Grundlage bildeten. Die beiden letzteren Kontinua unterschieden sich allein hinsichtlich des Vokals, welcher dem Sibilanten nachfolgte. Genauer gesagt wurde der Vokal des zweiten Kontinuums ambig gehalten, wohingegen der Vokal für das dritte Kontinuum (/sɛ–ʐɛ/) manipuliert wurde, sodass sich die Formantwerte schrittweise von Werten, die mit dem dentalen Sibilanten kompatibel waren, zu solchen Werten änderten, die mit denen des alveopalatalen Sibilanten zusammenpassten. Für das erste Kontinuum zwischen dem dentalen und dem retroflexen Sibilanten wurde der Vokal ambig gehalten, da zwischen diesen beiden Sibilanten grundsätzlich starke Ähnlichkeiten in den Werten des zweiten Formanten des Vokals, welcher /s ʂ/ nachfolgt, gefunden werden können. Dies kann durch die F2-Werte zu Beginn des nachfolgenden Vokals in den ursprünglichen CV-Sequenzen gezeigt werden, welche als Basis zur Erstellung der Kontinua verwendet wurden (für weitere Ergebnisse zu den sehr ähnlichen F2-Werten zu Beginn des Vokals, welcher /s ʂ/ folgt, siehe Kapitel 2). Für das zweite Kontinuum gilt, dass dieses ebenfalls mit ambigem Vokal generiert wurde, um zum einen die gleichen Bedingungen zu schaffen, wie sie für das erste Kontinuum gegeben waren, und um zum anderen den Einfluss der Vokaltransitionen im Vergleich mit dem dritten Kontinuum zu testen. Für dieses dritte Kontinuum wurden daher, neben den spektralen Parametern im Sibilanten selbst, ebenfalls die Formanten im nachfolgenden Vokal /ɛ/ manipuliert, sodass vor allem die niedrigen F2-Werte für die /sɛ/-Sequenz schrittweise in hohe F2-Werte für die /ʐɛ/-Sequenz verändert wurden.

Die drei Kontinua wurden mit dem Programm TANDEM-STRAIGHT (Kawahara et al., 2009) erzeugt. Mit diesem Programm ist es möglich verschiedene Merkmale im Sprachsignal zu manipulieren. Die Manipulation/das Morphen findet für die nachfolgenden Parameter statt: 1. Segmentdauer; 2. Verlauf der Intensität; 3. Verlauf der Grundfrequenz; und 4. Hüllkurve des Spektrums. Dabei ist es möglich diese vier Parameter zusammen oder auch getrennt voneinander zu manipulieren.

Vor dem tatsächlichen Morphen bestand der erste Schritt aus der Analyse der ursprünglichen Tokens, die aus der artikulatorischen Sprachdatenbank ausgewählt wurden. Dabei wurden zunächst die Sibilant-Vokal-Sequenzen in das Quell- und das Filtersignal getrennt.

Darauf folgend wurde eine Grundfrequenzanalyse des Quellsignals und anschließend eine Aperiodizitätsanalyse für stimmlose Lauteinheiten durchgeführt (mehr Details zum Morphen von Kontinua mit TANDEM-STRAIGHT vgl. Reubold, 2012). Im Anschluss daran wurden für das Morphen im Kontinuum zwischen /sɛ-ʃɛ/ und im Kontinuum zwischen /sɛ-çɛ/ im sibilantischen Teil sechs und für den vokalischen Teil drei Zeitanker markiert. Nachfolgend wurden für jeden dieser Zeitanker Frequenzanker definiert, um eine horizontale Überlagerung der Lautsequenzen bei der Erzeugung der Kontinua sicherzustellen. Grundsätzlich können diese verschiedenen Ankerpunkte als eine Art Segmentation angesehen werden, welche bestimmte akustische Markierungspunkte in den beiden, zu morphenden, Signalen darstellt. Diese Ankerpunkte werden eben auf der zeitlichen (Zeitanker) und auf der spektralen (Frequenzanker) Ebene gesetzt. Die Zeitanker stecken meist den Beginn und das Ende des Sibilanten und des Vokals in den CV-Sequenzen ab, während die Frequenzanker genutzt werden, um spektrale Eigenschaften, wie den Frequenzschwerpunkt im Sibilanten und beispielsweise die Formanten im Vokal zu markieren. In den hier verwendeten CV-Sprachsignalen wurden zwei Zeitanker im sibilantischen Teil zum Onset und Offset des Lautes gesetzt. Die restlichen vier zeitlichen Anker wurden dann dazwischen in äquidistanten Abständen in der sibilantischen Sequenz festgelegt. Die Frequenzanker im sibilantischen Teil der CV-Sequenzen wurden jeweils an den Flanken des spektralen Gipfels – Frequenz mit der höchsten Amplitude – als auch am Gipfel selbst zum Zeitpunkt des jeweiligen Zeitankers gesetzt. In der vokalischen Sequenz wurden mit zwei Zeitankern Onset und Offset markiert, während der dritte Zeitanker den zeitlichen Mittelpunkt des Vokals anzeigte. Die Endpunkte der beiden ersten Kontinua wurden erzeugt, in dem alle akustischen Eigenschaften in den Sibilanten aus den anfangs ausgewählten Tokens erhalten wurden, wohingegen der vokalische Teil genau am Mittelpunkt zwischen dem Paar /s-ʃ/ für das erste Kontinuum beziehungsweise zwischen /s-ç/ für das zweite Kontinuum gemorpht und damit ambig zwischen den jeweiligen Endpunkten der Sibilanten gehalten wurde. Zwischen den synthetischen Endpunkten (/s/ und /ʃ/ beziehungsweise /s/ und /ç/) wurden die spektralen Eigenschaften im sibilantischen Teil anschließend zu einem jeweils 22-Schritt-Kontinuum gemorpht. Es entstand damit ein Kontinuum zwischen /s/ und /ʃ/ beziehungsweise zwischen /s/ und /ç/, welches aus jeweils 22 Stimuli bestand. Dabei wurde der spektrale Schwerpunkt von /s/ nach /ʃ/ beziehungsweise /ç/ von Schritt zu Schritt gesenkt, während die Formanten im nachfolgenden Vokal ambig waren.

Das dritte Kontinuum wurde analog zu den beiden vorherigen erstellt, mit dem Unterschied, dass sowohl die akustischen Merkmale des Sibilanten als auch die akustischen Merkmale

des Vokals aus den ursprünglich ausgewählten Aufnahmen beim Morphen übernommen wurden. Dafür wurden zusätzlich zu den sechs temporalen Ankern im Sibilanten, sechs Zeitanker im Vokal gesetzt. Diese sechs Zeitanker im Vokal wurden wie folgt verteilt: jeweils ein Zeitanker, um Onset und Offset zu markieren, sowie vier Anker, welche in äquidistanten Abständen in der vokalischen Sequenz zwischen Onset und Offset verteilt wurden. Anschließend wurden spektrale Anker für alle 12 Zeitanker definiert, wobei im vokalischen Teil darauf geachtet wurde, dass die Frequenzanker mit den Formantwerten der jeweiligen Originalaufnahmen übereinstimmten. Dies erfolgte durch die Zuhilfenahme des Programms *praat* (Boersma & Weenik, 2013), in welchem die Formantwerte aus den Originalaufnahmen zu den Zeitpunkten, die vorher durch die temporalen Anker festgelegt wurden, extrahiert und anschließend in TANDEM-STRAIGHT übernommen wurden. Bei der Erstellung der Endpunkte wurden dann sowohl die sibilantischen als auch die vokalischen Merkmale der ursprünglichen Aufnahmen übernommen, sodass sich nicht nur der Sibilant in seinen Eigenschaften in 22 Schritten von /s/ nach /ɕ/ änderte, sondern auch der Vokal im Laufe der 22 Stimuli von einem, mit dem dentalen Laut kompatiblen Vokal, zu einem Vokal verändert wurde, dessen spektrale Merkmale zu dem alveolopalatalen Sibilanten passten.

Für die tatsächliche Auswahl der Stimuli wurde eine akustische Analyse durchgeführt. Hierfür wurden die M1-Werte der kompletten Sibilantsequenz berechnet. Anschließend wurde über eben diese M1-Werte in der Sibilantsequenz gemittelt, sodass man für jeden Stimulus einen gemittelten M1-Wert erhielt (vgl. Abb. 3.1-3.3). Aus diesen gemittelten M1-Werten wurde dann für jedes der drei Kontinua aus den 22 Stimuli, die elf Stimuli ausgewählt, welche in ihren gemittelten M1-Werten (nahezu) äquidistante Abstände zum nächsten Stimulus aufwiesen. Jeder der 11 CV-Stimuli wurde nachfolgend kopiert; Original und Kopie wurden dann zu einem CVCV-Logatom (C1V1=C2V2=Stimulus₁₋₁₁) konkateniert. Die erste und die zweite CV-Sequenz im CVCV-Zielwort enthielten somit immer die gleiche Stimulusnummer und waren damit komplett identisch. Im nächsten Schritt wurden die Stimuli in den Trägersatz „Ania woła [STIMULUS₁₋₁₁] aktualnie“ (wörtlich „Ania ruft [STIMULUS₁₋₁₁] momentan“) eingebettet.

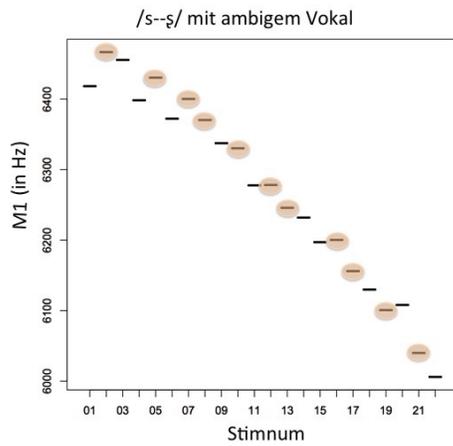


Abbildung 3.1: Stimuli Auswahl /s-ʃ/.

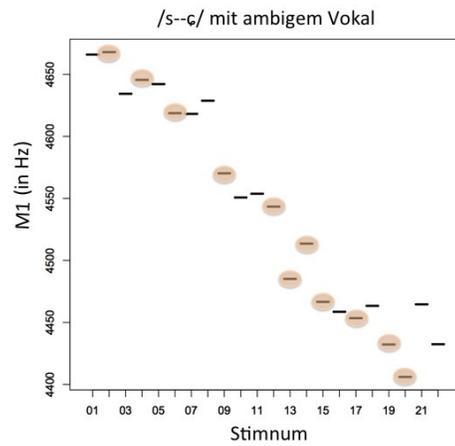


Abbildung 3.2: Stimuli Auswahl /s-ç/.

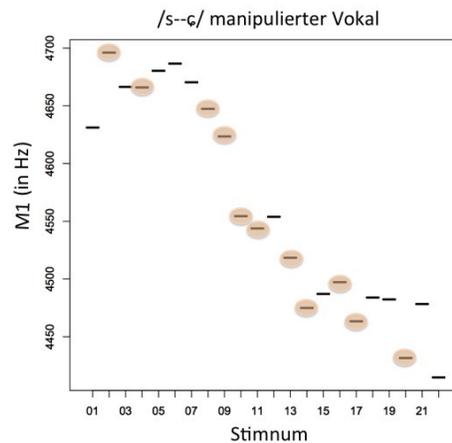


Abbildung 3.3: Stimuli Auswahl /s-ç/ manipulierter Vokal.

Die Trägersätze mit den eingebetteten Stimuli wurden mit Hilfe des in *praat* (Boersma & Weenik, 2013) implementierten Algorithmus *Pitch-Synchronous-Overlap-and-Add* (Moulines & Charpentier, 1990) in die zwei gewünschten prosodischen Kontexte – deakzentuiertes und akzentuiertes Zielwort – gebracht. Um diese beiden prosodischen Kontexte zu erstellen, wurden für die kompletten Sätze, d.h. Trägersatz mit jeweiligem CVCV-Zielwort, drei Parameter manipuliert:

- Dauer
- Pitch
- Amplitude

Für den akzentuierten Kontext galt für alle drei Kontinua, dass die Pitch vom Ende des Wortes „woła“, welches dem Zielwort voranging, bis zum zeitlichen Mittelpunkt im ersten Vokal der CVCV-Zielwort-Sequenz um 14 Hz anstieg und anschließend wieder bis zum Ende des zweiten Vokals der CVCV-Sequenz um 59 Hz abfiel. Die Dauer der beiden Vokale im Zielwort wurde jeweils auf 0.98 ms beziehungsweise auf 0.077 ms gelängt. Die Intensität des kompletten Satzes wurde im Verhältnis zu den Vokalen im Stimulus um 10 dB gesenkt. Durch diese Maßnahmen wurde sichergestellt, dass die Zielwort-Sequenz den prominentesten Teil im Trägersatz darstellte.

In der deakzentuierten Bedingung wurde zum zeitlichen Mittelpunkt des initialen Vokals /a/ im Wort „Ania“ die Pitch auf 270 Hz angehoben, die dann bis zum Auslaut dieses satzinitialen Wortes auf 216 Hz abfiel und weiter auch über das Zielwort hinweg niedrig blieb. Die Dauer der Vokale in der CVCV-Sequenz wurde auf 0.78 beziehungsweise auf 0.73 ms gekürzt. Das satzinitiale Wort „Ania“ hatte in dieser Bedingung eine Intensität von 81 dB, während die Intensität der restlichen Satzteile um 14 dB, gemessen an eben diesem satzinitialen Eigennamen, gesenkt wurde. Durch diese Maßnahmen wurde die Bedingung geschaffen, dass nur der satzinitiale Eigenname als prominent wahrgenommen wird, jedoch nicht das Zielwort. Diese Manipulationen führten demnach zu Unterschieden in der Satzakkzentuierung mit der Betonung der Stimulus-Sequenz im akzentuierten Kontext – eine Akzentuierung, die kompatibel mit einer Antwort auf die Frage „Was hat Ania gerufen?“ wäre –, sowie mit der Betonung des satzinitialen Eigennamens „Ania“ in der deakzentuierten Bedingung, wobei hier die Zielwort-Sequenz unbetont blieb – eine Akzentuierung, die eine Möglichkeit zeigt, auf die Frage „Wer hat gerade [CVCV] gerufen?“ zu antworten. In Abb. 3.4 sind die oben beschriebenen F_0 -Manipulationen für das akzentuierte und das deakzentuierte Zielwort exemplarisch dargestellt.

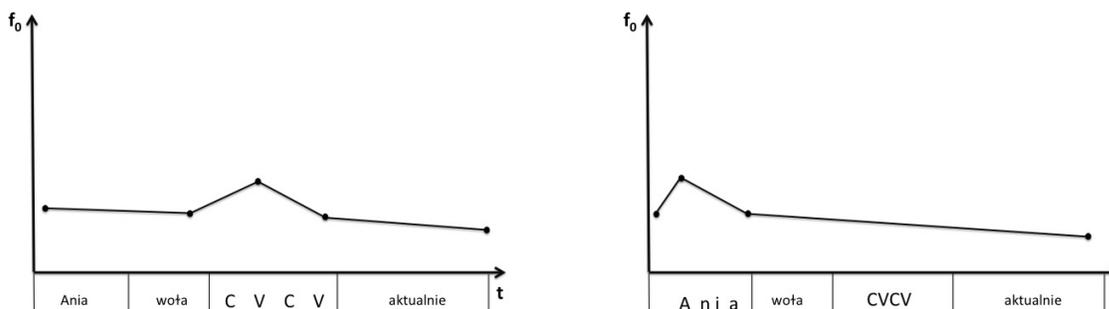


Abbildung 3.4: Exemplanische Darstellung der F₀-Manipulationen für den akzentuierten (links) und den deakzentuierten Kontext (rechts).

Alle sechs Kontinua (/sɛ-ɕɛ/ Vokal ambig & /sɛ-ʂɛ/ Vokal ambig & Vokal manipuliert /sɛ-ɕɛ/ × zwei Akzentuierungen) wurden zwei L1-Polnisch sprechenden Hilfskräften zur Beurteilung präsentiert. Der Eindruck der beiden Testhörer war, sowohl dass die CV-Sequenzen gut zu identifizieren waren als auch dass die Manipulation der Akzentuierung mit den Betonungsmustern des Polnischen übereinstimmten.

3.3.2 Datenerhebung und Hörer

Die Versuchspersonen, die an dem hier vorgestellten Perzeptionsexperiment teilnahmen, waren 13 L1-Polnisch sprechende Erwachsene, davon vier männlich und neun weiblich, im Alter zwischen 17 und 42 Jahren (Ø28.7). Die Versuchspersonen führten das Experiment nicht im Sprachlabor durch, sondern mit der Web-Applikation *Percy*, einem von Draxler (2011) entwickelten Programm zur Durchführung von webbasierten Perzeptionsexperimenten. Das Programm verschafft den Versuchspersonen die Flexibilität die Umstände ihrer Teilnahme an einem Experiment selbst zu bestimmen, da es über den Browser an einem von ihnen gewählten Computer an einem von ihnen selbst gewählten Ort durchgeführt werden kann. Vor dem tatsächlichen Experiment wurden die Versuchspersonen gebeten, Auskunft über die jeweiligen Experimentbedingungen zu geben, sowie weitere demographische Fragen zu beantworten. Die folgenden anonymen Angaben, die dazu dienten die Ergebnisse und vor allem mögliche abweichende Ergebnisse zu erklären, wurden von den Teilnehmern bei diesen Fragen gemacht:

- Alter der Versuchsperson
- Geschlecht der Versuchsperson
- Aufnahmeort (Büro, öffentlicher Ort, zu Hause etc.)
- Eingabegerät (Rechner, Notebook, Smartphone, Tablet etc.)
- Ausgabegerät (interne oder externe Lautsprecher, Kopfhörer, Ohrstöpsel)
- Ort der Einschulung
- Wohnort in Woiwodschaft

Die Versuchspersonen, die an diesem online-Perzeptionsexperiment teilnahmen, lebten teilweise in München und teilweise in Polen. Die Herkunftsregionen (Woiwodschaften) der Versuchspersonen waren: Niederschlesien, Lebus, Ermland-Masuren und Kleinpolen (für jeweils zwei der Probanden), sowie Masowien, Schlesien, Westpommern und Karpatenvorland (für jeweils eine Versuchsperson). Alle weiteren abgefragten Informationen zu Herkunft, Sprachgebrauch, sowie die Informationen zur Experimentdurchführung der Teilnehmer sind in Tabelle B.1 in Anhang B.1 aufgelistet.

3.3.3 Experimentdurchführung

Wie schon im vorangegangenen Abschnitt 3.3.2 beschrieben, wurden die Perzeptionsdaten anhand der online-Perzeptionsexperiment Web-Applikation *Percy* (Draxler, 2011) erhoben. Der Link, der die Teilnehmer zum Experiment führte, wurde durch zwei L1-Polnisch sprechende Hilfskräfte an potentielle Probanden weitergeleitet. Der Vorteil eines solchen online-Experiments liegt in der Freiheit der Versuchspersonen das Experiment an einem von ihnen gewählten Ort zu einer von ihnen gewählten Zeit durchzuführen. Jedoch kann ebendiese Freiheit der Teilnehmer auch zum Nachteil werden, da erst nach Abschluss des Experiments kontrolliert werden kann, wie das Experiment und ob es überhaupt im Ganzen durchgeführt wurde. In dem hier vorliegenden Fall wurde das Experiment von mehr als der Hälfte der ursprünglich 29 Versuchspersonen nicht komplett durchgemacht, was dazu führte, dass diese Probanden nicht in die Analyse aufgenommen werden konnten, sondern eben nur die 13 Teilnehmer, die schon in Abschnitt 3.3.2 genannt wurden.

Die Versuchspersonen hörten im Laufe des Experiments immer alle sechs Kontinua – deakzentuiert: /sɛ–ʂɛ/ (Vokal ambig), /sɛ–ɕɛ/ (Vokal ambig), /sɛ–çɛ/ (Vokal manipuliert); akzentuiert: /sɛ–ʂɛ/ (Vokal ambig), /sɛ–ɕɛ/ (Vokal ambig), /sɛ–çɛ/ (Vokal manipuliert). Jeder Stimulus wurde im Laufe des Experiments siebenmal wiederholt, was zu einer Gesamtanzahl von 462 Stimuli (3 Kontinua × 11 Stimuli × 2 prosodische Kontexte × 7 Wiederholungen) führte, die von den Versuchspersonen identifiziert und kategorisiert werden sollten. Eine männliche Versuchsperson brach das Experiment nach 460 gehörten Stimuli ab. Für diesen Probanden, sowie für eine weitere weibliche Versuchsperson musste außerdem aufgrund technischer Probleme eine weitere Beobachtung aus dem Datensatz gelöscht werden. Da diese beiden Probanden den Großteil der Stimuli angehört und kategorisiert hatten, wurden beide trotzdem in die Analyse aufgenommen. Die Stimuli wurden den Probanden in randomisierter Reihenfolge zur Bewertung in einem *forced choice identification task* präsentiert. Die Versuchspersonen hörten die jeweilige Audiodateien maximal einmal an und mussten dann aus einer der zwei auf dem Bildschirm in orthografischer Form präsentierten Antwortmöglichkeiten wählen, was sie gehört hatten. Für das Kontinuum zwischen /sɛ–ʂɛ/ waren die Antwortvorschläge, die die Versuchsperson auf dem Bildschirm zu sehen bekam SESE und SZESZE, während für die beiden /sɛ–ɕɛ/ Kontinua die Wahl zwischen SESE und SIESIE gestellt wurde¹. Der nächste Satz wurde den Versuchspersonen erst präsentiert, nachdem sie eine Auswahl zum vorher gehörten Stimulus abgegeben hatten.

3.3.4 Kategoriale Wahrnehmung

In der nachfolgenden Abbildung werden fünf verschiedene Formen von psychometrischen Kurven gezeigt, welche die Antworten von Hörern auf Kontinua abbilden. Für diese exemplarischen Kurven wird von zwei Kategorien ausgegangen: x und y . Die blaue Linie in Abb. 3.5 zeigt eine kategoriale Wahrnehmung des Kontrasts zwischen den Kategorien x und y , d.h. von Stimulus 1 bis Stimulus 6 gibt der Hörer ausschließlich x Antworten, während ab Stimulus 7 bis Stimulus 11 nur y -Antworten eliziert werden. Für die blaue Linie ist es möglich eine Kategoriengrenze zu berechnen, wobei dafür der Umkipppunkt von Kategorie x zu Kategorie y der Perzeptionskurve herangezogen wird. Die pinke Linie stellt eine kontinuierliche Wahrnehmung des Kontinuums zwischen den Kategorien x und y dar. Zwar werden die Endpunkte des Kontinuums von den Hörern identifiziert, jedoch erfolgt

¹Die Graphem-Phonem Beziehung ist im Polnischen eindeutig, d.h. der dentale Sibilant /s/ wird orthografisch mit dem Graphem <s> dargestellt, der retroflexen Frikativ wird mit dem Graphem <sz> abgebildet, der alveopalatale Laut /ɕ/ mit <si> und der Vokal /ɛ/ mit dem Graphem <e>; vgl. Gussmann, 2007.

die Wahrnehmung der einzelnen Stimuli des Kontinuums auf eine kontinuierliche Art und Weise. Die grüne Linie in Abb. 3.5 zeigt eine Wahrnehmung an, die durch eine komplette Kontrastneutralisierung geprägt ist, wobei keine Tendenzen in Richtung der einen oder anderen Kategorie zu erkennen ist. Die rote und die gelbe Linie repräsentieren ebenfalls eine komplette Kontrastneutralisierung, jedoch jeweils mit einer Tendenz hin zu Kategorie x (rote Linie) und zu Kategorie y (gelbe Linie).

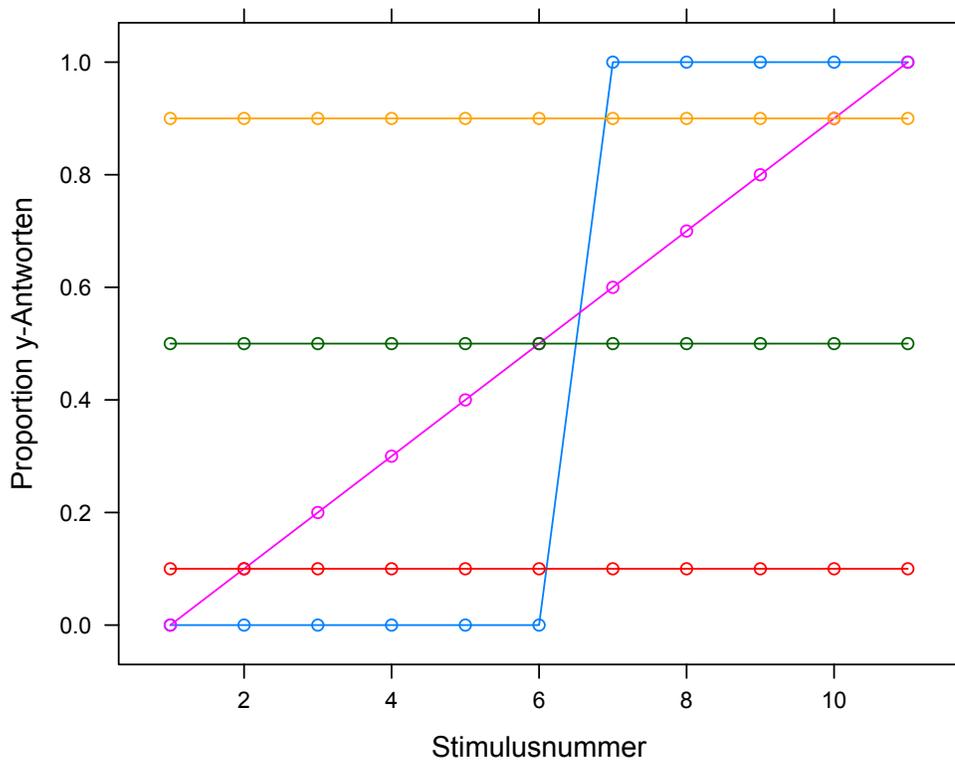


Abbildung 3.5: Schematische Darstellung fünf verschiedener psychometrischer Kurven.

3.3.5 Datenanalyse

Zur Analyse der Daten wurden sechs logistische Regressionsmodelle auf die Antworten der Probanden angewendet, sodass alle Kombinationen von Kontinuum (/sɛ-ʂɛ/ ambiger Vokal vs. /sɛ-ɕɛ/ ambiger Vokal vs. /sɛ-ɕɛ/ manipulierter Vokal) und prosodischem Kontext (akzentuiertes vs. deakzentuiertes Zielwort) miteinander kombiniert wurden. Die binäre Antwort zwischen /s/ und /ʂ/ beziehungsweise zwischen /s/ und /ɕ/ wurde als abhängige Variable gesetzt, während die Stimulusnummer ($1 < n < 11$) die unabhängige Variable

darstellte. Mittels dieser Analyse konnten psychometrische Kurven getrennt für die drei Kontinua (/sɛ-ʂɛ/ ambiger Vokal vs. /sɛ-çɛ/ ambiger Vokal vs. /sɛ-çɛ/ manipulierter Vokal) und separat für die beiden prosodischen Kontexte (akzentuiert vs. deakzentuiert) dargestellt werden.

Im Anschluss an die Berechnung dieser ersten Modelle wurden weitere Modelle gerechnet, welche eine Interaktion zwischen Stimulus und Hörer integrierten, sodass für jeden Hörer einzeln die Interzepts, die Neigungen und die Kategoriengrenzen analysiert werden konnten². Zunächst werden im nachfolgenden Ergebnisteil die Beobachtungen hinsichtlich der ersten Forschungsfrage dargestellt. Hierzu wird ausschließlich auf die Kontinua /sɛ-çɛ/ mit ambigem und mit manipuliertem Vokal eingegangen. Der zweite Ergebnisteil, befasst sich mit der potentiellen Instabilität des /sɛ-ʂ/ Kontinuums gegenüber dem vermeintlich stabileren Kontinuum zwischen /sɛ-çɛ/.

3.4 Ergebnisse

3.4.1 Ergebnisse I: Rolle der Formanttransitionen für den /s ç/-Kontrast

Nachfolgend wird die Analyse des /sɛ-çɛ/ Kontinuums in Bezug auf die Rolle der Formanttransitionen für die Kategorisierung der Sibilanten vorgestellt.

In Abb. 3.6 sind die psychometrischen Kurven zu den Kontinua /sɛ-çɛ/ mit ambigem Vokal (schwarz) und /sɛ-çɛ/ mit manipuliertem Vokal (rot) in den beiden prosodischen Kontexten (links Zielwort akzentuiert, rechts Zielwort deakzentuiert) dargestellt.

²Ursprünglich standen Perzeptionsdaten von 13 Versuchspersonen, die das Experiment komplett durchgeführt hatten, zur Analyse zur Verfügung. Für drei Probanden wurde jedoch festgestellt, dass sie den Sibilantkontrast im /sɛ-çɛ/ Kontinuum mit ambigem Vokal komplett neutralisierten. Für eine Versuchsperson wurde der Kontrast neutralisiert ohne eine Neigung in Richtung einer bestimmten Kategorie zu zeigen, d.h. die Stimuli zeigten bei diesem Probanden keinen Effekt auf die Perzeption. Für die beiden anderen Versuchspersonen wurde der Kontrast mit einer Bias hin zu /ç/-Antworten neutralisiert, d.h. diese Versuchspersonen gaben fast ausschließlich /ç/-Antworten (vgl. Abb. B.1-B.6 in Anhang B.2.)

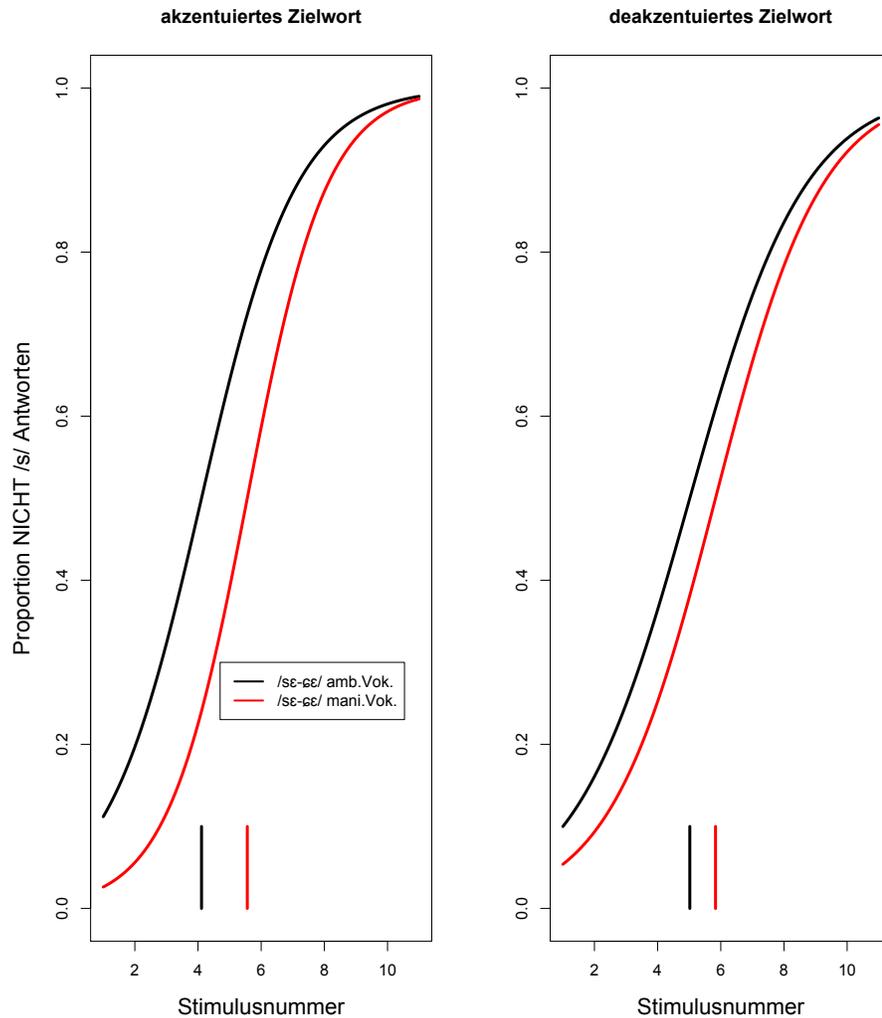


Abbildung 3.6: Psychometrische Kurven der Proportion an NICHT-/s/-Antworten als Funktion der Stimulus Nummer für die Kontexte akzentuiert (links) und deakzentuiert (rechts), separat für die Kontinua /sɛ-ɕɛ/ mit ambigem Vokal (schwarz), sowie /sɛ-ɕɛ/ mit manipuliertem Vokal (rot). Die vertikalen Linien am unteren Rand der Abbildung zeigen die Kategoriengrenzen bei denen /s/- und NICHT-/s/-Antworten gleich wahrscheinlich sind.

Genauer gesagt sind in dieser ersten Abbildung die Proportionen an NICHT-/s/-Antworten (y-Achse) als Funktion der Stimulusnummer (x-Achse) zu sehen. Betrachtet man die Kategoriengrenzen (Linien am unteren Rand der Abbildung) ist zu erkennen, dass die Information im Vokal scheinbar einen Einfluss auf die Kategorisierung der Kontinua hat. Die Kategoriengrenze für das Kontinuum mit ambigem Vokal ist nach links verschoben, was bedeutet, dass die Versuchspersonen mehr /ɕ/-Antworten gaben als für das Kontinuum,

für welches der Vokal ebenfalls manipuliert wurde. Für letzteres liegt die Kategoriengrenze relativ mittig im Kontinuum zwischen der Stimulusnummer 5 und 6. Für die beiden prosodischen Kontexte, in welchen die Zielwörter präsentiert wurden, kann nur ein kleiner Effekt auf die Kategoriengrenze des /sɛ-ɛɛ/ Kontinuums mit ambigem Vokal gesehen werden. Für dieses Kontinuum ist die Kategoriengrenze in der deakzentuierten Position, im Vergleich zur Kategoriengrenze in der Akzentuierung, ein Stück nach rechts verschoben. Quantifiziert wurden diese Beobachtungen mit einer Varianzanalyse mit Messwiederholungen mit der hörerspezifischen Kategoriengrenze als abhängige Variable, sowie mit Kontinuum (ambiger vs. manipulierter Vokal) und prosodischem Kontext (akzentuiert vs. deakzentuiert) als abhängige Variablen. Für dieses Analyse zeigte nur der Faktor Kontinuum einen signifikanten Effekt auf die Kategoriengrenze ($F[1, 9] = 20.0, p < 0.001$), nicht jedoch der prosodische Kontext. Es wurde keine Interaktion der beiden Faktoren gefunden.

Weiterhin kann man in dem Vergleich der beiden Kontinua in Abb. 3.6 erkennen, dass das Kontinuum, für welches der Vokal ambig (schwarz) gehalten wurde, eine flachere Perzeptionskurve zeigt. Diese Beobachtung wird auch in der nachfolgenden Abbildung (3.7), welche die hörerspezifischen Neigungswerte der psychometrischen Kurven der beiden Kontinua darstellt, bestätigt.

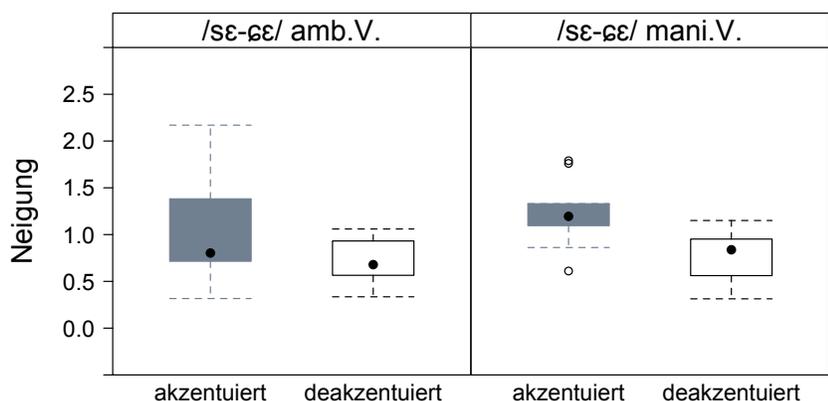


Abbildung 3.7: Hörerspezifische Neigungen der psychometrischen Kurven getrennt für den akzentuierten (grau) und den deakzentuierten Kontext (weiß) für die beiden /sɛ-ɛɛ/ Kontinua: mit ambigem Vokal im linken Panel, /sɛ-ɛɛ/ mit manipuliertem Vokal im rechten Panel. Die Rechtecke zeigen den Interquartilsabstand an, während die Punkte in der Mitte der Boxen die Mediane der Antwortverteilungen angeben.

In Abb. 3.7 sind die Neigungswerte der psychometrischen Kurven getrennt für das Kontinuum zwischen /sɛ-ɕɛ/ mit ambigem Vokal (linkes Panel) und das /sɛ-ʐɛ/-Kontinuum mit manipuliertem Vokal (rechtes Panel), sowie separat für die beiden prosodischen Bedingungen (akzentuiertes Zielwort in grau vs. deakzentuiertes Zielwort in weiß) dargestellt. Es ist zu sehen, dass das Kontinuum mit dem manipulierten Vokal leicht steilere Neigungswerte aufweist als das Kontinuum mit dem ambigem Vokal, was vor allem an den Medianen (schwarze Punkte in den Boxen) zu erkennen ist. Das bedeutet, dass die Versuchspersonen in der Kategorisierung der Sibilanten für das Kontinuum mit fehlenden Informationen im Vokal mehr Unsicherheiten zeigten als für das Kontinuum zwischen dentalem und alveolopalatalem Sibilanten mit manipuliertem Vokal. Für beide Kontinua sind für die deakzentuierte Bedingung, im Vergleich zur Akzentuierung, flachere Neigungen der psychometrischen Kurven zu finden. Diese Beobachtungen wurden mit einer Varianzanalyse mit den hörerspezifischen Neigungen der psychometrischen Kurven als abhängige Variable und mit Kontinuum und prosodischem Kontext als unabhängige Faktoren quantifiziert. Das Ergebnis dieser Analyse brachte einen signifikanten Einfluss des prosodischen Kontexts ($F[1, 9] = 13.0, p < 0.05$) auf die Neigungswerte der psychometrischen Kurven hervor. Zwar zeigte der Faktor Kontinuum keine signifikanten Effekte, dennoch war in der statistischen Analyse ein Trend dahingehend zu erkennen, dass auch das Kontinuum ($p = 0.07$) einen Einfluss auf die Neigungswerte hat. Hinsichtlich der Frage nach der Rolle der Transitionen für die Kategorisierung des /sɛ-ɕɛ/ Kontinuums sind außerdem die Versuchspersonen von Interesse, welche von der Analyse ausgeschlossen wurden, da sie in ihren Antworten eine Neutralisierung des Kontinuums zwischen dentalem und alveolopalatalem Sibilanten mit ambigem Vokal zeigten. Für die anderen beiden Kontinua zeigten diese Versuchspersonen größtenteils eine kategoriale Wahrnehmung der Sibilanten (vgl. Abb. B.1-B.6 in Anhang B.2).

3.4.2 Ergebnisse II: In-/Stabilität des /s ʒ/-Kontrasts

Im nachfolgenden Abschnitt werden die Ergebnisse hinsichtlich der Identifikation und Kategorisierung der Kontinua /sɛ-ʒɛ/ mit ambigem Vokal und /sɛ-çɛ/ mit manipuliertem Vokal dargestellt.

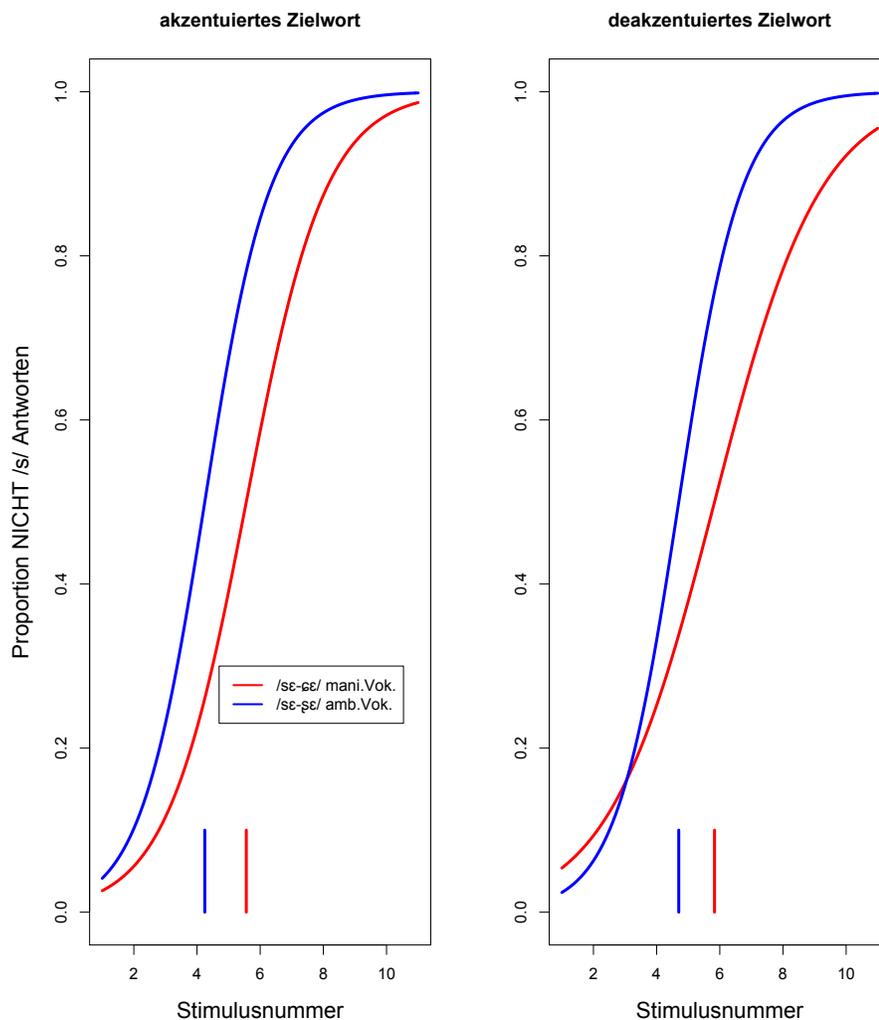


Abbildung 3.8: Psychometrische Kurven der Proportion an NICHT-/s/-Antworten als Funktion der Stimulus Nummer für die Kontexte akzentuiert (links) und deakzentuiert (rechts), separat für die Kontinua /sɛ-ʒɛ/ (blau), /sɛ-çɛ/ (rot) mit manipuliertem Vokal. Die vertikalen Linien am unteren Rand der Abbildung zeigen die Kategoriengrenzen bei denen /s/- und NICHT-/s/-Antworten gleich wahrscheinlich sind.

In Abb. 3.8 sind die psychometrischen Kurven der beiden Kontinua /sɛ-ʒɛ/ (blau) und /sɛ-çɛ/ mit manipuliertem Vokal (rot) in den beiden prosodischen Kontexten (akzentuiert

linkes Panel und deakzentuiert rechtes Panel) dargestellt. Analog zu Abb. 3.6 sind in dieser Abbildung die Proportionen an NICHT-/s/-Antworten (y-Achse) als Funktion der Stimulusnummer (x-Achse) zu sehen. Abb. 3.8 zeigt, dass die Kategoriengrenzen (kurze Linien am unteren Rand der Abbildungen) für das Kontinuum zwischen /sɛ-ʂɛ/ (blau), im Vergleich zum Kontinuum zwischen /sɛ-ʐɛ/ mit manipuliertem Vokal (rot), weiter nach links verschoben sind, was bedeutet, dass die Versuchspersonen für das Dental-Alveolopalatal Kontinuum mehr /s/-Antworten angaben, was auf die so genannte Kompensation für Koartikulation hinweist. Generell besagt diese, dass die Hörer koartikulatorische Effekte aus dem Signal heraus rechnen und der eigentlichen Quelle zuschreiben. Für den hier vorliegenden Fall bedeutet dies, dass im Kontinuum zwischen /sɛ-ʂɛ/ ein Teil der höheren Frequenzen dem hohen Vokal /ɛ/ zugeschrieben wird und nicht den höheren Frequenzen des dentalen Sibilanten. Die Hörer rechnen quasi einen größeren Anteil an hohen Frequenzen dem Vokal zu und weniger dem dentalen Sibilanten selbst, wodurch mehr /ʂ/-Antworten hervorkamen. In deakzentuiertem Kontext, d.h. in der Bedingung, in welcher das Zielwort deakzentuiert präsentiert wurde, bleibt diese Verschiebung des Dental-Retroflex Kontinuums nach links zwar erhalten, jedoch finden sich in der deakzentuierten Bedingung, verglichen mit dem akzentuierten Kontext, für ebendieses Kontinuum mehr /s/-Antworten (zu sehen an der leichten Verschiebung der Kategoriengrenze nach rechts). Für das Kontinuum zwischen dentalem und alveolopalatalem Frikativ (blau) ist so eine Änderung der Entscheidungsgrenze im Vergleich von akzentuiertem und deakzentuiertem Kontext in etwas geringerem Maße zu sehen. Quantifiziert wurden diese Beobachtungen mit einer Varianzanalyse mit der Kategoriengrenze als abhängige Variable, sowie mit prosodischem Kontext (akzentuiert vs. deakzentuiert) und mit Kontinuum (/sɛ-ʂɛ/ vs. /sɛ-ʐɛ/) als unabhängige Faktoren. Die Ergebnisse zeigten, dass nur Kontinuum ($F[1, 9] = 20.0, p < 0.001$) die Kategoriengrenze signifikant beeinflusste, während der prosodische Kontext keinen Effekt auf die Kategoriengrenze hatte. Eine Interaktion zwischen den beiden getesteten Faktoren Kontinuum und Kontext lag ebenfalls nicht vor. Dennoch zeigen die Mittelwerte der hörerspezifischen Kategoriengrenzen einen Unterschied zwischen den Kontinua und den Kontexten: der Sprung in den Entscheidungsgrenzen zwischen deakzentuiertem und akzentuiertem Kontext lag im Kontinuum zwischen /sɛ-ʂɛ/ bei 0.49, während im Vergleich dazu im Kontinuum /sɛ-ʐɛ/ mit manipuliertem Vokal ein Unterschied zwischen den Kategoriengrenzen von nur 0.21 festgestellt werden konnte, was, dem nicht signifikanten Effekt von Kontext auf die Kategoriengrenze zum Trotz, ein Hinweis darauf ist, dass es im Dental-Retroflex Kontinuum einen etwas größeren Effekt des prosodischen Kontexts auf die Entscheidungen der Pro-

banden gibt.

Ein weiterer Aspekt, der in Abb. 3.8 zu erkennen ist, ist dass die psychometrischen Kurven des Kontinuums zwischen dem dentalen und dem alveolopalatalen Sibilanten (mit manipuliertem Vokal) in beiden prosodischen Bedingungen flacher sind als die Kurven des Kontinuums zwischen dem dentalen und dem retroflexen Frikativ. Aus diesem Grund werden in Abb. 3.9 die hörserspezifischen Neigungswerte der psychometrischen Kurven separat betrachtet.

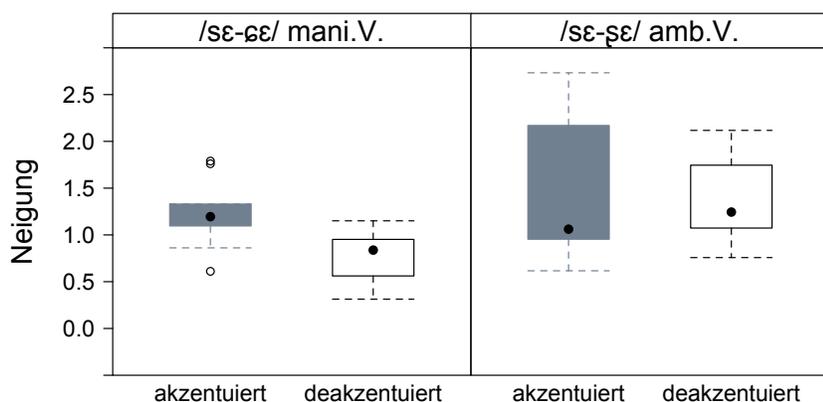


Abbildung 3.9: Hörserspezifische Neigungen der psychometrischen Kurven separat für den akzentuierten (grau) und den deakzentuierten Kontext (weiß) für die Kontinua /sɛ-ʂɛ/ (linkes Panel) und /sɛ-çɛ/ mit manipuliertem Vokal (rechtes Panel). Die Rechtecke zeigen den Interquartilsabstand an, während die Punkte in der Mitte der Boxen die Mediane der Antwortverteilungen angeben.

In Abb. 3.9 sind die hörserspezifischen Neigungswerte der psychometrischen Kurven zu sehen. Ein erster Punkt, welcher hier angemerkt werden soll, ist, dass für das Kontinuum zwischen dem dentalen und dem retroflexen Sibilanten, gefolgt von einem ambigen Vokal, die Ambiguität des Vokals scheinbar keinen Einfluss auf die Identifikation und Kategorisierung dieser zwei Frikative hat. Dieses Kontinuum zeigt, im Vergleich zum Kontinuum zwischen dem dentalen und dem alveolopalatalen Sibilanten mit manipuliertem Vokal, eine Tendenz zu leicht höheren Neigungswerten, was wiederum steileren Perzeptionskurven entspricht. Wie schon im Methodenteil in Abschnitt 3.3.1 aufgrund vorheriger Ergebnisse hinsichtlich der F2-Formantwerten zu Beginn des Vokals, welcher den Sibilanten /s ʂ/ nach-

folgt, beschrieben, ist diese Beobachtung auf die starke Ähnlichkeit der F2-Werte (niedrige Werte für /s ʂ/; hohe F2-Werte für /ʐ/) im Übergang von dentalem beziehungsweise vom retroflexen Sibilanten in den nachfolgenden Vokal zurückzuführen.

Für das Kontinuum zwischen dentalem und alveolopalatalem Sibilanten ist zu sehen, dass die Neigungswerte für die deakzentuierte Bedingung niedriger sind als die des akzentuierten Kontexts. Die Neigungswerte im /sɛ–ʂɛ/-Kontinuum verhalten sich entgegen der erwarteten Tendenz hin zu flacheren Kurven im deakzentuierten Kontext. Für dieses Kontinuum wird fast kein Unterschied in den Neigungswerten zwischen Deakzentuierung und Akzentuierung erkannt. Die Neigungswerte dieses Kontinuums zeigen größere Streuungen in den höheren Wertebereich, was auf eine größere Variation in den Neigungswerten zwischen den Hörern hinweist. Verwunderlich ist weiterhin, dass der Median der deakzentuierten Bedingung einen höheren Wert aufweist als der Median der Neigungswerte in der akzentuierten Position.

Um herauszufinden, ob die Kategorisierung der Stimuli geschlechtsspezifische Muster zeigt, wurde neben dem Einfluss des prosodischen Kontexts, auch die Rolle des Geschlechts der Versuchsteilnehmer mit in die Analyse aufgenommen. Dies geschah vor allem auch aus dem Grund, da in den Ergebnissen zur Produktion der Sibilanten in Kapitel 4 ein Unterschied zwischen den Sibilantproduktionen von männlichen und weiblichen Versuchspersonen gefunden werden konnte. Dieser geschlechtsspezifische Unterschied bestand in Kapitel 4 darin, dass die männlichen Probanden die Sibilantkategorien etwas weniger distinkt produzierten als die weiblichen Versuchspersonen (für mehr Details vgl. Kap. 4). In Abb. 3.10 sind die Neigungen der in Abb. 3.8 dargestellten psychometrischen Kurven getrennt für die beiden Kontinua /sɛ–ʂɛ/ (jeweils rechtes Panel) und /sɛ–ʐɛ/ (letzteres mit manipuliertem Vokal), getrennt für die weiblichen (links und rechts oben) und männlichen (links und rechts unten) Hörer und separat für die akzentuierte (graue Boxen) und deakzentuierte (weiße Boxen) Bedingung zu sehen. Für beide Geschlechter fällt hinsichtlich der Neigungswerte der Kontinua in den beiden prosodischen Kontexten auf, dass meist die akzentuierte Bedingung die steilere Neigung zeigt, während in deakzentuierter Position die Neigungen flacher werden.

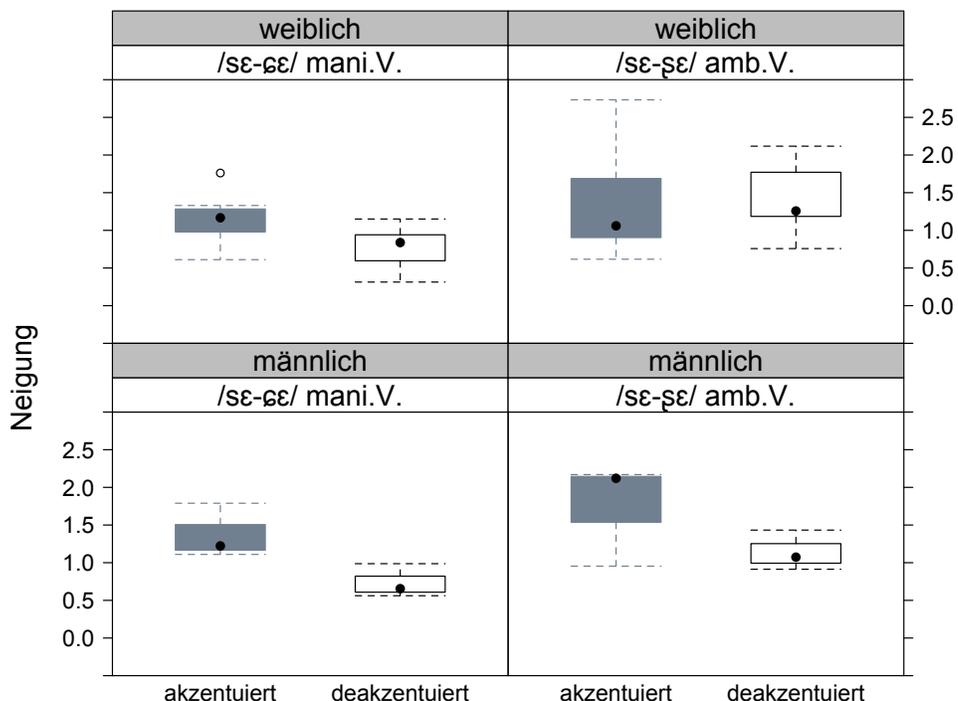


Abbildung 3.10: Hörerspezifische Neigungen der psychometrischen Kurven getrennt für weibliche (obere Reihe) und männliche Hörer (untere Reihe) und separat für den akzentuierten (linke Spalte) und den deakzentuierten Kontext (rechte Spalte) für die Kontinua /sɛ-ʂɛ/ (jeweils linkes Panel) und /sɛ-ɕɛ/ mit manipuliertem Vokal (jeweils rechtes Panel). Die Rechtecke zeigen den Interquartilsabstand an, während die Punkte in der Mitte der Boxen die Mediane der Antwortverteilungen angeben.

Diese Beobachtung trifft jedoch nicht auf die weiblichen Hörer und deren Neigungen für das /sɛ-ʂɛ/-Kontinuum mit manipuliertem Vokal zu. Für dieses Kontinuum ist, in dieser Hörergruppe zu sehen, dass der Median wider Erwarten leicht niedrigere Werte im akzentuierten Kontext, verglichen mit dem deakzentuierten Kontext zeigt, was für eine flachere psychometrische Kurve in der akzentuierten Bedingung spricht. Des Weiteren zeigen die weiblichen Sprecher für das Dental-Retroflex Kontinuum im akzentuierten Kontext die größte Streuung, was auf starke Unterschiede in den Neigungswerten der weiblichen Hörer hinweist. Außerdem ist zu sehen, dass der Abstand zwischen den Neigungswerten der beiden prosodischen Kontexte für die weiblichen Versuchspersonen für beide Kontinua kleiner ist. Für die männlichen Probanden scheinen die Unterschiede zwischen den Neigungswerten der beiden prosodischen Kontexte für das /sɛ-ʂɛ/ und das /sɛ-ɕɛ/-Kontinuum etwas

größer zu sein. Um die Beobachtungen aus den Abschnitten zu den Neigungswerten der psychometrischen Kurven zu quantifizieren wurde eine Varianzanalyse mit den Neigungen der psychometrischen Kurven als abhängige Variable, sowie mit Kontinuum, prosodischem Kontext und Geschlecht als unabhängige Variablen gerechnet. Das Ergebnis zeigte ausschließlich für Kontinuum ($F[1, 8] = 26.3, p < 0.001$) einen signifikanten Effekt auf die Neigungswerte der Perzeptionskurven, nicht aber für die Faktoren prosodischer Kontext und Geschlecht. Eine Interaktion zwischen den getesteten Faktoren wurde ebenfalls nicht gefunden.

Ein letzter Schritt in der Analyse der Perzeptionsdaten der polnischen Hörer war die Subtraktion der hörerspezifischen Neigungen des deakzentuierten Zielwortes von den Neigungen des akzentuierten Zielwortes (vgl. Abb. 3.10). Diese Berechnung diente dazu herauszufinden, ob es für ein Kontinuum und/oder für eine Hörergruppe in der Wahrnehmung schwächere oder auch stärkere Unterschiede in den Neigungswerten zwischen dem deakzentuierten und dem akzentuierten Zielwort gab.

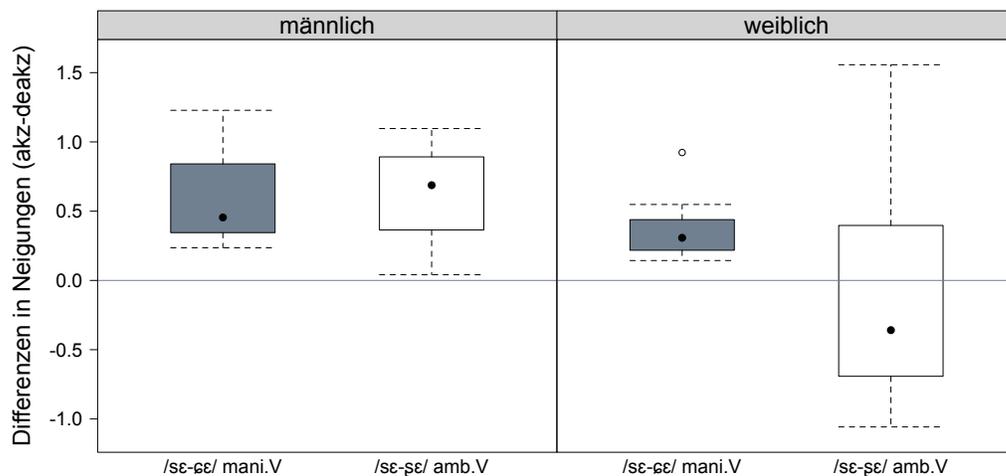


Abbildung 3.11: Differenzen zwischen den hörerspezifischen Neigungswerten des akzentuierten und deakzentuierten Kontexts getrennt für die Kontinua /sɛ-ʂɛ/ (weiße Boxen) und /sɛ-ʐɛ/ (graue Boxen) und separat für die männliche (linkes Panel) und weibliche (rechtes Panel) Versuchspersonengruppe. Der Punkt in der Mitte der Boxen gibt den Median der jeweiligen Neigungswerte an.

Betrachtet man die Mediane (schwarze Punkte in den Boxen) in Abb. 3.11, so ist zu erkennen, dass für die männlichen Versuchspersonen für beide Kontinua die Differenzwerte

zwischen den Neigungswerten der psychometrischen Kurven im positiven Wertebereich liegen. Dies bedeutet dass in dieser Versuchspersonengruppe die akzentuierte Bedingung steilere Neigungswerte aufwies als die deakzentuierte Bedingung. Ferner ist für die männlichen Probanden an den etwas höheren Medianwerten eine leichte Tendenz hin zu höheren Differenzwerten im potentiell instabilen Kontinuum zwischen /sɛ-ʂɛ/ zu sehen. Für das scheinbar stabile Kontinuum zwischen /sɛ-ɕɛ/ liegt der Median der Differenzwerte etwas tiefer, was für einen etwas geringeren Effekt des prosodischen Kontexts auf die Perzeption dieses Kontinuums spricht. Die Kontinua der weiblichen Versuchspersonengruppe weichen in den Differenzen ihrer Neigungswerte stark voneinander ab. Die Differenz zwischen den Neigungen der akzentuierten und deakzentuierten Bedingung liegt für das dental-alveolopalatal Kontinuum (grau) im positiven Wertebereich. Für dieses Kontinuum waren, vergleichbar mit den Differenzwerten der männlichen Probanden, die Neigungswerte des akzentuierten Kontexts höher als die der deakzentuierten Bedingung. Für dieses Kontinuum ist dennoch ein Unterschied zwischen den Geschlechtern festzustellen: die weiblichen Hörer haben niedrigere Differenzwerte als die männlichen Probanden, was auf einen weniger großen Einfluss des prosodischen Kontexts auf die Kategorisierung dieses Kontinuums für die weiblichen Hörer hinweist. Für das Kontinuum zwischen /sɛ-ʂɛ/ mit ambigem Vokal (weiß) zeigen die weiblichen Probanden ein weniger konsistentes Muster. Für die weibliche Versuchspersonengruppe liegen einige Differenzwerte im negativen Bereich, was anzeigt, dass hier einige der Neigungen ihrer psychometrischen Kurven für das akzentuierte Zielwort flacher waren als für den deakzentuierten Kontext. Diese Beobachtungen wurden mit einer Varianzanalyse mit der Differenz in den Neigungswerten zwischen akzentuierter und deakzentuierter Bedingung als abhängige Variable und mit den unabhängigen Faktoren Geschlecht und Kontinuum quantifiziert. Diese Analyse brachte jedoch für keinen der Faktoren einen signifikanten Einfluss auf die Differenzwerte der Neigungen hervor.

3.5 Diskussion

Diese Studie befasste sich mit mehreren perzeptiven Aspekten der drei Sibilanten im Standard-Polnischen. Die erste Forschungsfrage, welche hier fokussiert wurde, war die Notwendigkeit der Informationen im Vokal /ɛ/, welcher den Sibilanten nachfolgte, für die Kategorisierung der Sibilanten /s ɕ/. Die zweite und zentralere Frage dieser Studie befasste sich mit der Neutralisierung des /s ʂ/-Kontrasts, welche für einige dialektale Varietäten des Polnischen bestätigt wurde (Nowak, 2006a). Demnach lag das Hauptaugenmerk der

Studie auf der synchronen Instabilität des /s ʂ/-Kontrasts, verglichen mit dem scheinbar stabileren /s ʐ/-Kontrast. Dieser Punkt wurde in die beiden nachfolgenden Teilfragen unterteilt: (1) Spielt für die Kontrasterhaltung beziehungsweise für die Kontrastneutralisierung zwischen den Sibilanten /s ʂ/ und zwischen /s ʐ/ der prosodische Kontext eine Rolle, in welchem die Laute vorkommen?, und (2) Inwiefern wird die Perception des Kontrasts durch das Geschlecht der Hörer beeinflusst? Zur Klärung der oben genannten Fragen wurden Perzeptionsdaten von 10 L1-Polnisch sprechenden Versuchspersonen (sieben weiblich, drei männlich) präsentiert.

Hinsichtlich der ersten Forschungsfrage, nämlich inwiefern die Informationen im Vokal, welcher den Sibilanten folgt, notwendig für eine Kategorisierung der Kontinua zwischen den Sibilanten /s ʐ/ ist, kann gesagt werden, dass die Informationen im Vokal einen starken Einfluss auf die Identifikation des Dental-Alveopalatal Kontrasts haben. Dies konnte in den Abb. 3.6 und 3.7 gesehen werden, welche die flachere psychometrische Kurve für das Kontinuum zwischen /sɛ-ʐɛ/ mit ambigem Vokal zeigten. Im Gegensatz dazu war im Kontinuum zwischen dem dentalen und alveopalatalen Sibilanten, in welchem die Vokalinformationen ebenfalls manipuliert wurden, eine kategorialere Perzeptionskurve mit steileren Neigung zu erkennen. Weiterhin kann festgehalten werden, dass für das /sɛ-ʂɛ/ Kontinuum beziehungsweise für die Unterscheidung des dentalen und des retroflexen Sibilanten die Informationen im Vokal nur wenig bis keinen Effekt haben. Für dieses Kontinuum wurde der Vokal bei der Erstellung der Stimuli ambig gehalten (siehe Abschnitt 3.3.1), da in Kapitel 2 gezeigt wurde, dass die Transitionswerte des zweiten Formanten zu Beginn des Vokals, welcher diesen beiden Lauten nachfolgte, sehr starke Ähnlichkeiten aufweisen (weitere Evidenzen bezüglich der Ähnlichkeit der F2-Transitionen von /s ʂ/ in den nachfolgenden Vokal hinein, siehe Nowak, 2006 und F. Li, 2008). Dieser fehlenden Information im Vokal zum Trotz konnte für dieses Kontinuum eine kategoriale Wahrnehmung des Kontrasts gefunden werden. Dabei wies dieses Kontinuum sogar etwas steilere Neigungen verglichen mit dem /sɛ-ʐɛ/-Kontinuum mit manipuliertem Vokal auf (vgl. Abb. 3.8).

In Bezug auf die Instabilität des Dental-Retroflex Kontrasts und die vermeintliche Stabilität des Dental-Alveopalatal Kontrasts, wurden die Kategoriengrenzen und die Neigungen dieser beider Kontinua verglichen. Die Kategoriengrenze für das /sɛ-ʐɛ/ Kontinuum war, im Vergleich zum /sɛ-ʐɛ/ Kontinuum (mit manipuliertem Vokal), nach links verschoben, was bedeutet, dass die Versuchspersonen weniger Stimuli der dentalen Kategorie zuordneten. Die Versuchspersonen kompensierten scheinbar im Kontinuum zwischen

dentalem und retroflexem Sibilanten in der akzentuierten Bedingung für die koartikulatorischen Einflüsse des Vokals. Genauer gesagt gaben die Hörer in diesem Kontinuum mehr /ʃ/-Antworten, da sie die steigenden Schwerpunktfrequenzen der sibilantischen Sequenzen, nicht dem Sibilanten selbst, sondern dem Einfluss des Vokals zuschrieben. Für das /sɛ-çɛ/ Kontinuum mit manipuliertem Vokal, lag die Kategoriengrenze relativ genau bei der Hälfte des 11-Schritte-Kontinuums – zwischen Stimulusnummer fünf und sechs (vgl. Abb. 3.8). Diese mittige (und im Vergleich zum /sɛ-ʃɛ/ Kontinuum nach rechts verschobene) Lage der Kategoriengrenze für das Dental-Alveolopalatal Kontinuum ist ebenfalls durch mögliche Effekte der Koartikulation zu erklären. Die Hörer wissen, dass der alveolopalatale Sibilant einen starken koartikulatorischen Einfluss auf den nachfolgenden Vokal hat, d.h. hohe F2-Werte im Vokal im Anschluss an /ç/, während für den dentalen Laut /s/ niedrige Transitionswerte des zweiten Formanten gefunden werden können. Lassen die Versuchspersonen dieses Wissen mit in die Perzeption einfließen, so besteht die Möglichkeit die vermehrten /s/-Antworten für das /sɛ-çɛ/ Kontinuum dahingehend zu erklären, dass die Hörer eben die höheren Frequenzen im Vokal nicht mehr nur dem Vokal anrechnen, sondern dem alveolopalatalen Sibilanten (eben in dem Wissen, dass dieser die Formanten der umliegenden Vokale stark beeinflusst), sodass im Vergleich mit dem /sɛ-ʃɛ/ Kontinuum mehr Stimuli des 11-Schritte Kontinuums als dentaler Sibilant kategorisiert werden. Für die deakzentuierte Bedingung ist in Abb. 3.8 zu sehen, dass sich die Kategoriengrenze für das /sɛ-ʃɛ/ Kontinuum nach rechts, d.h. hin zu mehr /s/-Antworten verschiebt, was möglicherweise auf eine reduzierte Kompensation für Koartikulation in der deakzentuierten Bedingung hinweisen könnte. Für das Dental-Alveolopalatal Kontinuum ist ebenfalls eine leichte Bewegung der Kategoriengrenze nach rechts zu sehen, jedoch nicht in gleichem Maße, wie es für das Dental-Retroflex Kontinuum der Fall ist.

Beim Betrachten der Neigungswerte der psychometrischen Kurven zeigte sich in Abb. 3.9, dass die Neigungswerte für das Dental-Retroflexe Kontinuum, im Vergleich zu den anderen Kontinua, in beiden prosodischen Kontexten steiler sind. Außerdem sind die psychometrischen Kurven im deakzentuierten Kontext meist flacher als die der akzentuierten Bedingung. Eine Ausnahme besteht hier für die weiblichen Hörer, die für das /sɛ-ʃɛ/ Kontinuum in der deakzentuierten Bedingung im Vergleich zum akzentuierten Kontext teilweise leicht steilere Neigungswerte zeigten.

Hinsichtlich der Frage nach der Rolle der Geschlechter bei der Neutralisierung des Lautkontrasts zwischen /s ʃ/ wurde in Abb. 3.11 eine leichte Tendenz dahingehend beobachtet,

dass die männlichen Versuchspersonen einen größeren Unterschied zwischen den Neigungswerten der akzentuierten und der deakzentuierten Bedingung aufzeigen als die weiblichen Versuchspersonen. Die größte Differenz zwischen den Neigungswerten von Akzentuierung und Deakzentuierung konnte für das /sɛ–ɕɛ/-Kontinuum der männlichen Hörer gefunden werden. Die statistischen Analysen der Daten unterstütze diese letzte Beobachtung zu den Abbildungen 3.9 und 3.10 jedoch nicht.

Die grundsätzlich steileren Neigungswerte der weiblichen Versuchspersonen, welche vermeintlich auf eine etwas kategorialere Wahrnehmung der Sibilanten bei den Hörerinnen schließen lässt, passen mit den Produktionsdaten der L1-Polnisch sprechenden Versuchspersonen in Kapitel 4 zusammen. Für die Sprachproduktionsdaten, welche geschlechtsspezifisch analysiert wurden, konnte gezeigt werden, dass die weiblichen Versuchspersonen generell Sibilantkategorien produzierten, die eine etwas größere akustische Trennung zeigten als die der männlichen L1-Polnisch Sprecher. Die männlichen Probanden hingegen, produzierten Sibilanten, welche, verglichen mit den weiblichen Sprechern, in den untersuchten akustischen Parametern (DCT-Koeffizienten[1, 2] zum zeitlichen Mittelpunkt der Sibilanten, sowie Werte des zweiten und dritten Formanten zum zeitlichen Vokalonsset) etwas weniger stark divergierten. Verknüpfen wir die Produktionsdaten mit der hier vorliegenden Analyse der Sibilantperzeption, ist es möglich zu sagen, dass Frauen durch ihre vermutlich periphereren Kategorien in der Produktion, die Sibilanten ebenfalls distinktiver wahrnahmen.

Betrachtet man die Ergebnisse der Perzeptionsstudie vor dem Hintergrund der Neutralisierung des Lautkontrasts zwischen /s ʂ/ gibt es scheinbar nur in der Analyse der Kategoriengrenzen eine Andeutung auf eine größere Instabilität des Dental-Retroflex Kontrasts. Die Kategoriengrenzen zeigen für das /sɛ–ɕɛ/ Kontinuum eine etwas größere Verschiebung der Kategoriengrenze zwischen akzentuiertem und deakzentuiertem Kontext, verglichen mit dem /sɛ–ɕɛ/ Kontinuum mit manipuliertem Vokal. Dabei verschiebt sich die Kategoriengrenze auf eine Weise, die Hinweise auf einen möglichen Einfluss einer reduzierten Kompensation für Koartikulation in deakzentuierter Bedingung in Bezug auf einen potentiellen Merger des /s ʂ/-Kontrasts zulässt¹. Harrington et al. (2008, 2013, 2015) zeigten

¹Um eine tatsächliche reduzierte Kompensation für Koartikulation im deakzentuierten Kontext zu testen, ist es jedoch notwendig ein Perceptionsexperiment durchzuführen, in welchem die Hörer die sibilantischen Kontinua ebenfalls mit nachfolgendem Vokal /o/ kategorisieren müssen. Hier würden wir erwarten, dass in akzentuiertem Kontext für das /s–ʂ/ Kontinuum mehr /s/-Antworten gegeben werden würden, was der Kompensation für Koartikulation entspräche. In der deakzentuierten Bedingung sollte im /o/ Vokalkontext eine leichte Verschiebung hin zu mehr /ʂ/-Antworten zu erkennen sein. Der Abstand zwischen

Möglichkeiten auf, wie es in verschiedenen Altersgruppen und auch in prosodisch schwachen Konstituenten zu einem geringeren Maß an Kompensation für Koartikulation im Hörer kommen kann, was wiederum Hinweise auf Lautwandelprozesse bringen kann. Zum einen konnte hinsichtlich der /u/-Frontierung im Britischen Englisch weniger Kompensation für Koartikulation für die jüngere, und im Lautwandel fortgeschrittenere Generation, im Vergleich zu den älteren Versuchspersonen gefunden werden (Harrington et al., 2008). Und zum anderen konnte gezeigt werden, dass die Kompensation für Koartikulation in prosodisch schwachen Kontexten im Vergleich zur akzentuierten Bedingung reduziert werden kann (Harrington et al., 2013, 2015). Vor dem Hintergrund dieser Studien wäre es möglich eine Erklärung für die Neutralisierung des Dental-Retroflex Kontrasts in der möglicherweise geringeren Kompensation für Koartikulation in deakzentuierter Position zu finden. Harrington et al. (2013) beschrieben die Möglichkeit eines Lautwandels infolge schwächerer Kompensation in deakzentuiertem Kontext wie folgt: für den Fall, dass die Kompensation für Koartikulation in einem unbetonten Kontext geschwächt ist, jedoch in der Produktion das Maß an Koartikulation in akzentuierter und deakzentuierter Bedingung unverändert bleibt, könnte es dazu kommen, dass die Koartikulation nicht mehr mit ihrer Quelle, sondern mit dem von der Koartikulation betroffenen Laut in Verbindung gebracht wird. In den vorliegenden Daten würde das bedeuten, dass die höheren Frequenzen nicht mehr dem Einfluss des Vokals /ε/, sondern eher dem Sibilanten an sich zugeschrieben werden würden. Diese Gedanken passen in das Modell, welches Lautwandel durch fehlerhafte Perzeption beschreibt (Ohala, 1993, 2012). Doch ist es möglich, dass das Ergebnis trotzdem etwas von Ohala's (1993; 2012) originalem Modell zu Lautwandel abweicht, und zwar in die von Harrington et al. (2013) beschriebene Richtung. Harrington und Kollegen (2013) fanden in ihrer Studie zur polysyllabischen Kürzung in der Sprachproduktion kaum einen Einfluss der prosodischen Gegebenheiten auf die Koartikulation, während jedoch in der Perzeption die Kategorisierung des /a-a:/-Kontinuums Unterschiede in betonter und unbetonter Position aufwies. Daraus wurde geschlossen, dass in einem möglichen Lautwandelprozess das Verhältnis zwischen Hörer und Sprecher nicht immer im Vordergrund stehen muss, sondern, dass auch die Kontexteffekte für die Variabilität von Lauten vordergründig sein können. Kompatibel mit diesen Überlegungen wurde in Kapitel 2 dieser Arbeit in den akustischen und artikulatorischen Daten zu den polnischen Sibilanten, weder ein Effekt der Sprechge-

den beiden Perzeptionskurven in den verschiedenen Vokalkontexten würde dadurch minimiert werden, was dann wiederum auf eine weniger starke Kompensation für Koartikulation hinweisen könnte. In der deakzentuierten Bedingung wären die Hörer unsicherer hinsichtlich der Kompensation für Koartikulation. Dieses hypothetisch aufgeführte Ergebnis eines weiteren Perzeptionsexperiment müsste verifiziert werden, um den möglichen Effekt der reduzierten Kompensation für Koartikulation zu bestätigen.

schwindigkeit gefunden, noch konnte eine auffallende Überlappung der /s ʂ/-Kategorien berichtet werden. Artikulatorisch wurde in der vorangehenden physiologischen Studie damit grundsätzlich eine gute Trennung der drei Sibilanten festgestellt. In der Perception jedoch konnte eine Variabilität der Kategoriengrenze hinsichtlich der beiden prosodischen Kontexte für das Dental-Retroflex Kontinuum gefunden werden. Für das /s-ʐ/ Kontinuum wurde eine Variation der Kategoriengrenze nur in geringerem Maß beobachtet.

In der Analyse der Neigungen der psychometrischen Kurven wurde beobachtet, dass in den meisten Fällen die Neigungen des deakzentuierten, verglichen mit den Neigungswerten des akzentuierten Kontexts, flacher waren. Eine Ausnahme bildeten die Neigungswerte der weiblichen Versuchspersonen im /sɛ-ʂɛ/ Kontinuum, die in deakzentuierter Position steiler waren als für die akzentuierte Bedingung. Diese Beobachtung spricht für generell mehr Unsicherheit der Hörer in der Kategorisierung der Sibilantkategorien in der deakzentuierten Bedingung. Jedoch konnte keine weitere Tendenz zu flacheren Neigungen im potentiell instabilen /sɛ-ʂɛ/ Kontinuum verglichen mit dem scheinbar stabilen /sɛ-ʐɛ/ Kontinuum gefunden werden. Eher war der umgekehrte Fall im Vergleich von de- und akzentuierter Bedingung zu beobachten, nämlich generell flachere psychometrische Kurven für das /sɛ-ʐɛ/ Kontinuum mit manipuliertem Vokal, verglichen mit den grundsätzlich steileren Kurven für /sɛ-ʂɛ/. Dieses unerwartete Ergebnis könnte mit Vermutungen aus Iskarous et al. (2011) hinsichtlich der perzeptiven Stabilität des /s ʃ/-Kontrasts trotz einer Tendenz zur Neutralisierung dieses Kontrasts in Sibilant-Konsonant-Klustern mit Erhalt des postalveolaren Frikativs /ʃ/ in einigen Varietäten des Englischen (weitere Details zur /s/-retraction siehe Baker et al., 2011 und Stevens & Harrington, 2016) beschrieben werden. Iskarous et al. (2011) merken an, dass das Maß an Variation im ersten spektralen Moment für den /s ʃ/-Kontrast – trotz häufig beobachtbarer Neutralisierung des /s ʃ/ Kontrasts in sCV-Sequenzen in der Produktion – nicht groß genug ist, um einen Effekt auf die Kategorisierung des Kontrasts in der Perception zu haben. Für den /s ʂ/-Kontrast im Polnischen, dessen Neutralisierung in der Standardsprache des Polnischen quasi nicht zu beobachten ist, ist die Möglichkeit einen Effekt auf die Neigungswerte der Perzeptionskurven zu finden demnach ebenfalls als sehr gering zu betrachten. Das fehlende Maß an Variabilität und die Deutlichkeit des Kontrasts in den Frequenzschwerpunkten im Friktionsanteil – hoher Schwerpunkt für /s/ und sehr niedrige Werte für /ʂ/ – würden in diesem Fall die steilen Kurven für das /sɛ-ʂɛ/ Kontinuum erklären. Weiter brachten die Ergebnisse zur Perception der drei Sibilanten /s ʂ ʐ/ bei Mandarin-Chinesischen und Taiwan Mandarin Hörern, wobei letztere Varietät das Merkmal der Neutralisierung des Dental-Retroflex

Kontrasts aufweist, in Chang et al. (2013) hervor, dass der /s ʃ/-Kontrast sowohl von den Hörern der Standardsprache (Mandarin-Chinesisch) als auch von den Hörern der dialektalen Varietät (Taiwan Mandarin), kategorial wahrgenommen wurde. Niedzielski (1999) zeigte, dass die Erwartungen der Hörer einen Einfluss auf ihre Perzeption haben können. Im Falle der vorliegenden Studie wäre es somit denkbar, dass die Versuchspersonen eine Standardsprachliche Ausspracheform und damit eine /s ʃ/-Kontrasterhaltung erwarteten und damit eine gute Kategorisierung des potentiell instabilen Kontrasts zeigten. Im Gegensatz dazu besteht die Möglichkeit, dass die flacheren Neigungen der psychometrischen Kurven im /s-ɛ/ Kontinuum durch zwei Aspekte zustande kamen. Zum einen wurden im Falle dieses Kontinuums im Vergleich zum /s-ʃ/ Kontinuum zwei Merkmale manipuliert – der Frequenzschwerpunkt im Sibilanten und die Formanten im Vokal /ɛ/, welcher den Sibilanten folgte. Auch wenn sich zeigte, dass für den Dental-Alveolopalatal Kontrast die Manipulation des Vokals notwendig ist, um die Sibilanten kategorisieren zu können (siehe Ergebnisse zu den Perzeptionskurven des /sɛ-ɛɛ/ Kontinuums mit ambigem Vokal vs. mit manipuliertem Vokal in Abb. 3.6 und Abb. 3.7), so lässt dieses Kontinuum im Vergleich zum /sɛ-ʃɛ/ Kontinuum, für welches nur der sibilantische Frequenzschwerpunkt manipuliert wurde, durch die zwei manipulierten Parameter eventuell mehr perzeptive Variabilität zu. Auch könnten die flacheren Kurven im Dental-Alveolopalatal Kontinuum hinsichtlich der Ergebnisse aus Lee-Kim's (2015) Studie zur Wahrnehmung des alveolopalatalen und des retroflexen Sibilanten im Mandarin-Chinesischen erklärt werden. In ihrer Studie rief der Vokal /ɛ/ auch mit herausgeschnittenen Transitionen eine Tendenz zur /ɛ/-Wahrnehmung bei Hörern hervor. Trotz der herausgeschnittenen Transitionen konnten erst durch die zusätzlichen Informationen im Sibilanten mehr /ʃ/-Antworten elizitiert werden. Chiu (2009) fand, dass der polnische Sibilant /s/ mit nachfolgendem Vokal /a/, welcher aus einer /ca/ Sequenz extrahiert und an den dentalen Laut angefügt wurde, von Taiwan Mandarin Probanden relativ schwer der Kategorie /s/ zugeordnet werden konnte, da die Transitionen einen sehr starken Einfluss auf die Wahrnehmung des vorangehenden Sibilanten hatten. In der hier vorliegenden Studie sollte mitunter getestet werden, inwiefern die polnischen Hörer zur Wahrnehmung beziehungsweise zur Kategorisierung des Dental-Alveolopalatal Kontrasts die Informationen in den Transitionen des nachfolgenden Vokals benötigen. Es wurde zwar festgestellt, dass die /s ɛ/-Kategorisierung durch passend manipulierte Informationen im Vokal erheblich verbessert wurde. Jedoch könnte trotzdem durch die perzeptiven Parameter, welche in zwei verschiedene Richtungen wiesen, die Möglichkeit zur größeren Unsicherheit in der Kategorisierung des Kontinuums /sɛ-ɛɛ/ mit manipuliertem Vokal im

direkten Vergleich mit dem /sɛ–ʂɛ/ Kontinuum bei den Hörern bestanden haben: der Vokal /ɛ/ an sich, welcher für /ɕ/ Antworten spricht und niedriger werdende Frequenzen im Frikationsanteil der CV-Sequenzen, die die Wahrscheinlichkeit für ein dentales /s/ erhöhen. Es könnten diese beiden entgegengesetzten perceptiven Cues gewesen sein, welche die Antworten der Versuchspersonen unsicherer beziehungsweise weniger kategorial werden ließen. Vor allem in der Gegenüberstellung mit dem Kontinuum zwischen /s ʂ/, in welchem die Informationen im nachfolgenden Vokal ambig gehalten wurden, sodass die Hörer hier nur einen variierenden perceptiven Cue – niedriger werdende Schwerpunktfrequenzen von /s/ nach /ʂ/ – zur Verfügung hatten, anhand welchem sie die Stimuli den zwei sibilantischen Kategorien zuordnen sollten. Weiter wurde berichtet, dass der alveolopalatale Sibilant einen Frequenzschwerpunkt aufweist, welcher zwischen dem Wert des dentalen und dem niedrigen Wert des retroflexen Frikativs liegt (F. Li 2008; Nowak 2006a; Żygis und Hamann 2003). Dabei merkte auch Flemming (2002) an, dass der retroflexe Sibilant, verglichen mit dem alveolopalatalen Laut, durch die involvierte Lippenrundung seinen Schwerpunkt in niedrigeren Frequenzen zeigt. Dieser daraus resultierende kleinere Unterschied in den Schwerpunktfrequenzen zwischen /s ɕ/ gegenüber der stärkeren Distinktion zwischen den hohen Frequenzen des dentalen Lautes /s/ und den sehr niedrigen im /ʂ/, kann ebenfalls einen Einfluss auf die Kategorisierung der beiden Kontinua gehabt haben. Es könnte sein, dass die sehr gute perceptiv Kategorisierung des scheinbar instabilen Kontrasts zwischen dem dentalen und dem retroflexen Sibilanten für das Standard-Polnische durch eine Art *auditory enhancement* (Kingston & Diehl, 1994, 1995) sichergestellt wurde: ein größerer Abstand zwischen den Kategorien durch zusätzliche Lippenrundung im retroflexen Sibilanten, was wiederum mit niedrigeren Frequenzschwerpunkten für den Sibilanten /ʂ/ und einer größeren akustischen Distanz zu /s/ verbunden ist.

Die größere Verschiebung der Kategoriengrenze bei weiterhin steileren Neigungen für das vermeintlich instabile Kontinuum vom akzentuierten hin zum deakzentuierten Kontext, d.h. mehr /s/-Antworten in der deakzentuierten, verglichen mit mehr /ʂ/-Antworten in der akzentuierten Bedingung, und der damit verbundenen Wahrscheinlichkeit eines Lautwandels, der durch einen Zusammenfall des /s ʂ/-Kontrasts zugunsten des dentalen Sibilanten gekennzeichnet ist, wäre ebenfalls kompatibel mit den Beschreibungen von Lautwandelprozessen durch die Exemplar Theorie (Johnson, 1997; Pierrehumbert, 2001, 2002). In dieser Theorie wird angenommen, dass der Hörer die wahrgenommenen Objekte (auf Lexem- und Segment-Ebene) mit all ihren feinen phonetischen Details in einer Art Exemplar-Wolke

im mentalen Lexikon speichert. Ist der Hörer neuen, veränderten Exemplaren ausgesetzt, werden diese ebenfalls seiner Exemplar-Wolke zugefügt, was dann zu einer langsamen Verschiebung der jeweiligen Lautkategorie führen kann. Betrachten wir die Verschiebung der Kategoriengrenze hin zu mehr /s/-Antworten in der deakzentuierten Bedingung so könnte die Möglichkeit bestehen, dass die Exemplar-Wolke der Hörer, welche die Kategorie /s/ darstellt, immer mehr Varianten mit leicht gesenktem Frequenzschwerpunkt in einem deakzentuierten Kontext aufnimmt, sodass der Hörer in der deakzentuierten Positionen diese, hin zu niedrigeren Frequenzen, veränderten Exemplare ebenfalls als gute /s/ Varianten akzeptiert. Die Wolke für den Sibilanten /ʃ/ hingegen wird Schritt für Schritt kleiner, da mehr und mehr Exemplare mit etwas niedrigeren Frequenzen der dentalen Kategorie zugeschrieben werden. Hypothetisch ist eine solche Verschiebung in Abb. 3.12 dargestellt. Die grauen Kreise zeigen die Exemplar-Wolken von /ʃ/ (links) und /s/ (rechts) in akzentuierter Bedingung an. Die Sibilanten sind für die Standard-Polnischen Hörer klar kategorial voneinander getrennt. In den grünen Kreisen sind die beiden Kategorien in deakzentuiertem Kontext dargestellt. Hier werden mehr Exemplare mit niedrigeren Frequenzen, welche in der Akzentuierung der /ʃ/-Kategorie zugeordnet wurden, der /s/ Exemplar-Wolke zugeschrieben. Die /s/-Wolke wird größer, während die /ʃ/-Exemplar-Wolke etwas abnimmt. Durch diesen Zuwachs der /s/-Kategorie und die Reduzierung der Exemplare in der /ʃ/-Wolke und dem zeitgleichen Beibehalten der perceptiven Distanz zwischen den Kategorien, könnte eine kategoriale Wahrnehmung der beiden Sibilanten im Hörer beibehalten werden, ohne dass Unsicherheiten in der Kategorisierung aufkämen, welche die Neigungen der Perzeptionskurven hin zu flacheren Kurven in deakzentuiertem Kontext verändern würden. Im weiteren Verlauf würde die /ʃ/-Wolke Stück für Stück kleiner werden, während in der Sprachproduktion die Variabilität der /s/-Exemplare möglicherweise zunehmen und der generelle Frequenzschwerpunkt der dentalen Kategorie ein Stück gesenkt würde.

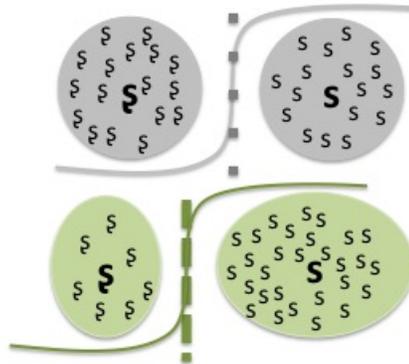


Abbildung 3.12: Hypothetische Verteilung und Verschiebung der Sibilantkategorien /s ʂ/ im Sinne der Exemplartheorie. Graue Kreise stellen die kategoriale Wahrnehmung der Sibilanten in akzentuierter Position dar. Grüne Kreise zeigen die deakzentuierte Bedingung mit vermehrten /s/ Antworten an. Der Abstand der beiden Kategorien bleibt in de- und in akzentuierter Position gleich, sodass zwar die Kategoriengrenzen verschoben werden, jedoch der Kontrast an sich nicht verändert wird und die Perzeptionskurven nicht flacher werden.

Labov's (1990, 2001) sozio-phonetische Studie zur Geschlechterrolle im Verlauf von Lautwandel brachte eine Asymmetrie in bestimmten Lautwandelprozessen zwischen Frauen und Männern hervor. Während Frauen scheinbar eher Lautwandel *from below*, d.h. durch interne phonetische Faktoren herbeigeführte Veränderungen der Sprache mit ihren neuen Aussprachevarianten annehmen und verbreiten, ist die Sprache der männlichen Mitglieder der Sprachgemeinschaft oftmals eher von traditionelleren, dialektal gefärbten Aussprachemustern (Milroy & Milroy, 1985) betroffen. In dem Fall der hier vorliegenden Studie wäre die Hypothese zur Geschlechterrolle für eine mögliche /s ʂ/-Neutralisierung im Polnischen, dass für die männlichen Versuchspersonen die weniger kategoriale Wahrnehmung gefunden

werden sollte. In Abb. 3.10 ist für das /sɛ–ʂɛ/-Kontinuum zu sehen, dass die männlichen Versuchspersonen in der akzentuierten Bedingung steilere Kurven und flachere psychometrische Kurven in der Deakzentuierung zeigten als die weiblichen Hörer, welche, wie an den Differenzwerten der Neigungen in Abb. 3.11 zu sehen, teilweise sogar höhere Neigungswerte für das Dental-Retroflex Kontinuum in der Deakzentuierung aufwiesen. Es wäre denkbar, dass dieser Unterschied in den Neigungswerten der weiblichen und männlichen Hörer mit den Produktionsdaten aus Kapitel 4 in Zusammenhang steht. Wie schon in einem der vorangehenden Absätze erwähnt, zeigten die weiblichen Sprecherinnen in den Produktionsdaten der Sibilanten eine größere Trennung der Kategorien /s ʂ ɕ/. Die männlichen Sprecher zeigten jedoch Kategorien, die etwas weniger unterschiedlich waren als die der polnischen Frauen. Wodurch die stärkere Distinktion der Kategorien in den weiblichen Sprecherinnen möglicherweise verantwortlich für die steileren Neigungen dieser Versuchspersonengruppe in der Perzeption ist.

Zusammenfassend kann bezüglich der hier vorliegenden Ergebnisse der Perzeptionsstudie gesagt werden, dass die Informationen in den Transitionen im post-sibilantischen Vokal, für die /s ɕ/-Kategorisierung notwendig sind. Dieses Ergebnis geht mit Studien zu den Mandarin-Chinesischen (Chiu, 2009; Lee-Kim, 2015) und den Sibilanten des Polnischen (Nowak, 2006a) einher, welche zeigten, dass die Nutzung der perzeptiven Cues zur Sibilantunterscheidung abhängig von der jeweiligen Sprache (Wagner et al., 2006) ist. Weiterhin waren die Ergebnisse hinsichtlich der Bedingungen, welche einen Zusammenfall des /s ʂ/-Kontrasts herbeiführen könnten, weniger deutlich. Es konnte in den Kategorien Grenzen für das potentiell instabile /s–ʂ/ Kontinuum mehr Variation gefunden werden als für das /s–ɕ/ Kontinuum. Diese Verschiebung bringt Hinweise auf Lautwandel-Theorien, welche die fehlerhafte Interpretation des Sprachsignals durch den Hörer größtenteils für Lautwandel verantwortlich machen. Auch könnten die Ergebnisse im Sinne der Exemplar-Theorie interpretiert werden, welche einen potentiellen /s ʂ/ Merger durch eine Verschiebung der Kategorisierung der /s ʂ/-Varianten in prosodisch schwacher Position erklären würde. Außerdem konnten meist niedrigere Neigungen für die Stimuli beobachtet werden, welche den Hörern im deakzentuierten Kontext präsentiert wurden – verglichen mit den steileren Kurven für die Stimuli in akzentuierter Position. Diese Ergebnisse gehen einher mit Studien zur größeren Variabilität und zur stärkeren Reduktion der Sprachproduktion und damit auch der Perzeption in hypoartikulierter Sprache (Lindblom, 1990). Darüber hinaus konnte die Analyse der Neigungswerte der psychometrischen Kurven jedoch keine

weiteren Tendenzen hinsichtlich einer Neutralisierung des Dental-Retroflex Kontrasts bei Hörern des Standard-Polnischen aufzeigen. Der Vergleich von männlichen und weiblichen Hörern ließ vermuten, dass die weniger distinkten Kategorien von Männern in der Produktion, deren Perzeption hin zu flacheren Neigungen ihrer Perzeptionskurven im prosodisch schwachen Kontext hervorbrachten. Die weiblichen Teilnehmer verhielten sich in der Perzeption ebenfalls kompatibel mit den Produktionsdaten (vgl. Kap. 4): sie zeigten sowohl stärker getrennte Kategorien in der Produktion als auch steilere Kurven in der Perzeption. Diese Ergebnisse sind mit Labov's (1990) Studie zur Rolle des Geschlechts im Laufe eines Lautwandels im Fortschritt in Einklang zu bringen.

Weiterführende Studien zur Perzeption der Sibilanten im Polnischen sind notwendig, um die Rolle der Wahrnehmung hinsichtlich des Mergers des /s ʂ/-Kontrasts zu klären. Interessante Erweiterungen dieser Perzeptionsstudie könnten die Perzeption von Stimuli aus natürlicher Sprache einbinden, sodass mögliche Artefakte, welche durch die synthetisch erzeugten Stimuli zustande gekommen sein könnten, ausgeschlossen werden können. Auch wäre die Untersuchung einer direkten Verbindung zwischen Produktion und Perzeption bezogen auf einen Zusammenfall der Sibilanten aufschlussreich. Dabei wäre ein direkter Vergleich von Produktions- und Perzeptionsdaten derselben Versuchspersonen interessant, um Muster in der Produktion in einen Zusammenhang mit der Perzeption von Sibilanten der jeweiligen Probanden zu bringen. Außerdem könnte auch eine Untersuchung von polnischen Dialektsprechern und Hörern, die den /s ʂ/-Kontrast tatsächlich neutralisieren, aufschlussreich sein für eine Erklärung der Umstände, welche zu ebendiesem Merger führen können.

Kapitel 4

Variation im Spracherwerb: Sibilantproduktionen von Kindern und Erwachsenen

4.1 Zusammenfassung

Ziel der Studie war es, zu untersuchen, ob in kindlichen Sprachproduktionsmustern im Laufe des Spracherwerbsprozesses Erklärungen für die Neutralisierung des /s ʂ/-Kontrasts, wie sie in einigen dialektalen Varietäten des Standard-Polnischen beobachtet wurde, gefunden werden können. Um Erklärungsansätze diesbezüglich zu erhalten, wurden akustische Daten von kindlichen und erwachsenen L1-Polnisch Sprechern erhoben. Dabei produzierten die Probanden in einem Bildbenennungsexperiment polnische ein- und zweisilbige Echtwörter mit den wortinitialen Sibilanten /s ʂ ʐ/ gefolgt von dem Vokal /ɛ/. In der akustischen Analyse wurde die Distanz der drei Sibilanten zueinander hinsichtlich der spektralen Informationen im Friktionsanteil der Sibilanten und der Formantwerte zum zeitlichen Onset des Vokals, welcher den Sibilanten nachfolgte, gemessen. Die Ergebnisse zeigten eine Separierung der Sibilantkategorien bezogen auf die untersuchten Parameter für die beiden Versuchspersonengruppen. Doch konnten kleine Unterschiede in den Distanzen zwischen den Sibilanten in der separaten Betrachtung der Kinder und der männlichen und weiblichen Probanden aufgezeigt werden. Während für die weiblichen Sprecher die größte Distanz und die geringste Überlappung zwischen den Kategorien sowohl im Friktionsgeräusch der Sibilanten als auch in den F2 und F3-Werten zu Beginn des nachfolgenden Vokals gefunden wurde, schienen die kindlichen und die männlichen Sprecher die Kategorien etwas

weniger unterschiedlich zu produzieren. Der größte Unterschied zwischen den Versuchspersonengruppen konnte dabei in der Analyse der Werte des zweiten und dritten Formanten im Onset des Vokals, welcher den Sibilanten nachfolgte, gefunden werden. Hier überlappen die kindlichen und männlichen Sprecher stark in ihren /s ʂ/ Kategorien, während die weiblichen Sprecher eine gute Trennung der beiden Laute zeigten. Diese Ergebnisse könnten damit im Sinne von Lautwandelmodellen interpretiert werden, welche eine schrittweise und graduelle Veränderung der Lautmuster vorhersagen. Weiterhin lassen die Daten vermuten, dass es sich bei der Neutralisierung des /s ʂ/-Kontrasts, um einen Prozess handelt, welcher geschlechtsspezifisch interpretiert werden sollte, und zwar dahingehend, dass männliche Mitglieder der Sprachgemeinschaft Lautwandel anführen, welche regionale und weniger innovative Sprachmuster hervorbringen.

4.2 Einleitung

Das Ziel der vorliegenden Studie war es, zu untersuchen, ob während des kindlichen Spracherwerbs bestimmte Muster in der Sprachproduktion auftreten, welche einen Beitrag zur Erklärung einer diachronen Lautneutralisierung leisten könnten. Getestet wurde dies anhand der drei im Standard-Polnischen kontrastierenden Sibilanten /s ʂ ʐ/, produziert von L1-Polnisch sprechenden Kindern, verglichen mit Sibilantproduktionen einer L1-Polnisch sprechenden Kontrollgruppe von erwachsenen Sprechern. Die spezifischere Frage der hier vorgestellten Studie war, ob ein solcher Vergleich von kindlichen und erwachsenen Sibilantproduktionen möglicherweise Aufschlüsse über die Neutralisierung des /s ʂ/-Kontrasts im Sibilantsystem einiger polnischer Dialekte geben kann.

Die Idee einer Verbindung zwischen Missartikulationen während des kindlichen Erstspracherwerbs und diachronem Lautwandel reicht bis in die frühe Zeit der Linguistik zurück (Jakobson, 1941; Kiparsky, 1988; Schleicher, 1971) und wird bis in die heutige Lautwandel- und Spracherwerbsforschung kontrovers diskutiert (Beckman et al., 2014; Foulkes & Vihman, 2015; Fromkin, Rodman & Hyams, 2011). Die frühe Sichtweise der Junggrammatiker hinsichtlich einer solchen Verbindung war es, dass sich Lautwandel in den Missartikulationen von Kindern im Laufe des Spracherwerbs wiederfinden (Schleicher, 1971), sowie dass Universalien einer phonologischen Grammatik existieren, welche sowohl Lautwandel als auch den Spracherwerb lenken (Jakobson, 1941). Spezieller argumentiert Locke (1983) zu Gunsten der Existenz einer solchen Verbindung, dass kindliche Sprachproduktions- und

Missartikulationsmuster, wie zum Beispiel Klusterreduktionen, finale Entstimmungen und Tilgungen von finalen Konsonanten, ebenfalls in bekannten historischen Lautwandelprozessen beobachtet werden können.

Dieser sehr frühen Sichtweise der Junggrammatiker zu den unterschiedlichen Verbindungen zwischen kindlichem Spracherwerb und diachronem Lautwandel stehen neuere Studien von Foulkes und Vihman (2015) und Beckman et al. (2014) gegenüber. Dabei wird von eben diesen neueren Studien solch ein linearer Zusammenhang zwischen Missartikulationen von Kindern und diachronem Wandel weitestgehend abgelehnt. Inkelas und McAllisterByun (2013) fassen vier Gründe aus mehreren Studien für das Ablehnen eines direkten Zusammenhangs zwischen Lautwandel und Spracherwerbsmustern zusammen. Der erste Aspekt der dabei genannt wird, stammt aus der Studie von Aitchison (2003), in welcher auf das Fehlen von Hinweisen aufmerksam gemacht wird, dass erwachsene Sprecher kindliche Varianten übernehmen, wodurch die Frage aufkommt wie sich ein Lautwandel dann über die ganze Sprachgemeinschaft ausbreiten kann. Der zweite Punkt betrifft die Fehlermuster von Kindern, da diese nicht notwendigerweise einheitlich sind, sodass systematische oder gar universale Muster zwischen Kindern nicht immer gefunden werden können Kiparsky (1988). Dieser Punkt führt zu der Überlegung, welche der vielen kindlichen Sprachvariationen dann aber zu einem potentiellen Lautwandel heranwachsen. Auch die Verschiedenheit von kindlichen Missartikulationen und Lautwandel bei erwachsenen Sprechern wird als Gegenargument zu einer linearen Verbindung von Lautwandel und kindlichen Spracherwerbsmustern angeführt (Foulkes & Vihman, 2015). Ein Beispiel für diese Verschiedenheit zeigen Foulkes und Vihman (2015) anhand von Klusterreduktionen auf, denn diese konnten sowohl im Laufe der kindlichen Sprachentwicklung als auch in historischen Lautwandel attestiert werden. Bei der Reduktion von Konsonantklustern bei Kindern, bestehend aus einer Plosiv-Liquid-Sequenz, wird in der kindlichen Sprachproduktion eher der Plosiv erhalten. Für Klusterreduktionen im Zuge von Lautwandelprozessen wird jedoch ein umgekehrtes Muster gefunden, beispielsweise in einem Lautwandelprozess, welcher im Slawischen zu beobachten ist. Hier bleibt entweder das Cluster komplett erhalten oder aber nur der Liquid. Ein weiterer Punkt, welcher als Argument gegen die Verbindung von Lautwandel und Fehlern im Laufe des Spracherwerbs dient, ist die Vergänglichkeit von kindlichen Sprachmustern/Fehlern. Kinder behalten in den meisten Fällen ihre Fehler nicht über die Spracherwerbsphase hinweg bei. Die Fehlermuster verändern sich mit fortschreitendem Spracherwerb und weichen korrekten Aussprachemustern (Vihman, Ferguson & Elbert, 1986).

Ein weiterer Aspekt mit einer etwas anderen Sichtweise als die oben aufgeführten frühen Ideen zu der Verbindung von Lautwandel und Spracherwerb wird erstmals von Sweet (1888) aufgeführt. Er nimmt an, dass Lautwandel aus inkorrektem Lernen beziehungsweise aus fehlerhaftem Imitieren auf Seiten der Kindern resultiert, welches in der unterschiedlichen Anatomie des Vokaltrakts von Kindern und Erwachsenen begründet liegt.

In eine ähnliche Richtung gehen die Gedanken von Sully (1896) mit der Folgerung, dass Lautwandel durch die inkorrekte Imitation von Sprachmustern in aufeinanderfolgenden Generationen entsteht. Diese ersten Gedanken zu fehlerhafter Imitation und von einer Generation zur nächsten Generation schrittweisen Verstärkung von veränderten Artikulationen, die zunächst feine Veränderungen sind und deren akustische Konsequenzen auditiv nicht wahrnehmbar sind, finden sich auch in neueren Studien und Ideen zu Spracherwerb von Foulkes und Vihman (2015) und zu Lautwandel von Labov (2001) wieder. Für viele Studien existiert der Beitrag von Kindern an Lautwandel aus einer etwas anderen Perspektive. Dabei wird oft auf den Unterschied zwischen der Übermittlung und der Verstärkung von Lautwandelmustern und die darauf bezogene Rolle von Kindern hingewiesen, wie es schon von Hockett in seinem Artikel aus dem Jahre 1950 erstmals zur Diskussion gestellt wurde. Dabei wird angenommen, dass im Kindesalter grundsätzlich Lautwandel in jeder Generation passieren kann, jedoch auf zwei verschiedene Arten und Weisen. Kinder, deren soziale Kontakte noch nicht über ihre Bezugsperson und ihre Familie hinausgehen (Kinder jünger als drei Jahre alt) übernehmen laut Hockett (1950), Roberts (1994), Eckert (1997) und Foulkes und Vihman (2015) bestehende Muster von ihren Bezugspersonen. Kinder, deren soziales Netz langsam erweitert wird und über den engen Familienkreis hinauswächst (Kinder älter als drei Jahre) können an den Innovationen von neuen Sprachmustern beteiligt sein. Dieser Gedanke geht einher mit den bekannten Begriffen *transmission* (Dt. Übermittlung) und *incrementation* (wörtlich übersetzt *schrittweise Verstärkung*), die vor allem durch sozio-phonetische Studien von Labov (1990) geprägt wurden. In der Phase der *transmission* übernimmt das Kind das phonologische System seiner Bezugsperson, was je nach Sprachmuster der Bezugsperson auch ein eher konservatives System sein kann. Im Alter zwischen vier und sechs Jahren beginnt das Kind sein soziales Netzwerk zu erweitern – zum Beispiel durch den Eintritt in den Kindergarten und den damit verbundenen Kontakt zu anderen Kindern und anderen Erwachsenen, um nur eines von vielen Beispielen für die Vergrößerung des Bewegungsradius von Kleinkindern zu nennen – und nimmt damit eine größere Bandbreite an sprachlicher Variation in der Interaktion mit gleichaltrigen Kindern und anderen Erwachsenen auf. Im Laufe der kindlichen Entwicklung und dem vermehr-

ten Kontakt mit anderen Personen kommt es auch zu der Möglichkeit, dass sich einige der Heranwachsenden von der Mehrheit abgrenzen wollen, was sie durch zum Beispiel eine in die Richtung des Lautwandels fortgeschrittenere Sprachproduktion aufzeigen. Durch dieses Verhalten von Kindern besteht die Möglichkeit, dass sich die Bandbreite an Sprachvariationen erweitert, und dass sich deren Verteilung schrittweise weiter in Richtung des Lautwandels verschiebt. Das phonologische System dieser Generation stabilisiert sich, laut Labov (2001), im jungen Erwachsenenalter und wird dann von ebendieser Generation an ihre Kinder weitergegeben. Parallel dazu entwickeln sich andere Kinder in dem fortschreitenden Lautwandel noch ein Stück weiter, sodass das phonologische System, welches von den Erwachsenen an ihre Kinder weitergegeben wird, wieder etwas konservativer ist als das System, dem deren Kinder dann ausgeliefert sind, wenn sie im Übergang vom Säuglingshinein ins Kleinkindalter aus dem engen Familienumfeld heraustreten (für mehr Details vgl. Labov, 2001). Die Studien von William Labov sind durch ihr sogenanntes *apparent-time* Design bekannt, welches ermöglicht, einen potentiellen Lautwandel im Fortschritt (und nicht wie zu Zeiten der komparativen Methode längst schon vergangene Lautwandel) zu untersuchen. Labov's Vorgehensweise war es, Daten von zwei Altersgruppen zu einem bestimmten Zeitpunkt zu erheben, zu untersuchen und miteinander zu vergleichen. In dieser Art und Weise der Untersuchung von Sprachproduktionen (sowie auch von Daten zur Sprachperzeption) besteht die Möglichkeit, in den feinen phonetischen Details die auditiv noch nicht wahrnehmbaren, schrittweisen Veränderungen der Sprachproduktion in Richtung Lautwandel von Generation zu Generation zu beobachten.

Nehmen wir diese ersten Ideen zur Verbindung von Lautwandel und Spracherwerb und deren nachfolgenden Veränderungen durch die neueren Studien, lässt sich zwar sagen, dass man vermutlich nicht von einem linearen Zusammenhang der beiden Prozesse – Lautwandel und Spracherwerb – sprechen kann, dass jedoch Kinder nichtsdestoweniger einen Einfluss auf die *incrementation*, d.h. die Verstärkung und Verbreitung von sprachlichen Variationen und Innovationen haben können. Auch besteht die Möglichkeit, anhand einer methodologischen Vorgehensweise, welche der Methode eines *apparent-time* Designs folgt, in den Sprachmustern von Kindern möglicherweise die Richtung eines bestimmten Lautwandels erkennen zu können.

Wie im ersten Absatz dieser Einleitung schon erwähnt, werden in der hier vorgestellten Studie mögliche Erklärungen und potentielle Vorstufen für einen Zusammenfall eines in

der Standardsprache bestehenden phonemischen Kontrasts der Sibilanten einiger dialektaler Varietäten des Polnischen innerhalb kindlicher Sprachmuster gesucht. Was aber ist interessant an Sibilanten an sich und vor allem an den Sibilanten des Polnischen?

Sprachübergreifende Studien zu akustischen und artikulatorischen Merkmalen von Sibilanten haben gezeigt, dass sich Sprachen in den akustischen Merkmalen, die zur Sibilantkategorisierung genutzt werden, unterscheiden. Dabei sind die von diesen Sprachen zur Sibilantunterscheidung verwendeten Merkmale abhängig vom jeweiligen Kontrast und den damit verbundenen verschiedenen Artikulationsparametern. Eine Sprache, anhand welcher es möglich ist, einen Teil der vielfältigen Eigenschaften von Sibilanten zu beschreiben, da hier nicht nur der Artikulationsort bei der Sibilantproduktion eine große Rolle spielt, sondern auch die Zungenform, ist das Standard-Polnische. In dieser Sprache findet man die in den Sprachen der Welt eher seltene Situation vor, dass drei Sibilanten miteinander kontrastieren: dental /s/, retroflex /ʂ/ sowie alveolopalatal /ɕ/. Der dentale und der retroflexe Sibilant werden dadurch produziert, dass sich die Zungenspitze dem Gaumen nähert und dadurch eine Verengung entsteht, welche die aus der Lunge strömende Luft verwirbelt. Ein markantes Unterscheidungsmerkmal zwischen dental /s/ und retroflex /ʂ/ ist die Lage dieser Konstriktion und der daraus resultierende Unterschied im spektralen Frequenzschwerpunkt. Während sich die Zungenspitze für den dentalen Sibilanten den Alveolen direkt hinter den oberen Schneidezähnen annähert, ist für den retroflexen Sibilanten die Konstriktion eher hinter den Alveolen, d.h. weiter hinten im Vokaltrakt, zu finden. Für den alveolopalatalen Laut findet sich in der anterior-posterior Dimension der Konstriktionsort zwischen der Verengung von /s ʂ/, genauer bewegt sich das Zungenblatt nach oben, hinter den Zahndamm, wobei ebenfalls der komplette Zungenrücken angehoben wird (während im Vergleich dazu /s ʂ/ einen eher flachen Zungenrücken zeigen), wodurch wiederum ein langer palataler Kanal entsteht (Ladefoged & Maddieson, 1996). Diese verschiedenen Artikulationen haben, wie schon oben erwähnt, unterschiedliche Auswirkungen auf die akustischen Eigenschaften der Sibilanten. Während die Lage der Konstriktion innerhalb der anterior-posterior Dimension im Vokaltrakt eher die spektralen Eigenschaften im Sibilanten selbst beeinflussen – je weiter vorne im Vokaltrakt und je kürzer dadurch der Bereich zwischen Verengung und Lippen, desto höher der spektrale Schwerpunkt –, werden durch die unterschiedlichen Zungenformen die Transitionen des zweiten Formanten in den nachfolgenden Vokal beeinflusst – je kürzer die *back cavity*, desto höher F2 (für mehr Details zur Akustik und Artikulation von Sibilanten siehe Kapitel 1).

Studien zu akustischen Merkmalen von polnischen Sibilanten haben gezeigt, dass die In-

formationen im Friktionsanteil der Laute ausreichen, um /s/ von /ʃ ɕ/ zu unterscheiden, nicht jedoch um letztere Laute zuverlässig voneinander zu trennen (Jassem, 1995; Nowak, 2006a). Ähnliche Ergebnisse brachte auch die physiologische Studie Kapitel 2 hervor. Die Daten aus den simultan erhobenen akustischen und artikulatorischen Aufnahmen (elektromagnetische Artikulografie) von neun L1-Polnisch sprechenden Erwachsenen wurden mittels einer Gaußschen Klassifikation von mehreren verschiedenen akustischen und artikulatorischen Merkmalen kategorisiert und zeigten dabei, dass für den Dreier-Kontrast im Polnischen, /s/ aufgrund seiner spektralen Eigenschaften (gemessen anhand des ersten und zweiten DCT-Koeffizienten) von /ʃ ɕ/ getrennt werden kann, wobei die Klassifikation für die beiden letzteren in diesen Parametern eine höhere Verwechslungsrate zeigte. Im Falle der F2 und F3-Transitionen in den nachfolgenden Vokal ergab die Klassifikation in Kapitel 2, dass der alveopalatale Sibilant /ɕ/ durch seine hohen Transitionen charakterisiert ist und in ebendiesem Merkmal klar von /s ʃ/ getrennt werden kann. Letztere zeigen niedrigere und überlappende F2-Werte im Übergang vom Sibilanten in den Vokal. Für das Mandarin-Chinesische, welches einen gleichartigen sibilantischen Dreier-Kontrast aufweist, zeigte F. Li (2008) ähnliche Ergebnisse, gemessen jedoch anhand anderer akustischer Parameter. In ihrer Studie wurde für die männlichen erwachsenen Mandarin-Chinesisch Sprecher eine Trennung von /s/ und /ʃ /, bezogen auf die jeweiligen Werte des ersten spektralen Moments (kurz M1) gefunden, während auch hier die beiden letzteren Laute eher ähnliche M1-Werte zeigten. Die F2-Transitionen in den nachfolgenden Vokal hinein trennen /ɕ/ von /s ʃ/, wobei der dentale und der retroflexe Laut wiederum in ihren niedrigen F2-Werten überlappen. Die weiblichen Sprecherinnen des Mandarin-Chinesischen zeigten in F. Li's (2008) Studie ähnliche F2-Muster verglichen mit den männlichen Versuchsteilnehmern. Der Unterschied zwischen den zwei Geschlechtern lag in den M1-Werten der drei Sibilanten: für die weiblichen Sprecher konnten mittels des ersten spektralen Moments alle drei Sibilanten unterschieden werden, d.h. es wurde hier keine Überlappung von /s ʃ/ gefunden. Ähnliche geschlechtsspezifische Unterschiede in der Sibilantproduktion von polnischen erwachsenen Sprechern fanden Żygiś und Hamann (2003). In deren Studie konnte für die weiblichen Versuchspersonen der retroflexe Laut /ʃ/ durch einen niedrigeren Frequenzschwerpunkt vom alveopalatalen Sibilanten /ɕ/ getrennt werden, während für die männlichen Versuchspersonen stärkere Ähnlichkeiten in den M1-Werten der beiden Laute /ʃ ɕ/ kennzeichnend waren.

Anhand der Ergebnisse dieser vorherigen Studien kommen wir zu der ersten Frage der

hier vorgestellten Studie, welche die rein akustischen Daten der Sibilanten betrifft: es soll hier zunächst herausgefunden werden, ob die Ergebnisse aus den akustischen und artikulatorischen Analysen in Kapitel 2 durch die hier vorgestellten rein akustischen Daten von L1-Polnisch sprechenden Kindern und Erwachsenen repliziert werden können. Vor allem soll damit die Möglichkeit geprüft werden, ob die Akustik der Sibilanten durch die EMA Sensoren in der vorherigen Studie beeinflusst wurde, wie es von Meenakshi, Yarra, Yamini und Ghosh (2014) gefunden wurde. Vor allem die Spektren der alveolaren Frikative wurden, laut Meenakshi et al. (2014), von den Sensoren, die bei EMA-Aufnahmen auf den Lippen, der Zunge und dem Kiefer der Versuchsperson befestigt sind, perturbiert. Jedoch war der Grad der jeweiligen Perturbation durch die Sensoren sprecherspezifisch, da verschiedene Versuchspersonen unterschiedlich lange brauchten, um sich an die Sensoren auf der Zunge zu gewöhnen. Bei den EMA-Aufnahmen in Kapitel 2 wurde grundsätzlich von den Versuchsleitern¹ darauf geachtet, dass nur Versuchspersonen in die Analyse inkludiert wurden, die sich relativ schnell an die Sensoren gewöhnten und bei denen damit von den Versuchsleitern rein auditiv keine Störung der Produktion der Sibilanten wahrgenommen wurde.

Für diese erste Forschungsfrage der hier vorliegenden Studie wird, den oben genannten Studien zur Akustik polnischer Sibilanten folgend, damit angenommen, dass L1-Polnisch Sprecher mittels der unterschiedlichen spektralen Merkmale im Friktionsgeräusch des Sibilanten selbst /s/ von /ʂ ɕ/ klar zu unterscheiden sein sollten, und dass basierend auf den F2-Werten zum Onset des nachfolgenden Vokals /ɛ/ von den beiden anderen Sibilanten zu trennen sein sollte.

Wie die oben aufgeführten akustischen und artikulatorischen Charakteristika zu den verschiedenen Sibilanten zeigen, ist die Kategorisierung des retroflexen Sibilanten im Polnischen (siehe alle Sprecher in Kap. 2) und im Mandarin-Chinesischen (siehe Sibilantproduktionen der männlichen Probanden in F. Li, 2008), also in Sprachen, die zwischen drei Sibilanten unterscheiden, abhängig von mehreren Merkmalen: niedrige Werte des spektralen Frequenzschwerpunktes, um den retroflexen Sibilanten von dental /s/ zu unterscheiden (da /s ʂ/ in F2 überlappen) und niedrige F2-Transitionen, um /ʂ/ vom alveolopalatalen Laut /ɕ/ trennen zu können (da /ʂ ɕ/ in ihren spektralen Werten im Friktionsgeräusch ähnlich sind).

Interessanterweise ist es eben der retroflexen Sibilant, der in einigen dialektalen Varietäten

¹Ein Versuchsleiter war eine L1-Polnisch sprechende Hilfskraft mit linguistischen Kenntnissen, welche die Aussprache der Sibilanten beurteilen konnte.

des Polnischen, sowie auch im Mandarin-Chinesischen von einer Neutralisierung betroffen ist (Duanmu, 2006; Gussmann, 2007; Nowak, 2006a; Sussex & Cubberley, 2006). Dabei wurde in den bisher bekannten Studien zu eben dieser Kontrastneutralisierung beobachtet, dass der retroflexe Sibilant eher mit dem dentalen Sibilanten zusammenfällt und der dentale Laut allein neben dem alveolopalatalen Sibilanten erhalten bleibt (für das Polnische vgl. Nowak, 2006a; Sussex & Cubberley, 2006; für das Mandarin-Chinesische vgl. Beckman et al., 2014; Duanmu, 2006). Dies führt zur zweiten Frage dieser Studie, nämlich, welche möglichen Gründe für die Instabilität des retroflexen Sibilanten gefunden werden können. Wie im vorherigen Abschnitt erläutert, liefert die Akustik dieses Dreier-Kontrasts einen Hinweis darauf. Genauer gesagt benötigt dieser Sibilant zwei akustische Merkmale, um klar von den beiden anderen Sibilanten separiert werden zu können: das Friktionsgeräusch, um ihn von /s/ zu trennen und die F2-Transitionen der umliegenden Vokale hinsichtlich einer Trennung von /ç/. Dieser Erklärungsmöglichkeit aus der Akustik folgen die Ergebnisse aus der artikulatorischen Studie in Kapitel 2, die zeigt, dass eine weitere mögliche Erklärung für die Variation und Instabilität des retroflexen Sibilanten im Verhältnis zwischen Artikulation und Akustik gefunden werden kann. Gemessen an den Zungenspitzenbewegungen, genauer gesagt an der vertikalen und horizontalen Position und -orientierung der Zungenspitze, wurde eine sehr gute Trennung der drei Sibilanten gefunden, während jedoch in der akustischen Analyse der Daten für den retroflexen Sibilanten, je nach gemessenem Parameter, immer Ähnlichkeiten mit einem der anderen beiden Sibilanten gezeigt wurden. Der artikulatorische Aufwand beziehungsweise die relativ komplexe Geste des retroflexen Frikatives resultierte in nur sehr geringen akustischen Auswirkungen – hauptsächlich niedrige Werte des dritten Formanten in den nachfolgenden Vokal –, was zu der Annahme führte, dass ebendieses nicht-lineare Verhältnis zwischen Artikulation und Akustik möglicherweise ein weiteres Indiz für die größere Instabilität des retroflexen Frikatives sein könnte (vgl. Kap. 2). Für die hier vorliegende Studie stellt sich die Frage, ob die Möglichkeit besteht, im Spracherwerb weitere Hinweise auf die Instabilität und Variabilität des retroflexen Lautes zu finden.

Grundsätzlich ist aus den wenigen Studien, die es zum Erwerb von Sibilanten bei Kindern gibt, herauszulesen, dass diese Lautklasse zu den von Kindern relativ spät erlernten Lauten zählt. Kent (1992) begründet die Schwierigkeit für Kinder, Frikative zu erlernen damit, dass Frikative eine präzise Zungenposition und Form benötigen, um den Luftstrom zum und durch den Konstriktionsort zu leiten, damit das typische Friktionsgeräusch ent-

stehen kann.

Studien zum kindlichen Erstspracherwerb gehen davon aus, dass die Artikulation von Kindern in der Anfangsphase ihrer Lautproduktionen durch globale und weniger präzise Gesten gekennzeichnet ist (Nittrouer, 2006; Nittrouer et al., 1989). Erst in den späteren Stadien des Lauterwerbs werden die statischen Aspekte der Lautproduktion fokussiert. Genauer formuliert in Nittrouer's (1989, 2006) Modell zum kindlichen Spracherwerb und der Sprachentwicklung bedeutet dies, dass die Bewegungen hin zu einem bestimmten Artikulationsort noch vor dem genauen Grad der Konstriktion und dem genauen Timing der Bewegungen beherrscht werden. Begründet liegt das, laut Nittrouer et al. (1989), darin, dass das Kind anfänglich zwischen Worten beziehungsweise zwischen Silben kontrastiert und nicht zwischen den einzelnen sprachspezifischen Phonemen. In ihrer Studie zu Sprachproduktionen von drei-, vier-, fünf- und sieben-jährigen englischsprachigen Kindern zeigten sie, dass im Verlauf der Sprachentwicklung die akustische Unterscheidung, gemessen anhand spektraler Merkmale in den englischen Sibilanten /s ʃ/, zunimmt und sich damit der Kategorisierung von erwachsenen Sprechern angleicht. Beim koartikulatorischen Einfluss der Vokale auf die Sibilantproduktionen der Kinder verhält es sich jedoch genau gegenteilig: mit steigendem Alter und Entwicklungsstand verringert sich der koartikulatorische Effekt auf die Sibilanten. Diese Ergebnisse aus der Produktion werden weiter gestützt durch die Studien zur Perzeption von Frikativ-Vokal Silben von Nittrouer und Studdert-Kennedy (1987) und Nittrouer (2006). Hier wurde gezeigt, dass Kinder im Alter von drei bis fünf Jahren in einem /s-ʃ/ Kontinuum eher sensitiv auf Veränderungen in den Transitionen reagieren und dieses Merkmal vorrangig zur Sibilantkategorisierung nutzen, verglichen mit den spektralen Veränderungen im Sibilant selbst. Dieses Muster in der Perzeption ändert sich schon bei Kindern im Alter von sieben Jahren, deren Perzeption Ähnlichkeiten mit den Mustern der erwachsenen Versuchspersonen zeigt, nämlich eine stärkere Gewichtung der spektralen Eigenschaften im Vergleich zu den Transitionen bei der Wahrnehmung von Sibilanten.

Ferguson und Farwell (1975), Menn (1975) und Clark und Bowerman (1986) sahen in solchen Spracherwerbsmustern einen weiteren Faktor, der die sprachliche Variabilität innerhalb eines Kindes und auch die Variabilität zwischen Kindern steigern könnte. Für sie zeigte die Variation im phonologischen Spracherwerb eine Art Strategie, um artikulatorische Schwierigkeiten bewältigen zu können. Auch Nittrouer (1995) geht davon aus, dass vom Erwerbszeitpunkt eines Lautes auf seine „Schwierigkeit“ geschlossen werden kann. Genauer gesagt nimmt sie an, dass Kinder Laute graduell lernen – angefangen mit den

leichteren Lauten und später hin zu den artikulatorisch schwierigeren. Jedoch zeigen verschiedene Studien zum Lauterwerb in verschiedenen Sprachen, dass auch hier nicht von universalen Mustern ausgegangen werden kann (Beckman et al., 2014; F. Li & Munson, in Druck). Zieht man diese Überlegungen für die polnischen Sibilanten heran, lässt sich vermuten, dass der retroflexe Laut im Polnischen aufgrund seiner artikulatorischen Komplexität als einer der letzten Laute gelernt wird, was auch von Łukaszewicz (2006, 2007) genauso beschrieben wird. Der Erwerb des retroflexen Sibilanten liegt, laut Łukaszewicz (2006, 2007), im Alter zwischen vier und fünf Jahren und ist damit einer der letzten erworbenen Laute des Polnischen.

Mit den oben aufgeführten Punkten thematisch verbunden, weisen weitere Befunde auf die hohen Anforderungen einiger Laute an Kinder im Spracherwerb aus einer etwas anderen Sichtweise hin – nämlich aus Beobachtungen von Lautersetzungen. So ist für das Englische bekannt, dass der postalveolare Sibilant /ʃ/ oft von Kindern vorverlagert wird, und durch den alveolaren Laut /s/ ersetzt wird. Für die Sibilanten /s ç/ im Japanischen zeigt sich ein anderes Muster. Hier wird der dentale Sibilant stereotypisch von Kindern durch sein alveolopalatales Gegenstück ersetzt. Zhu und Dodd (2000) finden für den Erwerb des Dreier-Kontrasts im Mandarin-Chinesischen, dass alle drei Sibilanten relativ früh gelernt werden, wobei der alveolopalatale Sibilant, welcher im Alter von ungefähr drei Jahren erworben wird, den beiden anderen Lauten im Erwerb voraus geht, die dann im Alter von 5 Jahren folgen. Studien zu Ersetzungen in den Mandarin-Chinesischen Sibilanten bringen eher kontroverse Ergebnisse, sodass hier kein so klares Ersetzungsmuster zu finden ist, wie es für die englischen und japanischen Sibilanten aufgezeigt wurde (weitere Details hierzu siehe, Beckman et al., 2014, sowie Li et al., 2009).

Wie zuvor erwähnt, ist die Literatur zum kindlichen Sibilanterwerb und noch dazu zum Erwerb dieser Lautklasse in Sprachen mit dem sehr seltenen Dreier-Kontrast, wie er im Polnischen zwischen /s ʃ ç/ gefunden werden kann, eher rar. Eine Studie, die sich mit den akustischen Eigenschaften der Sibilanten im Mandarin-Chinesischen beschäftigt, die akustisch und artikulatorisch starke Ähnlichkeiten mit den polnischen Sibilanten zeigen, ist die Studie von F. Li (2008). In dieser Studie wurden Sprachproduktionsdaten erwachsener Sprecher mit denen von Kindern im Alter zwischen drei und sieben Jahren verglichen, wobei gefunden wurde, dass Kinder zwar sprachspezifische Kategorisierungsmerkmale nutzen, um Sibilanten in der Produktion zu trennen – ähnlich den Mustern, welche für die weiblichen Versuchspersonen gefunden wurden –, jedoch divergieren die Kategorien der Kinder

nicht in dem Maße, wie es für die Erwachsenen gefunden wurde, und das sowohl in Bezug auf die gemessenen spektralen Parameter (M1) als auch auf die Transitionen des zweiten Formanten in den nachfolgenden Vokal hinein.

Betrachten wir die oben aufgeführten Abschnitte in Bezug auf Sibilanten und im Speziellen zu Sibilantproduktionen von Kindern und Erwachsenen, kommen wir zur zweiten und dritten Frage dieser Studie. Die zweite Frage lautet, ob es möglich ist die Ergebnisse aus F. Li's Studien (F. Li, 2008; F. Li & Munson, in Druck) zum Vergleich von Sibilantproduktionen bei Kindern und Erwachsenen im Mandarin-Chinesischen zu replizieren und damit zu zeigen, dass L1-Polnisch sprechende Kinder sprachspezifische Muster im spektralen Frikationsgeräusch der Sibilanten und in den F2-Werten zum Onset des Vokals, welcher den drei Sibilanten nachfolgte, zeigen, nur eben mit weniger starker Separierung der Kategorien bezogen auf die untersuchten Parameter verglichen mit der Kontrollgruppe der L1-Polnisch sprechenden Erwachsenen. Dabei werden in dieser Studie nicht nur die F2-Werte genauer betrachtet, sondern auch die F3-Werte zu Beginn des Vokals, welcher jeweils den drei Sibilanten folgt, um herauszufinden ob sich die F3-Werte der polnischen Kinder kompatibel mit den Ergebnissen zu den F2-Werten in Kapitel 2 zeigen.

Die dritte Frage soll klären, ob es gerade für den /s ʃ/-Kontrast Muster, wie zum Beispiel mehr Variation in der Produktion und/oder stärkere Überlappung in den hier untersuchten Parametern in den kindlichen Produktionsdaten gibt, die dann mögliche Erklärungsansätze für die Instabilität des retroflexen Sibilanten und für die diachrone Neutralisierung des /s ʃ/-Kontrasts in einigen polnischen dialektalen Varietäten liefern können.

4.3 Parameter zur Analyse von Sibilanten bei Kindern

In den frühen Studien zur Sprachproduktion von Kindern beobachtete man meist ein einzelnes Kind über einen längeren Zeitraum hinweg und notierte dessen sprachliche Äußerungen für die sogenannten *diary studies*. Bei den Beschreibungen in diesen Tagebüchern wurden rein auditiv die verschiedenen Lautproduktionen der Kinder hinsichtlich des jeweiligen sprachspezifischen Phoneminventars kategorisiert (Darwin, 1887). Im Laufe der Entwicklung der Spracherwerbsforschung wurden diese *diary studies* hinsichtlich der Versuchspersonenanzahl ausgeweitet (Poole, 1934; Prather, Hedrick & Kern, 1975), wobei auch hier die Sprachproduktionen der Kinder weiterhin rein auditiv beschrieben wurden, d.h. die kindlichen Laute wurden in Hinblick auf die phonologischen Kategorien von Erwachsenen

transkribiert. Heutige Studien versuchen diese Art der rein beschreibenden Vorgehensweise bei der Untersuchung von kindlichen Sprachproduktionen aus den folgenden Gründen zu umgehen:

Fehlende Konsistenz und Subjektivität innerhalb eines und Variation zwischen verschiedenen Transkribierern, hervorgerufen durch verschiedenes linguistisches und phonetisches Wissen. Auch die Anzahl der verschiedenen Fremdsprachen, die ein Transkribierer beherrscht, beeinflusst so auf eine subjektive Art und Weise seine Transkription. In diesem Fall könnte Unklarheit darüber herrschen, ob die verschiedenartige Transkription durch tatsächliche verschiedene Produktionen und Artikulationen der Versuchsperson, oder aber durch die fehlende Konsistenz der Transkribierenden hervorgerufen wurde.

Fehlen von feinen phonetischen Details in der Transkription. Während der Transkription werden von den Transkribierern die kindlichen Äußerungen und Laute in die Kategorien von erwachsenen Sprechern eingepasst, was zur Folge hat, dass feine phonetische Details, welche verschiedene Entwicklungsschritte von Kindern im Laufe des Spracherwerbs aufzeigen könnten, unentdeckt bleiben.

Der Vorteil von akustischen Analysen von Sprachdaten gegenüber der Methode der Transkription ist, dass die oben genannten Punkte vermieden und feine phonetische Details sichtbar gemacht werden können. Akustische Analysen von Sibilantproduktionen von Erwachsenen (Heinz & Stevens, 1961; Hughes & Halle, 1956) sind schon wohlprobt und auch die Analysen von kindlichen Sprachproduktionen haben sich von der reinen Transkription zur akustischen Analyse geändert (Nissen & Fox, 2005; Nittrouer, 1995; Nittrouer et al., 1989; Soli, 1985). Dabei bediente man sich vor allem bei der Untersuchung von Sibilanten der Analyse von spektralen Momenten, welche eine Möglichkeit bietet, um die spektralen Eigenschaften des Friktionsgeräusches von Sibilanten zu untersuchen. Anhand dieser Methode wurden Sibilanten, produziert von Erwachsenen und Kindern, erfolgreich analysiert, beschrieben und unterschieden (Beckman et al., 2014; F. Li, 2008; Nissen & Fox, 2005; Nittrouer, 1995). Das **erste spektrale Moment**, gibt den Frequenzschwerpunkt an und ist geeignet, um den Artikulationsort hinsichtlich der anterior-posterior Dimension aufzuzeigen, da das erste spektrale Moment negativ mit der Länge des Teils des Vokaltrakts korreliert, welcher vor der Konstriktion liegt, d.h. mit der Länge des Bereichs von der Konstriktion bis zu den Lippen. Das **zweite spektrale Moment** beschreibt die Varianz des Spektrums, genauer gesagt inwiefern das Spektrum von der Schwerpunktfrequenz abweicht

(Jongman et al., 2000; Nissen & Fox, 2005). Das **dritte spektrale Moment** – die Schiefe –, zeigt die Art und die Richtung der Schräge des Spektrums an, d.h. dieser Parameter beschreibt ob die Energiekonzentration eher oberhalb (wie zum Beispiel für den alveolaren Sibilanten /s/ im Deutschen) oder eher unterhalb (wie zum Beispiel für den postalveolaren Laut /ʃ/ im Deutschen) des Schwerpunkts liegt. Das **vierte spektrale Moment** – die Kurtosis – gibt an inwiefern das Spektrum einer Normalverteilung ähnelt (Jongman et al., 2000). Neben den spektralen Momenten wurde für eine weitere Analysevariante gezeigt, dass sie eine geeignete Methode darstellt, um Sibilanten voneinander zu unterscheiden, nämlich die *Diskrete Cosinus Transformation* – kurz *DCT* (Guzik & Harrington, 2007; Harrington, 2010; Watson & Harrington, 1999). Ähnlich der Analyse der spektralen Momente wird auch bei der DCT-Analyse die Fülle an Informationen, die ein Spektrum beinhaltet, auf wenige Parameter, die wiederum die Form eines Spektrums beschreiben, reduziert. Anhand dieser Parameter besteht die Möglichkeit, Lautkategorien voneinander zu unterscheiden (Bukmaier, Harrington, Reubold & Kleber, 2014; Jannedy et al., 2015; Watson & Harrington, 1999). In Guzik und Harrington (2007), sowie in Kapitel 2 wurde für den Dreier-Kontrast der polnischen Sibilanten gezeigt, dass die DCT-Koeffizienten [1] und [2] eine Möglichkeit darstellen, /s ʃ ʑ/ produziert von erwachsenen Sprechern erfolgreich zu unterscheiden, wobei die gefundenen Muster der spektralen Verteilung dieser Laute mit den Ergebnissen aus Studien übereinstimmen, welche die Analyse der spektralen Momente von Sibilanten anwendeten (vgl. Analyse des Dreier-Kontrast in den Sibilanten des Mandarin-Chinesischen in Li, 2008). Die beiden DCT-Koeffizienten, die meist in der spektralen Analyse von Sibilanten näher betrachtet werden, können wie folgt beschrieben werden:

DCT [1]], welcher die Steigung des Spektrums betrachtet, unterscheidet /s/ von den beiden anderen polnischen Sibilanten durch eine ansteigende Steigung, welche in den niedrigeren Werten von *C1* für alveolare/dentale Sibilanten begründet liegt.

DCT [2]], welcher die Krümmung des Spektrums beschreibt, trennt /s/ von /ʃ ʑ/ dadurch, dass für den dentalen Laut eine Parabel-ähnliche Form des Spektrums gefunden werden kann, was durch die höheren *C2*-Werte für /s/ angezeigt wird.

Obwohl die Analyse der spektralen Momente gegenüber der DCT-Analyse für die Untersuchung von kindlichen Sibilantproduktionen eher erprobt ist, sollen die hier vorgestellten Daten der polnischen Kinder und Erwachsenen anhand der DCT-Analyse untersucht werden. Ein Grund dafür ist, dass die Daten dieser Studie mit den Ergebnissen aus Kapitel

2 vergleichbar gehalten werden sollen, vor allem um zu sehen ob eine Beeinflussung der Sibilantproduktion durch die Sensoren in der vorherigen artikulatorischen Studie stattgefunden hat (siehe Ausführungen in Abschnitt 4.2).

Bevor jedoch die DCT-Analyse tatsächlich in diese Studie aufgenommen wurde, wurde zunächst rein deskriptiv untersucht, ob sich die Kinderdaten bei der DCT-Analyse ähnlich den Erwachsenendaten verhalten, sodass ausgeschlossen werden konnte, dass etwaige Unterschiede zwischen den Sprechergruppen auf die Methode und nicht auf Produktionsunterschiede zurückzuführen sind (vgl. Abb. 4.1).

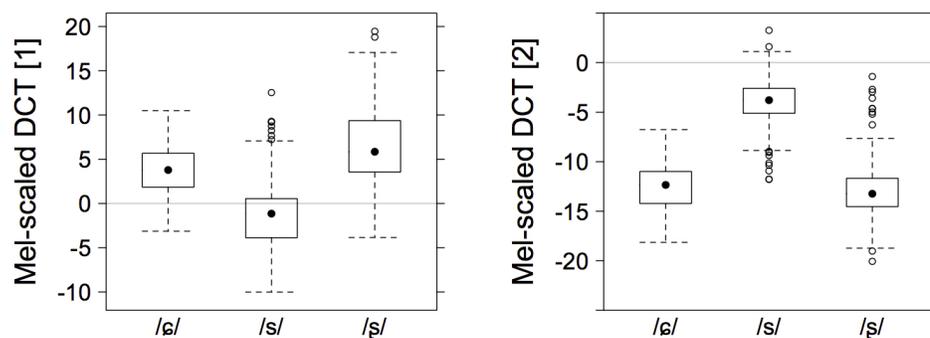


Abbildung 4.1: DCT-Koeffizient [1] und [2] (Steigung und Krümmung) zum zeitlichen Mittelpunkt der Sibilanten /s ɕ ʂ/ für L1-Polnisch sprechende Kinder. DCT [1] (links) und DCT [2] (rechts) zeigen eine gute Trennung der drei Sibilanten.

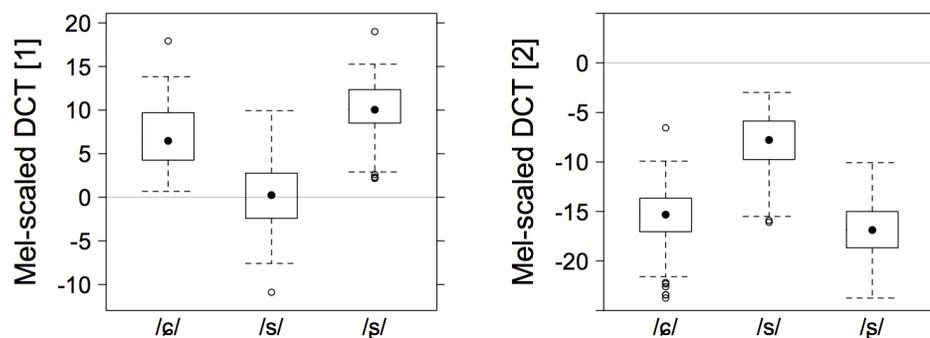


Abbildung 4.2: DCT-Koeffizient 1, 2 (Steigung und Krümmung) zum zeitlichen Mittelpunkt der Sibilanten /s ɕ ʂ/ für L1-Polnisch sprechende Erwachsene. DCT 1 (links) und DCT 2 (rechts) zeigen eine gute Trennung der drei Sibilanten.

In den Abbildungen 4.1 (Kinder) und 4.2 (Erwachsene) ist gut zu erkennen, dass sich die Sibilanten der Kinder als auch die der Erwachsenen in der DCT-Analyse beziehungsweise hinsichtlich der DCT-Koeffizienten mit früheren Studien zu dieser Lautklasse und mit Analysen anderer Parameter (z.B. Analyse der spektralen Momente) vergleichbar darstellen. In diesen beiden Koeffizienten ist zu erkennen, dass sich im Friktionsgeräusch der dentale Sibilant von /s̺ ɕ/ abhebt, während sich letztere in diesen Parametern eher ähneln.

Wie schon in der Einleitung (Abschnitt 4.2 hier und in Kapitel 2) aufgezeigt, sind die Transitionen des zweiten Formanten im Vokal, welcher den Sibilanten folgt, ausschlaggebend für die akustische Differenzierung des Dreier-Kontrasts /s̺ ɕ ʃ/. F. Li (2008) zeigte sowohl für das erste spektrale Moment im Frikativ als auch für die F2-Werte zum zeitlichen Vokalonsset, dass es möglich ist in den kindlichen Sibilantproduktionen im Mandarin-Chinesischen sprachspezifische Muster zu finden, die mit der erwachsenen Kontrollgruppe zu vergleichen sind (weitere Ergebnisse zu Analysen von Formanten bei Kindern siehe Nittrouer, 1995). Aufgrund dieser vorherigen Studien wird auch für die hier vorliegende Studie die Formantanalyse zum zeitlichen Vokalonsset angewendet, um zu testen ob polnische Kinder Muster in der Verteilung der F2-Werte zeigen, welche mit denen der erwachsenen Kontrollgruppe zu vergleichen sind. Auch für diesen Parameter wurde zu Beginn der Analyse deskriptiv getestet, ob die bisherige Vorgehensweise zur Messung des koartikulatorischen Einflusses sowohl für Erwachsene als auch für Kinder adäquat ist. In den vorherigen Studien zu den F2-Transitionswerten von den Sibilanten in den nachfolgenden Vokal hinein, wurde zur Messung der Koartikulation in den meisten Fällen der Vokalonsset als Referenzpunkt herangezogen.

Dass dieser Zeitpunkt für die Trennung der drei Sibilanten geeignet ist, zeigen die Abbildungen in 4.3, welche die Lobanov- und zeitnormalisierten F2-Werte vom Onset bis zum zeitlichen Mittelpunkt des Vokals aufzeigen. Für die erwachsenen Sprecher (rechts) ist hier wie zu erwarten zu sehen, dass sich vor allem der alveolopalatale Sibilant (schwarze Linien) von /s̺ ɕ/ (blaue und rote Linien) durch höhere Werte des zweiten Formanten abhebt. Im Falle der Kinder (Abb. links) ist zu beobachten, dass das Muster der F2 Verläufe generell durch mehr Variation, und eine nicht so klare Trennung des alveolopalatalen Sibilanten (schwarze Linie) von den beiden anderen geprägt ist. Dennoch zeigt dieser Vergleich, dass zu Anfang des Vokals bei Kindern die Verteilung der sibilant-spezifischen Transitionen der erwachsenen Verteilung am ähnlichsten ist, während später im Vokal mehr Variation und größere Unterschiede zwischen den Erwachsenen und den Kindern in den Formantstrukturen zu sehen sind.

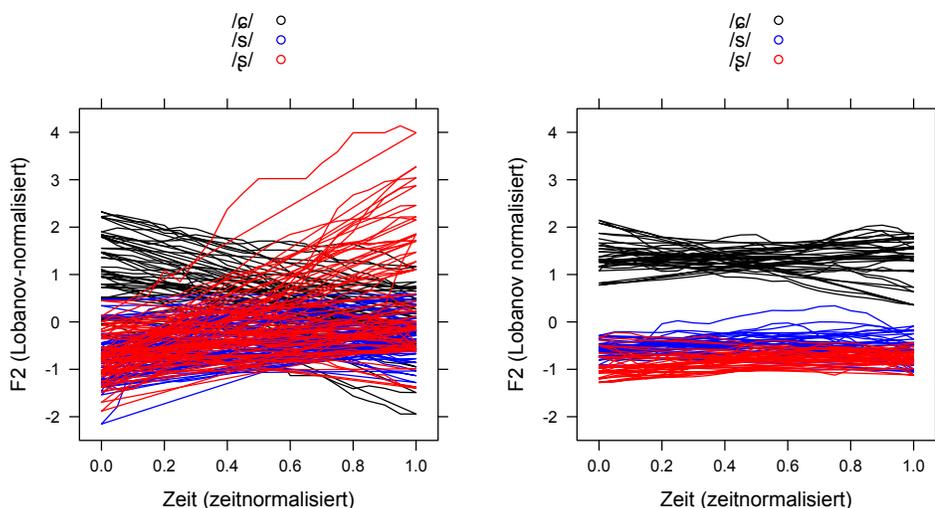


Abbildung 4.3: Gemittelte und zeitnormalisierte F2-Transitionen vom Onset bis zum Mittelpunkt des den Sibilanten /s/ (blau), /ʂ/ (rot), /ɕ/ (schwarz) nachfolgenden Vokals für L1-Polnisch sprechende Kinder (links) und Erwachsene (rechts).

Viele bisherige Studien, die sich mit den akustischen Merkmalen der drei Sibilanten /s ʂ ɕ/ beschäftigten, untersuchten meist die spektralen Merkmale im Sibilant selbst mittels der Berechnung von spektralen Momenten, wie auch die unterschiedlichen F2-Transitionen in die Vokale hinein, welche den Sibilanten folgten. Fant (1960) beschrieb aber, dass gerade für den retroflexen Sibilanten ebenfalls der dritte Formant als akustisches Merkmal von Bedeutung sein kann. Gordon et al. (2002) zeigte für die drei kontrastierenden Sibilanten in Toda, dass der retroflexe Frikativ niedrigere Werte des dritten Formanten in den nachfolgenden Vokal hinein aufweist. Damit vergleichbar wurde in Kapitel 2 dieser Arbeit ebenfalls berichtet, dass die Informationen im dritten Formanten dadurch zur Unterscheidung der drei Sibilanten beitragen können, dass dem retroflexen Laut niedrige F3-Werte folgen. Aus diesem Grund werden auch für die hier vorliegende Studie nicht nur die F2-Werte, sondern auch die Werte des dritten Formanten zum zeitlichen Onset des Vokals, welcher den Sibilanten /s ʂ ɕ/ nachfolgte, genauer betrachtet.

4.4 Experimentaufbau

4.4.1 Datenerhebung und Sprecher

Für dieses Sprachproduktionsexperiment wurden zwei L1-Polnisch Sprechergruppen rekrutiert: 22 Kinder, davon 14 weiblich und 8 männlich, im Alter zwischen 5 und 6 Jahren, sowie 10 Erwachsene, davon 7 weiblich und 3 männlich, im Alter zwischen 20 und 41 Jahren (\bar{x} 29). Die Gruppe der kindlichen Versuchsteilnehmer wurde von einer weiblichen L1-Polnisch sprechenden Hilfskraft in drei verschiedenen Kindergärten in Polen aufgenommen. Die Kindergärten in denen die Kinder rekrutiert wurden, befinden sich alle in der Kleinpolnischen Dialektregion. Die Kinder, die an der Studie teilnahmen, zeigten altersgerechte kognitive und vor allem auch altersgerechte sprachliche Entwicklungen auf, was durch die Erzieherinnen in den jeweiligen Kindertagesstätten bestätigt wurde.

Die Hilfskraft rekrutierte ebenfalls vier der erwachsenen Versuchspersonen (2 weiblich, 2 männlich) in dieser Region in Polen, wo die Aufnahmen in einem ruhigen Zimmer zu Hause durchgeführt wurden. Die restlichen sechs erwachsenen Teilnehmer wurden in München rekrutiert und entweder in einem ruhigen Zimmer bei ihnen zu Hause oder im Aufnahmestudio des Instituts für Phonetik und Sprachverarbeitung in München aufgenommen. Alle Probanden, die in München aufgenommen wurden, erfüllten die Bedingung, zum Aufnahmezeitpunkt nicht länger als zwei Jahre außerhalb von Polen gelebt zu haben. Die Versuchspersonen aus dieser Teilnehmergruppe stammten aus und lebten die meiste Zeit ihres Lebens in folgenden Dialektregionen in Polen: Großpolen, Kleinpolen, Schlesien und Niederschlesien. Den auditiven Einschätzungen der L1-Polnisch sprechenden Hilfskraft zufolge, waren alle Versuchspersonen Sprecher des Standard-Polnischen ohne regionale Färbung. Alle Teilnehmer, d.h. sowohl die Kinder als auch die Erwachsenen waren frei von Hör- und Sehschäden.

4.4.2 Sprachmaterial und Experimentaufbau

Das hier vorgestellte Produktionsexperiment ist der erste Teil eines dreiteiligen Gesamtexperiments, bestehend aus einem Produktions-, einem Imitations- und einem Perzeptionsteil. Alle Versuchspersonen durchliefen alle drei Experimentabschnitte innerhalb einer Sitzung. Die Dauer des Gesamtexperiments (Produktion, Imitation und Perzeption hintereinander) belief sich auf ungefähr 45 Minuten (inklusive Spielpausen bei den Kindern). Die Gesamtdauer war vor allem auf Grund der kindlichen Versuchsteilnehmer und deren eher

geringerer Konzentrationsfähigkeit relativ kurz gehalten. Die zeitliche Einschränkung des Experiments hinsichtlich der Belastbarkeit der Kinder brachte somit auch Einschränkungen bezüglich der Anzahl an Wiederholungen der Zielwörter (Produktions- und Imitationsexperiment) beziehungsweise Stimuli (Perzeptionsexperiment) mit sich.

Für das Produktionsexperiment wurden polnische Echt-Wörter elizitiert. Für die Kinder, die in diesem Alter noch unzureichende Lesekompetenzen zeigen, wurde das Produktionsexperiment als Bildbenennungsaufgabe aufgebaut. Um sicherzustellen, dass die Kinder die auf den Bildern dargestellten Wörter kennen, mussten sie die Bilder vor der eigentlichen Aufnahme benennen. Im Falle einer falschen Bildbenennung im Laufe dieser Testproduktionen wurden die Kinder von der Experiment-leitenden Hilfskraft korrigiert und dazu aufgefordert das korrekte Wort zu wiederholen.

Die Zielwörter waren ein- und zweisilbige Wörter des Polnischen mit wortinitialem Sibilanten (/s ʂ ɕ/) (vgl. Tab 4.1). Um den koartikulatorischen Einfluss zu kontrollieren, sowie um den Umfang des Experiments nicht durch mehrere Vokalkontexte auszudehnen, war der Vokal, der den jeweiligen Sibilanten folgte, immer der ungespannte, vordere Vokal /ɛ/.

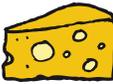
					
ser	serce	szecz	szeryf	siekacz	siedem
/ser/	/sertsɔ/	/ʂɛʂ/	/ʂɛrif/	/ɕɛkats/	/ɕɛdɛm/
Käse	Herz	Sechs	Sheriff	Schneidezahn	Sieben

Tabelle 4.1: Als Bildbenennungsaufgabe aufgebautes Sprachmaterial für das Produktionsexperiment der kindlichen Versuchspersonen.

Die Zielwörter für das Produktionsexperiment mit den kindlichen Versuchspersonen mussten Wörter sein, die Kindern im Alter zwischen 5 und 6 Jahren bekannt und die weiterhin bildlich darstellbar waren. Diese Einschränkungen brachten nur nahe Minimalpaare als Zielwörter hervor, die wie oben schon erwähnt in der Silbenanzahl zwischen ein- und zweisilbigen Wörtern variierten (vgl. Tab. 4.1).

Die Sprachaufnahmen wurden mit dem Programm *speechrecorder* (Version 2.6.14; vgl. Draxler & Jänsch, 2004), einem Audio Interface (M-Audio Fast Track) und einem Stereo Headset (beyerdynamic) durchgeführt. Das Produktionsexperiment der Kinder bestand aus fünf Wiederholungen der sechs Zielwörter. Die Bilder, die die verschiedenen Zielwörter darstellten, wurden den Kindern auf einem MacBookPro Computerbildschirm (eingebun-

den in das Aufnahmeprogramm *speechrecorder*) in randomisierter Reihenfolge präsentiert (vgl. Abb. 4.4 (a)). Die Präsentation des nächsten Bildes wurde bei den Kindern manuell durch die Versuchsleiterin geregelt, d.h. das nächste bildlich dargestellte Zielwort wurde den Kindern erst auf dem Bildschirm angezeigt, sobald sie das vorangegangene Zielwort produziert hatten. Dieses Vorgehen wurde gewählt, um Datenverluste aufgrund verspäteter Wortproduktionen oder aufgrund von Missartikulationen der Kinder zu vermeiden. Von den insgesamt 660 Zielwörtern, die von den an der Studie teilnehmenden 22 Kindern produziert wurden (6 Zielwörter \times 5 Wiederholungen \times 22 L1-Polnisch sprechende Kinder), wurden nach Ausschluss von missartikulierten Wörtern 429 Zielwörter analysiert. Versuchsperson *Kind007* wurde dabei komplett von der weiteren Analyse ausgeschlossen, da für dieses Kind der alveolopalatale Sibilant nur einmal realisiert wurde.

Die erwachsenen Probanden bekamen die isolierten Zielwörter als Leseaufgabe präsentiert. Die Zielwörter der erwachsenen Versuchspersonen unterschieden sich teilweise von denen der Kinder. Der Grund hierfür lag in einem der Ziele dieses Experiments, nämlich die rein akustischen Daten mit den artikulatorischen und akustischen Daten aus der EMA Studie in Kapitel 2 zu vergleichen. Das Korpus der erwachsenen Versuchsteilnehmer, welches einer Auswahl an Zielwörtern aus Kapitel 2 entspricht, enthielt die in Tab. 4.2 abgebildeten zweisilbigen Logatome und polnischen Echt-Wörter.

	Logatom	Echt-Wort
/sɛ/	Sese /sɛsɛ/	Serce /sɛrtɛ/ Herz
/ʂɛ/	Szesze /ʂɛʂɛ/	Szereg /ʂɛrɛk/ Reihe
/ɕɛ/	Siesie /ɕɛɕɛ/	Siekacz /ɕɛkatɕ/ Schneidezahn

Tabelle 4.2: Als Leseaufgabe aufgebautes Sprachmaterial für das Produktionsexperiment der erwachsenen Versuchspersonen.

Wie in der obigen Tabelle (4.2) zu sehen, waren die Zielwörter (Logatome wie auch die Echt-Wörter) der Erwachsenen allesamt zweisilbige Wörter mit wortinitialem Sibilanten, immer gefolgt von dem hohen ungespannten Vokal /ɛ/ (die Wahl des Vokals erfolgte anhand

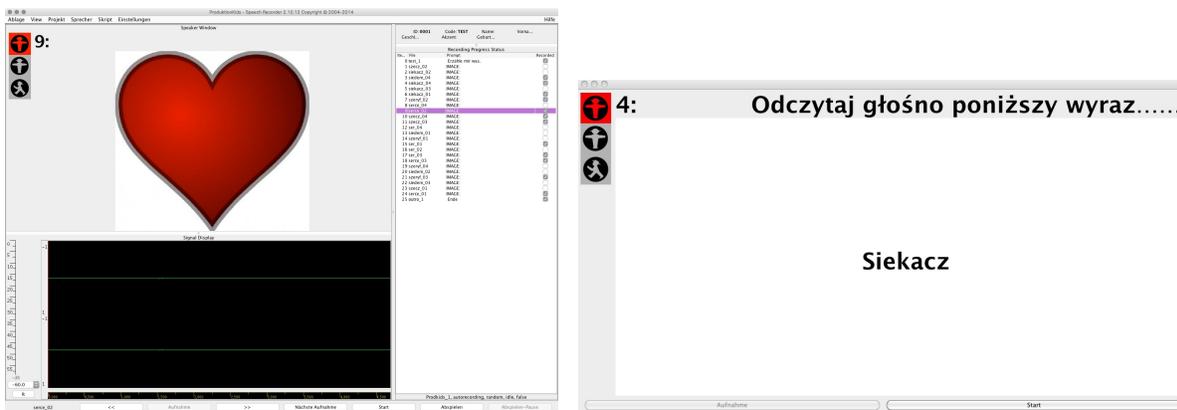
der Gründe, die für die Auswahl der Zielwörter im Korpus der Kinder angegeben wurden. Genauer gesagt wurde /ɛ/ als nachfolgender Vokal gewählt, um die koartikulatorischen Einflüsse auf den Vokal im Vergleich zwischen kindlicher und erwachsener Sibilantproduktion konsistent zu halten).

Die Sprachaufnahmen wurden bei den Erwachsenen ebenfalls mit dem Programm *speech-recorder* (Version 2.6.14; vgl. Draxler & Jänsch 2004), einem Audio Interface (M-Audio Fast Track) und einem Stereo Headset (beyerdynamics) durchgeführt. Das Produktionsexperiment der Erwachsenen bestand aus fünf Wiederholungen der sechs Zielwörter. Die isolierten Zielwörter der erwachsenen Versuchsteilnehmer wurden in orthografischer Form auf einem MacBookPro Computerbildschirm in randomisierter Reihenfolge präsentiert (vgl. Abb. 4.4 rechte Graphik). Die Präsentation des nächsten Wortes wurde bei den Erwachsenen automatisch geregelt: jedes Zielwort wurde den Teilnehmern 3.0 s lang angezeigt, bevor nach einer Pause von 0.2 s automatisch das nächste Wort eingeblendet wurde. Auf dem Bildschirm sahen die erwachsenen Versuchsteilnehmer durchgehend ihre Aufgabe für das Experiment in polnischer Sprache: „Odczytaj gło śno poniższy wyraz...“ (wörtlich übersetzt auf deutsch: „Bitte lies folgendes Wort laut vor...“) (vgl. Beispiel für das Zielwort *Siekacz* in Abb. 4.4 (b)). Von den insgesamt 300 Zielwörtern, die von den erwachsenen Versuchsteilnehmern produziert wurden (6 Zielwörter × 5 Wiederholungen × 10 L1-Polnisch sprechende Erwachsene), konnten nach Ausschluss von vier Missartikulationen 296 Zielwörter mit wortinitialem Sibilanten in die anschließende Analyse aufgenommen werden. Beide Altersgruppen zusammengenommen, wurden für dieses Produktionsexperiment 729

	Kinder	Erwachsene
/s/	164	96
/ʂ/	163	100
/ʑ/	106	100

Tabelle 4.3: Verteilung der insgesamt 729 Sibilantproduktionen in den kindlichen und den erwachsenen Versuchspersonen.

Wörter analysiert. Die genaue Verteilung der Sibilanten in den Altersgruppen ist in Tab. 4.3 aufgezeigt.



(a) Zielwortdarstellung Kinder

(b) Zielwortdarstellung Erwachsene

Abbildung 4.4: Darstellung der Zielwörter in Bildern für die jungen Probanden (a) und für die erwachsenen Teilnehmer als isolierte Wörter (b) im Verlauf der Produktionsexperimentdurchführung.

4.4.3 Datenanalyse

Die Audiodaten wurden mit einer Frequenz von 44.1 kHz digitalisiert. Alle Daten wurden mit dem *Munich Automatic Segmentation System* (MAuS, Schiel, 2004) automatisch segmentiert und phonetisch gelabelt. Die Segmentierungen wurden von einer L1-Polnisch sprechenden studentischen Hilfskraft kontrolliert und wenn notwendig manuell in der EMU-webApp (Winkelmann & Raess, 2014) korrigiert.

Nachdem eine 256 Punkte Fourier Transformation mit einer Frequenzauflösung von 40 Hz, einem Blackman Fenster von 5 ms und einer Fensterverschiebung von 5 ms auf die Zielwörter eines jeden Sprechers angewendet wurde, wurde das Powerspektrum eines jeden Frikatives extrahiert. Anschließend wurde das Spektrum zum zeitlichen Mittelpunkt zwischen dem akustischen On- und Offset der Sibilanten zur weiteren Analyse herangezogen, nachdem es in die mel-Skala konvertiert wurde. Für das mel-skalierte Spektrum wurden dann, analog zu der Vorgehensweise in Abschnitt 2.3.1 die DCT-Koeffizienten, C_1 (Steigung) und C_2 (Krümmung) berechnet. Genauer gesagt wurde für ein N -Punkte mel-skaliertes Spektrum, $x(n)$, welches sich in der Frequenz von $n = 0$ bis $N-1$ Punkten über den Frequenzbereich zwischen 500-5000 mel erstreckt, der m te DCT-Koeffizient C_m ($m = 1, 2$) berechnet (Formel zur Berechnung der DCT-Koeffizienten siehe 2.3.1). Um die akustische Distanz zwischen den drei Sibilanten genau quantifizieren zu können, wurde für jeden Sibilanten das logarithmisierte Verhältnis der Euklidischen Distanzen im $C_1 \times C_2$ -

Raum zu zwei sprecherspezifischen Ankerpunkten, nämlich den Zentroiden der Verteilungen von /s/ und /ʃ/ (E_s und $E_ʃ$) berechnet. Die dentalen und die retroflexen Sibilanten als Zentroiden zu wählen, lag darin begründet, dass Nowak (2006a) und Lisker (2001) für den alveolopalatalen Sibilanten einen spektralen Frequenzschwerpunkt zeigten, der zwischen dem Schwerpunkt des dentalen und dem des retroflexen Lautes lag. Das logarithmisierte Verhältnis der Euklidischen Distanzen d_{sib} wurde anhand der Formel (4.1) berechnet:

$$d_{sib} = \log(E_s/E_ʃ) = \log(E_s) - \log(E_ʃ) \quad (4.1)$$

Das Maß d_{sib} wurde berechnet, um einen zu zwei Ankerpunkten relativen Positionswert pro Sibilant zu erhalten, welcher dann wie folgt interpretiert werden konnte: positive Werte geben eine größere Nähe zum retroflexen Ankerpunkt an, negative Werte zeigen die Entfernung zum dentalen Ankerpunkt, während ein d_{sib} Wert von Null aufzeigt, dass der jeweilige Sibilant im akustischen Raum gemessen am ersten (Steigung) und zweiten (Krümmung) DCT-Koeffizienten vom dentalen und vom retroflexen Zentroiden gleich weit entfernt liegt.

Die koartikulatorischen Einflüsse der einzelnen Sibilanten wurden anhand der Formanten in den Vokal /ε/, welcher den jeweiligen Sibilanten folgte, quantifiziert (F1-F4; Pre-Emphase von -0.8, 20 ms Blackman Fenster mit einer Fensterverschiebung von 5 ms). Analog zur Vorgehensweise bei der Formantberechnung in Kapitel 2 wurde auch in diesem Kapitel bei der Berechnung der Formanten der nominale F1 Wert individuell für die einzelnen Sprechergruppen festgelegt. In Anlehnung an Hillenbrand et al. (1995)'s Artikel zu Formantwerten bei männlichen und weiblichen erwachsenen sowie bei kindlichen Sprechern des Amerikanischen Englisch wurde für die einzelnen Probandengruppen der hier vorliegenden Studie der nominale F1 Wert wie folgt angesetzt: 500 Hz für männliche erwachsene Sprecher, 600 Hz für die weiblichen erwachsenen Sprecher und 700 Hz im Falle der kindlichen Sprechgruppe (für Details zu den Auswirkungen von Veränderungen des nominalen F1 Wertes bei der Berechnung von Formanten siehe Abschnitt 2.3.4 dieser Arbeit).

Für die Analyse der koartikulatorischen Einflüsse wurden speziell die Werte des zweiten und dritten Formanten zum akustischen Onset des Vokals extrahiert. Der Vokalonsset wurde zum Zeitpunkt des positiven Nulldurchgangs der ersten periodischen Schwingung nach der Friktion des vorangegangenen Sibilanten markiert. Um Unterschiede zwischen den Sprechern zu reduzieren wurden die Formantdaten vor der weiteren Analyse Lobanov-normalisiert (für Details zur Lobanov-Normalisierung vgl. Abschnitt 2.3.4).

Um die Unterschiede beziehungsweise um die Überlappung der drei Sibilanten im Vergleich von Kindern und Erwachsenen in Bezug auf die drei untersuchten Parameter (d_{sib} im $C_1 \times C_2$ -Raum und F2 und F3-Werte zum zeitlichen Vokalansatz) zu messen, wurden für jeden Sprecher einzeln die jeweiligen Werte von d_{sib} zu den Zentroiden /s/ und /ʃ/ und die jeweiligen F2 und F3-Werte der Frikative voneinander subtrahiert, sodass für jeden Parameter (d_{sib} und $F2/F3$) ein Differenzwert (siehe $\text{Diff}_{d_{sib}}$ und $\text{Diff}_{F2/F3}$ unten) für jedes Sibilantenpaar pro Sprecher in die Analyse einging.

$$\text{Diff}_{d_{sib}} = \begin{vmatrix} d_{sib \text{ /dental/}} - d_{sib \text{ /alveolopalatal/}} \\ d_{sib \text{ /dental/}} - d_{sib \text{ /retroflex/}} \\ d_{sib \text{ /retroflex/}} - d_{sib \text{ /alveolopalatal/}} \end{vmatrix} \quad (4.2)$$

$$\text{Diff}_{F2/F3} = \begin{vmatrix} F2/F3_{\text{/alveolopalatal/}} - F2/F3_{\text{/dental/}} \\ F2/F3_{\text{/retroflex/}} - F2/F3_{\text{/dental/}} \\ F2/F3_{\text{/alveolopalatal/}} - F2/F3_{\text{/retroflex/}} \end{vmatrix}$$

4.5 Ergebnisse

Die nachfolgenden Ergebnisse zeigen Gemeinsamkeiten und Unterschiede in der Akustik der drei polnischen Sibilanten von Kindern und Erwachsenen. Zunächst wird hier auf die erste Forschungsfrage eingegangen, nämlich ob es möglich ist, die Ergebnisse aus der artikulatorischen Studie in Kapitel 2 zu replizieren: (1) Hohe Werte in den spektralen Merkmalen im Friktionsanteil der Sibilanten für /s/ während /ʃ ɕ/ im Friktionsanteil niedrigere Werte zeigen, (2) Hohe F2-Werte zum zeitlichen Onset des Vokals, welcher /ɕ/ folgt, während sich /s ʃ/ stark in den F2-Werten ähneln und (3) Separierung von /ʃ/ durch dessen niedrigere F3-Werte in den anschließenden Vokal im Vergleich zu /s ɕ/. Zunächst werden im nachfolgenden Abschnitt die spektralen Eigenschaften im Friktionsgeräusch der drei Sibilanten betrachtet.

In Abb. 4.5 sind die d_{sib} -Werte im $C_1 \times C_2$ Raum für jeden einzelnen Sibilanten zu den Zentroiden /s/ und /ʃ/ getrennt für die kindlichen und die erwachsenen Sprecher dargestellt. Die Gruppe der erwachsenen Sprecher wurde außerdem nach Geschlecht getrennt betrachtet. Der Grund hierfür liegt in den Ergebnissen von Żygiś und Hamann

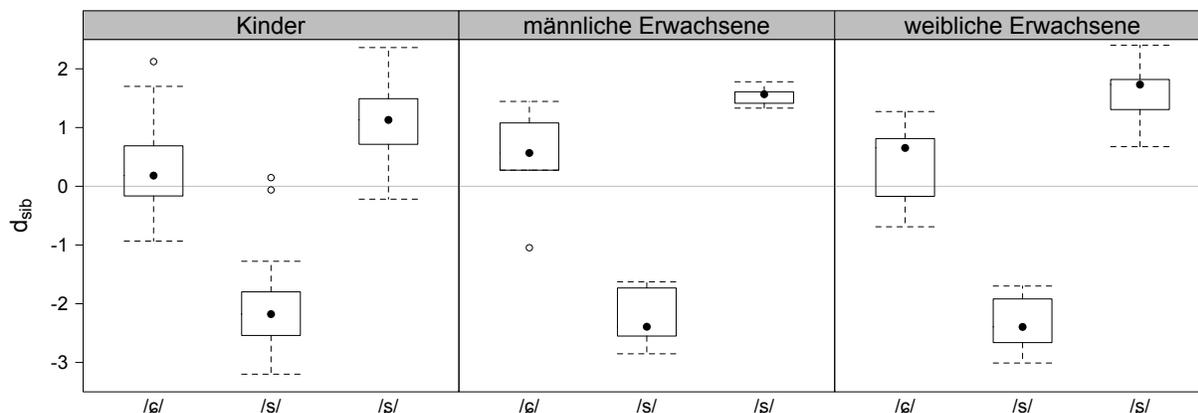


Abbildung 4.5: Logarithmisierte Verhältnisse der Euklidischen Distanzen zu den Mittelwerten von dental /s/ und retroflex /ʂ/ im $C_1 \times C_2$ Raum getrennt für die drei Sibilanten und separat für Kinder (links), männliche Erwachsene (mittig) und weibliche Erwachsene (rechts).

(2003) und Nowak (2006a) zu den polnischen Sibilanten und von F. Li (2008) zu den Sibilanten im Mandarin-Chinesischen, die jeweils in ihren Studien zeigten, dass die drei Sibilantkategorien der weiblichen Probanden anhand der spektralen Eigenschaften in den Sibilanten voneinander getrennt werden können. Im Gegensatz dazu standen die männlichen Versuchspersonen in diesen Studien, für die starke Ähnlichkeiten im ersten spektralen Moment von /ʂ ɕ/ gefunden werden konnten. Für die kindlichen Versuchspersonen wird für die Analyse der drei gewählten Parameter (d_{sib} im $C_1 \times C_2$ Raum, F2 und F3) keine Trennung nach Geschlecht vorgenommen. Laut F. Li et al. (2016) zeigen zwar englischsprachige Kinder im Vorschulalter geschlechtsspezifische Unterschiede in den akustischen Merkmalen des alveolaren Sibilanten /s/, woraus in ihrer Studie geschlossen wurde, dass die sprachliche Variation in Bezug auf geschlechtliche Sprech- und Stimmmerkmale sozialer Natur ist und Kinder damit schon im Vorschulalter ihre geschlechtliche Identität durch bestimmte Parameter aufzeigen können (siehe auch Fitch und Giedd, 1999; Lee, Potamianos und Narayanan, 1999; Perry, Ohde und Ahmead, 2001), jedoch wird aus Gründen des Umfangs und der Übersichtlichkeit dieses Kapitels auf eine geschlechtsspezifische Analyse der kindlichen Daten verzichtet (für einen Vergleich der männlichen und weiblichen kindlichen Sprecher siehe Anhang C.1 Abbildungen C.1-C.3).

Abb. 4.5 zeigt für die kindlichen und erwachsenen Versuchspersonen dieser Studie, dass

sowohl die Sibilanten der Kinder als auch die der Erwachsenen anhand des ersten (C_1 : Steigung) und zweiten (C_2 : Krümmung) DCT-Koeffizienten grundsätzlich gut voneinander unterschieden werden können. Dabei liegen die d_{sib} -Werte des alveolopalatalen Sibilanten / ζ / für alle Sprechergruppen zum Großteil oberhalb der Nulllinie, was bedeutet, dass der alveolopalatale Sibilant in den gemessenen spektralen Merkmalen generell etwas näher am retroflexen Laut liegt als am dentalen Frikativ. Jedoch ist für die L1-Polnisch sprechenden Kinder anhand der niedrigeren positiven Median-Werte zu erkennen, dass hier der alveolopalatale Laut nicht so nah am retroflexen Laut liegt, wie es für die erwachsenen Versuchspersonen der Fall ist. Des Weiteren ist eine Tendenz dahingehend zu erkennen, dass, verglichen mit den beiden erwachsenen Sprechergruppen, die Kategorien der kindlichen Probanden grundsätzlich etwas weniger distinktiv produziert wurden, d.h. für den dentalen und den retroflexen Sibilanten werden jeweils etwas weniger niedrige beziehungsweise etwas weniger hohe d_{sib} -Werte gefunden. Für die erwachsene Versuchspersonengruppe ist der alveolopalatale Sibilant stärker in die Richtung der retroflexen Kategorie verschoben, was an den höheren positiven d_{sib} -Werten dieses Lautes zu sehen ist. Am besten sichtbar wird diese Beobachtung an den Medianen, welche für / ζ / bei allen Sprechergruppen oberhalb der Nulllinie liegen, wobei eben für die Kinder die niedrigsten positiven Werte gefunden werden, während die Mediane der weiblichen und männlichen Erwachsene relativ ähnliche Werte für den Median des alveolopalatalen Lautes aufweisen. Jedoch kann für die erwachsene Sprechergruppe weiterhin beobachtet werden, dass die drei Sibilantkategorien der weiblichen und männlichen Versuchspersonen generelle Unterschiede in ihrer Verteilung aufzeigen: der retroflexe Sibilant der weiblichen erwachsenen Sprecher ist weiter in den positiven Bereich verschoben als der retroflexe Frikativ der männlichen Sprecher. Auch für die dentalen Laute sind leichte geschlechtsspezifische Unterschiede zu sehen: hier scheint es als sei der dentale Laute der weiblichen Probanden im Vergleich zum /s/ der männlichen Sprecher, etwas stärker in den negativen Wertebereich geschoben. Die Beobachtungen zu den spektralen Merkmalen der Sibilanten wurden in zwei gemischten Modellen – eins für die kindlichen Daten und das zweite Modell für die erwachsenen Probanden – quantifiziert. Für das Modell, welches auf die kindlichen Daten angewendet wurde, war der abhängige Faktor d_{sib} , während der unabhängige Faktor der Konsonant (drei Stufen: /s/ vs. / ζ / vs. / ζ /) war. Außerdem wurden die Faktoren Versuchsperson und Wort als Randomfaktoren in das Modell integriert. Dieses Modell, das auf die kindlichen Daten angewendet wurde, zeigte einen signifikanten Effekt des Faktors Konsonant auf die abhängige Variable d_{sib} ($\chi^2 [2] = 27.1, p < 0.001$). Post-hoc Tests zeigten für alle drei Stufen des Faktors Konsonant

signifikante Unterschiede ($/s\ \epsilon/$, $p < 0.001$; $/\text{ʃ}\ \epsilon/$, $p < 0.001$; $/\text{ʃ}\ s/$, $p < 0.001$). Das zweite Modell, welches auf die Daten der erwachsenen Sprecher gerechnet wurde, testete als abhängige Variable d_{sib} und als unabhängige Faktoren Konsonant (drei Stufen: $/s/$ vs. $/\text{ʃ}/$ vs. $/\epsilon/$) und Geschlecht (zwei Stufen: weiblich vs. männlich) und integrierte zusätzlich die Randomfaktoren Versuchsperson und Wort. Dieses Modell brachte weder eine Interaktion der Faktoren Geschlecht und Konsonant, noch konnte ein Effekt des Faktors Geschlecht festgestellt werden. Der Faktor Konsonant zeigte einen signifikanten Einfluss auf die abhängige Variable ($\chi^2 [3] = 31.3$, $p < 0.001$). Post-hoc Tests zeigten ebenfalls für alle drei Stufen des Faktors Konsonant signifikante Unterschiede in den d_{sib} -Werten ($/s/-/\epsilon/$, $p < 0.001$; $/\text{ʃ}/-/\epsilon/$, $p < 0.001$; $/\text{ʃ}/-/s/$, $p < 0.001$)

Die Statistik zeigt zwar signifikante Unterschiede zwischen allen Konsonanten für alle Sprechergruppen, dennoch können in Tab. 4.4, für welche die d_{sib} -Werte pro Konsonant und pro Sprechergruppe gemittelt wurden, kleine Unterschiede zwischen den Sprechergruppen und eine Tendenz zu ähnlicheren d_{sib} -Werten zwischen $/\text{ʃ}\ \epsilon/$ aufgezeigt werden.

	Kinder	männliche Erwachsene	weibliche Erwachsene
$/s/$	-2.1	-2.3	-2.4
$/\text{ʃ}/$	1.1	1.5	1.7
$/\epsilon/$	0.3	0.5	0.4

Tabelle 4.4: Gemittelte d_{sib} -Werte separat für die drei Sibilanten $/s\ \text{ʃ}\ \epsilon/$ und getrennt für kindlichen, die männlichen erwachsenen und die weiblichen erwachsenen Versuchspersonen.

In Tab. 4.4 sieht man die Beobachtungen zu Abb. 4.5 bestätigt. Zum einen wurden für alle Sprechergruppen positive Werte für den alveopalatalen Sibilanten gefunden, was bedeutet, dass dieser Laut in die Richtung des retroflexen Frikatives geschoben ist. Die Werte der Kinder ($d_{\text{sib}} = 0.3$) liegen dabei am nächsten an der Nulllinie, gefolgt von den weiblichen Versuchspersonen, die mit einem d_{sib} -Mittelwert von 0.4 ein Stück weiter von der Nulllinie entfernt liegen, wobei die männlichen Probanden den höchsten d_{sib} -Wert von 0.5 vorweisen und damit letztere in dem gemessenen Parameter d_{sib} im $C_1 \times C_2$ -Raum dem retroflexen Sibilanten am ähnlichsten sind. Für die erste Frage, nämlich ob eine Trennung von $/s/$ und $/\text{ʃ}\ \epsilon/$ in den spektralen Merkmalen im Friktionsanteil gefunden werden kann, während $/\text{ʃ}\ \epsilon/$ im Friktionsanteil Ähnlichkeiten zeigen, kann anhand der Beobachtungen in Abb. 4.5 und Tab. 4.4 bestätigt werden. Genauer gesagt kann somit das erste Ergebnis aus Kapitel 2 bezüglich der Verteilung der drei Sibilanten mittels der jeweiligen spektralen Eigenschaften im Sibilanten anhand der rein akustischen Daten dieser hier vorliegenden Studie repliziert werden.

Als nächstes sollte getestet werden ob es möglich ist, die Ergebnisse in Kapitel 2 bezüglich der F2-Werte zu replizieren, d.h. eine Überlappung von /s ʃ/ zum zeitlichen Vokalonsset und eine klare Trennung des alveopalatalen Sibilanten durch höhere F2-Werte zu diesem Zeitpunkt im Vokal.

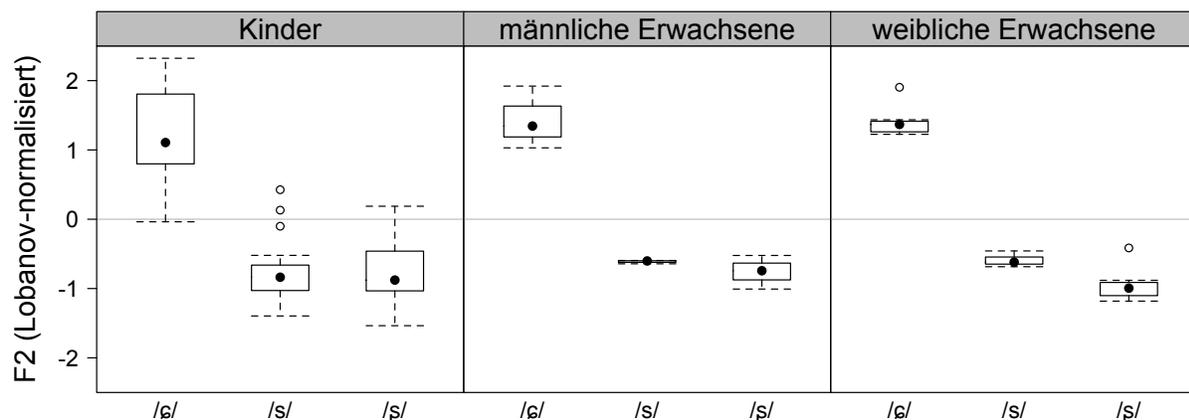


Abbildung 4.6: Lobanov-normalisierte F2-Werte zum zeitlichen Vokalonsset für den Vokal /ε/, welcher den jeweiligen Sibilanten /ç/ (links), /s/ (mittig) und /ʃ/ (rechts) folgte, getrennt für die kindlichen (links), die männlichen (mittig) und die weiblichen (rechts) erwachsenen Versuchspersonen.

Abb. 4.6 zeigt getrennt für die drei Sprechergruppen Kinder (links), männliche Erwachsene (mittig) und weibliche Erwachsene (rechts) die Lobanov-normalisierten F2-Werte zum zeitlichen Onset des Vokals, welcher den drei Sibilanten /ç/ (links), /s/ (mittig) und retroflex (rechts) folgt. Es ist zu erkennen, dass mittels der F2-Werte durchweg für alle Sprechergruppen der alveopalatale Frikativ /ç/ gut von den beiden anderen Sibilanten separiert werden kann. Weiterhin ist ebenfalls für alle drei Sprechergruppen zu beobachten, dass die F2-Werte von /s ʃ/ sehr nahe beieinander liegen. Am stärksten ausgeprägt ist diese Überlappung von /s ʃ/ für die kindlichen und die männlichen erwachsenen Probanden, bei denen die Mediane des dentalen und des retroflexen Frikatives fast auf einer Linie liegen. Etwas anders verhält es sich mit den weiblichen erwachsenen Versuchspersonen, deren F2-Werte für /s ʃ/ nicht überlappen, wie es für die Kinder und die männlichen Erwachsenen der Fall ist, sondern etwas getrennt voneinander liegen. Nichtsdestoweniger liegen auch die F2-Werte von /s ʃ/ der weiblichen erwachsenen Sprecher so nah beieinander, dass man

für alle drei Sprechgruppen sagen kann, dass das in Kapitel 2 gezeigte Ergebnis in dieser akustischen Studie für alle drei Sprechergruppen repliziert werden kann, d.h. eine klare Trennung von / ϵ / mit hohen F2-Werten zum Vokalansatz im nachfolgenden Vokal und sehr ähnliche F2-Werte für / s ζ /. Quantifiziert wurden diese Beobachtungen mit zwei gemischten Modellen. Ein Modell pro Altersgruppe mit F2 als abhängige Variable und Konsonant als fixen Faktor, sowie mit Wort und Versuchsperson als Randomfaktoren. Für die Gruppe der erwachsenen Sprecher war neben Konsonant, auch Geschlecht ein fixer Faktor. Das gemischte Modell der Kinder zeigte, dass Konsonant einen signifikanten Einfluss auf F2 hat ($\chi^2 [2] = 27.94$, $p < 0.001$). Post-hoc Tests zeigten signifikante Unterschiede in F2 zwischen / s ϵ / ($p < 0.001$), ebenso zwischen / ζ ϵ / ($p < 0.001$), nicht aber zwischen / s ζ /. Das Modell der erwachsenen Sprecher zeigte keine signifikante Interaktion zwischen Geschlecht und Konsonant. Die Faktoren Konsonant und Geschlecht wurden einzeln getestet, wobei das Ergebnis zeigte, dass F2 zwar signifikant durch den Faktor Konsonant ($\chi^2 [2] = 33.9$, $p < 0.001$) beeinflusst wurde, nicht jedoch von Geschlecht. Post-hoc Tests für den Faktor Konsonant zeigten, dass sich die Sibilanten in F2 signifikant voneinander unterscheiden (/ s ϵ /, $p < 0.001$; / ζ ϵ /, $p < 0.001$; / s ζ /, $p < 0.001$). Zwar zeigt die Statistik für die erwachsenen Versuchspersonen einen signifikanten Unterschied zwischen dental / s / und retroflex / ζ /, betrachtet man jedoch diese Sprechergruppe in Abb. 4.6 so wäre es möglich weiterhin von einer größeren Ähnlichkeit zwischen / s ζ / hinsichtlich der F2-Werte zu sprechen, während im Gegensatz dazu / s ϵ / sowie / ζ ϵ / im Vergleich zu ersterem Paar wesentlich unterschiedlichere Werte zeigen.

Als letzten Parameter zur Unterscheidung der Sibilanten werden die Werte des dritten Formanten in den Vokal hinein, welcher den jeweiligen Sibilanten / s ζ ϵ / folgt, betrachtet. Hier werden niedrige F3-Werte im Übergang vom retroflexen Sibilanten in den Vokal angenommen, welche diesen Laut von den beiden anderen (/ s ϵ /) trennen.

Abb. 4.7 zeigt getrennt nach Altersgruppe (Kinder vs. männliche Erwachsene vs. weibliche Erwachsene), sowie separat für den alveopalatalen (jeweils linke Box), den dentalen (jeweils mittige Box) und den retroflexen Sibilanten (jeweils rechte Box) die F3-Werte zum zeitlichen Onset des Vokals, welcher den drei Sibilanten folgt. Für alle drei Altersgruppen ist zu erkennen, dass der alveopalatale Frikativ die höchsten F3-Werte zum nachfolgenden Vokalansatz hat. Für die kindlichen Sprecher und die weiblichen Erwachsenen hat jeweils / ζ / die niedrigsten, während der dentale Sibilant / s / Werte zeigt, die

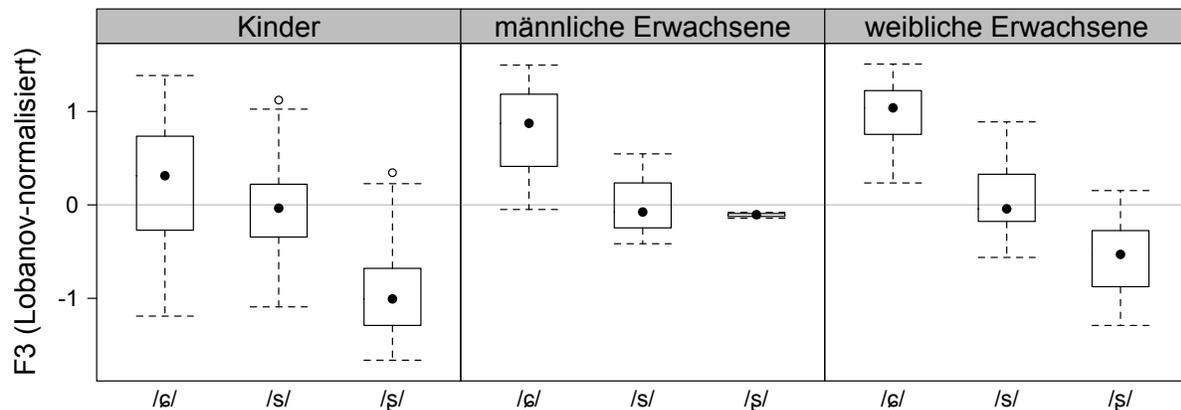


Abbildung 4.7: Lobanov-normalisierte F3-Werte zum zeitlichen Vokalonset für den Vokal /ε/, welcher den jeweiligen Sibilanten /ε/ (links), /s/ (mittig) und /ʃ/ (rechts) folgte, getrennt für die kindlichen (links), die männlichen (mittig) und die weiblichen (rechts) erwachsenen Versuchspersonen.

zwischen /ʃ ε/ liegen. Die männlichen Versuchspersonen dagegen weisen ein etwas anderes Muster hinsichtlich der F3-Werte, die /s ʃ/ folgen, auf: hier ist eine starke Überlappung in F3 zu erkennen. Hervorgerufen wird diese Ähnlichkeit in den F3-Werten des dentalen und des retroflexen Frikatives durch höhere F3-Werte des retroflexen Sibilanten /ʃ/. Für die männlichen Versuchspersonen sieht es so aus als näherte sich /ʃ/ an /s/ an. Quantifiziert wurde diese Beobachtung mit einem gemischten Modell pro Altersgruppe. Für das Modell, welches auf die Daten der kindlichen Probanden gerechnet wurde, wurde F3 als abhängige Variable, Konsonant als fixer Faktor, sowie die Randomfaktoren Versuchsperson und Wort integriert. Das Ergebnis dieses Modells zeigte einen signifikanten Einfluss von Konsonant auf die F3-Werte ($\chi^2 [2] = 10.6, p < 0.01$). Post-hoc Tests für die drei Stufen des Faktors Konsonant brachten signifikante Unterschiede zwischen den Konsonanten /ʃ s/ ($p < 0.01$) und zwischen /ʃ ε/ ($p < 0.05$), nicht jedoch zwischen den Konsonanten /s ε/. Das zweite Modell, welches auf die Daten der erwachsenen Sprecher angewendet wurde, integrierte F3 als abhängige Variable, Geschlecht und Konsonant als fixe Faktoren, sowie die Randomfaktoren Versuchsperson und Wort. Dieses Modell ergab eine signifikante Interaktion zwischen den Variablen Geschlecht und Konsonant ($\chi^2 [2] = 7.5, p < 0.05$). Post-hoc Tests zeigten signifikante Unterschiede zwischen den folgenden Paaren: /s/.männliche Erwachsene - /ε/.männliche Erwachsene ($p < 0.05$), /ʃ/.männliche Erwachsene - /ε/.männliche

Erwachsene ($p < 0.01$), /s/.weibliche Erwachsene - /ɕ/.weibliche Erwachsene ($p < 0.001$), /ɕ/.weibliche Erwachsene - /ʂ/.weibliche Erwachsene ($p < 0.001$), /ʂ/.weibliche Erwachsene - /s/.weibliche Erwachsene ($p < 0.01$). Für die Paarung /ʂ/.männliche Erwachsene - /s/.männliche Erwachsene konnte jedoch kein signifikanter Unterschied in den F3-Werten gefunden werden. Dieses Ergebnis bestätigt die Beobachtungen in Abb. 4.7, nämlich dass nur die männliche Sprechergruppe eine Überlappung der Sibilanten /s ʂ/ in den F3-Werten zeigt, während für die weiblichen und die kindlichen Sprecher der retroflexe Sibilant durch niedrige F3-Werte von /s ɕ/ getrennt werden kann.

Im zweiten Schritt dieser Studie wurde getestet, inwiefern sich die Kinderdaten von den Erwachsenendaten unterscheiden. Spezifischer war die Frage, ob Kinder die drei Kategorien generell weniger stark trennen als die Erwachsenen. Um dies zu testen wurden, wie in Abschnitt 4.4.3 beschrieben, die Differenzen, zum einen zwischen den einzelnen d_{sib} -Werten der Sibilanten, sowie zum anderen zwischen den jeweiligen F2 und F3-Werten der drei polnischen Frikative berechnet.

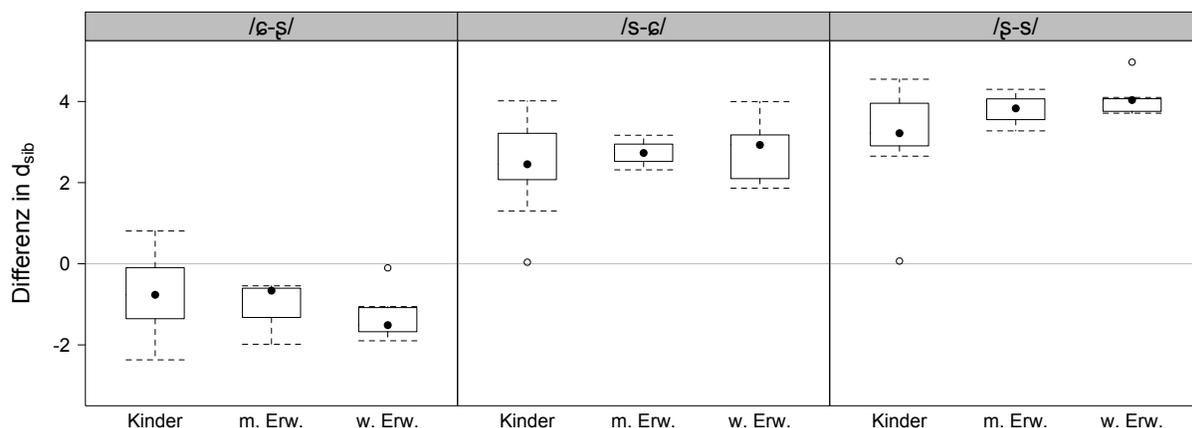


Abbildung 4.8: Differenzen der d_{sib} -Werte zwischen alveolopalatal und retroflex (linkes Panel), zwischen dental und alveolopalatal (Panel Mitte) und zwischen retroflex und dental (rechtes Panel) getrennt für die kindlichen (Boxen links), die männlichen (Boxen mittig) und die weiblichen (Boxen rechts) erwachsenen Sprecher.

In Abb. 4.8 werden zunächst die Differenzen in den d_{sib} -Werten der einzelnen Sibilantpaare, d.h. Differenz zwischen alveolopalatal und retroflex (linkes Panel), Differenz zwischen dental und alveolopalatal (Panel Mitte) und Differenz zwischen retroflex und dental

(rechtes Panel) getrennt für die Sprechergruppen (Kinder jeweils linke Box; männliche Erwachsene jeweils Box mittig; weibliche Erwachsene jeweils rechte Box) näher betrachtet. Es ist vor allem an den Medianen in Abb. 4.8 klar ersichtlich, dass die Differenzen zwischen den Sibilanten der Kinder, grundsätzlich für alle Sibilantpaare näher an der Nulllinie liegen als die der L1-Polnisch sprechenden Erwachsenen. Diese Beobachtung stützt auch die gemittelten d_{sib} -Daten in Tab. 4.4, die schon zeigten, dass für die Kinder grundsätzlich niedrigere (d.h. weniger hohe positive Werte für /s/ und weniger weit in den negativen Bereich reichende Werte für /ʂ/) d_{sib} -Werte gefunden werden, woraus geschlossen werden kann, dass hier die Kategorien nicht so gut getrennt sind, wie es beispielsweise für die weiblichen erwachsenen Probanden der Fall ist. Für die weiblichen Erwachsenen wird gefunden, dass diese Sprechergruppe für /s/ und /ʂ/ Werte zeigt, die verglichen mit den anderen Sprechergruppen am weitesten auseinander liegen (vgl. auch Tab. 4.4). Für alle Sibilantpaare liegen die Differenzwerte der weiblichen erwachsenen Versuchspersonen am weitesten von der Nulllinie entfernt (sowohl für den negativen Bereich für /s/ als auch für den positiven Bereich für /ʂ/), während die Kinder der Nulllinie am nächsten liegen, wobei die Differenzwerte der männlichen erwachsenen Probanden zwischen den beiden ersten angeordnet sind. Interessant ist die Beobachtung, dass der Sprung von den weiblichen erwachsenen Sprechern über die männlichen erwachsenen Sprecher hin zu den Kindern jeweils für die Differenzen in d_{sib} am größten ist, für welche die Differenz des retroflexen Sibilanten zu einem der beiden anderen Laute berechnet wurde, d.h. für die Differenzen zwischen den Sibilanten /ɕ-ʂ/ und /ʂ-s/. Die Unterschiede zwischen den Sprechergruppen in den Differenzwerten zwischen den Sibilanten /s-ɕ/ (also zwischen den Sibilanten, deren Kontrast nicht von der diachronen Neutralisierung in den polnischen dialektalen Varietäten betroffen ist) scheinen, im Vergleich zu den anderen Differenzen (/ɕ-ʂ/ und /ʂ-s/), nicht so ausgeprägt. Diese Beobachtungen wurden quantifiziert durch ein gemischtes Modell mit der Differenz in d_{sib} zwischen den Konsonanten als abhängige Variable und mit Konsonantpaar (drei Stufen: /ɕ-ʂ/ vs. /ʂ-s/ /s-ɕ/ vs. /s-ɕ/) und Sprechergruppe (drei Stufen: Kinder vs. männliche Erwachsene vs. weibliche Erwachsene) als fixe Faktoren, sowie mit Versuchsperson als Randomfaktor². Dieses Modell brachte keine Interaktion zwischen Konsonantpaar und Sprechergruppe hervor, jedoch zeigte die Analyse einen Trend hin zu einer Interaktion der beiden Faktoren Sprechergruppe und Konsonantpaar. Beim Testen der Faktoren,

²Für die Berechnung der Differenz zwischen den Werten der getesteten Parameter (d_{sib} , F2 und F3) wurde über den Faktor Wort gemittelt, da die Wortanzahl für die kindlichen Versuchspersonen, aufgrund von Datenverlusten durch Versprecher, nicht ausgeglichen waren. Das hatte zur Folge, dass in die gemischten Modelle, die auf die jeweiligen Differenzwerte gerechnet wurden, Wort nicht mehr als Randomfaktor integriert werden konnte.

konnte kein signifikanter Unterschied zwischen den Sprechergruppen festgestellt werden. Der Faktor Konsonantpaar zeigte hingegen den erwarteten signifikanten Einfluss auf die abhängige Variable Differenz zwischen den d_{sib} -Werten. Post-hoc Tests zeigten für alle Konsonantpaare Unterschiede in den Differenzwerten ($/s-ε/ - /ε-ε/$, $p < 0.001$; $/ε-s/ - /ε-ε/$, $p < 0.001$; $/ε-s/ - /s-ε/$, $p < 0.001$). Dieses Ergebnis, d.h. das Fehlen eines Effekts der Sprechergruppe auf die Differenzwerte, ist vermutlich der Überlappung der Boxen geschuldet, d.h. die Mediane der Versuchspersonengruppen sind zwar unterschiedlich weit von der Nulllinie entfernt, jedoch überlappen die Boxen der Versuchspersonengruppen stark miteinander, was einen signifikanten Effekt verhindert haben könnte.

Analog zu den Differenzen in den d_{sib} -Werten, wurden im nächsten Schritt die Differenzen zwischen den F2-Werten der einzelnen Sibilanten berechnet, um zu testen, ob die kindlichen Probanden die drei Sibilantkategorien in gleichem Maße in diesem Parameter trennen, wie es bei den L1-Polnisch sprechenden Erwachsenen der Fall ist.

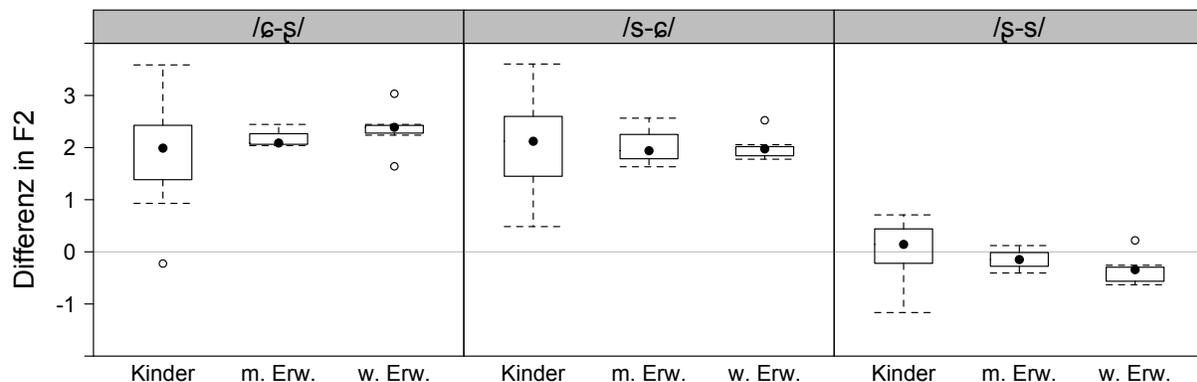


Abbildung 4.9: Differenzen der F2-Werte zum zeitlichen Onset des Vokals $/ε/$, der den Sibilanten $/s \ \epsilon \ \epsilon/$ folgt: Differenz zwischen Alveolopalatal und Retroflex linkes Panel, zwischen Dental und Alveolopalatal Panel Mitte und zwischen Retroflex und Dental rechtes Panel getrennt für die kindlichen (jeweils Boxen links), männlichen (jeweils Boxen mittig) und weiblichen (jeweils Boxen rechts) erwachsenen Versuchspersonen.

Abb. 4.9 zeigt getrennt für die Kinder (Box links), die männlichen (Boxen mittig) und die weiblichen (Boxen rechts) Erwachsenen die jeweiligen Differenzen in den F2-Werten zu Beginn des Vokals, welcher den einzelnen Sibilanten nachfolgte: Differenz zwischen $/ε/$ und $/ε/$ (Panel links), zwischen $/s/$ und $/ε/$ (mittig), sowie zwischen $/ε/$ und $/s/$ (Panel

rechts). In dieser Abbildung ist zu erkennen, dass die kindlichen Versuchspersonen nicht generell näher an der Nulllinie liegen, verglichen mit den erwachsenen Probanden. Betrachtet man die Mediane, ist dies insbesondere für die Differenz zwischen /s-ç/ nicht der Fall, für welche zu sehen ist, dass die Werte der Kinder weiter von der Nulllinie entfernt liegen. Was weiterhin auffällt, ist, dass die männlichen erwachsenen Probanden den kindlichen Differenzwerten ähnlicher sind, verglichen mit den weiblichen Probanden, deren Medianwerte etwas peripherer liegen. Dazu ist außerdem zu sagen, dass diese Beobachtungen vor allem auf die F2-Differenzen zwischen /ç/ und /ʃ/ (Panel links), sowie zwischen /s/ und /ʃ/ (Panel rechts) zutreffen, nicht jedoch auf die Differenzen in den F2-Werten zwischen /s/ und /ç/. Letztere Differenzwerte zeigen kaum Unterschiede zwischen männlichen und weiblichen erwachsenen Versuchspersonen bezogen auf den Median. Für die Kinder ist in diesem Kontrast der Median sogar weiter von der Nulllinie entfernt verglichen mit den beiden erwachsenen Versuchspersonengruppen. Diese Beobachtungen wurden durch ein gemischtes Modell mit der Differenz in den F2-Werten im Vokal, welcher den einzelnen Sibilanten nachfolgte als abhängige Variable, mit Konsonantpaar (drei Stufen: /ç-ʃ/ vs. /s-ç/ vs. /ʃ-s/) und Sprechergruppe (drei Stufen: Kinder vs. männliche Erwachsene vs. weibliche Erwachsene) als fixe Faktoren, sowie mit Versuchsperson als Randomfaktor, geprüft. Wie zu erwarten, zeigte dieses Modell einen signifikanten Einfluss von Konsonantpaar ($\chi^2 [2] = 114.1, p < 0.001$) auf die Differenz in den F2-Werten, nicht jedoch von Sprechergruppe und auch keine signifikante Interaktion der beiden fixen Faktoren. Post-hoc Tests für den Faktor Konsonantpaar brachten folgende Ergebnisse hervor: signifikanter Unterschied zwischen den Paaren /ʃ-s/ und /s-ç/ ($p < 0.001$), sowie zwischen den Paaren /ʃ-s/ und /ç-ʃ/ ($p < 0.001$), kein signifikanter Effekt zwischen den Paaren /s-ç/ und /ç-ʃ/.

In Abb. 4.10 sind die Differenzen der F3-Werte der drei Sibilanten dargestellt. Hier ist zu erkennen, dass die Paare /ç-ʃ/ und /s-ç/ für alle Sprechergruppen gut voneinander getrennt sind. Eine Ausnahme stellt hier die kindliche Versuchspersonengruppe dar, für welche die F3-Differenzwerte zwischen /s-ç/ sehr nah an der Nulllinie liegen, was dafür spricht, dass die F3-Werte dieser beider Sibilantkategorien relativ nah beieinander liegen. Weiterhin fällt auf, dass die F3-Differenzwerte zwischen /ʃ-s/ der männlichen Sprecher kaum von der Nulllinie entfernt liegen, d.h. es gibt hier kaum einen Unterschied in den F3-Werten zwischen /s ʃ/. Die beiden anderen Sprechergruppen, d.h. die Kinder und die weiblichen Erwachsenen, zeigen hingegen einen Unterschied in den F3-Differenzwerten der beiden Sibilanten /s ʃ/ auf. Ein gemischtes Modell mit den F3-Differenzwerten als abhän-

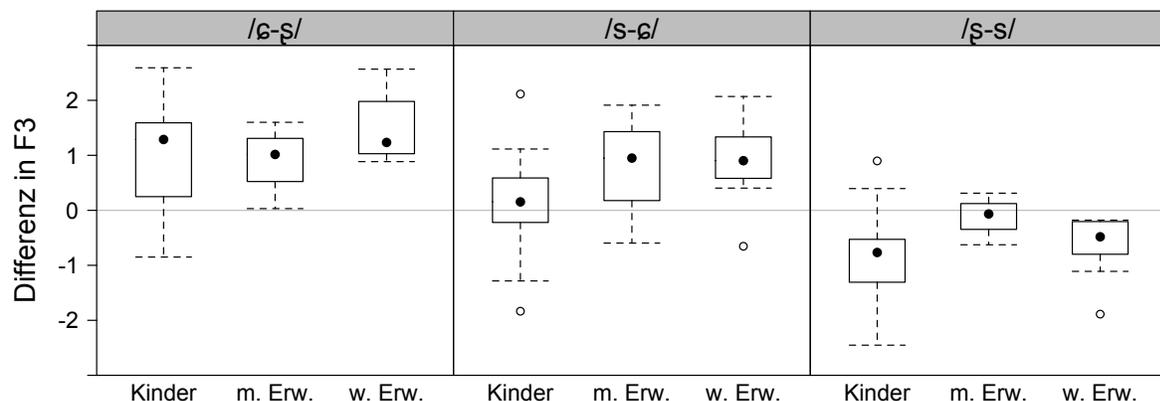


Abbildung 4.10: Differenzen in den F3-Werten zum zeitlichen Onset des Vokals /ɛ/, der den Sibilanten /s ʂ ɕ/ folgt: Differenz zwischen Alveolopalatal und Retroflex linkes Panel, zwischen Dental und Alveolopalatal Panel Mitte und zwischen Retroflex und Dental rechtes Panel getrennt für die kindlichen (jeweils Boxen links), die männlichen (jeweils Boxen mittig) und die weiblichen (jeweils Boxen rechts) erwachsenen Sprecher.

gige Variable und den fixen Faktoren Sprechergruppe (drei Stufen: Kinder vs. männliche Erwachsene vs. weibliche Erwachsene) und Konsonantpaar (drei Stufen: /ɕ-ʂ/ vs. /s-ɕ/ vs. /ʂ-s/), sowie mit Versuchsperson als Randomfaktor, brachte keine signifikante Interaktion zwischen den beiden fixen Faktoren und ebenfalls keinen signifikanten Einfluss von Sprechergruppe hervor. Ein signifikanter Unterschied konnte für die Konsonantpaare ($\chi^2 [2] = 62.4, p < 0.001$) gezeigt werden. Post-hoc Tests ergaben einen signifikanten Unterschied zwischen allen drei Konsonantpaaren (/ʂ-s/ - /s-ɕ/, ($p < 0.001$); /ʂ-s/ - /ɕ-ʂ/, ($p < 0.001$), /s-ɕ/ - /ɕ-ʂ/, ($p < 0.001$)).

Wie in der Einleitung dieses Kapitels aufgeführt, beschäftigte sich die dritte Frage dieser Studie mit der diachronen Neutralisierung des Dreier-Kontrasts in den Sibilanten des Polnischen, wie sie für einige dialektale Varietäten des Standard-Polnischen gefunden wurde. Zum einen sollten hier in einem Vergleich von kindlichen und erwachsenen L1-Polnisch sprechenden Versuchspersonen weitere Erklärungsansätze für die größere Instabilität des retroflexen Sibilanten gefunden werden. Zum anderen sollte nach möglichen Erklärungen gesucht werden, die gerade den Zusammenfall von dental /s/ und retroflex /ʂ/ in der hier beschriebenen diachronen Neutralisierung dieser Sibilanten, begründen könnten. In den Ergebnissen zu den Differenzen zwischen den d_{sib} -Werten der drei Sibilanten, sowie auch

zwischen den F2/F3-Werten zum zeitlichen Onset des Vokals, welcher den drei Sibilanten folgte, wurde beobachtet, dass die weiblichen Sprecher für die Differenzen zwischen den Sibilanten /ç š/ und zwischen den Sibilanten /ʃ s/ größere Differenzwerte zeigten als die kindlichen Probanden, mit Differenzwerten der männlichen Erwachsenen zwischen denen der beiden anderen Sprechergruppen. Im Gegensatz dazu bleiben die Differenzwerte in d_{sib} und in F2 zwischen den Sibilanten /s ç/ über die Sprechergruppen relativ konstant. Aus dieser Beobachtung heraus, wäre es möglich zu sagen, dass der Kontrast zwischen /ʃ s/ und der Kontrast zwischen /ç š/, welcher jeweils nur in einem Merkmal gefestigt ist (/ʃ s/ kontrastiert in den spektralen Merkmalen im Friktionsgeräusch; /ç š/ kontrastiert in den F2-Werten zum zeitlichen Vokalansatz), etwas instabiler ist als der Kontrast zwischen /s ç/, welcher auf zwei akustischen Merkmalen basiert. Klar ist jedoch auch vor diesem Hintergrund noch immer nicht, warum gerade der /s š/-Kontrast in den diachronen Neutralisierungen zusammenfällt. Für diese Frage sind die Ergebnisse aus den Daten der F2/F3-Werte (siehe Abb. 4.6 und 4.7) und den Differenzberechnungen zwischen den post-sibilantischen F2/F3-Werten im Vokalansatz (Abb. 4.9 und 4.10) genauer zu betrachten.

In den Abbildungen zu den F2-Werten ist zu sehen, dass die kindlichen und die männlichen erwachsenen Sprecher für die Kategorien /s š/ überlappende F2-Werte zeigen. Diese Beobachtung wird auch durch die F2-Differenzwerte zwischen den Sibilanten (Abb. 4.9) bestätigt. Während die Mediane der kindlichen Probanden wie auch die der männlichen Erwachsenen sehr nah an der Nulllinie liegen, zeigen die weiblichen Erwachsenen F2-Differenzwerte zwischen /s š/ mit dem größten Abstand zur Nulllinie auf. Der zweite Aspekt, der hier eine Rolle spielen könnte sind die Muster in den F3-Werten zum Vokalansatz im Anschluss an die jeweiligen Sibilanten. Es wurde in der Studie in Kapitel 2 gefunden, dass der retroflexe Sibilant einen niedrigeren F3-Wert zeigt, was seine Unterscheidung von /s ç/ unterstützt. Betrachten wir jedoch Abb. 4.7 so ist dies nur für zwei der drei Sprechergruppen verifizierbar. Für die kindlichen und die weiblichen erwachsenen Sprecher sind unterschiedliche F3-Werte für alle drei Sibilanten zu sehen, wobei die F3-Werte für /ʃ/ am niedrigsten sind. Die männlichen Erwachsenen zeigen jedoch starke Ähnlichkeiten in den F3-Werten von /s ç/. Quantifiziert wurden die Beobachtungen hinsichtlich der F2-Werte in drei gemischten Modellen, d.h. je ein Modell pro Sprechergruppe (Kinder, männliche Erwachsene und weibliche Erwachsene), welches dann einzig auf die F2-Werte der Konsonanten /s š/ gerechnet. Dabei waren in diesen Modellen die F2-Werte als abhängige Variable, Konsonant (zwei Stufen: /s/ vs. /ʃ/) als fixer Faktor und Versuchsperson und Wort als Randomfaktoren eingebunden. Während das Modell der weiblichen Sprechergruppe einen signifikanten Ef-

fekt des Faktors Konsonant ($\chi^2 [1] = 9.2, p < 0.01$) auf die F2-Werte zeigte, konnte für das Modell der kindlichen und das Modell der männlichen erwachsenen Sprechergruppe kein signifikanter Unterschied in den F2-Werten der beiden Konsonanten gefunden werden. Ebenso wurden drei gemischte Modelle – jeweils eins pro Sprechergruppe – allein auf die F3-Werte der Konsonanten /s ʂ/ angewendet. Jedes dieser Modelle integrierte die F3-Werte als abhängige Variable, Konsonant (zwei Stufen: /s/ vs. /ʂ/) als fixen Faktor, sowie Versuchsperson und Wort als Randomfaktoren. Für diese drei Modelle war das Ergebnis wie folgt: während für das Modell der kindlichen ($\chi^2 [1] = 13.8, p < 0.001$) und das Modell der weiblichen Sprechergruppe ($\chi^2 [1] = 6.6, p < 0.05$) ein signifikanter Effekt von Konsonant auf die F3-Werte gezeigt wurde, brachte das Modell der männlichen erwachsenen Sprechergruppe keinen signifikanten Einfluss von Konsonant auf die F3-Werte hervor. Diese Ergebnisse zeigen, dass der /s ʂ/-Kontrast für die kindlichen und die männlichen erwachsenen Probanden in den F2-Werten und für die männlichen erwachsenen Sprecher auch in den F3-Werten zum zeitlichen Vokalonset nicht aufrechterhalten wird.

4.6 Diskussion

Die hier vorliegende Studie beschäftigte sich mit drei generellen Fragen zu den Sibilanten des Polnischen. Für die erste Frage sollte geklärt werden, ob es möglich ist die rein akustisch erhobenen Sibilantproduktionen mit den artikulatorischen Aufnahmen (elektromagnetische Artikulografie) in Kapitel 2, in Bezug auf die spektralen Merkmale im Sibilanten als auch in den Formantwerten zum zeitlichen Onset des Vokals, welcher den Sibilanten nachfolgte, zu vergleichen und zu replizieren. Die zweite Frage befasste sich mit der Vergleichbarkeit der Sibilantproduktionen der L1-Polnisch sprechenden Kinder und den Produktionsdaten der erwachsenen L1-Polnisch sprechenden erwachsenen Versuchspersonen. Spezifischer war die Frage, ob in den fünf- bis sechsjährigen kindlichen Versuchspersonen sprachspezifische Verteilungen der sibilantischen Kategorien bezogen auf bestimmte spektrale Parameter im Friktionsgeräusch der Sibilanten und in den Formantwerten zum Onset des nachfolgenden Vokals zu finden sind, welche mit den Mustern der erwachsenen Probanden vergleichbar sind. In der dritten Frage sollte geklärt werden, ob es möglich ist in den Sprachmustern der kindlichen L1-Polnisch sprechenden Versuchspersonen Erklärungsansätze zu finden, welche Aufschlüsse über die Variabilität des retroflexen Sibilanten und vor allem über die Instabilität des retroflex-dental Kontrasts in der diachronen Neutralisierung der drei Sibilanten in einigen dialektalen Varietäten des Polnischen geben können. Nachfolgend werden diese

drei Fragen diskutiert, angefangen mit der ersten Frage oben in diesem Absatz.

In dieser hier vorliegenden rein akustischen Studie wurden in den spektralen Merkmalen im Sibilanten, gemessen anhand des logarithmisierten Verhältnisses der Euklidischen Distanzen jedes einzelnen Tokens zu den Zentroiden /s/ und /ʃ/, größere Ähnlichkeiten zwischen /ʃ ɕ/ gefunden, während der dentale Sibilant /s/ sehr gut von den beiden anderen separiert wurde. Bezogen auf die Messungen der F2-Werte zum zeitlichen Onset des Vokals /ɛ/, welcher den Sibilanten /s ʃ ɕ/ in den Zielwörtern dieser Studie folgte, wurde in Abb. 4.6 eine Überlappung der niedrigen F2-Werte von /s ʃ/ gefunden, während sich die hohen F2-Werte des alveolopalatalen Sibilanten stark von denen der beiden erst genannten unterschieden. Ähnliche Ergebnisse beziehungsweise damit vergleichbare Unterscheidungsmuster zwischen den polnischen Sibilanten, wurden auch in der physiologischen Studie in Kapitel 2 gefunden. Wir können hier somit zusammenfassen, dass die Ergebnisse aus der Studie von Meenakshi et al. (2014) zur Perturbation der Sprache, und vor allen Dingen der Artikulation der alveolaren Laute der Versuchspersonen durch die Sensoren, welche bei EMA-Aufnahmen auf den Artikulatoren der Probanden befestigt werden, nicht auf die hier und in Kapitel 2 erhobenen Daten zutreffen. Diskutiert vor dem Hintergrund der Perturbation der Sprachproduktion und den daraus resultierenden Konsequenzen für die akustischen Eigenschaften von Sprachlauten ist zu sagen, dass dieser Fund mit schon bekannten Studien zur Perturbation von Sprache und der artikulatorischen Kompensation für die Perturbationen vereinbar sind. Die Kompensation für die durch die Sensoren gestörte Artikulation funktioniert, laut bekannten Studien zur Perturbation der Artikulation, durch auditorisches Feedback und das nachfolgende adaptieren der Artikulation auf das Wahrgenommene, d.h. die Anpassung und Konfiguration der Sprachproduktion an die Störung der Artikulation (vgl. Hoole, 1987; Jones und Munhall). Die Versuchspersonen in Kapitel 2 schafften es, ihre Artikulation an die Sensoren anzupassen und die Artikulation neu zu konfigurieren, sodass die „typischen“ akustischen Eigenschaften der drei Sibilanten beibehalten werden konnten.

Weiterhin wurden in dieser Studie die akustischen Eigenschaften der drei Sibilanten /s ʃ ɕ/, produziert von kindlichen und von erwachsenen L1-Polnisch Sprechern verglichen, um die Hypothese zu testen, ob fünfjährige Kinder schon sprachspezifische Muster in ihren Sibilantproduktionen zeigen. Verglichen werden können die Ergebnisse dieser Studie mit den Ergebnissen aus F. Li (2008) und F. Li und Munson (in Druck), die in ihren Ex-

perimenten Sibilantproduktionen von L1-Mandarin sprechenden Kindern mit denen von L1-Mandarin sprechenden Erwachsenen verglichen. Dabei wurde gefunden, dass Kinder ab einem Alter von fünf Jahren fähig sind, den schwierigen Kontrast zwischen den drei Sibilanten /s ʃ ɕ/ wie er sowohl im Polnischen als auch im Mandarin-Chinesischen gefunden werden kann, in einer mit erwachsenen Sprechern vergleichbaren Art und Weise zu produzieren. Auffallend war im Vergleich der kindlichen und erwachsenen Sibilantproduktion in diesen Studien weiterhin, dass die Kategorien der Kinder in den spektralen Merkmalen im Sibilanten und in den Formantwerten nicht im gleichen Maße divergierten wie es für die erwachsenen Sprecher gefunden werden konnte. Ähnliche Ergebnisse für den Vergleich zwischen den Sprachmustern der Kinder und denen der Erwachsenen konnten auch für die polnischen Sprachproduktionsdaten, der hier vorliegenden Studie, festgestellt werden. Die L1-Polnisch sprechenden Kinder lagen in den d_{sib} -Werten im $C_1 \times C_2$ -Raum (vgl. Abb. 4.5 und 4.8) als auch in den F2 (vgl. Abb. 4.6 und 4.9) und F3-Werten (vgl. Abb. 4.7 und 4.10) in den nachfolgenden Vokal oftmals etwas näher an der Nulllinie und kontrastierten damit nicht so stark zwischen den Sibilanten, wie die erwachsene Kontrollgruppe. Die Muster der Separierung und der Gemeinsamkeiten zwischen den Sibilanten – stärkere Ähnlichkeit zwischen /ʃ ɕ/ und gute Trennung von /s/ im Friktionsanteil, sowie stärkere Ähnlichkeit zwischen /s ʃ/ und gute Unterscheidung von /ɕ/ in den F2-Werten – waren dabei ähnlich im Vergleich der polnischen Kindern und der polnischen Erwachsenen. Statistisch konnten die Beobachtungen zur stärkeren Überlappung der Kategorien in der kindlichen Sprechergruppe dennoch nicht verifiziert werden, was möglicherweise an der starken Streuung vor allem in den kindlichen Daten liegen könnte. Diese Streuung kann außerdem als ein weiteres Indiz für die Unterschiede im Ausmaß der Kontrastrealisierung zwischen den polnischen Kindern und Erwachsenen herangezogen werden. Zharkova, Hewlett und Hardcastle (2011) zeigten größere Variationen in der Koartikulation in CV-Sequenzen (C=/f/ und V=/i u a/) in den Produktionen ihrer kindlichen Probanden verglichen mit den erwachsenen Sprachproduktionen. Begründet wird diese größere Variabilität von Zharkova und Kollegen (2011) in den kindlichen Produktionen dadurch, dass Kinder scheinbar noch keine voll ausgeprägte Repräsentation der Sprachgesten aufweisen.

F. Li (2008) und Żygis und Hamann (2003) zeigten für die erwachsenen Sprecher geschlechtsspezifische Unterschiede in den akustischen Dimensionen der drei Sibilanten des Mandarin Chinesischen und des Polnischen. Leichte Hinweise auf solche geschlechtsspezifischen Unterschiede waren auch in der vorliegenden Studie, im Vergleich der weiblichen und

männlichen erwachsenen Sprecher zu erkennen. Grundsätzlich unterschieden sich /s ʃ ʒ/ der weiblichen Erwachsenen stärker als die der Kinder und die der männlichen erwachsenen Versuchspersonen. In den F2 und den F3-Werten in den nachfolgenden Vokal, wurden bezüglich einer weniger starken Distinktion der Kategorien der männlichen Probanden, die auffälligsten Muster beobachtet (vgl. Abb. 4.9 und 4.10). Die männlichen Sprecher zeigten in den F2 und F3-Differenzwerten für das Sibilantpaar /s ʃ/ kaum eine Abweichung von der Nulllinie, was für eine starke Überlappung der beiden Kategorien in diesen Parametern spricht. Die geschlechtsspezifischen Unterschiede zwischen den Sibilantkategorien, die in anatomischen, sowie in soziologischen Aspekten begründet liegen können, passen zu den bisherigen Studien, die spektrale Merkmale von weiblichen und männlichen Probanden verglichen. Henton (1995) zeigte beispielsweise, dass die weiblichen Versuchspersonen in ihrer Studie Vokale „more open-mouthed“ und damit phonetisch eindeutiger produzierten. Auch Simpson (2000) beschrieb, dass aufgrund anatomischer Gegebenheiten die Vokalkategorien bei Frauen weiter auseinander liegen als bei Männern. Für die Sibilantproduktionen von Frauen und Männern im Amerikanisch Englischen und im Deutschen fanden Weirich und Simpson (2015), dass ihre weiblichen Versuchspersonen den alveolaren Sibilanten /s/ mit einem Energieschwerpunkt in den höheren Frequenzen produzierten, was in einem größeren phonologischen Kontrast zu dem postalveolaren Gegenstück /ʃ/ resultierte. Neben den anatomischen Gründen (Simpson, 2000; Weirich & Simpson, 2015) für diesen Umstand sind auch in Labov's (1990) und in Eckert's (1989) soziolinguistischen Studien mögliche Erklärungen sozialer Natur für den größeren Kontrast in den Lautklassen der weiblichen Probanden aufgeführt: Frauen variieren stärker in ihrem Sprachstil und passen diesen stärker bestimmten Situationen an, sodass für die Ergebnisse dieser hier vorliegenden Studie angenommen werden könnte, dass die weiblichen Versuchspersonen eher die formale Standardsprache und damit stärker kontrastierende Sprache, in der eher formalen Aufnahmesituation produzierten, während die männlichen Versuchspersonen ihren Sprachstil nicht in gleichen Maße an die Situation anpassten.

Ein weiterer Aspekt, der bezogen auf Nittrouer's (1992; 2002) und Nittrouer und Studdert-Kennedy's (1987) Überlegungen und Ergebnissen zu Koartikulation in der Sprachentwicklung auffällt, ist das Maß der Koartikulation im Vergleich von Sibilantproduktionen von Erwachsenen und Kindern des Polnischen. Die hier vorgestellten Daten decken sich nicht direkt mit den vorherigen Erkenntnissen. Diese besagen, dass in den kindlichen Sprechern mehr Koartikulationsmuster gefunden werden sollten, da Kinder zunächst eher

die groben Bewegungen zu den jeweiligen Artikulationsstellen meistern und erst später das präzise Timing der Gesten und den genauen Grad der Konstriktion beherrschen. Gefolgt wird dies aus Beobachtungen von Vihman (1996) und Studdert-Kennedy (2000), die in ihren Studien zu kindlichen Sprachproduktionen zeigten, dass Kinder selten Fehler im Artikulationsort machten, was eben als Begründung dafür gesehen wurde, dass die globalen Bewegungen hin zu einem Artikulationsort schon recht früh erlernt werden. Der Ansicht von Nittrouer folgend, sollte für die Sprachproduktionsdaten der L1-Polnisch sprechenden Kinder angenommen werden können, dass sie /s ʂ/ mittels der F2-Werte mindestens im gleichen Maße trennen, wie es auch für die weiblichen Erwachsenen gefunden wurde. Die Muster, welche in den kindlichen Daten in der vorliegenden Studie gefunden wurden, sind jedoch nicht kompatibel mit Nittrouer's Ansichten, denn die kindlichen Probanden zeigten für zwei der drei Sibilantpaare im direkten Vergleich mit den polnischen Erwachsenen näher beieinander liegende F2-Werte zum Onset des nachfolgenden Vokals. Es wäre eine Möglichkeit zu argumentieren, dass die Unterschiede in den Formantwerten zwischen /s ʂ/ zu gering sind und Kinder damit weniger Grundlage haben, um diese Laute stärker mittels der Informationen im Vokal zu differenzieren. Jedoch spricht hier ein entscheidender Aspekt aus Nittrouer's Studien dagegen: Nittrouer und Kollegen (1992; 2002; 2006; 1987) fanden in Perzeptionsexperimenten zur Gewichtung von Transitionen bei der Wahrnehmung von Sibilanten des Englischen, dass bei Kindern der Einfluss von Transitionen schwankt, je nach dem wie stark unterschiedlich die jeweiligen Transitionen sind. Es wurde gezeigt, dass die Transitionen in den Kontinua zwischen /si – fi/ und /sa – fa/ relativ wenig von Kindern gewichtet werden, während im Kontinuum /su – fu/ eine starke Gewichtung der Transitionen beobachtet werden konnte. Trotz der Unterschiede in der Gewichtung der Informationen in den Transitionen der einzelnen Kontinua, wurde weiterhin gezeigt, dass Kinder im Vergleich mit Erwachsenen, Transitionen trotzdem grundsätzlich stärker zur perceptiven Kategorisierung von Sibilanten nutzten.

Das dritte wichtige Ergebnis in dieser Studie sollte einen Beitrag zum Verständnis der diachronen Lautneutralisierung der drei Sibilanten des Polnischen durch den Zusammenfall des /s ʂ/-Kontrasts leisten. Vor allem sollte versucht werden in den kindlichen Sprachproduktionen Evidenzen für die größere Instabilität des retroflexen Sibilanten in dieser Neutralisierung zu finden. Dabei konnten zwei Auffälligkeiten beobachtet werden. Beim Betrachten der Differenzen in den d_{sib} und in den F2-Werten zwischen den Sibilanten (vgl. Abb. 4.8 und 4.9) fällt zum einen auf, dass in der Differenz des retroflexen Frikatives jeweils

zu einem der beiden anderen Sibilanten, Kinder und Erwachsene am stärksten voneinander abweichen, d.h. für die Differenzen zwischen /ʃ-s/ und /ç-ʃ/ ist für Kinder grundsätzlich eine etwas weniger starke Trennung – die Differenzwerte der Kinder liegen für beide untersuchten Parameter etwas näher an der Nulllinie – zwischen diesen Sibilantpaaren zu finden. Während in den Differenzwerten für die beiden gemessenen Parameter zwischen /s-ç/ (Abb. 4.9 und 4.8 mittleres Panel) die kindlichen Werte eher den erwachsenen Werten gleichen. Zum zweiten ist auffällig, dass für die Differenzen zwischen den F2-Werten des dentalen und des retroflexen Sibilanten polnische Kinder und polnische männliche Erwachsene nicht von Null abweichen, was möglicherweise bedeutet, dass für diese beiden Sprechergruppen F2 als Unterscheidungsparameter zwischen /s ʃ/ als unwirksam eingestuft werden könnte. Ein Merkmal, das dennoch Hilfe bei der Unterscheidung des retroflexen Frikatives von den beiden anderen Sibilanten leisten könnte und damit zur Stabilität dieses Lautes beitragen könnte, sind die Werte des dritten Formanten zu Beginn des nachfolgenden Vokals. Hier wurde für die kindlichen wie auch für die weiblichen erwachsenen Probanden gezeigt, dass /ʃ/ durch seine niedrigen F3-Werte gut von /s ç/ separiert werden konnte, wobei dies für die männlichen erwachsenen Versuchspersonen nicht in gleichem Maße der Fall war. Hier war eine stärkere Ähnlichkeit der F3-Werte von /s ʃ/ zu beobachten. Vor allem hinsichtlich der F3-Werte ist die Abb. C.3 im Anhang C.1 interessant, in welcher die weiblichen und männlichen Kinder getrennt dargestellt sind, denn hier ist für die männlichen Kinder zu sehen, dass diese den retroflexen Sibilanten durch niedrige F3-Werte von /s ç/ trennen. Genauer gesagt lässt dies vermuten, dass das F3-Muster – die Überlappung von /s ʃ/ in den F3-Werten – der männlichen erwachsenen Sprecher erst später aufkommt.

Viele Studien zu Lautwandel haben gezeigt, dass Lautwandel nicht abrupt geschieht, sondern dass oft eine schrittweise und von Generation zu Generation beobachtbare Veränderung der Sprache zu finden ist. Akustische, perzeptive wie auch artikulatorische Studien brachten hervor, dass die feinen phonetischen Details für die Sprachverarbeitung eine große Rolle spielen (Hawkins, 2003). Auch für Lautwandel wurde gefunden, dass sowohl für Sprecher als auch für Hörer lautliche Details von großer Bedeutung sind (Harrington et al., 2008). Verschiedene Studien zu Lautwandel und dessen Verbindung zu ebendiesen phonetischen Details konnten beobachten, dass sich die untersuchten Alters- und/oder Dialektgruppen in einem kontinuierlichen/graduellen Lautwandel zunächst „nur“ in feinen phonetischen Details unterscheiden, wobei erst in späteren und fortgeschritteneren Lautwandelstufen diese dann die Wahrnehmung erreichen. Im Laufe dieses schrittweisen

Verändern von Lauten besteht die Möglichkeit, dass sich die phonetischen Innovationen zu neuen Laut- beziehungsweise Aussprachevarianten in der Sprachgemeinschaft wandeln (Harrington et al., 2008; Hawkins & Midgley, 2005; Wells, 1982). Labov's (2001; 2007) Modell zu Lautwandel via *incrementation* geht davon aus, dass Generationen von Kindern die sprachlichen Innovationen weiter in Richtung des Lautwandels und vor allem über die Ebene des elterlichen Systems hinaus voranbringen. Betrachten wir die Ergebnisse der kindlichen Produktionsdaten, wäre es eine Möglichkeit, dass die fünfjährigen L1-Polnisch sprechenden Kinder ihren erwachsenen Bezugspersonen einen kleinen Schritt bei der /s ʂ/ Neutralisierung voraus sind. Denn Kinder sollten zunächst durch den vermehrten Kontakt zur Mutter aufgrund der traditionellen Rollenverteilung in den Anfängen der kindlichen Entwicklung eher die Sprachmuster der Mutter übernehmen. Konkreter ausgedrückt, würde in einem solchen Fall vermutet werden, dass die L1-Polnisch sprechenden Kinder die Sibilantkategorien /s ʂ ʐ/ gemessen anhand der spektralen Merkmale im der Friktionsgeräusch des Sibilanten (erster und zweiter DCT-Koeffizient), wie auch mittels der F2 und F3-Werte in den nachfolgenden Vokal hinein, in einem Muster unterscheiden, welches den weiblichen Versuchspersonen ähnlicher ist. Jedoch können wir in dieser Studie beobachten, dass eher die männlichen erwachsenen Probanden und die Kinder ähnliche Unterscheidungsmuster zeigen. Es besteht selbstverständlich die Möglichkeit, die beobachteten Muster der Kinder dahingehend zu interpretieren, dass sie den noch nicht voll ausgereiften und noch immer in der Entwicklung befindlichen Sibilanterwerb bei fünfjährigen Kindern zeigen. In diesem Fall wäre davon auszugehen, dass sich die kindlichen Produktionsdaten mit steigendem Alter stärker den erwachsenen Mustern annähern. Eine zweite Möglichkeit wäre aber auch eine Interpretation hinsichtlich Labov's (2001; 2007) Lautwandelmodell. Mit Hinblick auf den generationsweisen Fortschritt in einem Lautwandel, ist es auch denkbar, dass die L1-Polnisch sprechenden Kinder in dieser Studie den nächsten Schritt in Richtung des Wandels, d.h. den Zusammenfall von /s ʂ/, anzeigen. Zusammen mit den Daten der männlichen Erwachsenen ist eine weitere – und rein spekulative – Vermutung, dass die diachrone Neutralisierung des Dreier-Kontrasts hin zu einem Zweier-Kontrast durch den Merger des dentalen und des retroflexen Lautes zugunsten von /s/ in einigen dialektalen Varietäten des Polnischen ein Lautwandel ist, welcher seinen Ursprung in der Sprache beziehungsweise in den Merkmalen von Sibilanten der männlichen Sprecher und Hörer hat. Ein Punkt, welcher die Überlegungen in diese Richtung stützt, sind die Verteilungen der Standardsprache und den nicht-hochsprachlichen Varianten zwischen den Geschlechtern. Dabei wird in einigen Studien darauf hingewiesen, dass Männer in ihrem

Sprachgebrauch häufiger zur nicht-hochsprachlichen Aussprachevariante tendieren, während Frauen oftmals eher die angeseheneren Formen der Standardsprache nutzen (Eckert, 1989; Labov, 1990). Frühe Studien, die Männer und Frauen in Lautwandelprozessen betrachteten, wiesen häufig darauf hin, dass sprachliche Innovationen oftmals von Frauen ausgehen (Labov, 1990). Dennoch wurden Evidenzen gefunden, dass sowohl Männer als auch Frauen Lautwandel anführen können, jedoch oftmals ebenfalls mit einer ähnlichen Verteilung bezogen auf die standardsprachlichen und nicht-hochsprachlichen Formen, d.h. Prestige-besetzte Innovationen werden häufig durch Frauen verbreitet, während Männer an der Ausbreitung der regionalen und nicht-hochsprachlichen Formen beteiligt sind. Trudgill (1972) zeigte, dass Männer in Norwich gerade in den Lautwandelprozessen vorgehen, welche die weniger hochsprachlichen Formen hervorbringen. Ebenso berichtete Labov (1968), dass Männer Lautwandelprozesse dominieren können. Für die Neutralisierung in den Mandarin-Chinesischen Sibilanten, die ebenfalls durch einen Merger des dentalen und des retroflexen Sibilanten gekennzeichnet ist, wobei der dentale Laut erhalten bleibt, fanden Chuang und Fon (2010) mit den oben beschriebenen geschlechtsspezifischen Verteilungen von Sprachmustern kompatible Ergebnisse. In diesen Studien tendierten eher die männlichen Versuchspersonen zur Kontrastneutralisierung, während die weiblichen Sprecher eher dazu neigten den Dreier-Kontrast beizubehalten. Zieht man diese Aspekte in Betracht, wäre es möglich, dass in diesen hier vorgestellten Daten der polnischen Kinder und der polnischen männlichen Erwachsenen die Richtung des Wandels im polnischen Sibilantsystem, so wie auch der mögliche Ursprung dieses Wandels beobachtet werden kann.

Zusammengefasst kann gesagt werden, dass die hier vorgestellten Ergebnisse Evidenzen beziehungsweise neue Einblicke zur Vergleichbarkeit von artikulatorischen (EMA Daten) und akustischen Studien hervorbrachte. Die Bedenken bezüglich einer Perturbation der Sprache durch die Sensoren, welche bei der elektromagnetischen Artikulografie auf den Artikulatoren der Versuchspersonen angebracht werden, sind nicht so stark, als dass die Ergebnisse in starkem Maße verfälscht würden. Dennoch ist bei einer solchen verallgemeinernenden Aussage auf individuelle Unterschiede in der Gewöhnung an die Sensoren zu achten. Des weiteren wurde gezeigt, dass polnische Kinder in ihren Sibilantproduktionen ähnliche Muster aufzeigten wie die Gruppe der polnischen erwachsenen Versuchspersonen. Trotz der Gemeinsamkeiten in den untersuchten Parametern – erster und zweiter DCT-Koeffizient, sowie F2/F3-Werte zum zeitlichen Vokalonset – konnten jedoch ebenfalls Unterschiede in den zwei Altersgruppen, wie auch bei der getrennten Betrachtung der männlichen und der

weiblichen erwachsenen Versuchspersonen festgestellt werden. Die Gemeinsamkeiten in den Produktionen weisen, vereinbar mit früheren Studien zu kindlichen Sibilantproduktionen, daraufhin, dass Kinder im Alter von fünf Jahren sprachspezifische Kategorisierungsmuster erlernt haben (F. Li, 2008; F. Li & Munson, in Druck). Doch sind die Kategorien noch nicht im selben Maße getrennt, wie es vor allem für die weiblichen erwachsenen Versuchspersonen der Fall ist. Des Weiteren waren Unterschiede für die männlichen und weiblichen erwachsenen Versuchspersonen im Maß der Kontrastrealisierung sowohl für die spektralen Merkmale im Sibilanten selbst wie auch in den F2 und F3-Werte zum zeitlichen Onset des nachfolgenden Vokals zu finden. Dabei waren es die weiblichen Versuchspersonen, die den Kontrast stärker realisierten als die männlichen Versuchspersonen. Diese geschlechtsspezifischen Unterschiede waren am stärksten für den /s ʂ/-Kontrast in deren F2 und F3-Differenzwerten ausgeprägt. Es wäre denkbar, dass diese alters- und geschlechtsspezifischen Unterschiede möglicherweise ein Indiz dafür sind, dass die Neutralisierung des /s ʂ/-Kontrasts ein Lautwandelprozess ist, welcher von männlichen Mitgliedern der Sprachgemeinschaft angeführt wird. Weiter könnte es sein, dass die geringeren Kategorisierungsmuster der Kinder nicht nur eine Entwicklungsstufe darstellen, welche zeigt, dass Kinder die Sibilanten noch nicht in einer erwachsenen Art und Weise erworben haben, sondern dass hier möglicherweise die Richtung und der schrittweise Fortschritt einer solchen Neutralisierung zu sehen ist. Um diese vielfältigen neuen Aspekte zu der diachronen Neutralisierung des Dreier-Kontrasts /s ʂ ʃ/ in einigen dialektalen Varietäten des Standard-Polnischen durch einen Zusammenfall der Kategorien /s ʂ/ zu verstehen und Erklärungen für ebendiese Neutralisierung zu finden, ist es notwendig die Forschung hinsichtlich kindlicher und geschlechtsspezifischer Sibilantproduktionen, dialektaler Realisierungsvarianten sowie auch perzeptiver Untersuchungen zu vertiefen und weiterzuverfolgen.

Kapitel 5

Zusammenfassung und Diskussion

Im Sibilantsystem des Standard-Polnischen und des Mandarin-Chinesischen findet man einen Dreier-Kontrast zwischen den folgenden Sibilanten: dental /s/, retroflex /ʂ/ und alveolopalatal /ç/. Bisherige Studien zu diesen drei Sibilanten in den beiden Sprachen zeigten, dass die unterschiedlichen artikulatorischen Merkmale der Sibilanten, vor allem Konsequenzen hinsichtlich der spektralen Schwerpunktfrequenzen im Friktionsgeräusch der Laute selbst und für die Transitionen des zweiten Formanten in den Vokal, welcher den Sibilanten folgt, hervorbringen (Bukmaier, Harrington & Kleber, 2014; F. Li et al., 2007; Proctor et al., 2012; Toda et al., 2010). Während der dentale Laut mit seiner kurzen *front cavity* und den dadurch bedingten hohen Schwerpunktfrequenzen gut von /ʂ ç/ getrennt werden kann, wobei letztere niedrigere spektrale Schwerpunktwerte aufweisen, zeigt /ç/ mit seinen hohen F2-Werten zum zeitlichen Onset des nachfolgenden Vokals eine klare Distinktion von /s ʂ/, welche in diesem Parameter miteinander überlappen. Für die Sibilanten /s ç/ existiert also jeweils ein markantes akustisches Merkmal – Friktionsgeräusch für /s/ und F2-Transitionen für /ç/ –, während der retroflexe Sibilant, sowohl das Friktionsgeräusch als auch die F2-Werte benötigt, um von /s/ und /ç/ separiert werden zu können. Interessanterweise ist es gerade der retroflexe Sibilant, für den in einigen Polnischen und ebenfalls in einigen Mandarin-Chinesischen dialektalen Varietäten eine Instabilität beschrieben wird. Genauer gesagt wurde beobachtet, dass der Kontrast zwischen dem dentalen und dem retroflexen Laut neutralisiert wird und damit neben dem alveolopalatalen Sibilant nur der dentale Laut bestehen bleibt.

Studien zur synchronen Variation zeigten auf, dass Sprachlaute aufgrund verschiedener Faktoren variieren können, welche in die Untergruppen der internen und der externen Faktoren kategorisiert werden können. Zu der intern-motivierten Variation zählen sprach-

liche Veränderungen, wie die gegenseitige koartikulatorische Beeinflussung nebeneinander vorkommender Segmente und die Variation aufgrund verschiedener prosodischer Kontexte. Für den Dreier-Kontrast im Sibilantensystem des Mandarin-Chinesischen wurde gezeigt, dass sowohl der Vokalkontext (Jeng, 2006) als auch verschiedene prosodische Kontexte (Chuang & Fon, 2010) die Distinktion der Sibilanten /s ʃ/ beeinflussen können, wobei zusätzlich die Faktoren – Geschlecht und dialektaler Hintergrund der Probanden – bei der Kontrastneutralisierung eine Rolle spielten (Chuang & Fon, 2010). Die letzteren zählen zur externen Variation und beinhalten damit Faktoren, die nicht linguistischer Natur sind, also eben das Geschlecht und den dialektalen Hintergrund des Sprechers/Hörers, sowie auch das Alter und die Identifikation einzelner Sprecher und verschiedener Sprachgemeinschaften mit einer bestimmten Gruppe. Für den Faktor Geschlecht fanden Chuang und Fon (2010), dass die weiblichen Versuchspersonen den Kontrast zwischen den Sibilanten /s ʃ/ grundsätzlich aufrechterhielten. Im Gegensatz dazu standen die männlichen Probanden, die den Dental-Retroflex Kontrast in Richtung des dentalen Sibilanten neutralisierten.

Während sich bisher schon einige Studien mit den Sibilanten und deren Neutralisierung im Mandarin-Chinesischen beschäftigten, existieren hinsichtlich der Neutralisierung der Sibilanten /s ʃ ʒ/ im Polnischen nach dem Wissen der Autorin dieser Arbeit bisher keine experimentellen Studien, die den Einfluss verschiedener Faktoren auf die synchrone In-/Stabilität der Laute /s ʃ/ untersuchten.

Das Ziel der vorliegenden Arbeit war daher herauszufinden, ob in der Sprachproduktion und der Perzeption von Sprechern/Hörern des Standard-Polnischen Charakteristika gefunden werden können, welche die Neigung hin zu einer scheinbaren Instabilität des retroflexen Sibilanten erklären könnten. Daneben sollte geprüft werden, ob Erklärungsansätze dafür hervorgebracht werden können, dass genau der Dental-Retroflex Kontrast einer Neutralisierung unterliegt und nicht etwa der /ʃ ʒ/ oder gar der /s ʒ/-Kontrast. Dieser Aspekt ist vor allem interessant, wenn man bedenkt, dass McGuire (2007) in seiner Studie zur Wahrnehmung der Sibilanten des Polnischen für die L1-Englisch sprechenden Versuchspersonen beobachten konnte, dass die englischen Hörer die größten Schwierigkeiten bei der perceptiven Kategorisierung des /ʃ ʒ/-Kontrasts und nicht etwa des /s ʃ/-Paares, hatten. Vor dem Hintergrund der bisher bekannten Studien und der verschiedenen Fragestellungen wurden artikulatorische und akustische Daten der Sibilanten in der Studie in Kapitel 2, Perzeptionsdaten in Bezug auf die Wahrnehmung von /s ʃ ʒ/ in verschiedenen prosodischen Kontexten in Kapitel 3 und rein akustische Daten der drei Frikative produziert von kindlichen und erwachsenen L1-Polnisch Sprechern im vierten und letzten experimentellen

Kapitel dieser Arbeit erhoben.

Zunächst werden in den nachfolgenden Abschnitten die experimentellen Kapitel 2-4 einzeln zusammengefasst und anschließend die daraus hervorgegangenen Ergebnisse, in Hinblick auf bisherige Erkenntnisse zu der Verbindung von synchroner Variation und diachronem Wandel, sowie hinsichtlich aktueller Lautwandeltheorien diskutiert.

5.1 Artikulation der Sibilanten /s ʂ ʐ/

In Kapitel 2 wurden verschiedene artikulatorische Merkmale und deren akustische Konsequenzen in Bezug auf die Kategorisierung der Standard-Polnischen Sibilanten untersucht. Dafür wurden anhand der elektromagnetischen Artikulografie (EMA), Zungenbewegungsdaten von neun L1-Polnisch sprechenden Erwachsenen im Laufe der Produktion der Sibilanten /s ʂ ʐ/ erhoben. Ferner wurde untersucht, ob die Faktoren Sprechgeschwindigkeit (schnell vs. langsam), Vokalkontext (/ɛ/ vs. /a/ vs. /ɔ/) und lexikalische Frequenz (Echtwörter vs. Logatome) einen Einfluss auf die Artikulation und die Stabilität der Sibilanten haben.

Die Ergebnisse brachten zum einen hervor, dass keiner der getesteten Faktoren – Sprechgeschwindigkeit, Vokalkontext und lexikalische Frequenz – einen Einfluss auf die Artikulation und die anschließende akustische Kategorisierung der Sibilanten hatte. Zum anderen konnte beobachtet werden, dass die drei Sibilanten hinsichtlich der untersuchten physiologischen Merkmale, gut voneinander separiert werden können. Genauer gesagt zeigten die Daten der vertikalen Zungenspitzenorientierung im sibilantischen Teil der analysierten CV-Sequenz (C=/s ʂ ʐ/; V=/ɛ a ɔ/), dass für den alveolopalatalen Sibilanten die Zungenspitze eher nach unten orientiert ist, während sie für /ʂ/ nach hinten-oben gedreht ist. Hinsichtlich der horizontalen Zungenspitzenposition konnte für den retroflexen Sibilanten gefunden werden, dass dessen Artikulation am weitesten rückverlagert ist, während die des dentalen Sibilanten am weitesten vorne gelegen ist. Der alveolopalatale Sibilant positioniert sich in der anterior-posterior Dimension zwischen dem dentalen und dem retroflexen Sibilanten. Diese physiologische Trennung der Sibilanten konnte jedoch nicht im selbem Maße in der Akustik der Laute aufrechterhalten werden. Die akustische Analyse zeigte eine Ähnlichkeit von /ʂ ʐ/ im Friktionsgeräusch der Sibilanten auf, was durch die Anwendung der Gaußschen Klassifikation der akustischen Daten bestätigt werden kann.

te. Genauer gesagt wurde der dentale Sibilant zu 96% richtig klassifiziert, während in der Klassifikation die Sibilanten /ʃ ʒ/ häufig miteinander verwechselt wurden.

In der artikulatorischen Analyse der Sibilanten zum zeitlichen Onset des Vokals, welcher den Sibilanten nachfolgte, wurde die horizontale Position und die vertikale Orientierung des Zungenrückensensors näher betrachtet. Hier konnte ebenfalls ein gute Trennung der Sibilanten aufgezeigt werden und zwar insofern, als dass der alveolopalatale Sibilant hinsichtlich der vertikalen Position des Zungenrückens wesentlich höhere Werte zeigte als die beiden anderen Sibilanten. Daraus wurde geschlossen, dass der koartikulatorische Einfluss dieses Sibilanten auf den nachfolgenden Vokal stärker ist als der Einfluss von /s ʒ/. Ein Unterscheidungsmerkmal zwischen /s ʒ/ stellt die horizontale Position des Zungenrückens im Laufe der Artikulation der Sibilanten dar, denn, ähnlich der horizontalen Zungenspitzenposition, zeigt der retroflexe Sibilant auch für den Zungenrücken zum nachfolgenden Vokalonsset eine Rückverlagerung. Für den dentalen Laut ist der Zungenrücken zu Beginn des Vokals dagegen weiter vorne positioniert. In die Analyse der akustischen Charakteristika der drei Sibilanten zum Zeitpunkt des akustischen Vokalonssets, wurden nicht nur die Werte der F2-Transitionen eingeschlossen, sondern auch die Werte des dritten Formanten, da, wie schon anfänglich in den Abschnitten 1.4.3 und 1.4.4 in der Einleitung dieser Arbeit beschrieben, letztere vor allem für die Klassifikation des retroflexen Sibilanten von Bedeutung sein können. Hinsichtlich der F2 und F3-Transitionen der Sibilanten wurden folgende Beobachtungen gemacht: (1) die Vokale, welche dem alveolopalatalen Sibilanten folgten, wiesen hohe F2-Transitionswerte auf, (2) die Vokale, welche /s ʒ/ folgten, zeigten niedrigere und überlappende F2-Transitionen, und (3) die F3-Werte waren zu einem gewissen Grad für die Separierung des retroflexen Lauts /ʒ/ von /s ʒ/ von Nutzen. Die Gaußsche Klassifikation der Sibilanten anhand der F2 und F3-Transitionswerte unterstützte diese Beobachtungen, d.h. während der alveolopalatale Sibilant zu 91% richtig klassifiziert wurde, zeigten die beiden Frikative /s ʒ/ eine deutliche Verwechslung miteinander.

Die Ergebnisse der artikulatorischen und akustischen Analyse zeigten sowohl für den Frikationsanteil des Sibilanten selbst als auch für das zeitliche Onset des nachfolgenden Vokals, jeweils eine gute artikulatorische Separierung, welche jedoch keine so klaren Auswirkungen auf die Akustik der drei Laute, und vor allem nicht für den retroflexen Sibilanten, brachte. Diese Überlegungen gehen mit Lindblom's (1998) Modell einher, welches den retroflexen Sibilanten als einen artikulatorisch komplexen und aufwändigen Laut beschreiben würde, wobei diese artikulatorische Komplexität nur einen geringen Beitrag zur akustischen und perzeptiven Hervorhebung dieses Frikatives leistet.

Die Idee der geringeren akustischen und vor allem perzeptiven Saliens wurde anschließend in der Perzeptionsstudie in Kapitel 3 dieser Arbeit untersucht.

5.2 Perzeption der Polnischen Sibilanten /s ʂ ʐ/ in verschiedenen prosodischen Kontexten

Im dritten Kapitel der Arbeit sollte die perzeptive Stabilität der Sibilanten und vor allem der Einfluss des prosodischen Kontexts auf die Kategorisierung der Sibilanten geprüft werden. Die Hypothese für dieses Experiment war, dass der /s ʂ/-Kontrast in deakzentuiertem Kontext weniger kategorial wahrgenommen wird als der zwischen /s ʐ/. Um Aufschlüsse hinsichtlich der perzeptiven In-/Stabilität der polnischen Sibilanten zu erhalten, wurden synthetische Kontinua zwischen dem potentiell instabilen /s ʂ/-Kontrast und zwischen dem vermeintlich stabilen Sibilantenpaar /s ʐ/ erzeugt und in zwei verschiedenen prosodischen Bedingungen eingebettet, d.h. das Zielwort mit der initialen CV-Sequenz wurde den Hörern zum einen in deakzentuierter und zum anderen in akzentuierter Position präsentiert. Zusätzlich wurde für das /s-ʐ/ Kontinuum getestet, inwiefern die Informationen im Vokal ausschlaggebend für eine korrekte Identifikation und Kategorisierung der Sibilanten durch die Hörer sind. Zur Beantwortung dieser zweiten Fragestellung, wurden für das Dental-Alveolopalatal Kontinuum zwei Kontinua generiert, d.h. eines mit variierenden Schwerpunktfrequenzen in der sibilantischen Sequenz und ambig gehaltenem Vokal, sowie ein zweites in welchem ergänzend die Informationen im Vokal kompatibel mit der vorangehenden Sibilantsequenz verändert wurden, d.h. von Formantwerten, welche mit dem dentalen Sibilanten vereinbar waren, hin zu Werten, die zum alveolopalatalen Sibilanten passten. Alle Kontinua wurden in einem *forced choice task* von 10 Standard-Polnisch sprechenden Versuchspersonen kategorisiert.

Die erste Beobachtung hinsichtlich der Antworten der Hörer auf die Kontinua war, dass für das /s-ʐ/ Kontinuum, für welches zusätzlich die Information im Vokal variierte, eine kategorialere Wahrnehmung festgestellt werden konnte als für das /s-ʂ/ Kontinuum mit ambig gehaltenem Vokal.

Bezogen auf die Frage nach Hinweisen auf die /s ʂ/ Neutralisierung waren die Ergebnisse weniger klar. Der erwartete Effekt auf die Neigungswerte der Perzeptionskurve blieb aus, d.h. die Perzeptionskurven des /s-ʂ/ Kontinuums waren in beiden prosodischen Kontexten steiler als die des /s-ʐ/ Kontinuums, was auf mehr Unsicherheit der Hörer bei der Kategorisierung des vermeintlich stabileren Kontrasts hindeutete. Dieses Ergebnis könnte durch die

generelle Unterscheidung der Sibilanten in den Frequenzschwerpunkten erklärt werden, da der /s ʂ/-Kontrast hinsichtlich des spektralen Schwerpunkts stärker separiert wird als der /s ʃ/-Kontrast. Genauer gesagt wurde in Studien zur Produktion der polnischen Sibilanten gefunden, dass der retroflexe Sibilant die niedrigsten und der dentale Laut die höchsten Schwerpunktfrequenzen aufweist, während der Schwerpunkt des alveolopalatalen Sibilanten, zwischen dem von /s/ und /ʂ/ liegt (Nowak, 2006a; Żygis & Hamann, 2003). Außerdem könnte dieses Ergebnis mit den Überlegungen in Iskarous et al. (2011) zur *s-retraction* erklärt werden. Ihre Vermutung hinsichtlich der Perzeption des /s ʃ/-Kontrasts sind, dass dieser Kontrast, trotz der Neigung zur /s ʃ/ Kontrastneutralisierung in sCr-Klustern, durch das doch noch immer zu geringe Maß an Variabilität im Frequenzschwerpunkt, in der Perzeption stabil bleiben sollte.

Trotz des Fehlens eines Effekts des prosodischen Kontexts auf die Neigungswerte der Kontinua, konnte dennoch eine etwas stärkere Auswirkung der Deakzentuierung auf die Kategoriengrenze des Dental-Retroflex Kontinuums im Vergleich zum /s-ʃ/ Kontinuum beobachtet werden. Genauer gesagt wiesen die Versuchspersonen für das potentiell instabile /s-ʂ/ Kontinuum im deakzentuierten Kontext mehr Stimuli der dentalen Kategorie zu, während die Kategoriengrenze des /s-ʃ/ Kontinuums in den beiden prosodischen Kontexten weniger stark variierte. Die Verschiebung der Kategoriengrenze lässt Spekulationen zu, dass möglicherweise die fehlerhafte Interpretation des Gehörten (Ohala, 1993, 2012) zusammen mit einer eventuell verringerten Kompensation für Koartikulation in prosodisch schwachen Konstituenten (Harrington et al., 2013, 2015) eine Rolle bei der /s ʂ/ Neutralisierung spielen könnte.

Davon ausgehend, dass die Kategoriengrenze für den /s ʂ/-Kontrast in der deakzentuierten Bedingung verschoben wird, jedoch die Neigungen der Perzeptionskurven weiterhin sehr steil bleiben, bestünde die Möglichkeit, Ansätze für eine Erklärung der Neutralisierung des /s ʂ/-Kontrasts in der Exemplar-Theorie (Johnson, 1997; Pierrehumbert, 2001, 2002) zu finden. Die Exemplar-Wolke der dentalen Kategorie würde in der Deakzentuierung durch die Akzeptanz von Exemplaren mit niedrigerem Frequenzschwerpunkt als gute Exemplare der dentalen Kategorie, schrittweise vergrößert werden, während die Exemplare für den retroflexen Laut schrittweise reduziert würden. Dadurch könnte die perzeptive Distanz zwischen den Kategorien erhalten bleiben, sodass die Neigungswerte der Perzeptionskurven nicht von einer Veränderung hin zu flacheren Kurven betroffen sein müssten. Im weiteren Verlauf wäre eine größere Variabilität des dentalen Frikatives zu erwarten, da die neue Exemplar-Wolke dieser Kategorie, Varianten mit niedrigeren Frequenzen einschließen

würde.

5.3 Sibilantproduktionen im Spracherwerb und in erwachsener Sprache

Im vierten experimentellen Kapitel der Arbeit wurden Sibilantproduktionen von L1-Polnisch sprechenden Kindern mit den Produktionen von L1-Polnisch sprechenden Erwachsenen verglichen. Dies geschah vor dem Hintergrund der Frage, ob im Spracherwerb Hinweise auf eine größere Instabilität des retroflexen Sibilanten und spezieller des /s ʂ/-Kontrasts gefunden werden können. Die Überlegungen früherer Studien zur linearen Verbindung von Fehlermustern im Laufe des Spracherwerbs und diachrotem Lautwandel wurden zwar durch die Erkenntnisse heutiger Spracherwerbsforschung relativiert (Foulkes & Vihman, 2015), dennoch zeigte Labov (2001) in seinen *apparent time* Studien, dass im Vergleich von zwei nebeneinander lebenden Generationen in den feinen phonetischen Details die Richtung eines möglichen Lautwandels erkannt werden könnte. Für diese Studie wurden Sprachproduktionsdaten von polnischen Kindern in einem Bildbenennungsexperiment und von polnischen Erwachsenen, welche die Zielwörter laut vorlasen, erhoben. Insgesamt wurden Produktionsdaten von 22 Kindern und von 10 Erwachsenen in die Analyse eingeschlossen. Das erste Ergebnis dieser Studie bezog sich dabei nicht auf die zentrale Frage nach Erklärungsansätzen für die /s ʂ/ Neutralisierung, sondern auf den Vergleich von Sibilantproduktionen mit und ohne EMA-Sensoren. Meenakshi et al. (2014) beobachteten, dass die Sensoren, welche im Zuge der elektromagnetischen Artikulografie auf der Zunge der Probanden angebracht werden, besonders bei der Produktion von koronalen Lauten problematisch sein können. Der Vergleich der rein akustischen Produktionsdaten in Kapitel 4 mit den Daten aus der physiologischen Studie in Kapitel 2 brachte keinen Effekt der Sensoren auf die Produktion der Sibilanten hervor, d.h. die akustischen Merkmale der Sibilanten waren für die beiden Experimente vergleichbar. Kompatibel mit Ergebnissen aus Perturbationsstudien (Hoole, 1987; Jones & Munhall, 2000, 2003) kann gesagt werden, dass es für die Sprecher in der EMA-Studie (Kap. 2) möglich war, die Sibilantproduktionen anhand des auditiven Feedbacks dahingehend zu justieren, dass die typische akustische Verteilung der Kategorien beibehalten werden konnte.

Der Vergleich von kindlichen und erwachsenen Sibilantproduktionen zeigte, dass die kindliche Sprechergruppe den Dreier-Kontrast der polnischen Sibilanten zwar auf eine ähn-

liche Art und Weise produzierten, wie es für die Erwachsenen der Fall war, dennoch divergierten die Kategorien der Kinder, gemessen anhand der DCT-Koeffizienten [1, 2] im Sibilanten und den F2 und F3-Werten zum nachfolgenden Vokalonsset, etwas weniger wie die Sibilanten der erwachsenen Sprecher. Dieser Fund ist kompatibel mit F. Li's (2008) Studie zur kindlichen Produktion der Sibilanten /s ʃ ɕ/ im Mandarin-Chinesischen. Die Ergebnisse in Kapitel 4 unterstützen jedoch nicht die von Nittrouer et al. (1989) beschriebene Beobachtung, dass in der kindlichen Produktion von Sibilant-Vokal Sequenzen mehr Koartikulation zu finden ist als für die Sibilantproduktionen der erwachsenen Sprecher. In den Differenzwerten zwischen den F2-Werten zum zeitlichen Onset des Vokals, welcher den Sibilanten /s ʃ ɕ/ nachfolgte, war zu erkennen, dass die Kinder, verglichen mit den erwachsenen Sprechern, hinsichtlich der Unterschiede zwischen den Paaren /ɕ-ʃ/ und /ʃ-s/ näher an der Nulllinie lagen. Einzig für das Sibilantpaar /s-ɕ/ konnten höhere Werte für die kindlichen Sprecher gezeigt werden. Für die kindliche Sprechgruppe konnte im Vergleich mit den erwachsenen Probanden eine etwas größere Variation in den Differenzwerten zwischen den Sibilantpaaren sowohl in den Wertebereich hinein, der größere Unterschiede zwischen den Sibilanten bedeutet (d.h. höhere Differenzwerte), als auch in den Bereich hinein, der für eine geringere Distinktion der Sibilanten in Bezug auf die F2-Werte spricht (d.h. niedrigere Differenzwerte), beobachtet werden. Dieses Ergebnis ist mit Erkenntnissen aus Zharkova et al.'s (2011) Ultraschall-Studie zu vereinbaren, die zeigten, dass Kinder stärker im Grad der Koartikulation in CV-Sequenzen (C=/f/ und V=/i a u/) variierten. Die Schlussfolgerung, die Zharkova et al. (2011) daraus zogen, war, dass diese größere Variation bei Kindern durch die scheinbar noch unvollständigen Repräsentationen der Sprachgesten aufkommt.

In Bezug auf die größere Instabilität des retroflexen Sibilanten und der potentiellen Neutralisierung des /s ʃ/-Kontrasts und möglichen damit verbundenen Mustern in der kindlichen Sprache, konnte gefunden werden, dass die Differenzwerte zwischen /ʃ s/ und /ɕ ʃ/ für die kindlichen Probanden, vor allem verglichen mit den weiblichen Sprechern, etwas geringer ausfielen. Der Kontrast zwischen dem dentalen und dem alveolopalatalen Sibilanten hingegen, ist, was die spektralen Merkmale im Sibilanten anbelangt, denen der erwachsenen Sprecher ähnlicher. Für die F2-Werte konnten jedoch teilweise größere Differenzwerte zwischen /ɕ s/ für die kindlichen Probanden, verglichen mit den etwas niedrigeren Werten der Erwachsenen, beschrieben werden. Zusätzlich konnte, gemessen an den Differenzen zwischen den F2-Werten, für den Kontrast zwischen /s ʃ/ bei den kindlichen Sprechern nur ein sehr geringer Abstand zur Nulllinie gefunden werden, wo-

hingegen zumindest die weiblichen L1-Polnisch Sprecher in den F2-Werten der Sibilanten /s ʂ/ divergierten. Diese Ergebnisse könnten Hinweise auf die größere Instabilität des retroflexen Sibilanten und zusätzlich einer größeren Wahrscheinlichkeit eines /s ʂ/ Mergers sein. Studien zu Lautwandel haben gezeigt, dass Lautwandel kein kategorialer Prozess ist, sondern dass sich Sprache kontinuierlich von Generation zu Generation verändert. Die feinen phonetischen Details in der Sprache und deren Variationen können im Laufe eines Lautwandelprozesses von Bedeutung sein (Harrington et al., 2008; Hawkins & Midgley, 2005; Wells, 1982). Die kontinuierliche Veränderung von Sprache in einem Lautwandel ist mit Labov's (2001, 2007) Modell zu Lautwandel durch *incrementation* und der Rolle, welche Kinder in einem Lautwandel spielen, in Zusammenhang zu bringen. In seinem Modell zu Lautwandel via *incrementation* wird angenommen, dass die Generationen der Kinder die sprachlichen Variationen in Richtung des Lautwandels voranbringen und dabei über die Ebene des elterlichen Sprachsystems hinaustreten. Mit den Kenntnissen zu den feinen phonetischen Details in der Artikulation, der Akustik und der Perzeption von Sprache zusammen mit dem Modell zu Lautwandel, welcher durch die schrittweise Verstärkung der Lautwandelmuster über Generationen hinweg passiert, besteht damit die Möglichkeit, dass die kindlichen Sprachmuster nicht einfach nur einen noch nicht vollständig abgeschlossenen Sibilanterwerb aufzeigen, sondern dass schon in den Sibilantproduktionen der kindlichen Versuchspersonen die Richtung selbst und womöglich ein minimaler Schritt in ebendiese Richtung der /s ʂ/ Neutralisierung zu erkennen ist.

5.4 Die Sibilanten /s ʂ ʐ/ und der Faktor Geschlecht

Soziolinguistische Studien zeigten, dass das Geschlecht als extra-linguistischer Faktor ebenfalls einen Einfluss auf Lautwandel haben kann (Labov, 1990). Daher wurde in den Analysen der Perzeption der Sibilanten in Kapitel 3 und in der Studie zum Vergleich von kindlichen und erwachsenen Sibilantproduktionen auch der Faktor Geschlecht in die Auswertung eingeschlossen. In der Produktionsstudie in Kapitel 4 konnten geschlechtsspezifische Unterschiede im Maß der Distinktion der drei Sibilanten vor allem für zwei der drei untersuchten Parameter, nämlich für die F2 und F3-Werte, beobachtet werden. Die weiblichen L1-Polnisch Sprecher zeigten größere Unterschiede zwischen den Sibilantpaaren, während, im Vergleich dazu, die Sibilantkategorien der männlichen Probanden hinsichtlich der untersuchten Parameter nicht so stark divergierten. Vor allem zwischen dem Sibilantpaar /s ʂ/ zeigten die männlichen Sprecher eine weniger starke Separierung für diese Parameter.

Ein weiterer Fund, welcher mit den Beobachtungen aus der Produktionsstudie in Verbindung gebracht werden kann, kam in der Perzeptionsstudie in Kapitel 3 auf. In der geschlechtsspezifischen Analyse der Neigungen der Perzeptionskurven zeigten die männlichen Versuchspersonen etwas flachere Perzeptionskurven als die weiblichen Hörer. Geht man von einem Zusammenhang zwischen Sprachproduktion und Perzeption aus (Fowler & Galantucci, 2005), so lässt sich vermuten, dass die steileren Perzeptionskurven der weiblichen Versuchspersonen in den periphereren Sibilantkategorien ebendieser Versuchspersonengruppe in der Sprachproduktion begründet liegen könnten. Genauer gesagt wäre denkbar, dass Frauen mit ihren periphereren Sibilantkategorien in der Produktion, die Kategorien auch in der Perzeption etwas kategorialer wahrnehmen als die männlichen Versuchspersonen.

In Anbetracht dieser Ergebnisse, ist es für die Neutralisierung des /s ʃ/-Kontrasts – rein spekulativ gesprochen – möglich, dass die männliche Sprachgemeinschaft diesen Lautwandel bedingt. Diese Überlegungen gehen mit Ergebnissen aus Studien von Labov (1990, 2001) und Milroy und Milroy (1985) einher, in welchen die weibliche Sprachgemeinschaft eher Lautwandelformen voranbringen, welche sprachliche Innovationen beinhalten, während die Gruppe der Männer eher die regionalen und dialektalen Aussprachemuster verbreiten.

5.5 Fazit und Ausblick

In der modernen Sprachwandelforschung konnten Evidenzen dafür aufgezeigt werden, dass die synchrone Variation eng mit diachronem Wandel verknüpft ist, wobei sowohl die interne als auch die externe Variation eine Rolle in Lautwandelprozessen spielen kann. Für die interne Variation wird angenommen, dass sie meist den Grund/Auslöser für Lautwandel aufzeigt, während die extra-linguistischen Faktoren häufig mit der Verbreitung der neuen/innovativen Formen in Verbindung gebracht werden (Harrington et al., 2016). Für den Kontrast zwischen den Sibilanten /s ʃ/ im Mandarin-Chinesischen konnten bisher Einflüsse der prosodischen Position, des Vokalkontexts, des Geschlechts und des dialektalen Hintergrunds der Sprecher/Hörer auf das Maß an Kontrasterhaltung/Neutralisierung gefunden werden. Vor allem wurde in den bisherigen Studien für das Mandarin-Chinesische, ähnlich der Beschreibung der Lautneutralisierung im Polnischen, eine regionale Aufteilung zwischen Kontrasterhaltung und Neutralisierung beschrieben. Chang (2011) zeigte in einer Produktionsstudie, dass Taiwan-Mandarin Sprecher den Sibilantkontrast neutralisieren,

während die Standardsprachlichen Beijing-Mandarin Sprecher alle drei Sibilanten erhalten. Der Übergang vom Erhalten der Retroflexion hin zur Neutralisierung des /s ʂ/-Kontrasts war in Chang's (2011) Studie nicht kategorial. Viel mehr konnte eine graduelle Veränderung der vollständigen Retroflexion für die Beijing-Mandarin Sprecher hin zur /s ʂ/ Neutralisierung für die Taiwan-Mandarin Sprecher gefunden werden, wobei letztere im Grad der Neutralisierung viele Zwischenstufen durchliefen. Studien zum Einfluss des Vokalkontexts brachten weniger eindeutige Ergebnisse hervor: während Jeng (2006) ein stärkeres Maß an Neutralisierung im Kontext des hohen gerundeten Vokals /u/ fanden, wurde in Y. Li (2009) von einer größeren Neigung hin zur Neutralisierung im Kontext hoher Vokale berichtet. Hinsichtlich des prosodischen Kontexts konnten in Produktionsstudien geschlechts- und sprecherspezifische Unterschiede im Grad der Realisierung der Retroflexion gefunden werden, d.h. Frauen behielten den retroflexen Sibilanten bei, während der Kontrast für die meisten männlichen Probanden in beiden prosodischen Bedingungen neutralisiert wurde. Chung (2006) zeigte, dass der Zusammenfall des Sibilantkontrasts stark von der jeweiligen Gesprächssituation, dem Alter und des Bildungsniveaus der Sprecher abhängig ist, d.h. in informelleren Situationen, für ältere Generationen und für Sprecher mit geringem Bildungsniveau kann der /s ʂ/ Zusammenfall häufiger beobachtet werden, während in formellen Diskursen mit dem Fokus auf die Formalität der Umstände, beide Sibilanten erhalten werden. Für die Perzeption der Sibilanten wurde in der Studie von Chang (2012) gezeigt, dass für die Versuchspersonen, welche Sprecher des Taiwan-Mandarin waren, der Kontrast zwischen allen drei Sibilanten perzeptiv beibehalten wurde. Chuang und Fon (2011) berichtete jedoch, dass die Akzeptanz von Wörtern, in welchen der retroflexe Sibilant durch den dentalen Laut ersetzt wurde, in der Perzeption größer war als für die Experimentbedingung, in welcher der dentale Laut durch /ʂ/ substituiert wurde. Eine weitere Studie zur Perzeption der Mandarin-Chinesischen Sibilanten berichtete, dass trotz der kategorialen Wahrnehmung der Sibilanten durch die Taiwan-Mandarin Versuchspersonen die perzeptive Distanz zwischen /s ʂ/ für ebendiese Hörergruppe am geringsten ist (Shih & Kong, 2011).

Während für die Variation und die Neutralisierung des /s ʂ/-Kontrasts im Mandarin-Chinesischen der Zusammenhang von verschiedenen linguistischen und extra-linguistischen Faktoren und dem Maß an Kontrasterhaltung/Neutralisierung getestet wurde, sind die Studien zu ebendieser Kontrastneutralisierung im Polnischen eher rar.

Die vorliegende Arbeit stellt einen ersten Versuch dar, Erkenntnisse hinsichtlich der potentiellen Kontrastneutralisierung im Polnischen zu sammeln. In Hinblick darauf wurden

artikulatorische, rein akustische und perzeptive Daten von erwachsenen und kindlichen L1-Polnisch sprechenden Versuchspersonen zu den polnischen Sibilanten erhoben. Die Rekrutierung von Standardsprachlichen Versuchspersonen wurde durch Ohala's (1981, 2012) Überlegungen zum Ursprung von Lautwandel gestützt, in welchen davon ausgegangen wird, dass die synchrone Variation einen direkten Einfluss auf diachrone Lautwandel haben kann. Genauer gesagt beschreibt Ohala (2012) Variationen in der Sprachproduktion, die in all den verschiedenen Sprachen der Welt vorkommen können, wobei in sehr seltenen Fällen diese Muster in der Produktion, die Perzeption der Hörer dahingehend beeinflussen können, dass es zu einem Lautwandel in der Sprachgemeinschaft kommt.

In Kapitel 2 dieser Arbeit wurde die genaue Artikulation der Sibilanten im Polnischen betrachtet, wobei vermutet wurde, dass das nicht-lineare Verhältnis zwischen Artikulation und Akustik in der Produktion, vor allem des retroflexen Sibilanten, die Wahrscheinlichkeit einer /s ʂ/ Neutralisierung ansteigen lassen könnte, da die komplexe Artikulation des retroflexen Lautes nur eine sehr geringe Auswirkung auf die akustische Separierung von /ʂ/ und den beiden anderen Sibilanten des Polnischen hatte.

Die Untersuchung der Perzeption der drei Sibilanten in zwei verschiedenen prosodischen Bedingungen in Kapitel 3 brachte hervor, dass die Prosodie einen Einfluss auf die Perzeption sowohl des /s ʂ/ Kontrasts als auch des /s ɕ/ Kontrasts hat, d.h. für die beiden Kontinua wurden für den Großteil der Versuchspersonen flachere Perzeptionskurven in der deakzentuierten Bedingung gefunden. Die Prosodie hatte weiterhin einen stärkeren Einfluss auf den /s ʂ/-Kontrast und zwar in dem Sinne, dass die Kategoriengrenze in der deakzentuierten Position in die Richtung von vermehrten /s/-Antworten verschoben wurde. Der Effekt der verschiedenen prosodischen Bedingungen auf die Kategorisierung des /s-ɕ/ Kontinuums fiel dagegen geringer aus. Der Unterschied zwischen den männlichen und den weiblichen Hörern in der Perzeption der Sibilanten wies weiter auf eine stärkere Distinktion der Sibilanten in der Perzeption der weiblichen Probanden hin. Während die weiblichen Probanden steilere Perzeptionskurven zeigten, waren die Neigungen der männlichen Probanden grundsätzlich etwas flacher. Dieses Ergebnis in der Perzeption konnte mit den Ergebnissen aus der Produktionsstudie in Verbindung gebracht werden, welche ebenfalls geschlechtsspezifische Unterschiede in der Distanz zwischen den drei Sibilantkategorien offen legte.

Genauer zeigten die Produktionsdaten in Kapitel 4, dass, gemessen anhand spektraler Eigenschaften im Sibilanten selbst, sowie anhand der F2 und F3-Werte in den nachfolgenden Vokal, für die kindlichen und die männlichen erwachsenen Versuchspersonen, der Abstand zwischen den Sibilantkategorien geringer ausfällt als für die weiblichen Probanden, wobei

die stärksten Unterschiede zwischen den Sprechergruppen für das /s ʂ/-Sibilantpaar beschrieben werden konnten.

Zusammengefasst weisen die Erkenntnisse der experimentellen Studien dieser Arbeit auf eine Neigung zur Instabilität des retroflexen Sibilanten und ferner auf eine Prädisposition in Richtung der /s ʂ/ Neutralisierung hin. Die Ergebnisse lassen sich weiterhin hinsichtlich aktueller Lautwandeltheorien interpretieren. Zum einen stehen die Ergebnisse in Einklang mit Theorien, die von einem nicht-linearen Zusammenhang von synchroner Variation in der Produktion und der Kompensation für diese Variation in der Perzeption ausgehen (Ohala, 1981, 2012). In der Studie zur Perzeption der polnischen Sibilanten in Kapitel 3 wurde darauf hingedeutet, dass die Verschiebung der Kategoriengrenze des /s-ʂ/ Kontinuums im prosodisch schwachen Kontext möglicherweise eine geringere Kompensation für Koartikulation in der Deakzentuierung andeutet, was wiederum zu den Erkenntnissen aus Studien von Harrington et al. (2013) passen würde. Diese zeigten, dass es der prosodische Kontext sein kann, welcher die Kompensation für Koartikulation dahingehend beeinflussen kann, dass in prosodisch schwachen Positionen die Kompensation für Koartikulation verringert wird und somit die Bedingungen für Ohala's Theorie zu Misinterpretationen des Gehörten und Lautwandel gegeben wären. Zum zweiten konnten Hinweise auf Theorien aufgezeigt werden, die die Unterschiede in Bezug auf den Sprachstil und damit verbundene Reduzierungen in der lautlichen Kontrasterhaltung (Lindblom, 1990; Lindblom et al., 1995) in den Vordergrund stellen. In der Studie in Kapitel 3 wurden die Kontinua je nach prosodischem Kontext anders wahrgenommen, d.h. die steileren Perzeptionskurven in der akzentuierten Position und die flacheren Kurven in der Deakzentuierung deuteten an, dass die Versuchspersonen in der Kategorisierung der Sibilantkontinua im deakzentuierten Kontext Unsicherheiten zeigen. Außerdem können Aussagen in Bezug auf Theorien getroffen werden, die den Zusammenhang zwischen der Verbreitung von sprachlichen Innovationen in der Sprachgemeinschaft und verschiedenen extra-linguistischen Faktoren fokussieren Labov (1990, 2001, 2007). In den Kapiteln 3 und 4 wurde getestet, wie das Geschlecht (Kap. 3 und 4) und das Alter (Kap. 4) der Probanden den Kontrast zwischen den Sibilanten beeinflusst. Hier wurde konsistent mit Labov (1990) gezeigt, dass die weiblichen Probanden eher die Standard-sprachliche Formen verwenden, was sowohl durch die stärkere Kontrasterhaltung durch peripherere Sibilantkategorien in der Produktion und steilere/kategorialere Perzeptionskurven im Perzeptionsexperiment gezeigt werden konnte. Die mögliche größere akustische Distanz zwischen den drei Sibilanten bei weiblichen L1-Polnisch Sprechern wird auch als

Motivation in dem von Czaplicki, Żygis, Pape und Jesus (2016) beobachteten Lautwandel im Sibilantensystem des Polnischen angeführt. Dieser von jungen L1-Polnisch sprechenden Frauen angeführte Wandel, ist dadurch gekennzeichnet, dass die alveolopalatale Lautklasse in informeller Sprache eher als ein palatalisierter dentaler Sibilant produziert wird. In ihrer akustischen Analyse, welche einen Vergleich der Standard-Polnischen und der neuen Variante des Sibilanten /ɕ/ beinhaltete, wurden höhere *center of gravity*-Werte für die neue Variante des alveolopalatalen Sibilanten, sowie ebenfalls höhere Werte in den Transitionen des zweiten Formanten des vorangehenden Vokales gefunden. Weiterhin wurde in der Analyse der Sibilanten von Czaplicki und Kollegen (2016) gezeigt, dass durch die neue Variante des alveolopalatalen Sibilanten die akustische Distanz zu den beiden Sibilanten /s ɕ/ erhöht wird, was wiederum mit der Neigung hin zu einer kontrastreicheren Sprache für weibliche Sprecher in Verbindung gebracht werden könnte (Henton, 1995; Labov, 1990; Simpson, 2000; Weirich & Simpson, 2015).

Die männlichen Probanden in der hier vorliegenden Arbeit zeigten hingegen etwas weniger distinkte Kategorien in der Produktion der Sibilanten, sowie etwas flachere/weniger kategoriale Perzeptionskurven in der Wahrnehmung der Sibilanten. Für die kindlichen Probanden in Kapitel 4 wurden Ergebnisse gefunden, welche mit den Kategorisierung der männlichen Erwachsenen und nicht mit denen der weiblichen erwachsenen Probanden vergleichbar sind. Aufgrund dieser Ergebnisse könnte man vermuten, dass es möglicherweise die Sprachmuster und die Sprachverarbeitung der kindlichen und vor allem aber die der männlichen Sprachgemeinschaft ist, welche die Wahrscheinlichkeit für einen Zusammenfall des retroflexen und des dentalen Sibilanten erhöhen.

Ferner wurden Überlegungen aufgezeigt, die sich damit beschäftigten, inwiefern die Ergebnisse in eine Theorie eingepasst werden können, welche sowohl die Initiierung von sprachlichen Neuerungen als auch die Ausbreitung von Lautwandel im Rahmen der Repräsentation von lautlichen Exemplaren im Sprecher/Hörer, erklärt (Johnson, 1997; Pierrehumbert, 2001, 2002). Hier bieten wiederum die Ergebnisse aus der Perzeptionsstudie Raum für Spekulationen. Für das /s-ɕ/Kontinuum wurde im schwachen prosodischen Kontext zwar eine Verschiebung der Kategoriengrenze beobachtet, jedoch blieb der vermutete Effekt auf die Neigungswerte der Perzeptionskurve, nämlich flachere Kurven in der Deakzentuierung, aus. Ein Versuch diese Beobachtung zu erklären sah wie folgt aus: die Verschiebung von ursprünglichen /ɕ/-Exemplaren in die /s/-Wolke und die zeitgleiche Reduzierung der /ɕ/-Exemplar-Wolke erhält die perzeptive Distanz zwischen den beiden Kategorien und damit die Steilheit der Perzeptionskurven bei zeitgleicher Verschiebung der Kategoriengrenzen.

Grundsätzlich ist zu sagen, dass der Effekt der verschiedenen getesteten Faktoren dennoch relativ gering ausfiel, was möglicherweise mit Beobachtungen aus Guzik und Harrington (2007) in Zusammenhang zu bringen ist. In ihrer physiologischen Studie zur Assimilation des Artikulationsortes bei polnischen Sibilanten wurde angemerkt, dass initiale Konsonanten des polnischen diachron stabiler sind als Konsonanten in wortmedialer oder finaler Position. Der nicht so deutlich ausfallende Einfluss der Sprechgeschwindigkeit (Kap. 2), des Vokalkontexts (Kap. 2), der lexikalischen Frequenz (Kap. 2), sowie der Prosodie (Kap. 3) könnte damit teilweise auch der initialen Position der Sibilanten in den jeweiligen Zielwörtern geschuldet sein. Außerdem ist, vor dem Hintergrund der artikulatorischen Einschränkungen, die notwendig sind, um die typischen akustischen Eigenschaften der sibilantischen Lautklasse zu erhalten (Recasens & Espinosa, 2009; Recasens et al., 1997), anzunehmen, dass durch ebendiese Einschränkungen, die Variabilität der Sibilanten in den hier durchgeführten Produktionsstudien eher geringe Ausmaße hatte.

Hinsichtlich der Neutralisierung der Sibilanten in einigen polnischen dialektalen Varietäten wäre es von Interesse in zukünftigen Studien den nachfolgenden Fragestellungen nachzugehen:

- Bezogen auf die dialektale Verteilung der /s ʂ/ Neutralisierung wäre es lohnend artikulatorische, akustische, sowie perzeptive Daten von polnischen Dialektsprechern zu erheben, welche den Kontrast neutralisieren, um festzustellen, wie sich die verschiedenen Modalitäten im Zuge einer tatsächlichen Neutralisierung verhalten.
- Ein weiterführender Aspekt zur Neutralisierung des /s ʂ/-Kontrasts könnte die sogenannten *trading relations* betreffen. In der Sprachproduktion können multiple akustische Merkmale einen Kontrast erhalten (für Beispiele zu *trading relations* siehe Harrington et al., 2016, S. 69). Sind zwei akustische Merkmale in einer *trading relation*, kann das Maß von einem Merkmal invers mit dem anderen Merkmal variieren. In Bezug auf den Kontrast zwischen dem dentalen und dem retroflexen Sibilanten wäre die Überlegung möglich, dass bei einer tatsächlichen Neutralisierung der Kontrast durch ein anderes Merkmal, beispielsweise die Segmentlänge, beibehalten werden könnte. Chang (2011) vermutete so zum Beispiel längere Segmentdauern für stark retroflexen Sibilanten des Mandarin-Chinesischen. Ein Zusammenhang zwischen Grad der Retroflexion und der Segmentdauer konnte in Chang's Studie jedoch nur für eine Versuchsperson gefunden werden. Dennoch wären Überlegungen und Experimente in

diese Richtung sinnvoll, da durch alternative akustische Merkmale Ambiguitäten in der Bedeutung des Gesagten/Gehörten, welche durch die Neutralisierung des Kontrasts entstehen, aufgelöst werden könnten. Hierfür könnte geprüft werden, ob ein Dauerunterschied zwischen echten dentalen Sibilanten und dentalen Sibilanten die zugrundeliegend retroflex repräsentiert sind, existieren. Auch wäre denkbar, dass die Unterscheidung der Sibilanten durch die umliegenden Vokale aufrechterhalten werden könnte. Wie in Abschnitt 1.4.1 dieser Arbeit beschrieben, existieren für die Vokale des Polnischen keine in Qualität oder Quantität variierende Gegenstücke, sodass rein hypothetisch Qualitäts- und/oder Quantitätsunterschiede in den umliegenden Vokalen zur Auflösung von Ambiguitäten bei einer /s ʂ/ Neutralisierung beitragen könnten.

- Die Perzeption der Sibilanten wurde in Kapitel 3 nur in einem Vokalkontext untersucht. Für ein Folgeexperiment wäre es sinnvoll die Wahrnehmung der Sibilantkontinua in weiteren Vokalkontexten, vor allem in Hinblick auf die Überlegungen zu einer möglichen verringerten Kompensation für Koartikulation in prosodisch schwachen Kontexten in Kapitel 3, zu prüfen.
- Mit dem oberen Punkt verbunden steht die Überlegung zu einer rein akustischen Studie, in welcher die Sibilanten in verschiedenen Vokalkontexten produziert werden sollten. Zwar wurde in Kapitel 2 die Artikulation und die Akustik der Sibilanten nicht von den Vokalkontexten /ɛ a ɔ/ beeinflusst, jedoch könnte dieses Ergebnis den Sensoren geschuldet sein, welche auf der Zunge der Probanden befestigt wurden. Die artikulatorischen Einschränkungen für die Sibilantproduktion könnte durch die Sensoren noch verstärkt worden sein, d.h. um die akustischen Eigenschaften der Sibilanten trotz der Sensoren beizubehalten, würde die Variabilität der Artikulatoren weiter eingeschränkt werden, was wiederum verringerte koartikulatorische Effekte der umliegenden Vokale auf die Sibilanten als Konsequenz gehabt haben könnte.
- Die Erkenntnisse aus Guzik und Harrington (2007) hinsichtlich der größeren Stabilität von polnischen Sibilanten in initialer Position bringen zum einen eine Erklärung für die geringe Variabilität der Sibilanten in der vorliegenden Arbeit mit sich, zum anderen können die Erkenntnisse aus dieser früheren Studie als Motivation für weiterführende Experimente angesehen werden, in welchen die Produktion und die Perzeption der Sibilanten in verschiedenen Wortpositionen getestet werden könnten.
- In der Studie von M. Stevens und Harrington (2016) zur *s-retraction* wurde gezeigt, dass die dentalen Sibilanten aus sC/sCr-Klustern, die von männlichen Probanden pro-

duziert wurden, eher als die Sibilanten der weiblichen Probanden, der /ʃ/-Kategorie zugeordnet wurden. Man stellte die Vermutung an, dass die *s-retraction* ihre Wurzeln möglicherweise in der männlichen Sprache findet und das Hörer die Rückverlagerung des alveolaren Sibilanten im Kontext von sC/sCr-Klustern, in der kognitiven Repräsentation mit der Sprache der männlichen Sprachgemeinschaft verknüpfen. In Bezug auf die Sibilantneutralisierung des Polnischen wäre daher ebenfalls lohnend zu testen, wie die Produktionsdaten der männlichen L1-Polnisch Sprecher aus Kapitel 4, welche weniger starke Kontraste zwischen den Sibilanten und vor allem zwischen dem /s ʂ/-Sibilantpaar in den untersuchten Parametern F2 und F3-Werte zum zeitlichen Vokalansatz aufzeigten, von Hörern kategorisiert werden würden.

Diese Studie war der erstmalige Versuch die Standard-Polnischen Sibilanten und eine mögliche Neutralisierung des Dreier-Kontrasts hin zu einem Zweier-Kontrast in diesem Sibilantensystem, wie sie für einige regionale polnische Varietäten beschrieben wurde, zu untersuchen. Dabei wurde der Einfluss von verschiedenen internen und externen Faktoren auf die Stabilität der Standard-Polnischen Sibilanten /s ʂ ʐ/ in der Produktion und in der Perzeption getestet. Diese Studie konnte aufzeigen, dass für die Neutralisierung von Lautkontrasten nicht ein einziger Faktor eine Rolle spielt, sondern dass es das Zusammenspiel verschiedener Gegebenheiten ist, welches die Bedingungen für einen Lautwandel schaffen kann. Weiterführende Experimente sind sinnvoll und notwendig, um das bisherige Wissen zu Lautwandel und im Speziellen zu der Prädisposition zur Veränderung des polnischen Sibilantensystems mit neuen Erkenntnissen anzureichern.

Anhang A

Anhang zu Kapitel 2

A.1 Multitaper-Methode

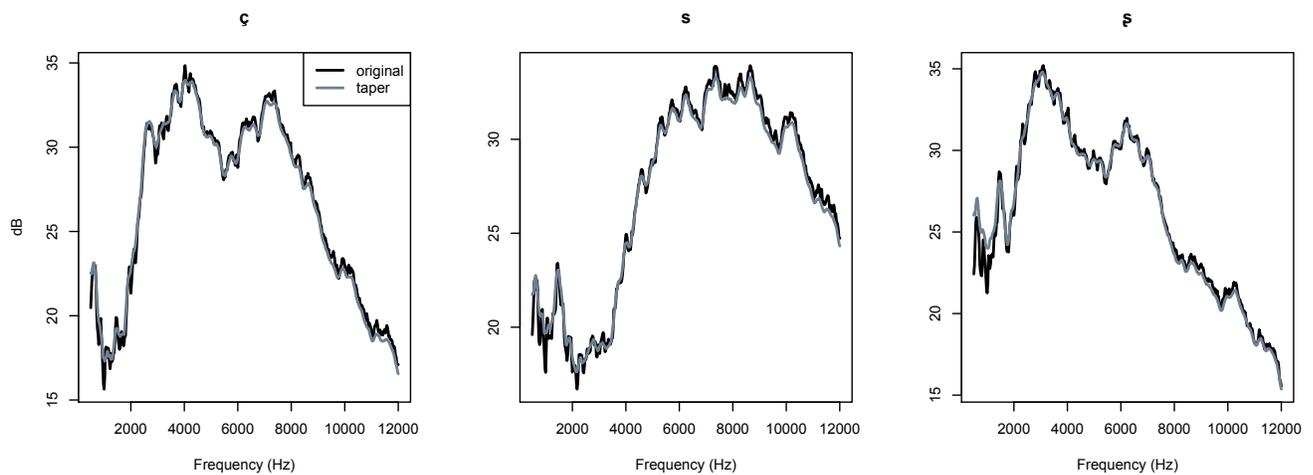


Abbildung A.1: Ensemble-gemittelte Spektren berechnet anhand der Methode in Abb. 2.5 (original) und anhand der Multitaper-Methode (taper) separat für alle drei Sibilanten.

Anhang B

Anhang zu Kapitel 3

B.1 Hintergrundinformationen zu den Teilnehmern des Perzeptionsexperiments

Vpn	Geschlecht	Alter	Region	Ausgabegerät	Eingabegerät
Vpn01	m	26	Masowien	Lautsprecher	Computer
Vpn02	w	32	Niederschlesien	Lautsprecher	Tablet
Vpn03	w	31	Lebus	interne Lautsprecher	Notebook
Vpn04	w	26	Lebus	interne Lautsprecher	Computer
Vpn05	w	25	Niederschlesien	Lautsprecher	Notebook
Vpn06	m	17	Ermland-Masuren	In-Ear Ohrhörer	Computer
Vpn07	w	28	Karpatenvorland	Kopfhörer	Computer
Vpn08	m	22	Kleinpolen	Kopfhörer	Computer
Vpn09	w	32	Ermland-Masuren	Lautsprecher	Computer
Vpn10	w	40	Kleinpolen	Lautsprecher	Computer
Vpn11	m	24	Westpommern	Lautsprecher	Notebook
Vpn12	w	22	Kleinpolen	Lautsprecher	Computer
Vpn13	w	42	Schlesien	In-Ear Ohrhörer	Computer

Tabelle B.1: Hintergrundinformationen zu allen 13 Versuchsteilnehmern, die ursprünglich an dem Perzeptionsexperiment teilgenommen haben.

B.2 Hörspezifische Perzeptionskurven

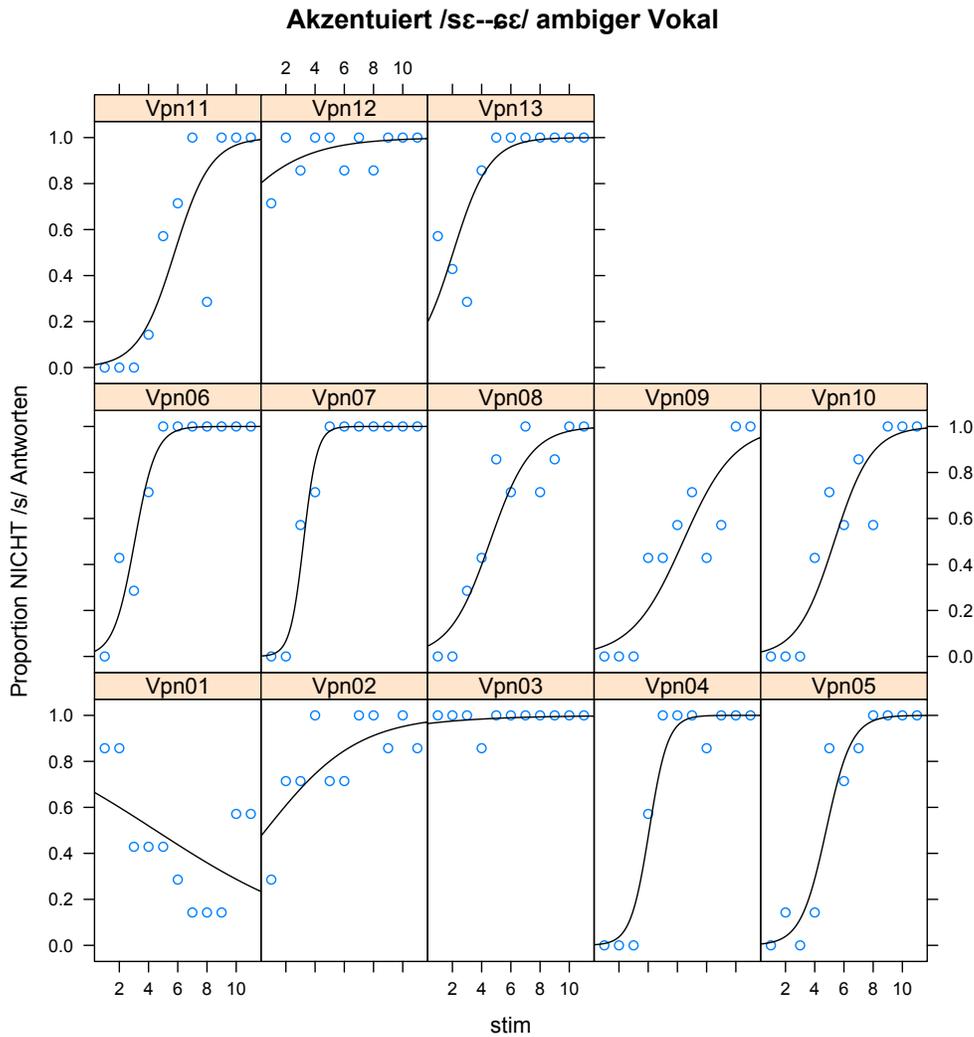


Abbildung B.1: Hörspezifische psychometrische Kurven der Proportion an NICHT-/s/-Antworten als Funktion der Stimulus Nummer für das /sɛ--æɛ/ Kontinuum mit ambigem Vokal und in akzentuierter Bedingung.

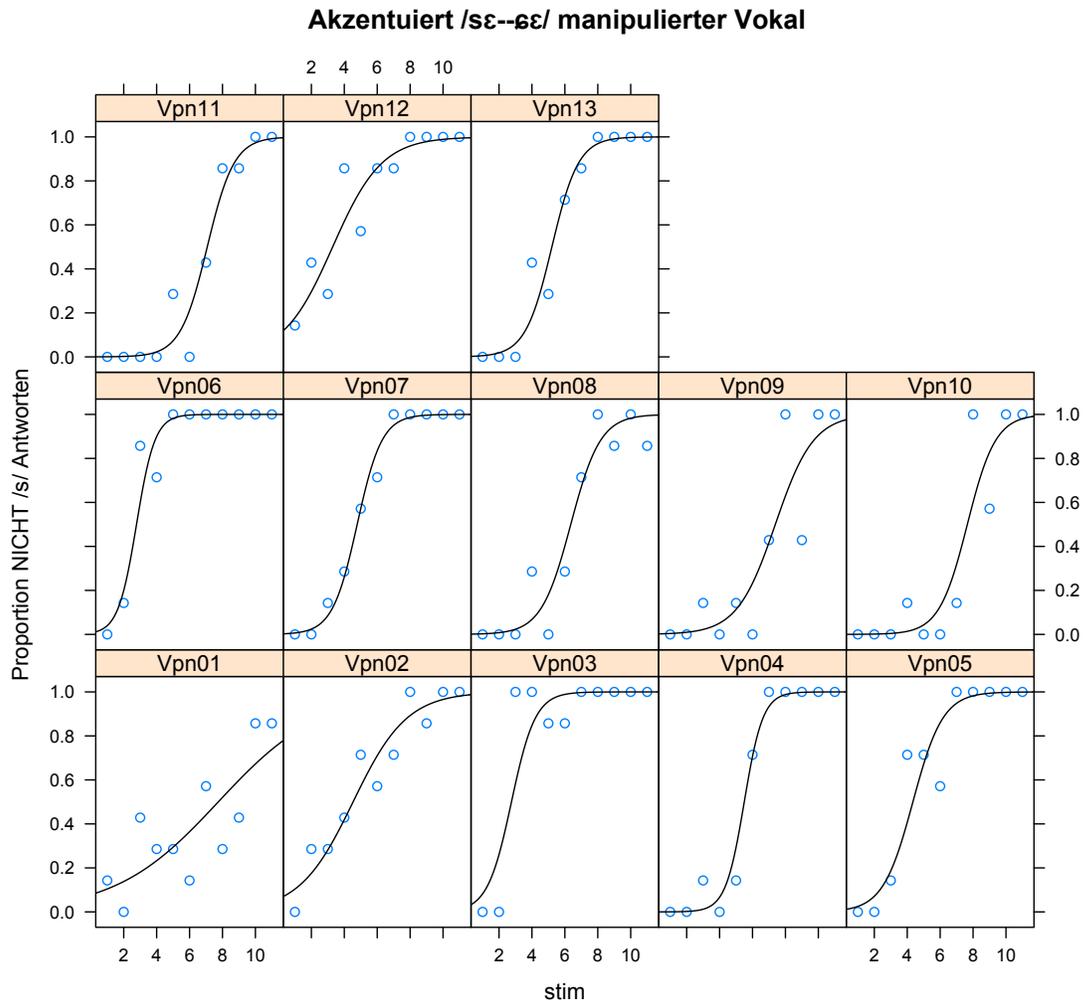


Abbildung B.2: Hörspezifische psychometrische Kurven der Proportion an NICHT-/s/-Antworten als Funktion der Stimulus Nummer für das /sɛ-æɛ/ Kontinuum mit manipuliertem Vokal und in akzentuierter Bedingung.

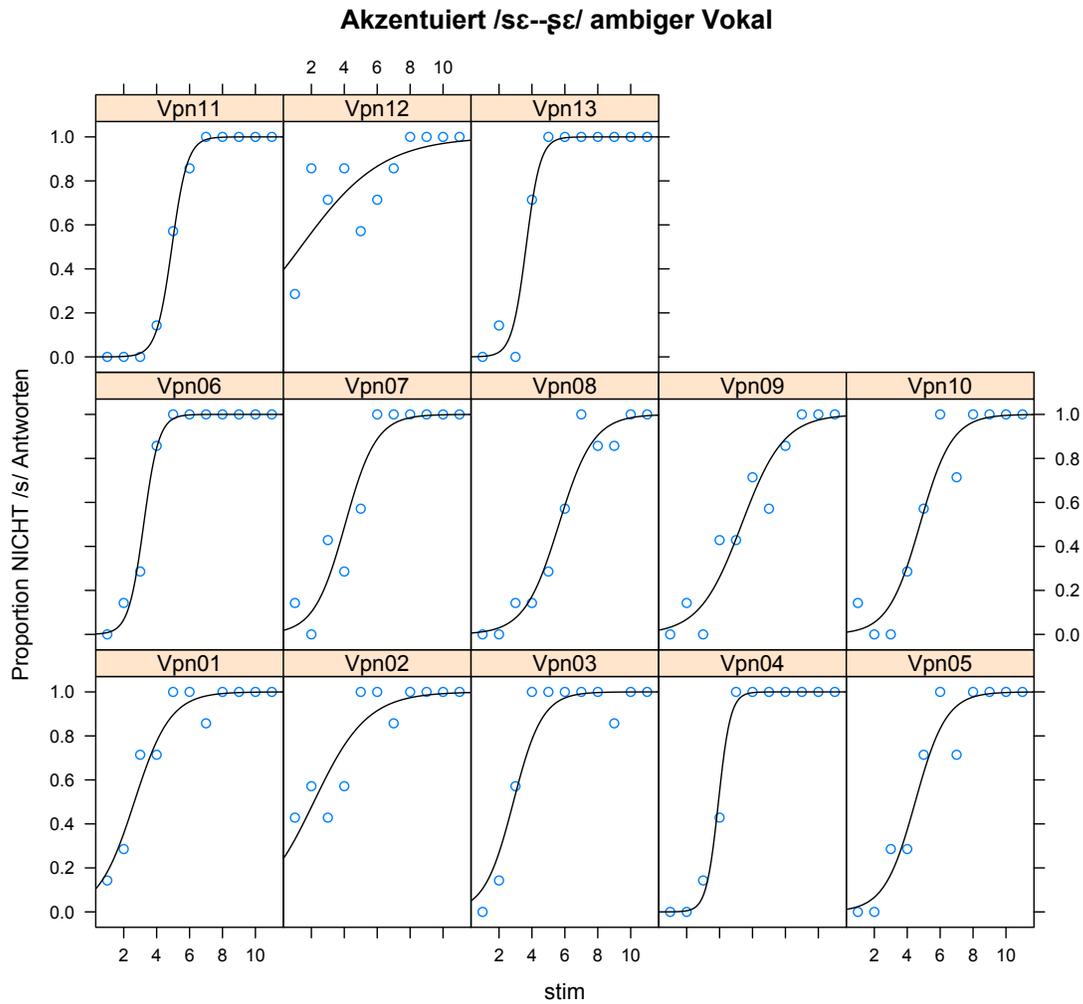


Abbildung B.3: Hörerspezifische psychometrische Kurven der Proportion an NICHT-/s/-Antworten als Funktion der Stimulus Nummer für das /sɛ–ʒɛ/ Kontinuum mit ambigem Vokal und in akzentuierter Bedingung.

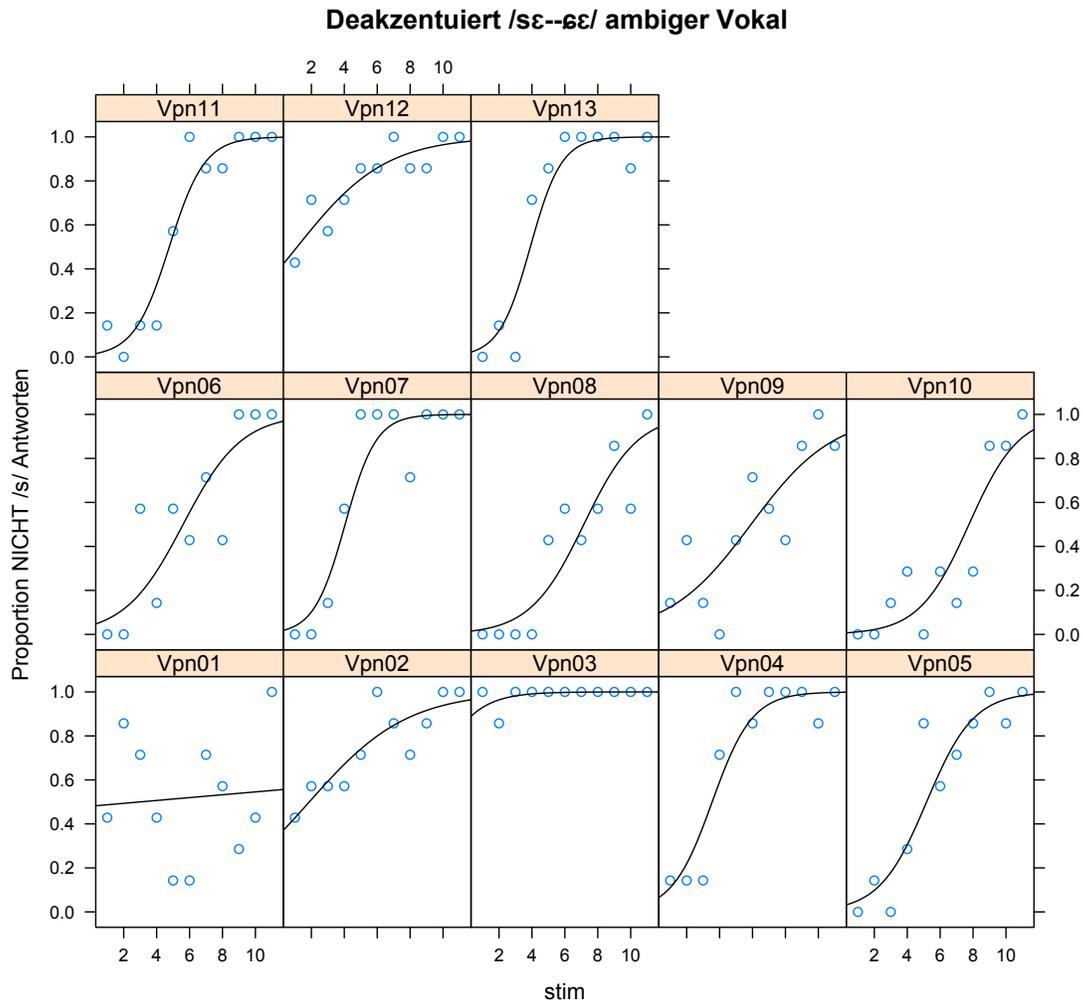


Abbildung B.4: Hörspezifische psychometrische Kurven der Proportion an NICHT-/s/-Antworten als Funktion der Stimulus Nummer für das /sɛ-œ/ Kontinuum mit ambigem Vokal und in deakzentuierter Bedingung.

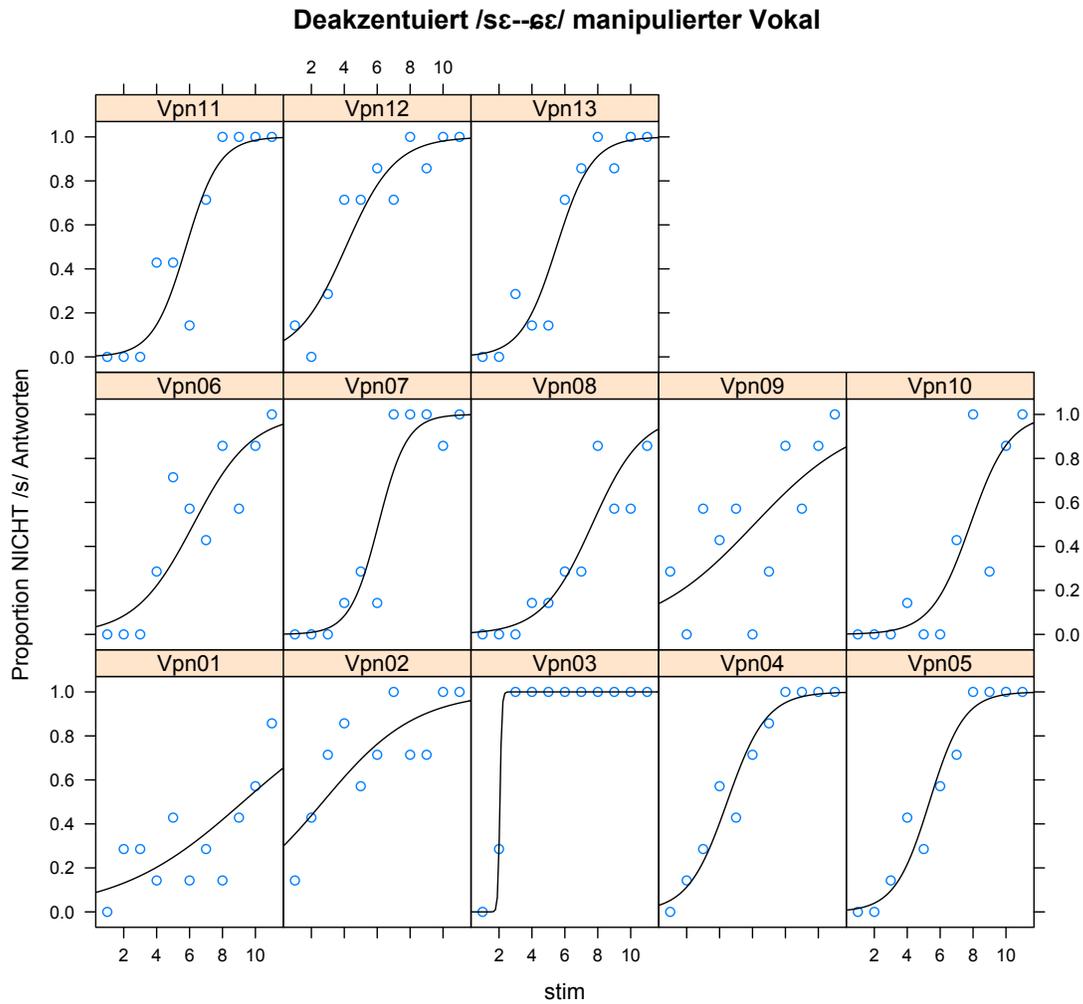


Abbildung B.5: Hörspezifische psychometrische Kurven der Proportion an NICHT-/s/-Antworten als Funktion der Stimulus Nummer für das /sɛ-ɛɛ/ Kontinuum mit manipuliertem Vokal und in deakzentuierter Bedingung.

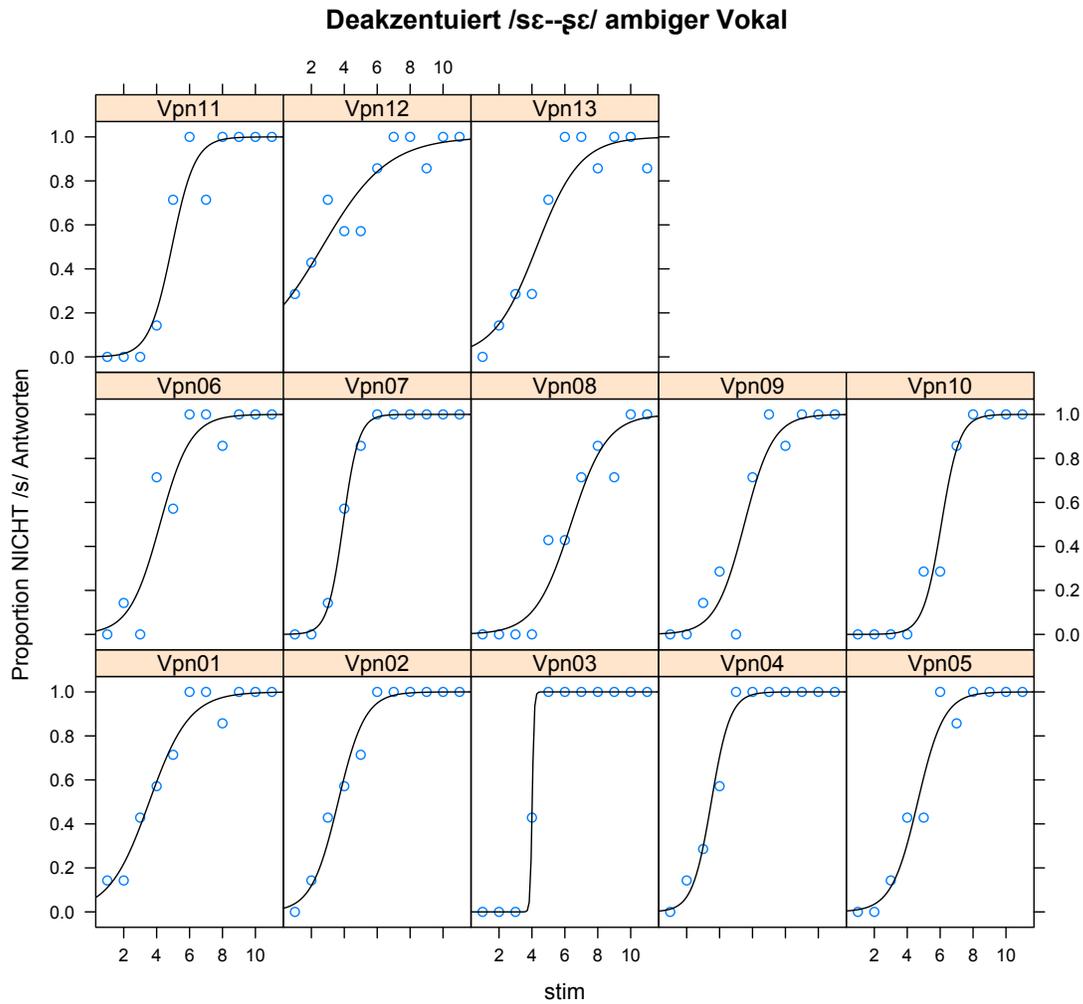


Abbildung B.6: Hörerspezifische psychometrische Kurven der Proportion an NICHT-/s/-Antworten als Funktion der Stimulus Nummer für das /sɛ–ʒɛ/ Kontinuum mit ambigem Vokal und in deakzentuierter Bedingung.

Anhang C

Anhang zu Kapitel 4

C.1 Produktionsdaten separat für die männlichen und weiblichen L1-Polnisch sprechenden kindlichen Versuchsteilnehmer

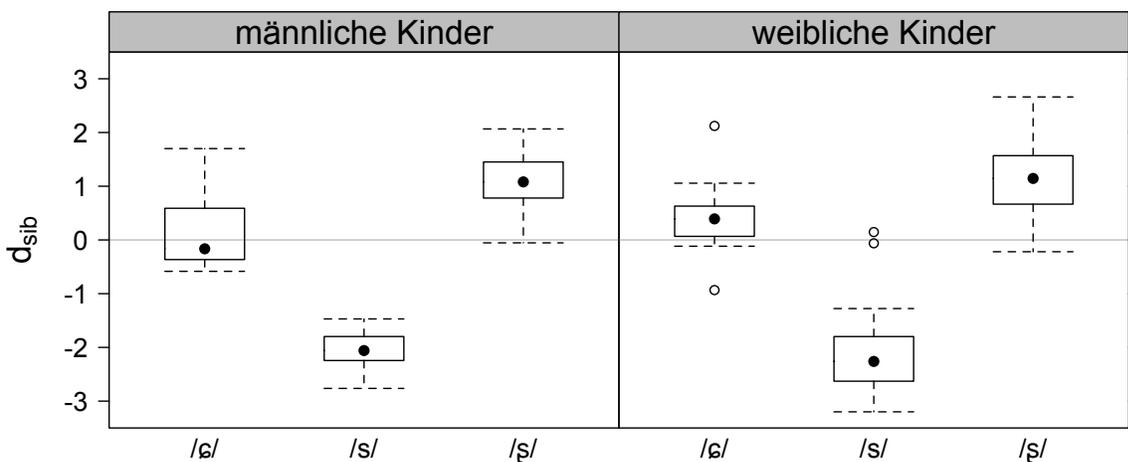


Abbildung C.1: Logarithmisierte Verhältnisse der Euklidischen Distanzen zu den Mittelwerten von Dental $/\text{s}/$ und Retroflex $/\text{ʂ}/$ im $C_1 \times C_2$ -Raum für die drei Sibilanten $/\text{ɕ}/$ (links), $/\text{s}/$ (mittig) und $/\text{ʂ}/$ (rechts), separat für die männlichen und weiblichen kindlichen Versuchspersonen.

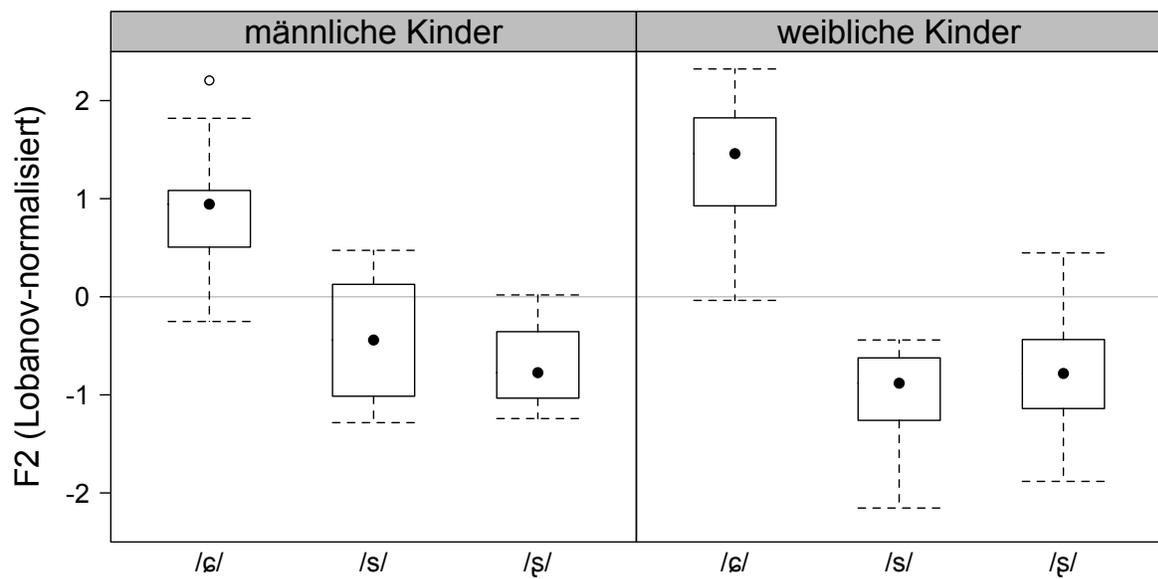


Abbildung C.2: Lobanov-normalisierte F2-Werte zum zeitlichen Onset des Vokals, welcher den Sibilanten /ç/ (links), /s/ (mittig) und /ʃ/ (rechts) folgt, separat für die männlichen und weiblichen kindlichen Versuchspersonen.

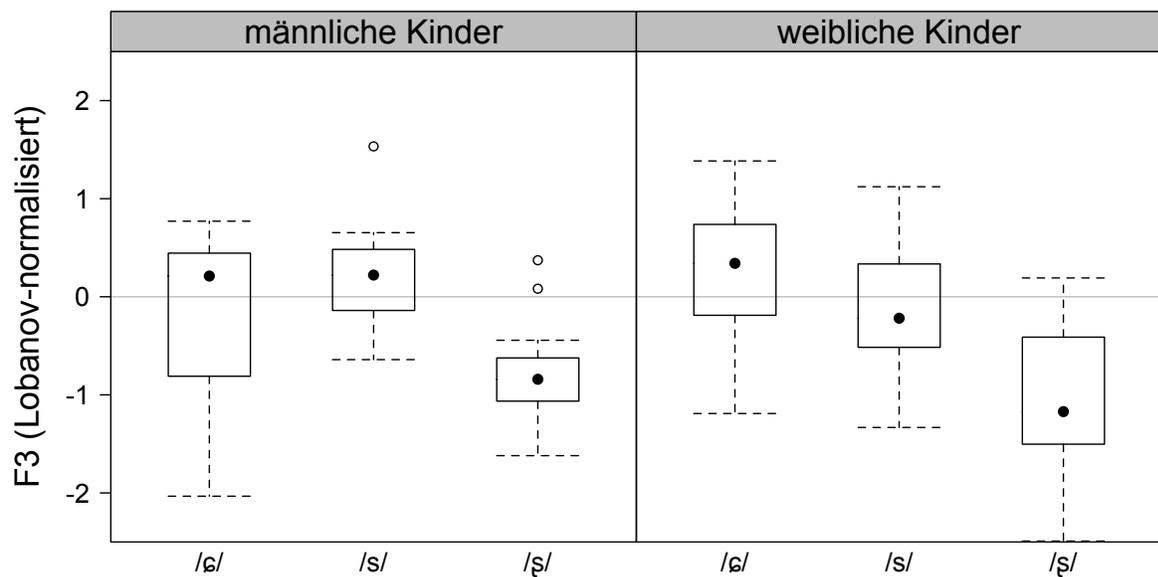


Abbildung C.3: Lobanov-normalisierte F3-Werte zum zeitlichen Onset des Vokals, welcher den Sibilanten /ɕ/ (links), /s/ (mittig) und /ʂ/ (rechts) folgt, separat für die männlichen und weiblichen kindlichen Versuchspersonen.

Anhang D

Publikationsliste

In den nachfolgenden Publikationen, welche vor der Abgabe dieser Dissertation veröffentlicht wurden, wirkte die Autorin der Dissertation als Mitautor mit.

Bukmaier, Véronique, Jonathan Harrington, Ulrich Reubold und Felicitas Kleber (2014). Synchronic variation in the articulation and the acoustics of the Polish threeway place distinction in sibilants and its implications for diachronic change. In *Proceedings of the 15th annual conference of the International Speech Communication Association*. Singapur.

Bukmaier, Véronique, Jonathan Harrington und Felicitas Kleber (2014). An analysis of post-vocalic /s-ʃ/ neutralization in Augsburg German: Evidence for a gradient sound change. *Frontiers in Psychology*, 5, 1–12.

Stevens, Mary, Véronique Bukmaier und Jonathan Harrington (2015). Pre-consonantal /s/-retraction. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow.

Bukmaier, Véronique und Jonathan Harrington (2016). The articulatory and acoustic characteristics of Polish sibilants and their consequences for diachronic change. *Journal of the International Phonetic Association*. Zugriff auf <http://dx.doi.org/10.1017/S0025100316000062>
Das zweite experimentelle Kapitel entstammt größtenteils dem letztgenannten Beitrag.

Literaturverzeichnis

- Aitchison, J. (2003). Psycholinguistic perspectives on language change. In B. Joseph & R. Janda (Hrsg.), *Handbook of Historical Linguistics* (S. 736–743). Oxford, Vereinigtes Königreich: Blackwell.
- Baker, A., Archangeli, D. & Mielke, J. (2011). Variability in American English s-retraction suggests a solution to the actuation problem. *Language Variation and Change*, 23, 347–374.
- Beckman, M., deJong, K., Jun, S. & Lee, S. (1992). The interaction of coarticulation and prosody in sound change. *Language and Speech*, 35 (1, 2), 45–58.
- Beckman, M., Li, F., Kong, E. & Edwards, J. (2014). Aligning the timelines of phonological acquisition and change. *Laboratory Phonology*, 5 (1), 151–194.
- Beddor, P. (2009). A coarticulatory path to sound change. *Language*, 85 (4), 785–821.
- Berlinska, A. (1998). Teaching mother tongue Polish. In W. Tulasiewicz & A. Adams (Hrsg.), *Teaching the mother tongue in a multilingual Europe* (S. 163–173). London, Vereinigtes Königreich: Cassell.
- Bladon, R., Clark, C. & Mickey, K. (1987). Production and perception of sibilant fricatives: Shona data. *Journal of the International Phonetic Association*, 17, 39–65.
- Bladon, R. & Nolan, F. (1977). A video-fluorographic investigation of tip and blade alveolars in English. *Journal of Phonetics*, 5, 185–193.
- Boersma, P. & Weenik, D. (2013). *Praat: doing phonetics by computer [computer program]. version 5.3.53*. Zugriff am 9.7.2013 auf <http://www.praat.org/>
- Bombien, L. (2011). *Segmental and prosodic aspects in the production of consonant clusters – on the goodness of cluster* (PhD. Thesis). Ludwig-Maximilians-Universität, München, Deutschland.
- Browman, C. & Goldstein, L. (1991). Gestural structures: Distinctiveness, phonological processes, and historical change. In I. Mattingly & M. Studdert-Kennedy (Hrsg.), *Modularity and the Motor Theory of Speech Perception* (S. 313–338). New Jersey,

- USA: Erlbaum.
- Bukmaier, V. & Harrington, J. (2016). The articulatory and acoustic characteristics of Polish sibilants and their consequences for diachronic change. *Journal of the International Phonetic Association*.
- Bukmaier, V., Harrington, J. & Kleber, F. (2014). An analysis of post-vocalic /s-ʃ/ neutralization in Augsburg German: Evidence for a gradient sound change. *Frontiers in Psychology*, 5.
- Bukmaier, V., Harrington, J., Reubold, U. & Kleber, F. (2014). Synchronic variation in the articulation and the acoustics of the Polish place distinction in sibilants and its implications for diachronic change. In *Proceeding of the 15th Annual Conference of the International Speech Communication Association*. Singapur.
- Bybee, J. & Hopper, P. (2001). *Frequency and the emergence of linguistic structure*. Cambridge, Vereinigtes Königreich: Cambridge University Press.
- Campbell, L. (2004). *Historical Linguistics* (2. Aufl.). Edinburgh University Press.
- Chang, Y. (2011). A corpus study of retroflex realizations in Beijing and Taiwan Mandarin. In *Proceedings of the 17th International Conference of Phonetic Sciences* (S. 440–443). Hong Kong.
- Chang, Y. (2012). *Variability in cross-dialectal production and perception of contrasting phonemes: The case of the alveolar-retroflex contrast* (PhD. Thesis). University of Illinois, Urbana, Illinois.
- Chang, Y. & Shih, C. (2015). Place contrast enhancement: The case of the alveolar and retroflex sibilant production in two dialects of Mandarin. *Journal of Phonetics*, 50, 52–66.
- Chang, Y., Shih, C. & Allen, J. (2013). Variability in cross-dialect perception of Mandarin alveolar-retroflex contrast. In *Proceedings of International Conference of Phonetics of the Languages in China*.
- Cheon, S. & Anderson, V. (2008). Acoustic and perceptual similarities between English and Korean sibilants: Implications for second language acquisition. *Korean Linguistics*, 14, 41–64.
- Chiu, C. (2009). Acoustic and auditory comparisons of Polish and Taiwanese Mandarin sibilants. *Canadian Acoustics*, 37 (3), 142–143.
- Chuang, Y. & Fon, J. (2010). The effect of prosodic prominence on the realization of voiceless dental and retroflex sibilants in Taiwan Mandarin spontaneous speech. In *Proceedings of the 5th International Conference on Speech Prosody*. Chicago, USA.

- Chuang, Y. & Fon, J. (2011). Cross-dialectal perception of voiceless dental and retroflex sibilant variants in Taiwan Mandarin. In *Proceedings of the 17th International Conference of Phonetic Sciences* (S. 496–499). Hong Kong.
- Chung, K. (2006). Hypercorrection in Taiwan Mandarin. *Journal of Asian Pacific Communication*, 16 (2), 197–214.
- Clark, E. & Bowerman, M. (1986). The Fergusonian impact: In honor of Charles A. Ferguson on the occasion of his 65th birthday. In J. A. Fishman, A. Tabouret-Keller, M. Clyne, B. Krishnamurti & M. Abdulaziz (Hrsg.), *From phonology to syntax* (Bd. 1, S. 51–68). Berlin, Deutschland: Mouton de Gruyter.
- Cole, J. & Hualde, J. (2013). Prosodic structure in sound change. In S. Chen & B. Slade (Hrsg.), *Grammatica et verba: Glamor and verve: Studies in South Asian, historical, and Indo-European linguistics in honor of Hans Heinrich Hock on the occasion of his seventy-fifth birthday*. (S. 28–46). Ann Arbor, New York: Beech Stave Press.
- Cole, J., Kim, H., Choi, H. & Hasegawa-Johnson, M. (2007). Prosodic effects on acoustic cues to stop voicing and place of articulation. *Journal of Phonetics*, 190–209.
- Czaplicki, B., Żygis, M., Pape, D. & Jesus, L. (2016). Acoustic evidence of new sibilants in the production of young Polish women. *Poznań Studies in Contemporary Linguistics*, 52, 1–42.
- Dart, S. (1991). *Articulatory and acoustic properties of apical and laminal articulations*. (PhD. Thesis). University California.
- Darwin, C. (1887). A biographical sketch of an infant. *Mind*, 2, 285–294.
- deJong, K. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America*, 97 (1), 491–504.
- deJong, K. (1998). Stress-related variation in the articulation of coda alveolar stops: Flapping revisited. *Journal of Phonetics*, 26, 283–310.
- Delattre, P., Liberman, A. & Cooper, F. (1962). Formant transitions and loci as acoustic correlates of place of articulation in American fricatives. *Studia linguistica*, 16, 104–121.
- Draxler, C. (2011). Percy - an html5 framework for media rich web experiments on mobile devices. In *Proceedings of the 12th Annual Conference of the International Speech Communication Association* (S. 3339–3340).
- Draxler, C. & Jänsch, K. (2004). Speechrecorder – a universal platform independent multichannel audio recording software. In *Proceedings of the 4th International Conference*

- Language Resources and Evaluation* (S. 559–562). Lisabon, Portugal.
- Duanmu, S. (2006). Chinese (Mandarin): Phonology. In K. Brown (Hrsg.), *Encyclopedia of Language and Linguistics* (2. Aufl., S. 351–355). Oxford, Vereinigtes Königreich: Elsevier Publishing House.
- Eckert, P. (1989). *Jocks and Burnouts: Social identity in the high school*. New York: Teach. Coll. Press.
- Eckert, P. (1997). Age as a sociolinguistic variable. In F. Coulmas (Hrsg.), *Handbook of Sociolinguistics* (S. 151–167). Oxford, Vereinigtes Königreich: Basil Blackwell.
- Evans, B. & Iverson, P. (2004). Vowel normalization for accent: An investigation of best exemplar locations in northern and southern British English sentences. *Journal of the Acoustical Society of America*, 115 (1), 352–361.
- Evers, V., Reetz, H. & Lahiri, A. (1998). Crosslinguistic acoustic categorization of sibilants independent of phonological status. *Journal of Phonetics*, 26, 345–370.
- Fant, G. (1960). *Acoustic theory of speech production*. The Hague: Mouton.
- Ferguson, C. & Farwell, C. (1975). Words and sounds in early language acquisition. *Language*, 51 (2), 419–439.
- Fitch, W. & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America*, 106 (3), 1511–1522.
- Flemming, E. S. (2002). *Auditory representations in phonology*. New York: Routledge.
- Forrest, K., Weismer, G., Milenkovic, P. & Dougall, R. D. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America*, 84, 115–123.
- Foulkes, P. & Vihman, M. (2015). First language acquisition and phonological change. In P. Honeybone & J. Salmons (Hrsg.), *The Oxford Handbook of Historical Phonology* (S. 1–34). Oxford, Vereinigtes Königreich: Oxford University Press.
- Fowler, C. & Galantucci, B. (2005). The relation of speech perception and production. In D. Pisoni & R. Remez (Hrsg.), *The Handbook of Speech Perception* (S. 633–652). Malden, USA: Blackwell.
- Fromkin, V., Rodman, R. & Hyams, N. (2011). *An introduction to language* (9. Aufl.). Boston: Wadsworth.
- Fujisaki, H. & Kunisaki, O. (1978). Analysis, recognition, and perception of voiceless fricative consonants in Japanese. *IEEE Transactions (ASSP)*, 26, 21–27.
- Garrett, A. & Johnson, K. (2012). Phonetic bias in sound change. In A. Yu (Hrsg.), *Origins*

- of sound change: Approaches to phonologization* (S. 51–97). Oxford University Press.
- Gordon, M., Barthmaier, P. & Sands, K. (2002). A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association*, 32, 141–174.
- Gussmann, E. (2007). *The phonology of Polish*. Oxford, UK: Oxford University Press.
- Guzik, K. & Harrington, J. (2007). The quantification of place of articulation assimilation in electropalatographic data using the similarity index (SI). *Advances in Speech-Language Pathology*, 9 (1), 109–119.
- Halle, M. & Stevens, K. (1997). The postalveolar fricatives of Polish. In S. Kiritani, H. Hirose & H. Fujisaki (Hrsg.), *Speech production and language: In honour of Osamu Fujimura* (S. 177–193). Berlin, Deutschland: Mouton de Gruyter.
- Hamann, S. (2002a). Postalveolar fricatives in Slavic languages as retroflexes. In S. Baauw, M. Huiskes & M. Schoorlemmer (Hrsg.), *OTS yearbook 2002* (S. 105–127). Utrecht, Niederlande: Utrecht Institute of Linguistics.
- Hamann, S. (2002b). Retroflexion and retraction revised. In T. A. Hall, B. Pompino-Marschall & M. Rochoń (Hrsg.), *Papers on phonetics and phonology: The articulation, acoustics and perception of consonants* (Bd. 28, S. 13–26).
- Harrington, J. (2007). Evidence for a relationship between synchronic variability and diachronic change in the Queen’s annual Christmas broadcasts. In J. Cole & J. Hualde (Hrsg.), *Laboratory Phonology 9* (S. 125–143). Berlin, Deutschland: Mouton de Gruyter.
- Harrington, J. (2010). *Phonetic analysis of speech corpora*. Chichester, England: Wiley-Blackwell.
- Harrington, J., Kleber, F. & Reubold, U. (2008). Compensation for coarticulation, /u/-fronting, and sound change in Standard Southern British: An acoustic and perceptual study. *Journal of the Acoustical Society of America*, 123, 2825–2835.
- Harrington, J., Kleber, F. & Reubold, U. (2013). The effect of prosodic weakening on the production and perception of trans-consonantal vowel coarticulation in German. *Journal of the Acoustical Society of America*, 134, 551–561.
- Harrington, J., Kleber, F., Reubold, U. & Siddins, J. (2015). The implications for sound change of prosodic weakening: Evidence from polysyllabic shortening. *Journal of Laboratory Phonology*, 6, 81–117.
- Harrington, J., Kleber, F., Reubold, U. & Stevens, M. (2016). The relevance of context and experience for the operation of historical sound change. In A. Esposito & L. Jain (Hrsg.), *Socially believable behaving systems – the quest for equipping machines with*

- human-level automaton intelligence* (S. 1–50). Springer Verlag.
- Harris, K. S. (1958). Cues for the discrimination of American English fricatives in spoken syllables. *Language and Speech*, 1, 1–7.
- Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics*, 31, 373–405.
- Hawkins, S. & Midgley, J. (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association*, 35, 183–199.
- Hay, J., Warren, P. & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, 34, 458–484.
- Heinz, J. & Stevens, K. (1961). On the properties of fricative consonants. *Journal of the Acoustical Society of America*, 33, 589–593.
- Henton, C. (1995). Cross-language variation in the vowels of female and male speakers. In *Proceedings of the 13th International Congress of Phonetic Sciences* (S. 420–423).
- Hickey, R. (2012). Internally and externally motivated language change. In J. M. Hernández-Compoy & J. C. Conde-Silvestre (Hrsg.), *The Handbook of Historical Sociolinguistics* (S. 401–421). Malden, USA: Wiley-Blackwell.
- Hillenbrand, J., Getty, L., Clark, M. & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97, 3099–3111.
- Hockett, C. F. (1950). Age-grading and linguistic continuity. *Language*, 26 (4), 449–457.
- Hombert, J.-M., Ohala, J. & Ewan, W. (1979). Phonetic explanations for the development of tones. *Language*, 55, 37–58.
- Hoole, P. (1987). Bite-block speech in the absence of oral sensibility. In *Proceedings of the 11th International Congress on Phonetic Sciences*. Tallinn, Estland.
- Hoole, P., Bombien, L., Kühnert, B. & Mooshammer, C. (2009). Intrinsic and prosodic effects on articulatory coordination in initial consonant cluster. In G. Fant, H. Fuijsaki & J. Shen (Hrsg.), *Frontiers in phonetics and speech sciences. Festschrift for Wu Zongji* (S. 275–286). Peking, China: Commercial Press.
- Hoole, P. & Mooshammer, C. (2002). Articulatory analysis of the German vowel system. In P. Auer, P. Gilles & H. Spiekermann (Hrsg.), *Silbenschnitt und Tonakzente* (S. 129–152). Tübingen, Deutschland: Niemeyer.
- Hoole, P. & Zierdt, A. (2010). Five-dimensional articulography. In B. Maassen & P. van Lieshout (Hrsg.), *Speech motor control: New developments in basic and applied research* (S. 331–349). OUP.

- Hu, F. (2008). The three sibilants in Standard Chinese. In R. Sock, S. Fuchs & Y. Laprie (Hrsg.), *Proceedings of the 8th International Seminar on Speech Production* (S. 105–108). Strasburg, Frankreich.
- Hughes, G. & Halle, M. (1956). Spectral properties of fricative consonants. *Journal of the Acoustical Society of America*, 28 (7).
- Inkelas, S. & McAllisterByun, T. (2013). *Ontogeny recapitulates phylogeny: Child speech development as a microcosmos of sound change*. (Handout Sound Change Actuation Workshop University Chicago)
- Iskarous, K., Shadle, C. & Proctor, M. (2011). Articulatory-acoustic kinematics: The production of American English /s/. *Journal of the Acoustical Society of America*, 129 (2), 944–954.
- Jakobson, R. (1941). *Child language, aphasia and phonological universals*. The Hague and Paris: Mouton de Gruyter.
- Janda, L. (2005). The Slavic languages. In K. Brown (Hrsg.), *Encyclopedia of Language and Linguistics* (2. Aufl., Bd. 11, S. 415–418). Oxford, Vereinigtes Königreich: Elsevier.
- Jannedy, S., Weirich, M. & Helmeke, L. (2015). Acoustic analyses of differences in [ç] and [j] in Hood German. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, Vereinigtes Königreich.
- Jassem, W. (1995). The acoustic parameters of Polish voiceless fricatives. *Phonetica*, 52, 252–258.
- Jeng, J. (2006). The acoustic spectral characteristics of retroflexed fricatives and affricates in Taiwan Mandarin. *Journal of Humanistic Studies*, 40 (1), 27–48.
- Jesus, L. & Shadle, C. (2002). A parametric study of the spectral characteristics of European Portuguese fricatives. *Journal of Phonetics*, 30, 437–464.
- Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J. W. Mullennix (Hrsg.), *Talker variability in speech processing* (S. 145–165). San Diego, USA: Academic Press.
- Jones, J. & Munhall, K. (2000). Perceptual calibration of f0 production: Evidence from feedback perturbation. *Journal of the Acoustical Society of America*, 108 (3), 1246–1251.
- Jones, J. & Munhall, K. (2003). Learning to produce speech with an altered vocal tract: The role of auditory feedback. *Journal of the Acoustical Society of America*, 113 (1), 532–543.
- Jongman, A. (1989). Duration of fricative noise required for identification of English

- fricatives. *Journal of the Acoustical Society of America*, 85, 1718–1725.
- Jongman, A., Wayland, R. & Wong, S. (1998). Acoustic characteristics of English fricatives: I. static cues. *Working Papers of the Cornell Phonetics Laboratory*, 12, 195–205.
- Jongman, A., Wayland, R. & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America*, 108 (3), 1252–1263.
- Kawahara, H., Takahashi, T., Morise, M. & Banno, H. (2009). Development of exploratory research tools based on TANDEM-STRAIGHT. In *Proceedings of APSIPAASC 2009: Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference* (S. 111-120). Japan: Asia-Pacific Signal and Information Processing Association.
- Keating, P. (1991). Coronal places of articulation. In C. Paradis & J. Prunet (Hrsg.), *The special status of coronals* (S. 29–48). London, UK: Academic Press.
- Kent, R. (1992). The biology of phonological development. In C. Ferguson, L. Menn & C. Stoel-Gammon (Hrsg.), *Phonological development: Models, research, implications* (S. 65–90). Timonium, MD: York Press.
- Kent, R. & Read, C. (1992). *The acoustic analysis of speech*. San Diego, USA: Singular Publishing Group Inc.
- Kingston, J. & Diehl, R. (1994). Phonetic knowledge. *Language*, 70, 419–454.
- Kingston, J. & Diehl, R. (1995). Intermediate properties in the perception of distinctive feature values. In B. Connell & A. Arvaniti (Hrsg.), *Phonology and phonetics: Papers in laboratory phonology IV* (S. 7–27). Cambridge, UK: Cambridge University Press.
- Kiparsky, P. (1988). Phonological change. In F. Newmeyer (Hrsg.), *The Cambridge survey of linguistics* (Bd. 1, S. 363–415). Cambridge, UK: Cambridge University Press.
- Kleber, F., Harrington, J. & Reubold, U. (2012). The relationship between the perception and production of coarticulation during a sound change in progress. *Language and Speech*, 55 (3), 383–405.
- Krakow, A. (1999). Articulatory organization of syllables: A review. *Journal of Phonetics*, 27, 23–54.
- Kudela, K. (1968). Spectral analysis of Polish fricative consonants. In W. Jassem (Hrsg.), *Speech analysis and synthesis* (S. 93–188). Warschau, Polen: Panstwowe Wydawnictwo Naukowe.
- Labov, W. (1963). The social motivation of a sound change. *Word*, 19 (3), 273–309.
- Labov, W. (1968). The reflection of social processes in linguistic structures. In J. Fishman (Hrsg.), *Readings in the sociology of language* (S. 240–251). DE GRUYTER.

- Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change*, 2 (2), 205–254.
- Labov, W. (2001). *Principles of linguistic change* (Bd. 2). Oxford, Vereinigtes Königreich: Blackwell.
- Labov, W. (2007). Transmission and diffusion. *Language*, 83.
- Ladefoged, P. (2001). *Vowels and consonants*. Malden, USA: Blackwell.
- Ladefoged, P. & Bhaskararao, P. (1983). Non-quantal aspects of consonantal production: A study of retroflex consonants. *Journal of Phonetics*, 11, 291–302.
- Ladefoged, P. & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford, Vereinigtes Königreich: Blackwell.
- LaRiviere, C., Winitz, H. & Herriman, E. (1975). The distribution of perceptual cues in English prevocalic fricatives. *Journal of Speech and Hearing Research*, 18, 613–622.
- Lee, S., Potamianos, A. & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America*, 105.
- Lee-Kim, S. (2015). The gradient effect of transitional magnitude: A source of the vowel context effect. In *Proceedings of the 18th International Congress of Phonetic Sciences*.
- Li, F. (2008). *The phonetic development of voiceless sibilant fricatives in English, Japanese and Mandarin Chinese* (PhD. Thesis). The Ohio State University.
- Li, F., Edwards, J. & Beckman, M. (2007). Spectral measures for sibilant fricatives of English, Japanese, and Mandarin Chinese. In *Proceedings of the 16th International Conference of Phonetic Science*.
- Li, F., Edwards, J. & Beckman, M. (2009). Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. *Journal of Phonetics*, 37, 111–124.
- Li, F. & Munson, B. (in Druck). The development of voiceless sibilant fricatives in Putongha-speaking children. *Journal of Speech, Language and Hearing Research*.
- Li, F., Munson, B., Edwards, J., Yoneyama, K. & Hall, K. (2011). Language specificity in the perception of voiceless sibilant fricatives in Japanese and English: Implications for cross-language differences in speech-sound development. *Journal of the Acoustical Society of America*, 129 (2), 999–1011.
- Li, F., Rendall, D., Vasey, P. L., Kinsman, M., Ward-Sutherland, A. & Diano, G. (2016). The development of sex/gender-specific /s/ and its relationship to gender identity in

- children and adolescents. *Journal of Phonetics*, 57, 59–70.
- Li, Y. (2009). Incomplete neutralization of alveolar and retroflexed sibilants in Taiwan Mandarin. In *Papers in Berkeley's linguistic society* (Bd. 35, S. 29–30). Berkeley, USA: University of California at Berkeley.
- Lieberman, A., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74 (6), 431–461.
- Lieberman, A., Harris, K., Hoffman, H. & Griffith, B. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54, 358–368.
- Lieberman, A. & Mattingly, I. (1989). The motor theory revised. *Cognition*, 21, 1–36.
- Lin, S., Beddor, P. & Coetzee, A. (2014). Gestural reduction, lexical frequency, and sound change: A case study of post-vocalic /l/. *Journal of Laboratory Phonology*, 5, 9–36.
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H and H theory. In W. Hardcastle & A. Marchal (Hrsg.), *Speech production and speech modeling* (S. 403–439). Dordrecht, Niederlande: Kluwer.
- Lindblom, B. (1996). Role of articulation in speech perception: Clues from production. *Journal of the Acoustical Society of America*, 99, 1683–1692.
- Lindblom, B. (1998). Systematic constraints and adaptive change in the formation of sound structure. In J. Hurford, M. Studdert-Kennedy & C. Knight (Hrsg.), *Approaches to the evolution of language: Social and cognitive bases*. Cambridge, Vereinigtes Königreich: Cambridge University Press.
- Lindblom, B., Guion, S., Hura, S., Moon, S.-J. & Willerman, R. (1995). Is sound change adaptive? *Rivista di Linguistica*, 7, 5–36.
- Lisker, L. (2001). Hearing the Polish sibilants [s ś š]: Phonetic and auditory judgments. In N. Grønnum & J. Rischel (Hrsg.), *Travaux du cercle linguistique de copenhagen xxxi. to honour eli fischer-jørgensen* (S. 226–238). Kopenhagen, Dänemark: C.A. Reitzel.
- Lisker, L. & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20 (3), 384–422.
- Lobanov, B. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America*, 49, 606–608.
- Locke, J. (1983). *Phonological acquisition and change*. New York: Academic Press.
- Luce, P. & Pisoni, D. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1–36.
- Łukaszewicz, B. (2006). Extrasyllabicity, transparency and prosodic constituency in the

- acquisition of Polish. *Lingua*, 116, 1–30.
- Lukaszewicz, B. (2007). Reduction in syllable onsets in the acquisition of Polish: Deletion, coalescence, metathesis and gemination. *Journal of Child Language*, 34 (1), 53–82.
- Maddieson, I. (1984). *Patterns of sounds*. Cambridge, Vereinigtes Königreich: Cambridge University Press.
- Malisz, Z. & Żygis, M. (2015). Voicing in Polish: Interactions with lexical stress and focus. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, Vereinigtes Königreich.
- Maniwa, K., Jongman, A. & Wade, T. (2009). Acoustic characteristics of clearly spoken English fricatives. *Journal of the Acoustical Society of America*, 125 (6), 3962–3973.
- Mann, V. & Repp, B. (1980). Influence of vocalic context on the perception of the [ʃ]-[s] distinction. *Perception and Psychophysics*, 28, 213–228.
- Mateus, M. H. & d’Andrade, E. (2000). *The phonology of Portuguese*. Oxford, UK: Oxford University Press.
- McGuire, G. (2007). English listeners’ perception of Polish alveolopalatal and retroflex voiceless sibilants: A pilot study. In *Uc berkeley phonology lab annual report* (S. 1–28). University of California Berkeley: eScholarship University of California.
- Meenakshi, N., Yarra, C., Yamini, B. K. & Ghosh, P. K. (2014, September). Comparison of speech quality with and without sensors in electromagnetic articulograph AG 501 recording. In *Interspeech 2014* (S. 935–939). Singapur.
- Menn, L. (1975). Counter example to ‘fronting’ as a universal of child language. *Journal of Child Language*, 2, 293–296.
- Michailovsky, B. (1975). On some Tibeto-Burman sound changes. In *Proceedings of the annual meeting of the berkley linguistics society* (Bd. 1, S. 322–332).
- Miller, G. & Nicely, P. (1955). Analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, 27, 338–353.
- Miller, J. (1981). Effects of speaking rate on segmental distinctions. In P. Eimas & J. Miller (Hrsg.), *Perspectives on the study of speech*. Hillsdale, NJ: Erlbaum.
- Milobog, M. & Garrett, P. (2011). Perceptions of and attitudes towards regional varieties of Polish: Views from two Polish provinces. *Language Awareness*, 20 (4), 275–293.
- Milroy, J. & Milroy, L. (1985). Linguistic change, social network and speaker innovation. *Journal of Linguistics*, 21, 339–384.
- Mitterer, H. & Ernestus, M. (2006). Listeners recover /t/s that speakers reduce: Evidence from /t/-lenition in Dutch. *Journal of Phonetics*, 34 (1), 73–103.

- Moon, S. & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America*, 96 (40).
- Moulines, E. & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication*, 9 (5-6), 453–467.
- Munson, B. (2007). Lexical access, lexical representation, and vowel production. In J. Cole & J. Hualde (Hrsg.), *Laboratory Phonology 9* (S. 201–227). Berlin, Deutschland: Mouton de Gruyter.
- Murray, I. R. & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on vocal emotion. *Journal of the Acoustical Society of America*, 93 (2), 1097–1107.
- Myers, D. (2002). *Social psychology* (7. Aufl.). New York: McGraw Hill Higher Education.
- Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, 18 (1), 62–85.
- Nissen, S. & Fox, R. (2005). Acoustic and spectral characteristics of young children's fricative productions: A developmental perspective. *Journal of the Acoustical Society of America*, 118 (4).
- Nittrouer, S. (1992). Age-related differences in perceptual effect of formant transitions within syllables and across syllable boundaries. *Journal of Phonetics*, 20, 1–32.
- Nittrouer, S. (1995). Children learn separate aspects of speech production at different rates: Evidence from spectral moments. *Journal of the Acoustical Society of America*, 97, 520–530.
- Nittrouer, S. (2002). Learning to perceive speech: How fricatives perception changes, and how it stays the same. *Journal of the Acoustical Society of America*, 112 (2), 711–719.
- Nittrouer, S. (2006). Children hear the forest (1). *Journal of the Acoustic Society of America*, 120 (4), 1799–1802.
- Nittrouer, S. & Studdert-Kennedy, M. (1987). The role of coarticulation effects in the perception of fricatives by children and adults. *Journal of Speech and Hearing Research*, 30, 319–329.
- Nittrouer, S., Studdert-Kennedy, M. & McGowan, R. (1989). The emergence of phonetic segments: Evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Language and Hearing Research*, 32, 120–132.

- Nowak, P. (2006a). The role of vowel transitions and frication noise in the perception of Polish sibilants. *Journal of Phonetics*, 34, 139–152.
- Nowak, P. (2006b). *Vowel reduction in Polish* (PhD. Thesis). University of California, Berkley.
- Ohala, J. (1971). The role of physiological and acoustic models in explaining the direction of sound change. In *Project on linguistic analysis reports* (Bd. 15, S. 25–40). Berkley: University of California at Berkley.
- Ohala, J. (1981). The listener as a source of sound change. In C. Masek, R. Hendrick & M. Miller (Hrsg.), *Papers from the parasession on language and behavior* (S. 178–203). Chicago, USA: Chicago Linguistic Society.
- Ohala, J. (1992). What's cognitive, what's not, in sound change. In G. Kellermann & M. Morrissey (Hrsg.), *Diachrony within synchrony: Language history and cognition* (S. 309–355). Frankfurt, Deutschland: Peter Lang.
- Ohala, J. (1993). Sound change as natur's speech perception experiment. *Speech Communication*, 13, 155–161.
- Ohala, J. (2012). The listener as a source of sound change. An update. In M. J. Solé & D. Recasens (Hrsg.), *The initiation of sound change. perception, production and social factors* (S. 21–36). Amsterdam/Philadelphia: John Benjamins.
- Padgett, J. & Żygis, M. (2003). The evolution of sibilants in Polish and Russian. In T. Hall & S. Hamann (Hrsg.), *ZAS Papers in Linguistics* (Bd. 32, S. 155–174). Zentrum für Allgemeine Sprachwissenschaft, Berlin.
- Percival, D. & Walden, A. (1993). *Spectral analysis for physical applications: Multitaper and conventional univariate techniques*. Cambridge, Vereinigtes Königreich: Cambridge University Press.
- Perry, T., Ohde, R. & Ahmead, D. (2001). The acoustic bases for gender identification from children's voices. *Journal of the Acoustical Society of America*, 109, 2988–2998.
- Peterson, G. & Barney, H. (1952). Control methods used in a study of the vowels. *Journal of the Acoustic Society of America*, 24, 175–184.
- Peterson, G. & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America*, 32, 693–703.
- Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (Hrsg.), *Frequency effects and the emergence of lexical structure* (S. 137–157). Amsterdam, Niederlande: John Benjamins.
- Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner

- (Hrsg.), *Laboratory Phonology VII* (S. 101–140). Berlin, Deutschland: Mouton de Gruyter.
- Poole, I. (1934). Genetic development of articulation of consonant sounds in speech. *Elementary English Review*, 11, 159–161.
- Port, R. & O’Dell, M. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics*, 13, 455–471.
- Prather, E., Hedrick, D. & Kern, D. (1975). Articulation development in children aged two to four years. *Journal of Speech and Hearing Disorders*, 53 (12).
- Proctor, M., Lu, L., Zhu, Y., Goldstein, L. & Narayanan, S. (2012). Articulation of Mandarin sibilants: A multi-plane realtime MRI study. In *Proceeding of the 14th Australasian International Conference on Speech Science and Technology*. Sydney, Australien.
- Recasens, D. (2015). The effect of stress and speech rate on vowel coarticulation in Catalan VCV sequences. *Journal of Speech, Language and Hearing Research*, 58, 1407–1424.
- Recasens, D. & Espinosa, A. (2009). An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in catalan. *Journal of the Acoustical Society of America*, 125, 2288–2298.
- Recasens, D., Pallarès, M. & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America*, 102, 544–561.
- Reidy, P. (2015). A comparison of spectral estimation methods for the analysis of sibilant fricatives. *JASA Express Letters*, 137 (4), 248–24.
- Repp, B. (1986). Some observations on the development of anticipatory coarticulation. *Journal of the Acoustical Society of America*, 79, 1616–1619.
- Reubold, U. (2012). *Über die Zusammenhänge zwischen Grundfrequenz und Vokalhöhe: Evidenzen aus longitudinalen Altersstimmenstudien, Perturbations- und Vokalerkennungsexperimenten* (PhD. Thesis). Ludwig-Maximilians-Universität, München, Deutschland.
- Reubold, U., Harrington, J. & Kleber, F. (2010). Vocal aging effects on f0 and the first formant: A longitudinal analysis in adult speakers. *Speech Communication*, 52, 638–651.
- Roberts, J. (1994). *Acquisition of variable rules: (-t,d) deletion and (ing) production in preschool children* (PhD. Thesis). University of Pennsylvania, Pennsylvania, USA.
- Rothstein, R. (2005). Polish. In K. Brown (Hrsg.), *Encyclopedia of Language and Lingui-*

- stics* (2. Aufl., Bd. 11, S. 672–675). Oxford, Vereinigtes Königreich: Elsevier Ltd.
- Rubach, J. (2005). Polish: Phonology. In K. Brown (Hrsg.), *Encyclopedia of Language and Linguistics* (2. Aufl., Bd. 11, S. 676–679). Oxford, Vereinigtes Königreich: Elsevier Ltd.
- Rubach, J. & Booij, G. (1990). Syllable structure assignment in Polish. *Phonology*, 7, 121–158.
- Ruch, H. (2013). Investigating a gradual metathesis: Phonetic and lexical factors on /s/-aspiration in Andalusian Spanish. In *University of pennsylvania working papers in linguistics* (Bd. 19, Kap. 19). Pennsylvania, USA: Penn Libraries University of Pennsylvania.
- Sawicka, I. (2001). *An outline of the phonetic typology of the Slavic languages*. Torun: Wydawnictwo Uniwersytetu Mikolaja Kopernika.
- Schiel, F. (2004). Maus goes iterative. In *Proceedings of the IV. International Conference on Language Resources and Evaluation* (S. 1015–1018). Lisabon, Portugal.
- Schleicher, A. (1971). Some observations made on children. Beiträge zur Vergleichenden Sprachforschung 2. In A. Bar-Adon & W. Leopold (Hrsg.), *Child language: A book of readings*. Prentice Hall.
- Shadle, C. (1991). The effect of geometry on source mechanisms of fricative consonants. *Journal of Phonetics*, 19, 409–424.
- Shadle, C. (2006). Acoustic phonetics. In K. Brown (Hrsg.), *Encyclopedia of Language and Linguistics. 2nd edition* (Bd. 9, S. 442–460). Oxford, Vereinigtes Königreich: Elsevier Ltd.
- Shadle, C. (2010). The aerodynamics of speech. In W. Hardcastle & J. Laver (Hrsg.), *The Handbook of Phonetic Sciences* (S. 39–80). Oxford, Vereinigtes Königreich: Blackwell.
- Shadle, C. & Mair, S. (1996). Quantifying spectral characteristics of fricatives. In *Proceedings of the International Conference on Spoken Language Processing* (S. 1517–1520). Philadelphia, USA.
- Shih, Y. & Kong, E. (2011). Perception of Mandarin fricatives by native speakers of Taiwan Mandarin and Taiwanese. In Z. Jing-Schmidt (Hrsg.), *Proceedings of the 23rd North American Conference on Chinese Linguistics* (S. 110–119).
- Simpson, A. (2000). Gender-specific differences in the articulatory and acoustic realization of interword vowel sequences in American English. In P. Hoole, M. Honda & C. Mooshammer (Hrsg.), *5th Seminar on Speech Production: Models and Data* (S.

- 209–212). Kloster Seeon, Deutschland.
- Soli, S. (1985). Second formants in fricatives: Acoustic consequences of fricative- vowel coarticulation. *Journal of the Acoustical Society of America*, 70, 976–984.
- Spencer, A. (1986). A non-linear analysis of vowel-zero alternations in Polish. *Journal of Linguistics*, 22, 249–280.
- Srivastava, A., Jermyn, I. & Joshi, S. (2007). Riemannian analysis of probability density functions with applications in vision. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*.
- Stevens, K. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. In P. Denes & E. D. Jr. (Hrsg.), *Human communication: A unified view* (S. 51–66). New York: McGraw Hill.
- Stevens, K. (1989). The quantal nature of speech. *Journal of Phonetics*, 17, 3–46.
- Stevens, K., Li, Z., Lee, C. & Keyser, S. (2004). A note on Mandarin fricatives and enhancement. In G. Fant, H. Fujisaki, J. Cao & Y. Xu (Hrsg.), *From traditional phonology to modern speech processing*. Beijing: Foreign Language Teaching and Research Press.
- Stevens, M. & Harrington, J. (2016). The phonetic origin of /s/-retraction and perceptual evidence from Australian English. *Journal of Phonetics*, 58, 118–134.
- Stone, M. & Lundberg, A. (1996). Three-dimensional tongue surface shapes of English consonants and vowels. *Journal of the Acoustical Society of America*, 96, 3728–3736.
- Stuart-Smith, J. (2007). Empirical evidence for the gendered production of /s/ in Glaswegian. In J. Cole & J. Hualde (Hrsg.), *Papers in Laboratory Phonology 9*. Mouton de Gruyter.
- Studdert-Kennedy, M. (2000). Imitation and the emergence of segments. *Phonetica*, 57, 275–283.
- Sully, J. (1896). *Studies of childhood*. New York: D. Appleton.
- Sumner, M. & Samuel, A. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language*, 60, 487–501.
- Sussex, R. & Cubberley, P. (2006). *The Slavic languages*. Cambridge, Vereinigtes Königreich: Cambridge University Press.
- Sweet, H. (1888). *A history of English sounds*. Oxford, Vereinigtes Königreich: Clarendon Press.
- Thomson, D. (1982). Spectrum estimation and harmonic analysis. In *Proceedings of the IEEE* (Bd. 70, S. 1055–1096).

- Toda, M., Maeda, S. & Honda, K. (2010). Formant-cavity affiliation in sibilant fricatives. In S. Fuchs, M. Toda & M. Žygis (Hrsg.), *Turbulent sounds. an interdisciplinary guide* (S. 343–374). Berlin, Deutschland: Walter de Gruyter.
- Torgersen, E. & Kerswill, P. (2004). Internal and external motivation in phonetic change: Dialect levelling outcomes for an English vowel shift. *Journal of Sociolinguistics*, 8, 24–53.
- Trudgill, P. (1972). Sex, covert prestige and linguistic change in the Urban British English of Norwich. *Language in Society*, 1, 179–195.
- Vihman, M. (1996). *Phonological development*. Oxford, Vereinigtes Königreich: Blackwell.
- Vihman, M., Ferguson, C. & Elbert, M. (1986). Phonological development from babbling to speech: Common tendencies and individual differences. *Psycholinguistics*, 7, 3–40.
- Vorperian, H. & Kent, R. (2007). Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language and Hearing Research*, 50, 1510–1545.
- Wagner, A., Ernestus, M. & Cutler, A. (2006). Formant transitions in fricative identification: The role of native fricative inventory. *Journal of the Acoustical Society of America*, 120 (4), 2267–2277.
- Watson, C. & Harrington, J. (1999). Acoustic evidence for dynamic formant trajectories in Australian English vowels. *Journal of the Acoustical Society of America*, 106, 458–468.
- Weinreich, U., Labov, W. & Herzog, M. (1968). Empirical foundations for a theory of language change. In W. Lehmann & Y. Malkiel (Hrsg.), *Directions for historical linguistics* (S. 97–195). Austin: University of Texas Press.
- Weirich, M. & Simpson, A. (2015). Gender-specific differences in sibilant contrast realizations in English and German. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, Vereinigtes Königreich.
- Wells, J. (1982). *Accents of English, vol 1: An introduction*. Cambridge, Vereinigtes Königreich: Cambridge University Press.
- Whalen, D. H. (1991). Perception of the English /s/-/ʃ/ distinction relies on fricative noises and transitions, not on brief spectral slices. *Journal of the Acoustical Society of America*, 90 (4), 1776–1785.
- Winkelmann, R. & Raess, G. (2014). Introducing a web applicator for labeling, visualizing speech and correcting derived speech signals. In N. Calzolari et al. (Hrsg.), *Proceedings of the 9th International Conference on Language Resources and Evaluation* (S.

- 26–31). European Language Resources Association (ELRA).
- Yu, A. (2011). Contrast reduction. In J. Goldsmith, J. Riggle & A. C. L. Yu (Hrsg.), *The Handbook of Phonological Theory* (2. Aufl., S. 291–318). Oxford, Vereinigtes Königreich: Blackwell Publishing.
- Yu, A., Abrego-Collier, C., Phillips, J., Pillion, B. & Chen, D. (2015). Investigating variation in English vowel-to-vowel coarticulation in a longitudinal phonetic corpus. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, Vereinigtes Königreich.
- Zharkova, N., Hewlett, N. & Hardcastle, W. J. (2011). Coarticulation as an indicator of speech motor control development in children: An ultrasound study. *Motor Control*, 15, 118–140.
- Zhu, H. & Dodd, B. (2000). The phonological acquisition of Putongha (Modern Standard Chinese). *Journal of Child Language*, 27, 3–42.
- Zydorowicz, P. (2010). Consonant clusters across morpheme boundaries: Polish morpho-notactic inventory and its acquisition. *Poznań Studies in Contemporary Linguistics*, 46 (4), 565–588.
- Żygis, M. & Hamann, S. (2003). Perceptual and acoustic cues of Polish coronal fricatives. In *Proceedings of the 15th International Conference of Phonetic Sciences* (S. 395–398).
- Żygis, M. & Padgett, J. (2010). A perceptual study of Polish fricatives, and its implications for historical sound change. *Journal of Phonetics*, 38, 207–226.
- Żygis, M., Pape, D. & Czaplicki, B. (2012). Dynamics in sibilant systems: Standard Polish and its dialects. In *Phonetik und Phonologie* (Bd. 8). Jena, Deutschland: Friedrich-Schiller-Universität Jena.

Danksagung

Auf dem Weg zur Fertigstellung dieser Arbeit wurde ich von vielen Menschen begleitet, die mich auf die verschiedensten Arten und Weisen unterstützten und all denen ein großes Dankeschön gebührt.

Allen voran danke ich meinem Betreuer, Jonathan Harrington, für seine Unterstützung. Seine Kommentare, Anregungen und Ideen waren immer eine große Hilfe und von unschätzbarem Wert. Ein Doktorvater, der für Fragen, Diskussionen, wie auch für neue Ideen jederzeit offen war.

Felicitas Kleber gebührt ein ganz ganz großes Dankeschön für ihre Unterstützung, ihren Zuspruch und die vielen vielen klärenden Diskussionen, die immer ein Wegweiser von größtem Wert für mich waren.

Ich danke Ulrich Reubold für seine immer offene Tür und seine unsägliche Geduld bei vielen Fragen meinerseits und Erklärungen seinerseits. Von allergrößtem Wert waren für mich seine hilfreichen Kommentare und Anmerkungen beim Korrekturlesen dieser Arbeit. Einer zweiten Korrekturleserin, habe ich nicht nur für ihre schnellen Korrekturen und Kommentare zu dieser Arbeit zu danken, sondern auch für die vielen Ablenkungen außerhalb der Bürotür. Danke Sandra Peters für die tollen Gespräche und die gemeinsamen sportlichen und kulinarischen Ausflüge!

Phil Hoole danke ich für die Hilfe bei den artikulatorischen Experimenten. Ein Dankeschön geht auch an Christoph Draxler für die schnelle Umsetzung des Online-Perzeptionsexperiments. Für die Unterstützung bei der Suche nach Versuchspersonen, für das Labeln und Segmentieren der Aufnahmen und die Hilfe mit der polnischen Sprache, danke ich Joanna Kaczmarek, Aleksander Wiatr, Nora Eichfeld, Rebecca Stegmeier und Susanne Waltl.

Meinen Bürokollegen Manfred Pastätter und Tom Lentz danke ich für ihr offenes Ohr und die vielen hilfreichen Diskussionen.

Für die immer sehr unterhaltsamen Kaffee-„Pausen“ früh am Morgen und die zahlreichen Latex-Tipps und Tricks, danke ich Thomas Kisler.

Bei Andrea habe ich mich für eine sehr wertvolle und kostbare Freundschaft zu bedanken. Anderl, vielen, vielen Dank für die Nervennahrung und die vielen Aufmunterungen.

Meiner Familie danke ich von ganzem Herzen dafür, dass sie einfach immer für mich da sind! Ohne eure Unterstützung in allen Lebenslagen und das in mich gesetzte Vertrauen hätte ich es nicht bis hierher geschafft. Tausend Dank für alles!

Martin danke ich für seine unsägliche Geduld mit mir, seine Zuversicht und die Ablenkungen, die mir immer wieder Kraft für die nächsten Etappen gaben. Du bist von unschätzbarem Wert!