

INAUGURAL-DISSERTATION
zur Erlangung der Doktorwürde der Philosophie (Dr. phil.)

des Fachbereichs Germanistik und Kunstwissenschaften
der Philipps Universität Marburg

**Dynamik und Stabilität der Repräsentationen im Sprachkontakt
Italienisch-Deutsch:
Eine psycholinguistische und phonetische Analyse**

Überarbeitete Fassung
vorgelegt von
Alexia Despina Leonidou
aus
Forlì (Italien)

Marburg/Lahn im September 2026

Vom Fachbereich Germanistik und Kunstwissenschaften der
Philipps-Universität Marburg als Dissertation angenommen am 21.10.2025

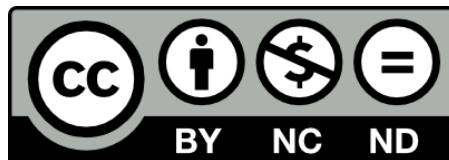
Tag der mündlichen Prüfung/Disputation am 06.11.2025

Erstgutachter: Prof. Dr. Christina Kauschke

Zweitgutachter: Prof. Dr. Ulrike Domahs

Drittgutachter: Prof. Dr. Birgit Alber (Freie Universität Bozen/Bolzano)

Originaldokument gespeichert auf dem
Open-Access-Repositorium der Philipps-Universität Marburg
<https://open.uni-marburg.de/>



Dieses Werk steht unter einer
Creative Commons
Namensnennung – Nicht-kommerziell – Keine Bearbeitungen
4.0 International
Lizenz

Die vollständige Lizenz finden Sie unter
<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode.de>

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

Einleitung	1
1 Terminologische Klärungen	3
1.1 Einleitung	3
1.2 Die Bezeichnung der L1 (Erstsprache)	4
1.3 Die Bezeichnung einer nach der L1 erworbenen Sprache (LX)	6
1.4 Sprachbiographische Begriffe im Englischen	8
1.4.1 Sprachbiographische Begriffe zur Bezeichnung der L1	8
1.4.2 Sprachbiographische Begriffe zur Bezeichnung der LX	10
1.4.2.1 <i>Bilingualism vs Second Language Learning</i>	10
1.4.2.2 <i>Second Language Learning vs Foreign Language Learning</i>	11
1.4.2.3 <i>Multilingualism vs Plurilingualism</i>	11
1.5 Sprachbiographische Begriffe im Deutschen	11
1.5.1 Sprachbiographische Begriffe zur Bezeichnung der L1	11
1.5.2 Sprachbiographische Begriffe zur Bezeichnung der LX	13
1.5.2.1 <i>Bilingualismus vs Bilingualität</i>	13
1.5.2.2 <i>Bilingualität vs Zweisprachigkeit</i>	14
1.5.2.3 <i>Zweisprachigkeit vs Fremdsprachigkeit</i>	14
1.5.2.4 <i>Mehrsprachigkeit vs Plurilingualismus/Multilingualismus</i>	15
1.6 Sprachbiographische Begriffe im Italienischen	16
1.6.1 Sprachbiographische Begriffe zur Bezeichnung der L1	16
1.6.2 Sprachbiographische Begriffe zur Bezeichnung der LX	18
1.6.2.1 <i>Bilinguismo</i>	19
1.6.2.2 <i>Lingua seconda vs lingua straniera vs lingua etnica</i>	20

1.7	Synoptische Betrachtung: Terminologische Konvergenzen und Divergenzen	20
2	Die Phonologische Dimension des LX-Erwerbs: Sprachkontakt und Interferenzphänomene	22
2.1	Die Organisation von phonologischen Systemen im multilingualen Gehirn	22
2.2	Wahrnehmung und Produktion von LX-Lautkategorien: Mechanismen und Modelle zur Klärung der <i>Cross-Linguistic Influence</i> (CLI)	24
2.2.1	<i>Cross-Linguistic Influence</i> (CLI) im LX-Erwerb: Natur und Mechanismen	24
2.2.2	Verhältnis zwischen Perzeption und Produktion im LX-Erwerb	26
2.2.3	<i>Cross-Linguistic Influence</i> (CLI) im LX-Erwerb: Relevante Einflussfaktoren	29
2.2.4	Theoretische Modelle des LX-Erwerbs	32
2.2.4.1	<i>Spezifische Modelle aus den Second Language Acquisition Studies</i>	33
2.2.4.2	<i>Modelle aus den Third or Additional Language Acquisition Studies</i>	35
2.3	Wahrnehmung und Produktion von Vokalen im LX-Erwerb	36
2.3.1	Die <i>Consonant-Vowel-Hypothesis</i>	36
2.3.2	Die Wahrnehmung und Produktion eines LX-Vokalsystems: Forschungsstand und -richtungen	37
3	Untersuchte Vokalsysteme vis-à-vis und Forschungsstand zu Wahrnehmung und Produktion im DaF-Erwerb	48
3.1	Die Vokalsysteme des Italienischen und des Deutschen im Vergleich	48
3.1.1	Der Status von /ə/ im Deutschen	55
3.1.1.1	<i>Vokalreduktion und Elision im Gebrauchsstandard</i>	56
3.1.2	Das Phonem /y:/ im Deutschen	57
3.1.3	Die Vokale [ə] und [y] in regionalen Varietäten des Italienischen	58
3.2	Artikulatorische Zielwerte und Transferhypothesen	60
3.3	Standarddeutsch für Deutschlerner:innen	62
3.4	Forschungsstand zur Wahrnehmung von deutschen Vokalen durch LX-Lerner:innen	63

3.4.1	Ausgangssprache Englisch	64
3.4.2	Ausgangssprache Polnisch	67
3.4.3	Ausgangssprache Türkisch	68
3.4.4	Ausgangssprache Italienisch	69
3.5	Forschungsstand zur Produktion von deutschen Vokalen durch LX-Lerner:innen	71
3.5.1	Ausgangssprache Englisch	71
3.5.2	Ausgangssprache Polnisch	73
3.5.3	Ausgangssprache Türkisch	75
3.5.4	Romanische Ausgangssprachen: Französisch und Italienisch	75
3.5.4.1	<i>Die Konstellation Französisch-Deutsch</i>	75
3.5.4.2	<i>Die Konstellation Italienisch-Deutsch</i>	76
3.6	Abschließende Anmerkungen und übergeordnete Forschungsfrage	78
3.7	Forschungsdesiderat und Zielsetzung der Studie	79
4	Eine psycholinguistische Studie zur Wahrnehmung deutscher Vokale	81
4.1	Forschungsfragen und Hypothesen	81
4.2	Methodisches Vorgehen	85
4.2.1	Versuchspersonen: Beschreibung der drei Stichproben	86
4.2.1.1	<i>Experimentalgruppe: Hörer:innen mit L1 Italienisch und Deutschkenntnis</i>	86
4.2.1.2	<i>Kontrollgruppe 1: Hörer:innen mit L1 Italienisch ohne Deutschkenntnis</i>	88
4.2.1.3	<i>Kontrollgruppe 2: Hörer:innen mit L1 Deutsch</i>	89
4.2.2	Aufbau des Experiments und Stimuli	89
4.2.3	Prozedere	93
4.2.3.1	<i>Testung der Sprachkenntnisse vor der Experimentphase</i>	93
4.2.3.2	<i>Die Experimentphase</i>	94
4.2.3.3	<i>Die Erfassung der Daten zu Sprachkompetenz, -dominanz und Sprachrepertoire</i>	95
4.2.4	Datenanalyse	97
4.3	Ergebnisse	98

4.3.1	Akkuratheit	98
4.3.1.1	<i>Ergebnisse zur perzeptiven Assimilation von /ə/ zu /e:/ oder /ɛ/</i>	102
4.3.1.2	<i>Ergebnisse zur perzeptiven Assimilation von /y:/ zu /u:/ oder /i:/</i>	107
4.3.1.3	<i>Übergreifende Analyse: Akkuratheit und Faktoren der Sprachkompetenz, Sprachdominanz und des Sprachrepertoires</i>	111
4.3.1.4	<i>Akkuratheit und Task-bedingte Variablen</i>	112
4.3.2	Reaktionszeiten	114
4.3.3	Eine explorative Studie zum Einfluss der dialektalen Hintergrundkenntnisse auf die Wahrnehmung des Deutschen: Akkuratheit und Reaktionszeiten	115
4.3.3.1	<i>Wahrnehmung von /ə/</i>	115
4.3.3.2	<i>Wahrnehmung von /y:/</i>	115
4.4	Diskussion	116
5	Eine phonetische Studie zur Produktion deutscher Vokale	122
5.1	Forschungsfragen und Hypothesen	122
5.2	Methodisches Vorgehen	124
5.2.1	Versuchspersonen: Beschreibung der drei Stichproben	126
5.2.1.1	<i>Die Experimentalgruppe: Sprecher:innen mit L1 Italienisch und Deutschkenntnis</i>	126
5.2.1.2	<i>Kontrollgruppe 1: Sprecher:innen mit L1 Italienisch</i>	127
5.2.1.3	<i>Kontrollgruppe 2: Sprecher:innen mit L1 Deutsch</i>	128
5.2.2	Aufbau des Experiments und Stimuli	128
5.2.2.1	<i>Die Aufgabe "Wegbeschreibung"</i>	128
5.2.2.2	<i>Die Aufgabe "Nacherzählung"</i>	129
5.2.2.3	<i>Die Aufgabe "Vorlesen"</i>	130
5.2.3	Prozedere	131
5.2.4	Datenanalyse	132
5.3	Ergebnisse	137
5.3.1	Allgemeine kontrastive Übersicht über die Vokalsysteme	137

5.3.2	Ergebnisse zur Realisierung von /ə/	139
5.3.2.1	<i>Realisierung von /ə/ in Kontrast zu /e:/ und /ɛ/</i>	140
5.3.2.2	<i>Realisierung von /ə/ in verschiedenen Positionen</i>	143
5.3.2.3	<i>CLI-Phänomene mit dem L1-phonologischen System (Italienisch) bei der Aussprache von /ə/</i>	146
5.3.2.4	<i>Der Einfluss der Sprachkompetenz und -dominanz sowie des Sprachrepertoires auf die Realisierung von /ə/</i>	147
5.3.2.5	<i>Der taskbedingte Einfluss auf Realisierung von /ə/</i>	148
5.3.3	Ergebnisse zur Realisierung von /y:/	150
5.3.3.1	<i>Realisierung von /y:/ in Kontrast zu /ʏ/, /i:/, /ɪ/, /u:/ und /ʊ/</i>	150
5.3.3.2	<i>Realisierung von /y:/ in Abhängigkeit vom linken und rechten phonologischen Kontext</i>	155
5.3.3.3	<i>CLI-Phänomene mit dem L1-phonologischen System (Italienisch) bei der Aussprache von /y:/</i>	160
5.3.3.4	<i>Der Einfluss der Sprachkompetenz und -dominanz sowie des Sprachrepertoires auf die Realisierung von /y:/</i>	162
5.3.3.5	<i>Der taskbedingte Einfluss auf Realisierung von /y:/</i>	163
5.3.4	Synopse der wesentlichen Ergebnisse	165
5.4	Diskussion	169
6	Übergreifende Diskussion	178
6.1	Einleitung zu den durchgeführten Studien	178
6.2	Übergreifende Diskussion: Gemeinsamkeiten und Unterschiede bei der Wahrnehmung und Produktion von /ə/ und /y:/	178
6.3	Einordnung der Ergebnisse innerhalb des theoretischen Frameworks	181
6.3.1	Einordnung in Modelle der L2-Wahrnehmung und -Produktion	181
6.3.2	Einordnung in L3-Modelle und Ausblick auf eine integrierte LX-Phonologieforschung	182
7	Fazit, Limitationen und Ausblick	184
7.1	Fazit	184
7.2	Limitationen und Ausblick	185
	Danksagung	188

Literatur	190
Software	204
Plagiarismuserklärung	205
Appendix	206

Abbildungsverzeichnis

Abbildung 1	Schematische Darstellung des PME-Filterungsprozesses	34
Abbildung 2	Vokaltrapeze des (betonten) standarditalienischen Vokalsystems in Anlehnung an Bertinetto (2010) und des standarddeutschen Vokalsystems (orale Monophthonge) adaptiert nach Kleiner & Knöbl (2015)	51
Abbildung 3	F2/F1-Grafik der deutschen Vokale, realisiert durch Lernende mit L1 Italienisch, in Anlehnung an Missaglia & Sendlmeier (1999, S. 92, Abb. 5)	61
Abbildung 4	Darstellung der Hypothesen anhand der kontrastierten Systeme des Standarditalienischen (vgl. z.B. Maturi, 2014) und des Standarddeutschen (vgl. z.B. Wiese 2000).	85
Abbildung 5	Beispiele für das Layout des Experiments	91
Abbildung 6	Darstellung der Assimilationsmuster nach den Hypothesen und der Kontrollkontraste	92
Abbildung 7	Durchschnittliche Akkuratheitswerte der Probandengruppen mit Standardfehlern (SE) und Trendverlauf	98
Abbildung 8	Durchschnittliche Akkuratheit nach Probandengruppen und Versuchsblock	100
Abbildung 9	Durchschnittliche Akkuratheit im Versuchsblock /ə/ für Kontraste und gleiche Abfolgen pro Probandengruppe	102
Abbildung 10	Durchschnittliche Akkuratheit in der Perzeption der Kontraste /ə-ɛ/, /ə-e:/ und /ə-i:/ nach Probandengruppe	104
Abbildung 11	Durchschnittliche Akkuratheit in Abhängigkeit von der Stimulusrichtung im Versuchsblock /ə/	106
Abbildung 12	Durchschnittliche Akkuratheit im Versuchsblock /y:/ für Kontraste und gleiche Abfolgen pro Probandengruppe	108
Abbildung 13	Durchschnittliche Akkuratheit in der Perzeption der Kontraste /y:-u:/, /y:-i:/ und /y:-e:/ nach Probandengruppe	109
Abbildung 14	Verteilung der Reaktionszeiten in allen Gruppen (Lernende, KO_ITA, KO_DEU) für beide Blöcke.	114
Abbildung 15	F2/F1-Grafik des Vokalraums der Lerner durch Angabe der Medianwerte und F2/F1 Grafik des Vokalraums der Lerner durch Angabe der Einzelwerte ausgewählter Vokale mit Hervorhebung der Datenstreuung für /ə/ und /y:/	136

Abbildung 16	F2/F1-Grafik des Vokalraums der Lernerinnen durch Angabe der Medianwerte und F2/F1 Grafik des Vokalraums der Lernerinnen durch Angabe der Einzelwerte ausgewählter Vokale mit Hervorhebung der Datenstreuung für /ə/ und /y:/	138
Abbildung 17	Vergleichende F2/F1-Grafik der Realisierung ausgewählter deutscher Vokale durch Angabe der Medianwerte für die Stichprobe der Frauen aus der Gruppe der Lerner:innen und Sprecher:innen mit L1 Deutsch	139
Abbildung 18	Vergleichende F2/F1-Grafik der Realisierung ausgewählter deutscher Vokale durch Angabe der Medianwerte für für die Stichprobe der Frauen aus der Gruppe der Lerner:innen und dem CLIPS-Korpus	139
Abbildung 19	Dispersionsellipsen (95% Konfidenzniveau) und Medianwerte für /ə/, /e:/ und /ɛ/ für die Lerner:innen (Frauen Männer)	141
Abbildung 20	Dispersionsellipsen (95%-Konfidenzniveau) und Medianwerte für /ə/, /e:/ und /ɛ/. für die Lernerinnen und die Sprecherinnen mit L1 Deutsch	142
Abbildung 21	Datenstreuung und Medianwerte für F1 und F2 bei Frauen und Männern über alle Positionen hinweg	143
Abbildung 22	Vergleich der Datenstreuungen und Medianwerte (F1, F2) für alle Positionen innerhalb der weiblichen Stichprobe für alle Positionen zwischen den Lernerinnen und den Sprecherinnen mit L1 Deutsch	145
Abbildung 23	Kontrastierende Datenstreuungen (F1, F2) und Medianwerte für die italienischen und die deutschen Vokale.	147
Abbildung 24	Datenstreuung für die F1- und F2-Werte für die Lerner:innen (F M) über alle Aufgaben hinweg	149
Abbildung 25	Dispersionsellipsen (95% Konfidenzniveau) mit Medianwerten zu F3/F2 für /y:/, /ʏ/, /i:/, /ɪ/ und /y:/, /ʏ/, /u:/, /ʊ/ für die Gruppe der Lernerinnen	151
Abbildung 26	Dispersionsellipsen (95% Konfidenzniveau) mit Medianwerten zu F3/F2 für /y:/, /ʏ/, /i:/, /ɪ/ und /y:/, /ʏ/, /u:/, /ʊ/ für die Gruppe der Lerner	152
Abbildung 27	Dispersionsellipsen (95% Konfidenzniveau) mit Medianwerten zu F3/F2 für /y:/, /ʏ/, /i:/, /ɪ/ für die Gruppe der Lernerinnen und Sprecherinnen mit L1 Deutsch	154

Abbildung 28	Dispersionsellipsen (95% Konfidenzniveau) mit Medianwerten zu F3/F2 für /y:/, /ʏ/, /u:/, /ʊ/. für die Gruppe der Lernerinnen und Sprecherinnen mit L1 Deutsch	154
Abbildung 29	Datenstreuung und Medianwerte für die F2- und F3-Werte über alle linken phonologischen Kontexte innerhalb der Stichprobe der Lerner:innen (F M)	156
Abbildung 30	Vergleich der Datenstreuungen und Medianwerte bei der Realisierungen von /y:/ hinsichtlich der F2- und F3-Werte für die geteilten linken Kontexte innerhalb der Stichprobe der Lernerinnen und Sprecherinnen mit L1 Deutsch	158
Abbildung 31	Datenstreuung und Medianwerte für die F2- und F3-Werte für die italienischen und deutschen Vokale	161
Abbildung 32	Datenstreuung und Medianwerte für die F2- und F3-Werte für die Lerner:innen (F M) über alle Aufgaben	164

Tabellenverzeichnis

Tabelle 1	Übersichtliche schematische Darstellung der Grundpfeiler des NRV-Frameworks mit Fokus auf die Wahrnehmung bei Erwachsenen.	40
Tabelle 2	Distinktive Merkmale des Italienischen, adaptiert nach Heinz & Schmid (2021, S.47, Tabelle 11)	48
Tabelle 3	Distinktive Merkmale der deutschen Vokale, adaptiert nach Wiese (2000, S. 20, Abbildung 6)	50
Tabelle 4	Mittelwerte zu F1 und F2 der deutschen Vokale nach Sendlmeier & Seebode (2006) und Zimmerer & Reetz (2011)	53
Tabelle 5	Medianwerte zu F1, F2 und F3 der deutschen Vokale nach Pätzold & Simpson (1997)	54
Tabelle 6	Medianwerte zu F1 und F2 der deutschen Vokale nach Pätzold & Simpson (1997)	54
Tabelle 7	Beispiele für alle kodierten Positionen im Datensatz	136
Tabelle 8	Anzahl an Beobachtungen, Medianwerte (F2) und IQR für beide Gruppen	149
Tabelle 9	Überblick der zentralen statistischen Ergebnisse der Produktionsstudie	166

Einleitung

Im Vergleich zur ursprünglichen, breiter angelegten Planung des Promotionsprojekts fand eine thematische und methodische Modifikation statt. Das anfängliche Forschungsdesign war darauf ausgerichtet, das Zusammenspiel zwischen dem italienischen und deutschen Vokalsystem im Hinblick auf die Perzeption und Produktion der Phoneme /ə/ und /y:/ bei bilingualen Kindern mit Sprachentwicklungsstörungen in Südtirol zu untersuchen. Die Ausgangshypothese besagte, dass die Erwerbsmerkmale dieser Zielgruppe bezüglich der genannten Vokale Rückschlüsse auf Mechanismen des Zweitspracherwerbs erlauben, die auch bei erwachsenen Lerner:innen mit der L1 Italienisch zu beobachten sind.

Ausschlaggebend für diese Neuausrichtung war die erste experimentelle Datenerhebung an der Universität Verona im Rahmen der zugrundeliegenden binationalen Kooperation. Die hohe Beteiligung erwachsener italienischer Deutschlerner:innen führte zu einer quantitativ und qualitativ substanziellen Datenbasis, die den ursprünglichen Status einer bloßen Kontrollgruppe überschritt. Die Auswertung dieser Daten verdeutlichte ein Forschungsdesiderat bezüglich des Erwerbs deutscher Vokale durch italienische Erwachsene. Bislang mangelt es an systematischen empirischen Untersuchungen, die den Zusammenhang zwischen Perzeption und Produktion unter Einbeziehung linguistischer Hintergrundvariablen akustisch-phonetisch beleuchten. Aus diesem Grund konzentriert sich die vorliegende Dissertation auf die detaillierte psycholinguistische und phonetische Analyse erwachsener Lerner:innen, um eine fundierte empirische Basis für die italienisch-deutsche *Interlanguage*-Forschung zu schaffen.

Über den rein wissenschaftlichen Beitrag zum aktuellen Forschungsstand hinaus wurde diese Arbeit in der Überzeugung verfasst, dass ihre Ergebnisse auch außerhalb des akademischen Rahmens und des rein theoretischen Diskurses von Relevanz sind. Ich bin zutiefst davon überzeugt, dass die Integration experimentalphonetischer Erkenntnisse in die Sprachdidaktik ein enormes Potenzial birgt – insbesondere für Institutionen, die junge Erwachsene auf ihrem oft mühsamen Weg des Fremdspracherwerbs begleiten. Da der Spracherwerb in den meisten Fällen den entscheidenden ersten Schritt zur beruflichen und sozialen Integration darstellt, hat ein mangelndes Beherrschen der lautlichen Ebene – was sowohl die effiziente Perzeption als auch die Verständlichkeit der eigenen Produktion umfasst – Konsequenzen von erheblichem Gewicht.

Obwohl Ansätze wie das *High Variability Phonetic Training* (vgl. z.B. Uchihara, 2025; Mora, 2021) die Validität einer engeren Verzahnung von Phonetik und Didaktik eindrucksvoll belegen, findet eine solche systematische Anwendung in der Praxis der (deutschen und italienischen) Sprachschulen bisher kaum statt. Dabei geht es in der vorliegenden Reflexion nicht um die bloße mechanische Übernahme spezifischer Trainingsmethoden, sondern vielmehr um die

Vermittlung metalinguistischer Bewusstheit. Das Verständnis der (sozio-)linguistischen Faktoren, die für Interferenzen verantwortlich sind, sowie die Erkenntnis, wie diese in unterschiedlichen Kontexten – von der Wortliste bis hin zur Spontansprache – variieren, verleiht den Lerner:innen ein wertvolles Instrument der Selbstwahrnehmung. Auch wenn sich die vorliegende Dissertation primär im Bereich der experimentellen Phonetik verortet und keine didaktische Abhandlung im engeren Sinne darstellt, versteht sie sich als Impulsgeberin für diesen Transfer. Ich hoffe daher, dass diese Einleitung dazu beiträgt, die Arbeit auch unter diesem Blickwinkel zu betrachten.

Die Arbeit ist wie folgt aufgebaut: Die ersten drei Kapitel stellen den theoretischen Rahmen und den Stand der Forschung dar. **Kapitel 1** behandelt grundlegende terminologische und konzeptuelle Fragen. **Kapitel 2** widmet sich dem Erwerb des phonologischen bzw. vokalischen Systems einer Nicht-L1, mit einem Fokus auf universelle und spezifische, sprachliche und nicht-sprachliche Mechanismen. **Kapitel 3** bietet eine kontrastive Analyse des deutschen und italienischen Vokalsystems und ordnet das Lernen des Deutschen als Fremdsprache in aktuelle Forschungsdiskussionen ein.

Die darauffolgenden drei Kapitel sind der Darstellung der empirischen Studien gewidmet: **Kapitel 4** beschreibt die psycholinguistische Studie zur Wahrnehmung, während **Kapitel 5** die soziophonetische Studie zur Produktion darstellt. **Kapitel 6** integriert die Ergebnisse beider Studien in einer übergreifenden Perspektive und diskutiert die theoretischen Implikationen. **Kapitel 7** fasst schließlich die zentralen Befunde zusammen, benennt Limitationen der Untersuchung und skizziert mögliche zukünftige Forschungsrichtungen an der Schnittstelle zwischen experimenteller Phonetik und angewandter Linguistik.

1. Terminologische Klärungen

1.1 Einleitung

Zum besseren Verständnis der Studien, die in dieser Dissertationsschrift präsentiert werden, sollen vorab einige wesentliche Grundbegriffe definiert werden, die sich auf den Erwerb einer ersten sowie weiterer Sprachen beziehen. Diese terminologische Schärfung erscheint aus zwei Gründen unabdingbar. Zum einen ist der fachwissenschaftliche Diskurs durch eine teils divergierende Verwendung identischer Termini geprägt, was zu begrifflichen Unschärfen führen kann. Eine Auseinandersetzung mit dieser Heterogenität dient daher der Sensibilisierung der Leserschaft sowie der Reflexion über die Notwendigkeit einer einheitlicheren Terminologie.

Zum anderen – und dies betrifft die spezifische Natur des vorliegenden Forschungsprojekts – manifestierte sich diese Ambiguität auch im empirischen Teil der Arbeit. Während der Datenerhebung zeigte sich, dass die definitorische Unschärfe nicht auf die Fachliteratur beschränkt blieb, sondern auch bei der Rezeption durch ein Laienpublikum (hier: jugendliche Lerner:innen) zutage trat. Die Konfrontation mit polysemen Begriffen wie *Zweitsprache* führte bei den Probanden zu Irritationen, die eine explizite Klärung im Versuchskontext erforderten. Dies kann als Resonanzeffekt mangelnder wissenschaftlicher Einheitlichkeit oder spezifischer Formen des Wissenstransfers auf nicht-akademische Zielgruppen gedeutet werden.

Vor diesem Hintergrund wird die englische Terminologie als Ausgangspunkt der internationalen wissenschaftlichen Kommunikation berücksichtigt, aber auch äquivalente Begriffe im Deutschen und Italienischen, da diese im Rahmen der durchgeführten Experimente als Instruktionssprachen fungierten.

Der Aufbau dieses Kapitels folgt einer deduktiven Logik, um eine eindeutige terminologische Basis zu gewährleisten: Zunächst werden die für diese Arbeit konstitutiven Arbeitsdefinitionen für die L1 (Erstsprache) sowie für weitere Sprachen (insbesondere das Konzept "LX") festgelegt. Auf dieser Basis erfolgt im Anschluss eine vertiefte Auseinandersetzung mit spezifischen sprachbiografischen Termini. Der diachrone Zugriff erfolgt dabei selektiv: Eine historische Herleitung wird spezifisch für jene Konzepte vorgenommen, die – wie etwa *native speaker* – eine weitreichende begriffsgeschichtliche Tiefe aufweisen. Dies dient dazu, die oben genannten Irritationen aufzulösen, die oft aus dem Nachhall historisch gewachsener Konnotationen resultieren, während rein fachsprachliche Konstrukte jüngerer Datums primär synchron behandelt werden.

Hinsichtlich der Makrostruktur dieses Kapitels wurde bewusst eine einzelsprachliche Gliederung (Englisch, Deutsch, Italienisch) anstelle einer thematischen Bündelung gewählt. Diese Trennung dient dazu, die systemische Geschlossenheit der jeweiligen nationalen Forschungstraditionen zu

wahren und die Begriffe in ihrem spezifischen kulturellen Kontext zu belassen. Etwaige inhaltliche Redundanzen bei universellen Basiskonzepten werden zugunsten dieser systematischen Klarheit und des besseren Nachschlagecharakters in Kauf genommen, wobei explizite Querverweise diese auf ein Minimum reduzieren. Eine abschließende synoptische Betrachtung am Ende des Kapitels (Abschnitt 1.7) führt die Fäden zusammen und arbeitet die zentralen Divergenzen und Konvergenzen im direkten Vergleich heraus.

1.2 Die Bezeichnung der L1 (Erstsprache)

Die Festlegung der "Muttersprache" eines Individuums stellt eine grundlegende Voraussetzung für jede Untersuchung zum Sprachkontakt dar. Das Hauptkriterium ist ein chronologisches, da die erste Sprache, die ein Neugeborenes wahrnimmt,¹ untersucht wird. Zusätzlich sind die Dominanz und die Kompetenz in der jeweiligen Sprache von entscheidender Bedeutung für die Definition der L1 der Sprecher:innen.

In Bezug auf den Terminus *Kompetenz* sei an dieser Stelle auf Treffers-Daller (Treffers-Daller, 2019, S. 376–377) hingewiesen. Unter sprachlicher Kompetenz wird in diesem Kontext die Gesamtheit der Kenntnisse der Sprecher:innen in den Sprachen ihres Repertoires verstanden. In Übereinstimmung mit den Ausführungen von Treffers-Daller (2019), die sich wiederum auf Bachman & Palmer (2010) bezieht, resultiert Kompetenz nicht nur aus der Kenntnis der Grammatik der Sprache. Dies umfasst zudem strategische Planungsfähigkeiten im Gebrauch der Sprache selbst (vgl. Bachman & Palmer, 2010). Die Dominanz, d.h. die Vorherrschaft einer Sprache zu anderen Sprachen des Repertoires, betrifft sowohl die reine Grammatikkompetenz als auch die metasprachliche und strategische (vgl. Treffers-Daller, 2019). Von besonderer Relevanz ist dabei, dass Kompetenz und Dominanz kontextuelle Variablen sind, d. h. zeitlichen und räumlichen Schwankungen unterliegen. Methoden der Einordnung von Sprachkompetenz in Etikettierungskategorien (wie z. B. der Verweis auf den Gemeinsamen Europäischen Referenzrahmen für Sprachen (GER), vgl. (Europarat, 2010) können nur als deskriptives Hilfsmittel dienen, entbinden aber nicht von einer vertieften Untersuchung.

Die genaue Bestimmung der L1 erfordert daher eine sorgfältige und multidimensionale Analyse, um die sprachliche Identität und die damit verbundenen sozialen und kognitiven Prozesse vollständig zu erfassen.

In Dewaele et al. (2021) werden einige der gängigsten Missverständnisse über das Konzept der L1 aufgeführt, die im alltäglichen Diskurs vorkommen. Die Ursachen für die Verbreitung dieser Misskonzepte sind nicht exakt bestimmbar. Dennoch lässt sich aus ihnen ableiten, dass die Kommunikation außerhalb der Wissenschaft eine immer größere Rolle spielt.

¹ In Übereinstimmung mit den meisten Studien wird aus Konvention die Geburt als Ausgangs- und Referenzpunkt für die Anwendung des chronologischen Kriteriums festgelegt. Es ist jedoch bekannt, dass eine Form von akustischem "Imprinting" bereits während der Schwangerschaft stattfindet (vgl. Kisilevsky et al., 2003).

Eine der gängigsten Assoziierungen ist die, dass die L1 mit der Geburt in einer bestimmten sozialen Gruppe oder sogar Staatsbürgerschaft assoziiert wird und dass die Sprache durch Genetik vererbt wird. Des Weiteren führt der Terminus *Muttersprachler/-in* zu der Fehlinterpretation, dass es sich um ein binäres Merkmal handelt, nach dem die "Muttersprachlichkeit" entweder entsteht oder nicht entsteht. Diese Annahme schließt jegliche Möglichkeit eines Kontinuums aus und damit eine granuläre Betrachtung der einzelnen Kompetenzen der Sprecher:innen auf allen Analyseebenen.

Um den semantischen Ambiguitäten zu begegnen, dient der Terminus *L1* in der vorliegenden Arbeit als primärer konzeptueller Bezugsrahmen. Während die Begriffe *Erstsprache* und *Muttersprache* in den folgenden Kapiteln in ihrer jeweiligen sprachspezifischen und theoretischen Verwendung kritisch beleuchtet und kontrastiert werden, fungiert die L1 als die in dieser Arbeit eigenständig definierte Analysekategorie. Darunter wird diejenige Sprache verstanden, die chronologisch zuerst erworben wurde und im individuellen Repertoire den höchsten Grad an Dominanz und Kompetenz aufweist. Die L1 wird somit für den gesamten Argumentationsstrang dieser Arbeit wie folgt definiert:

- Chronologisch gesehen ist sie die allererste Sprache, die die Sprecher:innen gelernt haben. Es kann auch mehr als eine L1 geben, in diesem Fall werden sie als 1L1, 2L1, 3L1 usw. bezeichnet;
- Die Sprache, die während der ersten Sozialisation des Kindes und möglicherweise auch während der Adoleszenz die dominierende Rolle gespielt hat. Der Parameter der Dominanz wird hier jedoch als im Laufe der Zeit und in verschiedenen Kontexten eher instabil angesehen, weshalb die Definition der L1 in Bezug auf diesen Parameter bewusst allgemeiner gehalten ist als die anderen Parameter;
- Die Sprache, mit der sich die Person potenziell identifiziert. Auch hier sollte jedoch ein vorsichtiger Spielraum für mögliche Szenarien gelassen werden. Eine Person kann eine Sprache als ihre L1 betrachten, die sie zwar als erste gelernt hat, in der sie aber im Vergleich zu anderen, später erworbenen Repertoiresprachen keine hohen Kompetenzen aufweist.

In Bezug auf den Terminus *Sprachrepertoire* sei an dieser Stelle auf die Definition von (Blommaert & Backus, 2013, S. 23), in Ghyselen et al. (2020, S. 204) verwiesen:

unique "indexical biographies, [...] biographically assembled patchworks of functionally distributed communicative resources, constantly exhibiting variation and change" (Blommaert/Backus 2013: 23) (Ghyselen et al., 2020, S. 204)

Die hier vorgestellte Definition erscheint geeignet, da sie die Kompetenzanalyse primär auf der individuellen Ebene verortet – ein Fokus, der in früheren Ansätzen oft zugunsten kollektiver Normen vernachlässigt wurde. Gleichwohl darf nicht außer Acht gelassen werden, dass Sprache

stets auch als soziales Phänomen fungiert und im Sinne von Schmidt & Herrgen (2011, S. 131) Teil eines "Gesamtsprachsystems" ist. Für die methodische Vorgehensweise dieser Arbeit ergibt sich daraus die Notwendigkeit einer differenzierten Betrachtung. Die Bestimmung des Status einer Sprache im individuellen Repertoire erfolgt idealerweise durch eine Verschränkung von subjektiver Binnensicht und objektiver Außensicht. Die Selbsteinschätzung der Sprechenden ist dabei von hohem Wert, da sie Aufschluss über das Sprachbewusstsein und die biografische Relevanz der Sprache gibt. Sie allein stellt jedoch noch keinen hinreichenden Indikator für die faktische sprachliche Kompetenz dar, da sie durch psychologische Faktoren (z. B. sprachliche Unsicherheit oder Überschätzung) verzerrt sein kann. Um ein wissenschaftlich valides Bild zu erhalten, wird die subjektive Einschätzung daher durch die Bewertung einer externen Instanz flankiert und validiert. Erst diese komplementäre Betrachtung ermöglicht eine belastbare Einordnung der Repertoiresprache.

1.3 Die Bezeichnung einer nach der L1 erworbenen Sprache (LX)

Die Bezeichnung "zweite Sprache" wird in der Regel durch die Abkürzung L2 definiert. Der Studienbereich selbst, *Second Language Acquisition Studies*, trägt einen an sich mehrdeutigen Namen: *second language*. Der Begriff *second language* bzw. L2 wird häufig als Oberbegriff verwendet und umfasst daher auch dritte, vierte, fünfte usw. Sprachen. Diese weisen normalerweise unterschiedliche Kompetenzen auf und nehmen somit völlig unterschiedliche Positionen im Repertoire der Sprecher:innen ein. Diese Verallgemeinerung geht jedoch mit einer gewissen terminologischen Unschärfe einher (Dewaele et al., 2021, S. 35).

Das Hauptkriterium zur Definition einer L2 als Nicht-L1 ist das chronologische. Diese Überlegungen müssen jedoch durch qualitative Überlegungen zur Dominanz, Kompetenz und zur Qualität des Erlernens in den Repertoire-Sprachen ergänzt werden (Flege et al., 2021). Das Konzept der "kritischen Periode" (im Englischen *critical period*), welches sich auf das hypothetische Vorhandensein eines (oder mehrerer) Zeitfenster bezieht, in denen das Lernen besonders effektiv erscheint, ist traditionell von entscheidender Bedeutung (vgl. Birdsong, 2018). Allerdings besteht bis heute kein allgemeiner Konsens über diese Hypothese. Die Hypothese einer kritischen Periode wird durch neurolinguistische Befunde gestützt, die einen direkten Zusammenhang zwischen dem Zeitverlauf und dem Verlust der neuronalen Plastizität belegen. Es gibt jedoch zahlreiche Gegenbeispiele, die die Gültigkeit der von Lenneberg (1967) aufgestellten Hypothese in Frage stellen könnten (Hakuta et al., 2003). Das *Speech Learning Model* (Flege, 1995) musste ebenfalls einer Überarbeitung unterzogen werden, nachdem eine Vielzahl von Daten eine Bestätigung dafür lieferte, dass auch im Erwachsenenalter eine *native-like* Kompetenz erreicht werden kann. Seit dem 2021 ist es als *Speech Learning Model – revised* (SLM-r) bekannt. Die Definition des Zeitraums, der in die kritische Phase einbezogen wird, sowie die Art und Qualität des Lernens in der kritischen Phase (natürlich oder gesteuert, in einem immersiven und/oder didaktischen Kontext) stellen Aspekte dar, die sich als besonders kritisch erwiesen haben. Auch die Ebenen der

Sprachanalyse und die Lerngeschwindigkeit sind in diesem Kontext von Relevanz (Vanhove, 2013). In Bezug auf die Ebenen der Sprachanalyse und den Umfang, in dem sie innerhalb der Kritischen Periode betroffen werden, besteht die Auffassung, dass unterschiedliche Sprachanalyseebenen in verschiedenen Kritischen Perioden unterschiedlich betroffen werden könnten. Dies würde bedeuten, dass es entsprechende Kritische Perioden für alle Analyseebenen gibt. Was hingegen die Lerngeschwindigkeitsvariable angeht, wird sie ersichtlich, wenn man sich die Ansichten einiger Autoren wie Vanhove (2013) vor Augen führt. Diese besagen, dass Erwachsene eine L2 schneller als Kinder erlernen. Dies deutet auf eine größere Lernbereitschaft hin, die auf einer größeren Neigung zur Verwendung von Sprachlernmethoden basiert. Ein weiterer wesentlicher Faktor ist die Berücksichtigung der Kontextualität, in der soziolinguistische Phänomene des Bi- und Multilingualismus beobachtet werden. Hier ist insbesondere die Komplexität der heutigen Sprachgemeinschaften von Bedeutung, in denen die Koexistenz mehrerer Sprachen innerhalb desselben Repertoires zunehmend häufiger wird und Monolingualismus somit zur Ausnahme macht. Die Komplexität ist nicht nur auf die Anzahl der Sprachen im Repertoire der Einzelnen zurückzuführen, sondern auch auf die Gründe und Methoden ihres Erlernens und Gebrauchs im sozialen Kontext

Im Rahmen einer granularen Analyse eines sprachlichen Repertoires, insbesondere wenn dieses komplex ist, sind die folgenden Hauptfaktoren zu berücksichtigen:

Chronologische Erwerbsfaktoren:

- Frühe L2: eine Sprache, die in der Kindheit in einem natürlichen und ungesteuerten Kontext erworben wird, oft vergleichbar mit dem des L1-Erwerbs;
- Späte L2: eine Sprache, die deutlich später erworben wird, meist in einem formalen Kontext, mit stark unterschiedlichen Erwerbsmodalitäten als bei der L1;

Funktionsfaktoren:

- Dominante L2: eine Sprache, die die alltägliche Kommunikation dominiert, d.h. die am häufigsten verwendet wird;
- Nicht-dominante L2: eine Sprache, deren Nutzung auf bestimmte Kontexte eingeschränkt ist.

Die beiden skizzierten Makrokategorien bilden die Grundlage für eine vertiefte Analyse der Rolle der L2 im individuellen Sprachrepertoire. Die Analyse muss die verschiedenen Facetten der oben genannten Faktoren berücksichtigen, von denen die folgenden entscheidend sind:

- Der Einfluss des Lernbeginns auf die L2-Beherrschung: Frühe L2 neigen dazu, eine eigene Aussprache, Grammatik und Wortschatz zu entwickeln, die von der der L1 unterscheidbar ist. Späte L2 können hingegen Interferenzen aus der L1 auf allen Analyseebenen aufweisen;

- Die Rolle des Lernkontexts: Lernen in einem natürlichen ungesteuerten Kontext fördert den spontanen Erwerb einer L2, während Lernen in einem formalen und gesteuerten Kontext stärker auf den grammatikalischen bzw. strukturellen Aspekt fokussiert ist;
- Die Menge und Häufigkeit der L2-Exposition: Eine häufige und langfristige Exposition gegenüber der L2 ist für die Entwicklung einer guten Sprachbeherrschung unerlässlich;
- Die Lernmotivationen und allgemein die emotionalen Faktoren: Intrinsische Motivationen, wie das Interesse an der Kultur oder der Wunsch, mit Menschen anderer Sprachen zu kommunizieren, fördern tendenziell ein effektiveres und dauerhaftes Lernen. Emotionale Faktoren wie die der Identität des Individuums – entweder von dem Individuum selbst oder von Dritten ausgedrückt – beeinflussen den Lernprozess in positiver bzw. negativer Weise.

In dieser Arbeit wird die Bezeichnung “LX” gemäß (Dewaele et al., 2021):

The label 'LX' refers to any new language acquired after the age at which the first language(s) was/were acquired, which for many language researchers is after the age of 3, to various levels of proficiency. [...] The crucial advantage of the labels 'L1' and 'LX', combined with the noun 'user', is that it circumvents the negative ideological connotations that have become naturalized in the terms 'native' and 'second' in scholarly usage. (Dewaele et al., 2021, S. 35–36).

Im Folgenden wird auf die Verwendung des Begriffs *user* neben “X” verzichtet, um die Aufmerksamkeit nicht nur auf die aktiven (produktiven) Sprachkenntnisse zu beschränken.

Nachdem für diese Dissertation eine Festlegung auf die Bezeichnungen L1 und LX getroffen wurde, sollen im Folgenden spezifische Implikationen sprachbiografischer Termini im Englischen, Deutschen und Italienischen diskutiert werden.

1.4 Sprachbiographische Begriffe im Englischen

1.4.1 Sprachbiographische Begriffe zur Bezeichnung der L1

Der Begriff “native speaker” ist im Englischen weit verbreitet und bezeichnet der/die Sprecher/-in einer L1 (oft als “mother tongue/language” definiert). Die Begriffe “native” und “mother” sind vielfach missverstanden worden. Ein *native speaker* wird, insbesondere im nicht-akademischen Bereich, als der/die Sprecher/-in schlechthin dieser Sprache angesehen, als perfekter Repräsentant des Gebrauchs seiner eigenen Sprache und sogar der ihr zugehörigen Kultur (Dewaele et al., 2021, Neuhold, 2022, S. 36)

Die Gründe für die Verwendung dieser Begriffe in der englischen Sprache liegen ursprünglich in der kolonialen Vergangenheit der Anglo-Sachsen, in der nicht nur die englische wirtschaftliche Hegemonie, sondern auch eine gewisse ethnisch-kulturelle Überlegenheit etabliert wurde. In Dewaele et al. (2021) wird das Phänomen als «Anglo-Saxonism superiority over the Roman world» bezeichnet:

Emerging for the first time in the 16th Century in the context of reformation and giving ideological support for the break with Rome, it re-appeared in the USA of the "Great Migration", a country feeling threatened by a large number of non-English speaking (and often Catholic) immigrants. Thus, from the very beginning, the notion of the NS [native speaker, ADL] has been carrying with it ideas of superiority and exclusion. (Dewaele et al., 2021, S. 26).

Die Begriffe *native* und *mother*, die die Substantive *language*, *tongue* und *speaker* begleiten, beziehen sich in erster Linie auf den Besitz einer Sprache seit der Geburt, weil sie etymologisch mit der Figur der Mutter in Zusammenhang gebracht werden. Der Begriff *heritage language* dagegen ist mit ethnischen Konnotationen verbunden: eine *heritage language* (auf Deutsch "Herkunftssprache") ist eine Sprache, die aus einer kulturellen Tradition stammt und nicht der aktuellen vorherrschenden Kultur am aktuellen Wohnort entspricht.

Der Begriff *first language* wird häufig synonym mit *mother tongue* verwendet. Das synonyme Verhältnis kann jedoch nicht als selbstverständlich angesehen werden, da der Begriff *first* mehrere Bedeutungen haben kann. In Bezug auf das chronologische Kriterium lässt sich ableiten, dass der Erwerb der Sprache in der genannten Reihenfolge stattgefunden hat. Aus kognitionswissenschaftlicher Perspektive bezeichnet der Begriff *first* eine relative Dominanz zu einer oder mehreren Sprachen des Repertoires. Konkret bezeichnet er die Sprache, die im Durchschnitt die größte Rolle spielt. Auf der kognitiven Ebene ist damit auch die relative Kompetenz verbunden. Des Weiteren ist die emotionale Komponente zu berücksichtigen, welche sich auf eine mögliche Präferenz der Sprecher:innen für die betreffende Sprache bezieht und oft mit ideologischen Überlegungen einhergeht. Diese können sowohl von den Sprecher:innen selbst als auch von Dritten formuliert werden.

In der Literatur, die diese Terminologie verwendet, lassen sich zwei Haupttendenzen unterscheiden: Quellen, die die beiden Begriffe synonym verwenden, ohne die Assoziation selbst zu problematisieren (vgl. z.B. Seid, 2019; Atkinson, 1987), und Quellen, die für ihre terminologische Wahl und deren Verwendung argumentieren (vgl. z.B. Kecskes & Papp, 2000, S.1).

Eine Überlappung der beiden Begriffe ist insbesondere in komplexen Sprachgemeinschaften problematisch, in denen die im häuslichen Umfeld erlernte Sprache von der dominanten Sprache im Repertoire der Sprecher:innen und/oder eventuell in der Gesellschaft abweicht. Die Hauptursache für diese Diskrepanz ist in der überwiegenden Mehrheit der Fälle auf sozio-politische Gründe zurückzuführen (vgl. Seid, 2019). Laut Seid (2019) bezeichnet "mother tongue" die im häuslichen Umfeld erlernte Sprache (in diesem Fall Äthiopisch), während "first language" die außerhalb dieses Kontexts dominante Sprache (in diesem Fall Englisch) bezeichnet. In sprachlich homogenen Gemeinschaften ist eine Überlappung dagegen eher legitim (vgl. Kecskes & Papp, 2009; Yadav, 2014). Es ist jedoch wichtig, den chronologischen Aspekt nicht mit der Rolle der Dominanz einer Repertoiresprache zu verwechseln. In der Literatur finden sich Fälle wie der von Yadav (2014), in denen die Verwendung von *primary* anstelle von *first* zu einer Überschneidung

von Konzepten führt, die klar unterschieden werden sollten: Eine *primary language* ist nicht notwendigerweise die *first language* im chronologischen Sinne und umgekehrt.

1.4.2 Sprachbiographische Begriffe zur Bezeichnung der LX

Die dargestellten theoretischen Problematiken, insbesondere die zur Definition einer oder mehrerer Kritischen Periode und des damit verbundenen Zusammenspiels kontextueller Variablen, bilden den Hintergrund für einen terminologischen Gebrauch, der auch in diesem Fall zu zahlreichen Unklarheiten führen kann. Einerseits werden die Begriffe *L2* und *second language* im wörtlichen und daher chronologischen Sinne verstanden, andererseits werden sie oft auch als übergeordneter Begriff verwendet, um jede nach der *L1* erlernte Sprache ohne weitere Spezifikationen zu bezeichnen (vgl. Myers-Scotton, 2006; Gut et al., 2023). Zweitens resultiert aus dem ersten Punkt eine mangelnde Differenzierung, sodass der Begriff *second language* (*L2*) eine Vielzahl von Ausprägungen annehmen kann, die durch das Zusammenwirken der Faktoren a) Lernalter, b) Kompetenz, c) Lernmodus, d) Lernkontext und e) Lernmotivation entstehen (vgl. Myers-Scotton, 2006). Drittens ist aufgrund der Komplexität des Lernphänomens selbst und seiner möglichen Konfigurationen sowie der noch bestehenden Meinungsverschiedenheiten im theoretischen Bereich eine Differenzierung der Begriffe, die scheinbar dasselbe Konzept bezeichnen, jedoch einer genaueren Spezifikation bedürfen, schwierig. In der Folge werden daher Konzepte unterschieden, die in der Literatur vielfach kontrovers diskutiert werden.

1.4.2.1 Bilingualism vs Second Language Learning

In der Literatur werden diese beiden Begriffe zumeist als zwei verschiedene Konzepte definiert, die jedoch in Abhängigkeit vom jeweiligen Analysekontext variieren können. In der Regel wird der Begriff *Bilingualism* verwendet, um ein individuelles oder soziales Phänomen zu beschreiben, wobei in diesem Kontext das Adjektiv *societal* verwendet wird. Im Falle des individuellen Phänomens, das hier von Interesse ist, handelt es sich um den Erwerb von zwei Sprachen im Kindesalter, der in der Regel im familiären Umfeld stattfindet. Daher ist es tendenziell von dem Begriff *Second Language Learning* zu unterscheiden, sowohl aufgrund eines chronologischen als auch eines kontextuellen Kriteriums. Letzteres besagt, dass die Zweitsprache (im weiteren Sinne) eine Sprache ist, die zu einem späteren Zeitpunkt in der Entwicklung und in einem formellen Kontext erlernt wird. Diese Definition könnte angesichts von Begriffen wie *early second language acquisition* (vgl. de Houwer, 2009) oder der einfachen Überlegung, dass es in einigen mehrsprachigen Gemeinschaften üblich ist, bereits im Kindergarten eine zweite Sprache (im chronologischen Sinne) zu unterrichten, ins Wanken geraten. In diesem Fall stimmen das chronologische und das kontextuelle Kriterium überein.

Eine Analyse könnte vereinfacht werden, wenn dem Begriff *second* ein funktionaler Wert zugewiesen wird. Demnach ist die *second language* eine Sprache, die nicht wesentlich für die

tägliche Interaktion in der eigenen sprachlichen Gemeinschaft ist, jedoch potenziell für andere Zwecke nützlich sein kann. In diesem Fall könnte der Terminus *secondary language learning* vorgeschlagen werden.

1.4.2.2 *Second Language Learning vs Foreign Language Learning*

Die Schwierigkeit, diese beiden Konzepte zu unterscheiden, hängt, wie im vorherigen Fall, von der Perspektive bzw. dem Kriterium ab, nach dem die Begriffe *second* und *foreign* interpretiert werden. Die größere Ambiguität des Begriffs *second* ergibt sich aus seiner zeitlichen und funktionalen Bedeutung. Der Begriff *foreign* hingegen bezieht sich ausschließlich auf die qualitativ-funktionale Dimension. Eine Fremdsprache ist demnach eine Sprache, die nicht die Amtssprache des Landes ist, in dem sie erlernt wird. Die Unterscheidung der beiden Konzepte kann anhand des chronologischen Kriteriums einfach vorgenommen werden. Das funktionale Kriterium, das *second* im Sinne von *secondary* definiert, führt jedoch zu einer tendenziellen Überlappung der beiden Konzepte, wie sie in der Literatur beobachtet werden kann.

1.4.2.3 *Multilingualism vs Plurilingualism*

Obwohl die beiden Begriffe *multilingualism* und *plurilingualism* auf den ersten Blick synonym erscheinen, lassen sie sich in erster Linie dadurch unterscheiden, dass der Begriff *multilingualism* ein soziales Phänomen definiert, während *plurilingualism* ein individuelles Phänomen beschreibt. Diese Differenzierung wird jedoch nicht durchgängig beibehalten, sodass die Begriffe häufig synonym verwendet werden. In beiden Fällen wird z.T. fälschlicherweise angenommen, dass die Kompetenz in den betreffenden Sprachen auf allen Analyseebenen hoch oder muttersprachlich sei.

1.5 Sprachbiographische Begriffe im Deutschen

1.5.1 Sprachbiographische Begriffe zur Bezeichnung der L1

Die Problematik des Begriffs *Muttersprache* spiegelt die bereits für den englischen Terminus *mother tongue* dargelegten Konnotationen wider (vgl. Abschnitt 1.4.1). Auch im Deutschen verbergen sich dahinter Ideologien, die genetische oder familiäre Faktoren überbewerten. Eine für den deutschen Sprachraum spezifische historische Dimension eröffnet jedoch Haugen (1991), der auf den mittelalterlichen germanischen Sprachraum als Ursprung des Begriff verweist: genauer gesagt, wird Straßburg als Stammort der ersten Erwähnung des Wortes "Muttersprache". Straßburg ist in der Tat ein Ort, dessen Lage beredt für die Geschichte dieses Begriffs ist: Seine Lage an der Grenze zwischen dem lateinischen und dem deutschen Sprachraum war sicherlich eine wesentliche Ursache für die Entstehung des sprachlichen Kontakts zwischen den beiden Welten (vgl. Haugen, 1991, S. 76). Wie erwartet in derartigen politischen und sozio-linguistischen Kontexten, führt der Sprachkontakt zu einer Definition des Sprachgebrauchs im Großen und im

Kleinen: von der territorialen Dimension bis hin zur eher begrenzten Dimension der einzelnen Haushalte gemischter Paare. Dabei wurde jedem Elternteil eine bestimmte Sprache zugeordnet. Im vorliegenden Fall war die "Muttersprache" die germanische Sprache der Ehefrau, die die Aufgabe hatte, sich um die Erziehung der Kinder zu kümmern (vgl. ebd.). Der Begriff hat im Laufe der Zeit eine zunehmend exklusive Konnotation angenommen, und die Vorstellung, dass eine Person, die ihre "Muttersprache" perfekt beherrscht, über herausragende Kompetenzen in dieser Sprache verfügt, hat sich zunehmend verbreitet. Dieses Vorurteil ist im Bereich des Deutschunterrichts als Fremdsprache (Deutsch als Fremdsprache, DaF) noch verbreitet, insbesondere im schulischen Bereich (und somit nicht im akademischen). Dabei handelt es sich insbesondere um Sonderfälle, die sich häufig auf Dienstleistungen beziehen, die privaten Fremdsprachenunterricht anbieten und "Muttersprachlichkeit" als die "schönste Visitenkarte" definieren² oder Vereine, die nicht-muttersprachlichen Deutschlehrer: innen Unterstützung anbieten, ohne besondere Qualifikationen von der Seite der deutschen Muttersprachler: innen zu erfordern³. In der deutschsprachigen Literatur findet sich trotz geringerer Quantität eine größere Auseinandersetzung mit der terminologischen Frage. Zudem existiert eine Vielzahl an Studien im glottodidaktischen Bereich, welche den Gebrauch durch Fachleute untersuchen.

Ein Beispiel dafür ist die Studie von (Triulzi et al., 2023). Die Autoren verfolgen das Ziel, die studentischen Begriffsnutzungen im Kontext von Migration und Deutsch als Zweitsprache zu analysieren und stellen die folgende Problematik dar:

[...] Andererseits scheint für diese Herausforderung bei den Studierenden nur zum Teil Sensibilität vorhanden zu sein und die Problematik bestimmter Begriffsnutzungen wenig reflektiert zu werden. Unsicherheiten werden so beispielsweise eher durch das Angebot mehrere Begriffe begegnet ("Zu den Herkunftssprachen bzw. Familiensprachen gehören (...) S7TT), statt Unzulänglichkeiten der Begrifflichen Konzepte selbst direkt zu thematisieren. Die hier untersuchten Begriffe scheinen von den Studierenden teils als austauschbar empfunden zu werden, was insbesondere dann problematisch wird, wenn das Konzept Sprachigkeit mit weiteren Begriffen [...] die dafür gar nicht geeignet sind [...] (Triulzi et al., 2023, S. 77)

König (2016) entwirft eine prototypische Definition von *Muttersprache*, indem sie die verschiedenen Aspekte hervorhebt, die beim Erwerb der L1 eine zentrale Rolle spielen (chronologisch, qualitativ, kontextuell, funktional und emotional)⁴:

[...] Prototypisch wird angenommen, dass Deutsch als Muttersprache (DaM) als erste Sprache von Lebensbeginn an natürlich und ungesteuert erworben wird und zumindest zum Zeitpunkt des Erwerbsbeginns von zentraler kommunikativer Bedeutung für das Individuum ist (vgl. Ahrenholz 2008 in

² Siehe z.B. die Blog-Seite des Dienstes "Superprof" unter "Lehrer mit Deutsch als Muttersprache": <https://www.superprof.de/blog/deutsch-als-fremdsprache-unterrichten-ausbildung/> (zuletzt aufgerufen am 09.07.2024).

³ Siehe z.B. die Webseite des Vereins "Wir Muttersprachler E.V.": <https://www.wir-muttersprachler.de/impressum/> (zuletzt aufgerufen am 09.07.2024).

⁴ Königs Definition basiert auf der Voraussetzung, dass die zu beschreibenden Sprachgemeinschaften eine gewisse Homogenität aufweisen. Es ist jedoch zu hinterfragen, ob diese Bedingung für die Beschreibung des Profils vieler derzeit existierender Sprachgemeinschaften tatsächlich erfüllt ist. Daher wird Königs Definition als ein "idealer" Ausgangspunkt betrachtet, der je nach Bedarf an die einzelnen Fälle angepasst werden sollte.

König 2016, ADL) [...] Als *Muttersprache* wird die Sprache eines Individuums bezeichnet, die es mit Mitgliedern einer kulturell homogenen Gemeinschaft als Erstsprache gemeinsam hat und zu der es auf dieser Grundlage eine spezifische, auch affektive Bindung empfindet. (König, 2016)

Laut Triulzi et al. (2023) können alternative Begriffe zur Bezeichnung einer L1 verwendet werden, die sich von den Begriffen *Muttersprache* oder *Erstsprache* unterscheiden. Dazu zählen die Begriffe *Herkunftssprache* und *Familiensprache*. Die wechselseitige Austauschbarkeit dieser Begriffe ist in vielen Fällen nicht eindeutig. Des Weiteren wird davon ausgegangen, dass sie nicht vollständig mit dem hier präferierten Begriff der Erstsprache übereinstimmen. Es handelt sich um zwei Begriffe, die sich auf eine Sprache "der Herkunft" und eine im familiären Kontext gesprochene Sprache beziehen. Letztere wird auch als Familiensprache bezeichnet und ist somit nicht im schulischen oder gesellschaftlichen Kontext zu verorten. Obwohl sie sich auf eine Sprache "der Herkunft" und eine im familiären Kontext gesprochene Sprache beziehen, sind die Begriffe vorwiegend im Bereich der DaF-Lehre verbreitet. Insbesondere zur Beschreibung von Sprecher:innen, die mindestens zwei Sprachen in ihrem Repertoire haben, diese funktional (Familiensprache) verwenden und unterschiedliche Kompetenz- und Dominanzniveaus in den (mindestens) zwei Repertoiresprachen aufweisen, werden sie verwendet.

1.5.2 Sprachbiographische Begriffe zur Bezeichnung der LX

Die Beschreibung einer nicht-L1-Sprache (LX), wie sie beispielsweise im Englischen erfolgt, stellt sich als eine Aufgabe dar, die auf den ersten Blick einfach erscheint, bei näherer Betrachtung jedoch komplex ist. Die Beschreibung des Profils einer nicht-L1-Sprache, wie sie im Repertoire der Sprecher:innen erscheint, ist eine komplexe Aufgabe. Die im Deutschen zu diesem Zweck verwendete Terminologie weist aufgrund ihres uneinheitlichen Gebrauchs in der Literatur einige Unklarheiten auf. Diese Inkonsequenz ist auf dieselben theoretischen Unsicherheiten zurückzuführen, die auch den anglophonen Raum betreffen. Des Weiteren ist zu berücksichtigen, dass die heutigen Sprachgemeinschaften zunehmend komplexer werden. Dies ist darauf zurückzuführen, dass die Sprecher:innen immer komplexere und variabelere Sprachrepertoires aufweisen, sowohl in räumlicher als auch in zeitlicher Hinsicht.

Neben allgemeineren kontextuellen Faktoren ist es empfehlenswert ebenso die individuellen Merkmale der Sprecher:innen zu berücksichtigen, wie die Sprachlernprädisposition, die verwendeten Lernstrategien und die erhaltene sprachliche Unterstützung. Im Folgenden wird der Versuch unternommen, die Unterschiede zwischen den am häufigsten verwendeten Begriffen zur Definition einer LX zu erörtern.

1.5.2.1 Bilingualismus vs Bilingualität

Ähnlich wie im englischen Diskurs (vgl. Abschnitt 1.4.2.1) existiert auch im Deutschen eine terminologische Unterscheidung zwischen dem sozialen und dem individuellen Phänomen. Während im Englischen jedoch oft *bilingualism* für beides verwendet wird, bietet das Deutsche

mit dem Begriffspaar *Bilingualismus* (gesellschaftlich/makroskopisch) und *Bilingualität* (individuell/mikroskopisch) eine präzisere Differenzierungsmöglichkeit an (vgl. Schneider, 2015).

Zwar finden sich in der Literatur auch abweichende Verwendungen – etwa bei Müller (2006), wo “Bilingualismus” auf den individuellen Erwerb bezogen wird, oder unscharfe Begriffe wie *frühe Zweisprachige* (vgl. Nitsch, 2007, S. 55) – doch folgt diese Arbeit konsequent der terminologischen Schärfe Schneiders. Bilingualität wird somit als der zutreffende Terminus für das Individuum definiert und – im Einklang mit Müller et al. (2023) – als ein spezifischer Fall von Mehrsprachigkeit betrachtet.

In dieser Arbeit wird sowohl die von Schneider (2015) vorgenommene Differenzierung als auch die von (Müller et al., 2023) vertretene Position berücksichtigt, wonach der deutsche Terminus *Bilingualität* als eine Unterkategorie bzw. ein spezifischer Fall von Mehrsprachigkeit zu betrachten ist.

1.5.2.2. *Bilingualität vs Zweisprachigkeit*

Die Differenzierung zwischen diesen beiden Begriffen ist nicht evident, da die Termini *Zweitsprachigkeit* und *Zwei(t)sprache/zweite Sprache* nicht durchgängig einheitlich definiert werden. Das zentrale Problem ist daher dasselbe wie beim englischen Begriff *second* (im chronologischen oder funktionalen Sinne): Es ist von entscheidender Bedeutung, welches der beiden Kriterien als das relevanteste erachtet wird. Im Falle des chronologischen Kriteriums ist es von entscheidender Bedeutung, einen Bezugspunkt zu bestimmen, ab dem die Bilingualität idealerweise endet und es legitim wird, von Zweisprachigkeit zu sprechen.

Die Relevanz des “Wann” wird ebenfalls in Chilla & Fox-Boyer (2016) betont, indem die Autorinnen in ihrem Leitfaden für Angehörige, Betroffene und Fachleute die Frage aufwerfen: Wer ist zweisprachig/bilingual?:

[...] Wer oder was zweisprachig ist, wird jedoch ganz unterschiedlich beantwortet. Die Begriffe “Bilingualität” oder “Zweisprachigkeit” haben für viele Menschen ganz unterschiedliche Bedeutung, je nachdem, welche Aspekte der Zweisprachigkeit in den Vordergrund gerückt werden. Es ist zum einen die Frage nach dem *Wann*, also dem Alter, in dem mit je einer der beiden Sprachen begonnen wurde. Zweitens stellt sich die Frage nach dem *Wie*, also der Art und Weise, wie er seine beiden Sprachen erwirbt. Dann wird auch die Qualität, mit der die erworbenen oder gelernten Sprachen beherrscht werden, betrachtet. Des Weiteren spielt eine Rolle, in welchen Kontexten zwei Sprachen verwendet werden und schließlich inwieweit eine individuelle Identifikation mit den beiden Sprachen besteht und welche Haltung man ihnen gegenüber einnimmt [...] (Chilla & Fox-Boyer, 2016, S. 12)

1.5.2.3 *Zweisprachigkeit VS Fremdsprachigkeit*

Die Abgrenzung folgt im Wesentlichen der funktionalen Unterscheidung, die bereits in Abschnitt 1.4.2.2 diskutiert wurde: Die Zweitsprache dient der unmittelbaren Handlungsfähigkeit im Zielland

(Überlebensfunktion), während die Fremdsprache primär im gesteuerten Unterricht ohne direkten Umgebungsbezug erworben wird (vgl. Elsner, 2020).

Wie von Elsner (2020) betont, erweist sich die Unterscheidung zwischen Zweitsprache und Fremdsprache anhand dieser Kriterien jedoch als nicht immer einfach. Dies zeigt sich besonders deutlich in spezifischen Szenarien wie der Sprachlandschaft Südtirols, die zu einer Verwirrung hinsichtlich der Terminologie führen können. In dieser einzigartigen Konstellation treffen Italienisch, Deutsch und Ladinisch aufeinander und koexistieren. Die Lehre dieser Sprachen fängt schon im Kindergarten an. In der Literatur wird darauf verwiesen, dass Deutsch als "dritte Sprache" unterrichtet werden kann (Ricci Garotti, 2017): es handelt sich um die dritte Sprache, die in der Reihenfolge des Erwerbs nach dem Italienischen und einer Varietät des Deutschen erlernt wird. Es lässt sich argumentieren, dass es sich hierbei um ein gesteuertes Lernen handelt, das dem Zweitsprachenlernen ähnlich ist, wie von Elsner (2020) beschrieben. Das in Südtirol im schulischen Kontext unterrichtete Deutsch ist Standarddeutsch, dessen Verwendung im Laufe des Lebens von grundlegender Bedeutung sein wird – also keineswegs eine "fremde Sprache".

1.5.2.4 Mehrsprachigkeit vs Plurilingualismus/Multilingualismus

Auch im letzten Begriffspaar lässt sich eine uneinheitliche Behandlung der Bedeutungen und ihrer Beziehungen zueinander feststellen. Auf lexikografischer Ebene besteht keine Einigkeit darüber, ob diese Begriffe als Synonyme betrachtet werden können und inwieweit. In einer verallgemeinernden Betrachtung lässt sich ableiten, dass der Begriff *Mehrsprachigkeit* auf andere Lemmata verweist und mit ihnen in Verbindung gebracht werden kann. In erster Linie ist hier der Begriff *Multilingualismus* zu nennen. Des Weiteren werden die Begriffe *Multilingualismus*, *Mehrsprachigkeit* und *Plurilingualismus* synonym verwendet (vgl. Flinz, 2014).

Der Hauptaspekt, der eine Differenzierung oder Überschneidung der Begriffe bedingt, ist der der Dimensionen, in denen sich die Verhältnisse zwischen mehreren Sprachen ausdrücken. In (Lüttenberg, 2010) wird ersichtlich, dass der Begriff *Mehrsprachigkeit* zwei verschiedene Hauptkonnotationen aufweisen kann:

- **Mehrsprachigkeit als Zustand**, in Bezug auf Glück (2000). Der Zustand kann sowohl auf individueller als auch auf sozialer Ebene verstanden werden (siehe den Verweis auf das Lemma "Diglossie" in Bußmann (1990), in Riehl (2014, S. 20) wird zudem zwischen "Triglossie" und "Polyglossie" unterschieden, je nachdem, wie viele Sprachen involviert sind);
- **Mehrsprachigkeit als Prozess**, im Einklang mit Weinreich (1964, in Lüttenberg, 2010), der von der «practice of alternatively using of two languages» spricht (Weinreich, 1964). Das Konzept wird ähnlich in (Oksaar, 2003):

[...] Mehrsprachigkeit ist die Fähigkeit eines Individuums, hier und jetzt zwei oder mehr Sprachen als Kommunikationsmittel zu verwenden und ohne weiteres von der einen in die anderen umzuschalten, wenn die Situation es erfordert. [...] (Oksaar, 2003, S. 31)

Die kommunikative Dimension wird bei Oksaar besonders betont. Allerdings wird postuliert, dass die Kompetenz mehrsprachiger Sprecher:innen sowohl auf aktiver (produktiver) als auch auf passiver (rezeptiver) Ebene Berücksichtigung finden sollte. Folglich können mehrsprachige Sprecher:innen auch dann als solche definiert werden, wenn sie nicht in der Lage sind, seine Sprachkompetenzen in einer LX aktiv zu nutzen. Da diese Fähigkeiten aufgrund ihrer Komplexität nur schwer mess- und kontrollierbar sind und zudem besonders anfällig für andere (auch äußere) Faktoren sind, wird es als ratsam erachtet, ihre Rolle bei der Definition von mehrsprachigen Sprecher:innen sowie mögliche gegenseitige Einflüsse nicht auszuschließen.

1.6 Sprachbiografische Begriffe im Italienischen

1.6.1 Sprachbiographische Begriffe zur Bezeichnung der L1

Wie im Deutschen stellt auch die italienische Terminologie (*madrelingua*, *lingua madre*) weitgehend ein Lehnkonstrukt dar, das die semantische Fracht des englischen *mother tongue* übernimmt (vgl. Abschnitt 1.4.1). Im Gegensatz zum deutschen Sprachraum fehlt im Italienischen jedoch eine historisch verankerte, kritische Auseinandersetzung mit der Begriffsgenese.

Eine explizite Problematisierung der Terminologie ist in der italienischen Forschungsliteratur ein Desiderat. Während sich die englisch- und deutschsprachige Linguistik intensiv mit den ideologischen Implikationen von *native speaker* und *Muttersprachler* auseinandersetzt, werden die Begriffe im italienischen Diskurs vorwiegend pragmatisch verwendet. In der angewandten Linguistik und Glottodidaktik (vgl. z.B. Marelllo, 2009; Ceracchi, 2007) dienen sie primär zur operativen Abgrenzung gegenüber dem Erwerb des Italienischen als Fremdsprache, ohne dass die zugrundeliegenden Konzepte selbst einer vertieften theoretischen Dekonstruktion unterzogen werden.

Diese theoretische Leerstelle führt dazu, dass Definitionen häufig direkt aus der Lexikografie übernommen werden. Bei der Durchsicht der Online-Versionen der Wörterbücher *Garzanti* (*Dizionari Garzanti Linguistica*, 2025), *Hoepli* (*Grande Dizionario Hoepli*, 2022) und *Il Sabatini Coletti* (*Il Sabatini Coletti - Dizionario della lingua italiana*, 2018) kristallisieren sich zwei konstante Überlegungen heraus: Einerseits die Tatsache, dass die L1 seit der Kindheit erlernt wird, andererseits ein beständiger Bezug bzw. eine Assoziation zum Herkunftsland. Ebenso konstant ist der Bezug auf den Singular, also auf eine einzige L1.

Eine bemerkenswerte Ausnahme in der ansonsten eher pragmatisch orientierten Debatte stellt der soziolinguistische Ansatz von Berruto (2008) dar. Wenngleich dieser Beitrag bereits vor einigen

Jahrzehnten publiziert wurde, erweist sich seine kritische Auseinandersetzung mit dem Begriff des *native speaker* (in Anlehnung an Ballmer, 1981) vor dem Hintergrund des komplexen italienischen Sprachpanoramas als erstaunlich aktuell.

Besondere Relevanz gewinnt Berrutos Analyse durch die Postulierung eines Kontinuums, das dem Konzept der "Muttersprachlichkeit" eine deutlich höhere Variabilität zuschreibt, als dies im heutigen, oft vereinfachenden Laiendiskurs der Fall ist. Dieses Kontinuum (vgl. Berruto, 2003, S. 7–8) spannt sich zwischen zwei Extrempolen auf: dem *good native speaker* à la Ballmer (einem monolingualen Idealtypus, der heute kaum noch anzutreffen ist) und dem *non-native speaker*. Dazwischen verortet Berruto differenzierte Kategorien wie:

- den ***ex-native speaker*** (Sprecher:innen, die durch *language attrition* Kompetenzen in der Erstsprache verlieren);
- den ***semi-native speaker*** (im Kontext des bilingualen Spracherwerbs, vgl. Berruto et al., 1988, S. 12);
- den ***quasi-native speaker*** (analog zum englischen *near-native*), der das Italienische nach der primären Sozialisation als Zweit- oder Fremdsprache erworben hat (vgl. Berruto, 1988, S. 8).

Während Berruto diese Differenzierung auf einer theoretisch-konzeptuellen Ebene vornimmt, liefert die quantitative sprachbiographische Forschung den empirischen Nachweis dafür, dass das Konzept der *Muttersprache* in der konkreten Perzeption der Sprecher:innen als ideologischer und identitätsstiftender Filter operiert. In diesem Kontext ist der sprachpolitische Ansatz von Dell'Aquila & Iannàccaro (2006; vgl. auch Dal Negro, 2011) zu verorten. Anhand umfangreicher Datenbasen aus multilingualen Kontaktträumen (u. a. der *Survey Ladins*) zeigen die Autoren, dass die unreflektierte Zuordnung des Begriffs *Muttersprache* auf der emischen Ebene der Sprecher:innen oft in einer klaren Divergenz zur tatsächlichen, ethisch beobachtbaren primären Sozialisationssprache steht. Die "muttersprachliche" Deklaration fungiert folglich nicht als rein biographische Deskription, sondern als symbolisches Bekenntnis zu einer Affiliationsgemeinschaft. Dieses sprachkognitive Muster, welches die Autoren als ein historisch sedimentiertes Erbe der europäischen Nationalstaatenideologie des 19. Jahrhunderts identifizieren (vgl. auch Bonfiglio, 2010), verdeutlicht die Notwendigkeit, den Begriff im Spannungsfeld zwischen akademischer Taxonomie und gesellschaftlicher Wahrnehmung zu verhandeln.

Dieser Übergang von der theoretischen Modellierung zur Analyse der sprachlichen Alltagswirklichkeit und deren Vermittlung prägt auch die aktuelle wissenschaftliche Divulgation. Hierbei steht das Bestreben im Vordergrund, das Bewusstsein für die Pluralität und Variabilität von Erstsprachkompetenzen einer breiteren Öffentlichkeit zugänglich zu machen. In diesem Rahmen ist der Beitrag von Gallina (2021) zu verorten, die sich explizit der Frage widmet, wie der Begriff *lingua madre* heute verstanden und vermittelt wird:

[...] Come si può evincere dalla parola, possiamo dire in prima battuta che è la lingua che ereditiamo dalla nostra famiglia, anche se le situazioni possono essere molto più complesse, soprattutto nei contesti plurilingui. [...] L'espressione 'lingua madre' è in parte sovrapponibile a espressioni come 'home language', 'lingua di origine', 'prima lingua', 'L1'. In tutti questi casi, il parlante è definito come '**madrelingua**' [sic, ADL] o '**parlante nativo**' [sic, ADL], per sottolineare il fatto che quell'individuo parla quella lingua perché l'ha sentita fin dalla propria nascita (e anche durante la gestazione!): si tratta della lingua con cui compie le prime interazioni, che spesso utilizza per imparare a leggere, a scrivere e a far di conto, anche se, come abbiamo visto, molti individui a scuola utilizzano una lingua che non dominano e che non è la loro lingua madre. Eppure, **non tutti i nativi sono uguali [sic, ADL]**, nel senso che ciascun individuo impara la lingua in un certo contesto, segue percorsi educativi e di vita che possono essere molto eterogenei, ha conoscenze linguistiche che non coincidono mai perfettamente con quelle di un altro individuo [...]»⁵ (Gallina, 2021)

Gallina (2021) weist darauf hin, dass, was den Prozess des Erstspracherwerbs betrifft, die individuellen Unterschiede zwischen L1-Sprecher:innen beträchtlich sind. Diese Unterschiede manifestieren sich in einzigartigen Wegen, die von jedem Individuum beschritten werden. Es lässt sich vermuten, dass diese Sensibilität gegenüber der Komplexität der Definition der L1-Sprecher:innen auch auf die idiosynkratische Einzigartigkeit der italienischen Sprachlandschaft zurückzuführen ist. Die italienische Amtssprache ist Italienisch, jedoch wird Standarditalienisch nicht als Erstsprache erlernt. Stattdessen wird von einer äußerst granularen Konstellation von "lokalen Italienisch-Varietäten" gesprochen, die sich in einem Kontinuum von stärker und weniger stark lokalen Merkmalen befinden.

1.6.2 Sprachbiographische Begriffe zur Bezeichnung der LX

Auch die italienischsprachigen Quellen zum Erwerb einer LX zeigen eine Terminologie, die stark an das Englische angelehnt ist. Dennoch finden sich auch spezifische Neukonzeptionen. Eine Literaturrecherche verdeutlicht, dass die Beiträge in italienischer Sprache quantitativ geringer sind

⁵ Übersetzung ins Deutsche: «Der Begriff 'Muttersprache' lässt sich zunächst dadurch definieren, dass er die Sprache bezeichnet, die eine Person von ihren Eltern erbt. In mehrsprachigen Kontexten können die Situationen jedoch viel komplexer sein. Der Ausdruck 'Muttersprache' überlappt teilweise mit Begriffen wie 'home language', 'Sprache der Herkunft', 'Erstsprache' und 'L1'. In all diesen Fällen wird der Sprecher als 'Muttersprachler' oder 'native speaker' bezeichnet, um zu betonen, dass diese Person diese Sprache spricht, weil sie sie seit ihrer Geburt gehört hat (und auch während der Schwangerschaft!). Es handelt sich um die Sprache, in der die ersten Interaktionen durchgeführt werden, ist oft auch die Sprache, in der Lesen, Schreiben und Rechnen erlernt wird. Dies ist jedoch nicht zwangsläufig der Fall, wie wir gesehen haben. Viele Menschen verwenden in der Schule eine Sprache, die sie nicht beherrschen und die nicht ihre Muttersprache ist. Dennoch ist zu berücksichtigen, dass nicht alle Muttersprachler in gleicher Weise geprägt sind. Diesbezüglich ist festzuhalten, dass jeder Einzelne die Sprache in einem spezifischen Kontext erwirbt, Bildungs- und Lebenswege verfolgt, die sich erheblich unterscheiden können, und sprachliche Kenntnisse aufweist, die nie exakt mit denen eines anderen Individuums übereinstimmen. [...]» (Übersetzung von ADL).

als im Englischen oder Deutschen und fast ausschließlich auf die Glottodidaktik – speziell den Unterricht des Italienischen als Zweit- oder Fremdsprache – ausgerichtet sind.

Diese Fokussierung resultiert aus der Notwendigkeit, auf zwei soziökonomische Phänomene zu reagieren: die Migration nach Italien und die sogenannte *delocalizzazione* (Verlagerung) italienischer Unternehmen ins Ausland (vgl. Balboni, 2014). Im ersten Fall zielt die Didaktik auf den Erwerb von Kompetenzen für einen Aufenthalt in Italien ab. Im zweiten Fall, bedingt durch Unternehmensgründungen in Osteuropa oder dem Maghreb, hat sich der Status des Italienischen gewandelt: Es ist nicht mehr nur ein ‚intellektuelles Luxusgut‘ (vgl. Balboni, 2014), sondern eine pragmatische Notwendigkeit. Da diese Szenarien unterschiedliche Bedürfnisse implizieren, werden die Begriffe *lingua seconda*, *lingua straniera* und *bilinguismo* in der Literatur nicht immer trennscharf verwendet. Im Folgenden werden diese Konzepte daher für die Zwecke dieser Arbeit präzisiert.

1.6.2.1 *Bilinguismo*

Der Begriff *bilinguismo* weist im Italienischen ähnliche Ambiguitäten auf wie im Englischen. Er wird häufig als Oberbegriff verwendet, ohne explizit zwischen der gesellschaftlichen und individuellen Ebene zu differenzieren – eine Unterscheidung, die im Italienischen primär durch das Begriffspaar *multilinguismo/plurilinguismo* geleistet wird (siehe Abschnitt 1.6.2.2).

In der Literatur wird *bilinguismo* zudem oft von *diglossia* abgegrenzt. Historisch relevant ist die Überwindung der Gleichsetzung von Bilingualismus mit „*equilinguismo*“ (perfekte, ausgewogene Kompetenz). Wie Milazzo (2015, S. 11) unter Verweis auf Weinreich (1964) darlegt, gilt diese idealisierte Vorstellung zumindest auf akademischer Ebene als obsolet.

Analog zur anglofonen Literatur wird *bilinguismo* heute multifaktoriell klassifiziert. Milazzo (2015) schlägt folgende, für den italienischen Diskurs relevante Unterscheidungen vor, wobei ein Kontinuum zwischen den Kategorien angenommen wird:

- ***bilinguismo primario vs bilingualismo secondario*** (primäre VS sekundäre Bilingualität): Die primäre Bilingualität entwickelt sich spontan in der Umgebung, in der die Sprache gesprochen wird, während der sekundäre Bilingualismus das Ergebnis des Lernens der Sprache im Herkunftsland ist;
- ***bilinguismo individuale/isolato vs bilingualismo collettivo*** (individuelle/isolierte Bilingualität): die kollektive Bilingualität lässt sich in einen bilingualen Kontext einbetten, während die individuelle/isolierte Bilingualität bezeichnet den gegenteiligen Fall;
- ***bilinguismo additivo vs bilingualismo sottrattivo*** (additive VS subtraktive Bilingualität): jeweils Bilingualität, die zu einer Bereicherung der eigenen Sprachkenntnisse führt, bzw. Bilingualität, die zu einer Verschlechterung der eigenen L1 führt;

- ***bilinguismo strumentale vs bilinguismo integrativo*** (instrumentelle Bilingualität VS integrative Bilingualität): jeweils wenn die Lernmotivation utilitaristischen Zwecken dient und wenn sie der Notwendigkeit dient, mit Muttersprachlern zu kommunizieren.

Milazzo zufolge definieren diese Begriffe keine absoluten Kategorien, sondern dienen lediglich der allgemeinen Beschreibung der betreffenden Art von Bilingualität. Zwischen ihnen wird ein Kontinuum angenommen. Dies ist einer der seltenen Fälle in der Literatur, in denen sich die Autoren nicht kategorisch der Bilingualität und der allgemeinen Beschreibung des Sprachverhaltens nähern, sondern eine Terminologie vorschlagen, die trotz ihrer Allgemeinheit eine gewisse Flexibilität und Anpassungsfähigkeit ermöglicht.

1.6.2.2 *Lingua seconda VS lingua straniera VS lingua etnica*

Eine im Vergleich zum Englischen und Deutschen besonders scharfe terminologische Abgrenzung findet sich in der italienischen Glottodidaktik bei Balboni (2014). Während in anderen Traditionen die Begriffe oft synonym oder fließend gebraucht werden, wird hier eine strikte Unterscheidung anhand des Lernkontexts vorgenommen:

- ***lingua seconda (L2)***: Bezeichnet das Italienische, das *in Italien* (z. B. von Migrant:innen oder Studierenden) in einem immersiven Umfeld erworben wird, in dem die Sprache für die alltägliche Interaktion notwendig ist.
- ***lingua straniera (LS)***: Bezeichnet das Italienische, das *im Ausland* (z. B. in Schulen) erlernt wird, wo das italienophone Umfeld fehlt und der Input auf den Unterricht beschränkt bleibt.
- ***lingua etnica***: Beschreibt – analog zur *heritage language* – das Italienische in Gemeinschaften italienischer Abstammung, wobei hier oft ein Kompetenzabbau über Generationen hinweg beobachtet wird.

Diese kontextuelle Präzision erstreckt sich auch auf die Unterscheidung zwischen gesellschaftlicher und individueller Ebene. In Übereinstimmung mit Favata (2020, S. 277) und Balboni (2014) wird im Italienischen konsensfähig zwischen multilinguismo (Vielzahl von Sprachen in einem Gebiet/soziales Phänomen) und plurilinguismo (Präsenz mehrerer Sprachen im individuellen Repertoire) differenziert. Diese Unterscheidung deckt sich inhaltlich mit der im deutschen Abschnitt diskutierten Differenzierung (vgl. 1.5.2.1) und wird für diese Arbeit übernommen.

1.7 Synoptische Betrachtung: Terminologische Konvergenzen und Divergenzen

Die einzelsprachliche Analyse hat gezeigt, dass die Terminologien im Englischen, Deutschen und Italienischen zwar auf denselben Basiskonzepten fußen, jedoch in ihrer Nuancierung unterschiedliche Schwerpunkte setzen. Der hier vorgenommene Sprachvergleich bietet für die

vorliegende Arbeit einen entscheidenden Mehrwert, da er es ermöglicht, die jeweils präzisesten Konzepte für die Untersuchung zu operationalisieren:

1. Dominanz und Kritik (Englisch): Der englische Diskurs liefert die globalen Basiskategorien (*native speaker*, *L2*), dient aber gleichzeitig als warnendes Beispiel für historisch gewachsene, teils ideologisch aufgeladene Konnotationen, die kritisch reflektiert werden müssen.
2. Systemische Schärfe (Deutsch): Die deutsche Terminologie erweist sich insbesondere in der Unterscheidung von *Prozess* und *Zustand* als wegweisend. Die Differenzierung zwischen *Bilingualismus* (gesellschaftlich) und *Bilingualität* (individuell) bietet ein analytisches Werkzeug, das im Englischen oft fehlt.
3. Kontextsensibilität (Italienisch): Die italienische Tradition steuert mit der strikten Trennung von *lingua seconda* (Immersionskontext) und *lingua straniera* (Fremdsprachenkontext) sowie der Unterscheidung *plurilinguismo/multilinguismo* eine notwendige Kategorie bei, um Lernumgebungen präzise zu erfassen.

Zum Abschluss des terminologischen Überblicks wird in der vorliegenden Dissertation das chronologische Kriterium als primärer Verankerungspunkt für die Verwendung der Bezeichnungen "L1" bzw. "LX" zugrunde gelegt, welcher jedoch durch kontextuelle Faktoren – wie die Dominanz – ergänzt wird. Sofern erforderlich, wird zudem die Bezeichnung "L2" verwendet, um die chronologisch zweit erworbene Sprache zu kennzeichnen, sofern diese eindeutig von einer simultan oder sukzessiv erworbenen weiteren L1 abgrenzbar ist.

2. Die phonologische Dimension des LX-Erwerbs: Sprachkontakt- und Interferenzphänomene

2.1 Die Organisation von phonologischen Systemen im multilingualen Gehirn

Die mehrsprachige Kompetenz, verstanden als die aktive und passive Fähigkeit der Sprecher:innen mit mehr als einer Sprache in ihrem Repertoire umzugehen, ist dynamisch und hängt von einer Vielzahl sprachlicher und außersprachlicher Faktoren ab. Zunächst ist es wichtig zu verstehen, wie Sprachen im menschlichen Gehirn organisiert sind und miteinander interagieren. Dazu wurden verschiedene Modelle entwickelt. In Anlehnung an Cook (2002) sowie Francis (1999), wie von Escudero (2009) dargelegt, lassen sich drei grundlegende Hypothesen zur Verteilung sprachlicher Systeme im menschlichen Gehirn und folglich sprachlicher Repräsentationen unterscheiden: 1) Hypothese der getrennten Systeme (*separate systems*), 2) Hypothese der gemischten Systeme (*mixed view*) und 3) Hypothese der verbundenen Systeme (*connected view*).

Die Hypothese der getrennten Systeme (*separate systems*): Diese Hypothese besagt, dass Kategorien (u.a. phonologische) in den jeweiligen unabhängigen Sprachsystemen verteilt sind. Sie wurde durch eine Reihe von Belegen gestützt, die unter anderem aus neurolinguistischen Untersuchungen mit Patient:innen stammen, die unter Sprachstörungen leiden (Paradis et al., 1982; Fabbro, 2001). Es wird insbesondere angenommen, dass bilingual aufwachsende Personen mit Übersetzungsstörungen je nach Richtung der Übersetzung (Sprache A → Sprache B, Sprache B → Sprache A⁶), auf unterschiedliche, voneinander unabhängige Kanäle zurückgreifen. Im Bereich der Phonologie wird die Hypothese durch das *Second Language Linguistic Perception Model* (L2LP) von Escudero (2009) gestützt, nach dem die menschliche Wahrnehmung von Sprachlauten spezifisch für jede Sprache ist, so dass separate Wahrnehmungssysteme existieren, die für die Wahrnehmung unterschiedlicher Lautsysteme zuständig sind.

Die Hypothese der gemischten Systeme (*mixed view*): Diese Hypothese besagt, dass die Sprachsysteme innerhalb eines einzigen kognitiven Systems repräsentiert sind. Innerhalb des gemeinsamen Systems können zwei Szenarien unterschieden werden: 1) Es erfolgt keine Unterscheidung zwischen den repräsentierten Sprachen, sodass die Perspektive als *merged* bezeichnet werden kann; 2) Die Sprachen werden innerhalb des gemeinsamen Systems unterschiedlich "etikettiert". Diese Perspektive wird als *integrated* bezeichnet und entspricht der

⁶ Hier werden die beteiligten Sprachen mit Buchstaben statt mit Zahlen bezeichnet (z.B. L1, L2 usw.), da die Beziehungen nicht chronologisch, sondern funktional sind (vgl. Fabbro 2001, S.213).

von Escudero (2009) in Anlehnung an Cook (2002) formulierten Definition. Im Bereich der Phonologie kommt das *Speech Learning Model* (SLM) von Flege (1995) dieser Hypothese nahe (vgl. Escudero 2009, S. 26), insbesondere mit Bezug auf die vierte Aussage des SLM. Diese besagt, dass Bilinguale danach streben, Kontraste zwischen den phonetischen Kategorien der L1 und L2 aufrechtzuerhalten, da diese in einem gemeinsamen phonologischen Raum existieren (vgl. Flege, 1995, S. 239).

Diese Perspektive steht auch im Einklang mit den Prinzipien der *Dynamic Systems Theory* (vgl. Verspoor et al., 2008).

Die Hypothese der verbundenen Systeme (*connected view*): Diese Auffassung besagt, dass die Sprachsysteme weitgehend unabhängigen Systemen angehören, wobei jedoch nicht ausgeschlossen werden kann, dass sie bestimmte Merkmale oder Elemente teilen. (vgl. Escudero 2009, S. 26).

Bis heute konnte keine dieser Hypothesen vollständig bestätigt werden. Stattdessen liefern die in verschiedenen Zweigen und Analyseebenen der Linguistik entwickelten Modelle zunehmend differenzierte Evidenz, die insbesondere die ersten beiden Hypothesen tendenziell stützt. Allerdings lässt sich in jüngeren Ansätzen eine Tendenz zur Kombination und Integration von Modellen erkennen. Das Modell von Escudero (2009), die jüngste Überarbeitung des Modells von Flege (nun als *Speech Learning Model-Revised* bezeichnet, SLM-r, vgl. Flege et al., 2021) sowie die Studien von Gut (2010), Kopečková et al. (2016) und Gut et al. (2023) demonstrieren die Relevanz, mehr Aufmerksamkeit auf nicht strikt sprachliche Variablen zu richten, die sich aus dem Vergleich konkurrierender Systeme ergeben. Vielmehr sollte der soziale Kontext, in dem das Lernen stattfindet, stärker berücksichtigt werden. Im Gegensatz zu den traditionellen Ansätzen, die vor allem auf die Analyse von Daten auf Populationsebene ausgerichtet sind, rücken neuere Ansätze die individuelle Variation stärker in den Vordergrund. Der Fokus der jüngeren Ansätze liegt auf dem Sprachlernen einzelner Individuen oder kleiner, möglichst homogener Gruppen sowie auf den externen Faktoren, welche die Qualität dieses Lernens maßgeblich beeinflussen und somit auch den Endzustand determinieren. Um den Anforderungen dieses Perspektivwechsels gerecht zu werden, hat sich in den letzten zwei Jahrzehnten zunehmend die *Complex Dynamic Systems Theory of Language Development* (CDST) etabliert. Hierbei handelt es sich um ein Modell für den Erwerb von L2/L3 (vgl. Gut et al. 2023) bzw. LX, das auf einem ursprünglich aus der Mathematik entlehnten Modell, der *Dynamic Systems Theory* (DST, vgl. de Bot et al., 2007), basiert. Das Kernkonzept dieses Modells besagt, dass Sprachen und linguistische Analyseebenen ein gemeinsames Netzwerk bilden, das zeitlichen Schwankungen unterliegt. Diese Variabilität wird äußerst von jeglichen Veränderungen interner und externer Faktoren beeinflusst, was die Vorhersagbarkeit der Entwicklung solcher Systeme im Laufe der Zeit erheblich erschwert (vgl. Kopečková et al. 2016, S. 429 ; de Bot, 2012).

Ein hoher Fokus auf die Variabilität in der Repräsentation linguistischer Systeme und ihrer Elemente zeigt sich auch in einigen Forschungsergebnissen, die mithilfe von Neuroimaging-Techniken gewonnen wurden. Diese legen nahe, dass die zerebrale Repräsentation von Sprachen (getrennt oder gemeinsam) nicht als absolut betrachtet werden sollte, sondern vielmehr als von bestimmten Variablen beeinflusst. Dazu zählt das chronologische Alter, in dem die L2 erlernt wurde. Bloch et al. (2009) demonstrieren, dass das Alter des Erwerbs der L2 (in Bloch et al., 2009 als "L2" bezeichnet⁷) die Art und Weise beeinflusst, wie diese Sprache im menschlichen Gehirn repräsentiert wird. Je früher die L2 erlernt wird, desto eher ähnelt ihre sprachliche Repräsentation derjenigen der L1 und ist gemeinsam mit ihr aktiv. Wird die L2 hingegen zu einem späteren Zeitpunkt erlernt, unterliegt ihre neuronale Repräsentation einer größeren Variabilität (vgl. Bloch et al. 2009). In der Studie von Perani et al. (1996) werden widersprüchliche Ergebnisse präsentiert. Im Gegensatz zu Bloch (2009) konnte eine Überlappung der aktivierten Bereiche sowohl bei frühen als auch bei späten Bilingualen festgestellt werden. Laut Perani et al. (1996) ist hingegen die Sprachflüssigkeit (*fluency*) maßgeblich dafür verantwortlich, ob gemeinsame Gehirnareale aktiviert werden oder nicht. In einer vergleichbaren Studie gelangten Videsott et al. (2010) zu ähnlichen Resultaten. Diese Studie basiert auf einer fMRI-Untersuchung, in deren Rahmen eine Bildbenennungsaufgabe mit Sprecher:innen durchgeführt wurde, deren L1 Ladinisch war und die fortgeschrittenen Kenntnisse der italienischen sowie mittlere Kenntnisse der englischen Sprache aufwiesen. Es konnte eine Aktivierung gemeinsamer Areale, insbesondere der rechten präfrontalen Kortex, beobachtet werden. Dies galt für die beiden Sprachen, in denen die Sprecher:innen in Bezug auf die Genauigkeit bei der Benennung der im Task präsentierten Bilder am besonders fließend und kompetent waren (Ladinisch und Italienisch). Die dargestellte Übersicht verdeutlicht die nach wie vor bestehende Unsicherheit darüber, welche Variablen die Art und Weise der sprachlichen Repräsentation im multilingualen Gehirn am stärksten beeinflussen.

2.2 Wahrnehmung und Produktion von LX-Lautkategorien: Mechanismen und Modelle zur Klärung der *Cross-Linguistic Influence* (CLI)

2.2.1 *Cross-Linguistic Influence* (CLI) im LX-Erwerb: Natur und Mechanismen

Mit *Cross Linguistic Influence* (CLI) wird ein komplexes Bündel von Transferphänomenen bezeichnet, deren systematische Erforschung auf Vildomec (1963) zurückgeht. Sie beschreiben das Phänomen im Rahmen des *language transfer* als die Nutzung von Wissen aus der L1 (oder einer anderen Sprache) beim Erwerb einer LX (vgl. Selinker & Gass, 1983, S. 372). Odlin (1989, S.

⁷ Es sei darauf verwiesen, dass die überwiegende Mehrheit der Aussagen über Erwerbs- und Verarbeitungsmechanismen beim Erlernen einer "dritten Sprache" (im weiteren Sinne in der Literatur so bezeichnet) auf Studien an bilingualen Proband:innen basiert. Derzeit existiert lediglich eine begrenzte Anzahl von Studien, die mit Neuroimaging-Techniken an multilingualen Proband:innen durchgeführt wurden. Folglich bleibt das Forschungsfeld weitgehend unerforscht (vgl. z.B. das Review von van den Noort et al., 2013).

27) präzisiert dies als einen Einfluss, der aus Ähnlichkeiten und Unterschieden zwischen der Zielsprache und jeder anderen, zuvor (wenn auch unvollkommen) erworbenen Sprache resultiert. Eine noch spezifischere Definition liefert Sharwood-Smith (1994, S. 198), der CLI nicht nur als Einfluss der L1 auf die Performanz oder Entwicklung der Zielsprache versteht, sondern diesen Einfluss explizit auf jede weitere dem Lerner bekannte Sprache (*other tongue*) ausdehnt.

In all diesen Definitionen finden sich die Kernelemente des modernen CLI-Begriffs wieder: das Konzept des Transfers, seine Multidirektionalität (vgl. z.B. Kopečková et al., 2016; Gut et al., 2023) sowie die fundamentale Rolle der Ähnlichkeiten und Unterschiede zwischen den Sprachsystemen. Diese Konstellation erklärt, weshalb die Erforschung mehrsprachiger Interaktionen ursprünglich im Rahmen der *Second Language Acquisition Studies* (SLA) verortet war. Zweifellos gibt es Gründe für diese Sichtweise, zugleich aber auch wesentliche Unterschiede, deren systematische Beobachtung zur Herausbildung des noch jungen Forschungszweigs der *Third or Additional Language Acquisition* (TLA) beigetragen hat (de Angelis, 2007a). Wie de Angelis (2007b) betont, sind sich die meisten Wissenschaftler:innen einig, dass eine allgemeine Theorie des fremdsprachlichen Erwerbs nicht allein auf dem Verhalten in der L2 basieren kann, da dies die komplexen Interaktionen eines multilingualen Repertoires unzureichend abbildet.

Dabei ist zu beachten, dass sich die Forschung lange Zeit auf lexikalisch-semantische und morphosyntaktische Ebenen konzentrierte, während die phonologische Ebene erst in jüngerer Zeit verstärkt Beachtung fand. Überträgt man die von Ringbom (2002) für das Lexikon definierten Ebenen auf die Phonologie, lässt sich CLI hier differenziert betrachten: 1) als Einfluss, der auf der wahrgenommenen Ähnlichkeit der phonologischen Systeme basiert; 2) als Transfer auf segmentaler Ebene und 3) als Transfer phonologischer Regeln. Abhängig von der wahrgenommenen Nähe oder Distanz zwischen den Systemen (Psychotypologie) lässt sich dieser Transfer in zwei Typen unterteilen: 1) positiver Transfer, der als erleichternde Strategie wirkt, wenn Strukturen als ähnlich erkannt und wiederverwendet werden, und 2) negativer Transfer (Interferenz), der oft als *Coping*-Strategie auftritt, wenn das dominante System bei der Verarbeitung unbekannter Strukturen der Zielsprache "einspringt".

Die Untersuchung dieser Faktoren in einem mehrsprachigen Kontext ermöglicht Einblicke in Mechanismen, die in einem auf das Duo L1-L2 beschränkten Kontext verborgen bleiben würden (vgl. de Angelis, 2007b). Insbesondere im Bereich der Phonologie stützt die Forschung die Annahme, dass das Lautsystem eines Lernalters zu jedem Zeitpunkt das Resultat seiner gesamten sprachlichen Verarbeitungserfahrung ist („full history“, Leather, 2003, S. 34).

2.2.2. Verhältnis zwischen Perzeption und Produktion im LX-Erwerb

Das CLI betrifft nicht nur jede Ebene der sprachlichen Analyse, sondern wirkt sich auch auf alle Aspekte des Spracherwerbs aus: Wahrnehmung, Verarbeitung und Sprachproduktion. Die Annahme eines Zusammenhangs zwischen der Wahrnehmungs- und der Produktionsebene (der sogenannte *Perception-Production Link*, vgl. Isbell, 2016) ist nicht nur intuitiv, sondern wird auch durch theoretische Grundlagen und zahlreiche empirische Belege gestützt. Im Bereich der Phonologie ist die Verbindung zwischen diesen beiden Dimensionen seit langem bekannt und lässt sich auf Polivanovs Erstveröffentlichung aus dem Jahr 1931 zurückführen, die 1969 in einer Neuauflage erneut publiziert wurde (vgl. Polivanov, 1969; Cardoso, 2011). Vor diesem Hintergrund lassen sich die folgenden Definitionen von Wahrnehmung und Produktion formulieren, die direkt auf die von Isbell (2016) bereitgestellten Definitionen zurückgreifen:

Perzeption: In Anlehnung an Isbell (2016, S. 57) umfasst die Wahrnehmung das Erkennen eines Lautes als zugehörig zu einer phonologischen Kategorie während der Verarbeitung des auditiven Inputs. Voraussetzung für eine korrekte Erkennung ist dabei die Existenz einer entsprechenden Kategorie im zugrunde liegenden phonologischen System des Lernalters.

Produktion: Diese bezeichnet die Nutzung des oral-artikulatorischen Systems zur Aussendung eines auditiven Signals, das eine linguistisch (phonologisch) kodierte Botschaft trägt (vgl. ibidem).

Auf theoretischer Ebene unterstützen das SLM (Flege 1995), dessen überarbeitete Version (SLM-r, Flege et al., 2021), sowie das *L2 Perception Model* (vgl. Escudero, 2009) und das PAM-L2 (C. Best & Tyler, 2007) die Annahme eines solchen Abhängigkeitsverhältnisses, indem sie betonen, dass die Realisierung von Zielphonemen stark von ihrer Wahrnehmung abhängt.

Es gibt zahlreiche empirische Belege für die Verbindung zwischen Wahrnehmung und Produktion, nach denen die Wahrnehmung zudem der Produktion vorausgeht. Im Kapitel zur Überarbeitung des SLM-r (Flege et al., 2021) heben die Autoren die Existenz einer starken bidirektionalen Verbindung zwischen Wahrnehmung und Produktion hervor (cfr. Flege et al. 2021, S. 29), die nicht nur durch neurologische Beweise gestützt wird, sondern auch als ausschließlich menschlich angesehen wird (vgl. Flege et al. 2021).

Flege et al. (1995) berichten, dass italienische L1-Sprecher:innen, die in den USA leben, die Phoneme /ð/ und /θ/ in Wortanfangsposition verzerrt im Vergleich zu den englischen L1-Realisationen aussprechen, und zwar in direktem Verhältnis zu ihrem Alter beim Spracherwerb (AOL, vgl. Flege 1995). Das von Flege verwendete Akronym AOL (*Age of Learning*) wird an dieser Stelle bewusst übernommen, da es – im Einklang mit der in Kapitel 1 (Abschnitt 1.6.2.2) erörterten Unterscheidung von Lernkontexten – den gesteuerten Lernkontext

im Erwachsenenalter präzise vom natürlichen *Age of Acquisition* (AoA) abgrenzt (vgl. Flege, 1995, S. 254–256).

In Jia et al. (2006) wurde festgestellt, dass das Maß an Genauigkeit in der Wahrnehmung der Zielphoneme mit dem Maß an Genauigkeit in deren Produktion korreliert, was die Beobachtung von Flege et al. (1999) bestätigt. Einem in Isbell (2016) erwähnten Studie von Baker & Trofimovich (2006) zufolge, die den Erwerb des Englischen durch koreanische L1 Sprecher:innen untersucht, ist die Verbindung zwischen Wahrnehmung und Produktion notwendig: Insbesondere wäre die Wahrnehmung eine notwendige Bedingung für die Produktion (Isbell 2016, S. 58).

Trotz dieser Belege, die nur eine Auswahl all jener darstellen, die zugunsten des Wahrnehmungs-Produktions-Zusammenhangs angeführt werden könnten, gibt es zahlreiche Gründe, diese Annahme infrage zu stellen. Diese lassen sich in zwei mögliche (da belegte) Szenarien zusammenfassen: 1) Wahrnehmung ohne Produktion und 2) Produktion ohne Wahrnehmung.

Perzeption ohne Produktion: Diese Fallkonstellation entspricht einem durchaus häufig zu beobachtenden Szenario, in dem die Entwicklung sprachlicher Kompetenzen in einer LX durch eine Diskrepanz zwischen Wahrnehmung und Produktion gekennzeichnet ist. Im Bereich der Phonologie wird die Asymmetrie in der Entwicklung phonologischer – perzeptiver und produktiver – Kompetenzen beim Erlernen einer LX von Neufeld als *phonological asymmetry* bezeichnet (vgl. Neufeld, 1988). Belege dafür finden sich bereits im zuvor erwähnten Kapitel von Flege (1995). Der Autor verweist, unter Bezugnahme auf eine frühere Studie (Flege & Port, 1981), darauf, dass zahlreiche arabische L1-Sprecher:innen (Standardarabisch), die seit Jahren in den USA lebten, das Phonem /p/ nicht als solches erkennen konnten, sondern es als /b/ wahrnahmen. Dies war vor allem auf eine Realisierung des stimmlosen Plosivs /p/ mit einer relativ langen Stimmhaftigkeitsverzögerung (VOT) zurückzuführen, die von Sprecher:innen/Hörer:innen den vertrauteren VOT-Werten des Phonems /b/ im arabischen Konsonantensystem zugeordnet wurde.

Kim & Park (1995) stellten eine Asymmetrie in der Wahrnehmung und Produktion der englischen Liquidkonsonanten /l-r/ bei koreanischen Lerner:innen mit Englisch als L1 fest: Auch hier war die Produktion dieser Phoneme deutlich abweichender als die Wahrnehmung, die mit der von koreanischen L1-Sprecher:innen vergleichbar war. Ein ähnliches Muster belegen (Chan, 2014) für das Erlernen des englischen Konsonanten /ð/ bei Lerner:innen mit Kantonesisch als L1 und die Studie von (Derwing & Munro, 2015) für das Erlernen des Konsonantenkontrasts /l/-/r/ bei spanischen Lerner:innen mit Englisch als L1.

Obwohl die Evidenzen eine direkte Verbindung zwischen der Wahrnehmungs- und der Produktionsebene infrage stellen, sollte laut Neufeld (in Leather 2003, S. 27) die *phonological asymmetry* nicht unbedingt als permanenter Zustand verstanden werden. Vielmehr könnte sie als ein Status quo gesehen werden, der sich ändern kann und insbesondere durch bestimmte Umstände und (mehr oder weniger bewusste) Entscheidungen der Sprecher:innen motiviert ist. Beispielsweise kann sie mit den spezifischen kommunikativen Zielen der Sprecher:innen zusammenhängen, die ihn dazu veranlassen, bestimmte Aspekte der sprachlichen Produktion zu "depriorisieren", die für das Verständnis der Botschaft nicht entscheidend sind. Zudem besteht die Möglichkeit, dass dieses Ungleichgewicht auf die angeborenen Eigenschaften der Sprecher:innen zurückzuführen ist, die ihn als fähiger in der Wahrnehmung von Fehlern als in der korrekten Produktion bestimmter sprachlicher Einheiten kennzeichnen könnten (Neufeld 1988, 1997). Ebenso bedeutsam ist die Qualität des L1-Erwerbs selbst, wie Flege et al. (2021) mit der *Category Precision Hypothesis* betonen:

[...] the more precisely defined L1 categories are at the time of first exposure to an L2, the more readily the phonetic difference between an L1 sound and the closest L2 sound will be discerned and a new phonetic category formed for the L2 sound. [...] (Flege et al. 2021, S. 36)

Produktion ohne Perzeption: Zwischen den beiden Szenarien ist dies sicherlich der weniger intuitive Fall und erscheint in der Literatur auch nicht als häufiges Phänomen. Der einzige häufig zitierte Fall ist der von (Sheldon & Strange, 1982), der seinerseits auf einer früheren Studie von Brière (1966) basiert. Letztere stellt den ersten Fall in der Literatur dar, in dem ein Ungleichgewicht zwischen den beiden Dimensionen zugunsten der Produktion belegt ist. In der Studie von Sheldon und Strange, die die Wahrnehmung und Produktion des Kontrasts /l/-/r/ bei sechs erwachsenen japanischen Lerner:innen des amerikanischen Englisch untersuchte, beschränkte sich die Wahrnehmungsschwierigkeit der Lerner:innen auf die Wahrnehmung der beiden Phoneme in Konsonantenclustern, die aus prävokalischem Okklusiv + Liquid bestehen. In der Wortendposition hingegen stellte die Wahrnehmung kein Problem dar. Diese Uneinheitlichkeit führte zu der allgemeinen Annahme, dass für diese kleine Gruppe aus Lerner:innen und für diesen spezifischen Kontrast die Produktionsleistung der Wahrnehmungsleistung überlegen war. Die Ergebnisse dieser Studie stimmen daher mit den Feststellungen von Brière (1966) überein; ihre Interpretation wird jedoch weiterhin diskutiert (siehe beispielsweise in Isbell, 2016).

Ein wesentlicher Grund für divergierende Ergebnisse zum Wahrnehmungs-Produktions-Zusammenhang liegt in methodologischen Aspekten. Im phonologischen Bereich bezieht sich dies auf die Schwierigkeit, vergleichbare Messungen in Wahrnehmungs- und Produktionsstudien zu finden, die Wahl der Aufgabenstellung sowie die Stichprobengröße (vgl. Isbell 2016, S. 63), die oft keine großen Zahlen erreicht und daher die

Formulierung von Verallgemeinerungen erschwert. Beispielsweise umfasst die bekannte Studie von Sheldon & Strange (1982) eine Stichprobe von weniger als zehn Teilnehmer:innen. Im Gegensatz dazu weisen die Studien von Flege et al. (1999) und Tsukada et al. (2005) jeweils 72 Teilnehmer:innen auf, während Jia et al. (2006) insgesamt 77 Teilnehmer:innen zählt. Die Bedeutung von Ergebnissen aus kleinen Stichproben kann aus verschiedenen Blickwinkeln bewertet werden, die sowohl positive als auch negative Aspekte aufweisen.

Flege et al. (2021) fassen die Sachlage zum Zusammenhang zwischen Wahrnehmung und Produktion folgendermaßen zusammen:

[...] Mitterer, Reinisch, and McQueen (2018) observed that from the stand-point of spoken word recognition, there is no need to assume that production and perception must be very similar. [...] Most importantly, the observation of significant positive production-perception correlations did not demonstrate causality. The correlations could just as easily be interpreted to mean that production accuracy places an upper limit on perceptual accuracy as the reverse (see, e.g., Best, 1995). [...] (Flege et al. 2021, S. 29-30)

2.2.3 Cross-Linguistic Influence (CLI) im LX-Erwerb: Relevante Einflussfaktoren

Während Studien im morphosyntaktischen Bereich bereits recht fortgeschritten sind, bleiben die phonologischen Aspekte des CLI weitgehend unerforscht. Um relevante Einflussfaktoren zu identifizieren, ist es daher notwendig, Beobachtungen auf unterschiedlichen Ebenen der Sprachanalyse heranzuziehen. Gemäß der *Complex Dynamic Systems Theory of Language Development* (CDST) gibt es im Wesentlichen zwei Arten von Einflussfaktoren auf den Spracherwerb (sei es L1 oder eine weitere Sprache): interne / systemische und externe / umweltbezogene (*environmental*) Faktoren. Ihre Wechselwirkung führt zu Veränderungen, die kontinuierlich, diskontinuierlich oder sogar chaotisch sein können (vgl. van Dijk & van Geert, 2007). Diese Veränderungen betreffen nicht nur ein, sondern potenziell mehrere linguistische Subsysteme gleichzeitig und beruhen auf schwer vorhersagbaren und komplexe Mechanismen. Die Vorstellung, dass das Lernen zu jedem Zeitpunkt extrem sensibel gegenüber Veränderungen ist, liegt der CDST zugrunde. Dem CDST-Ansatz zufolge kann bereits eine leichte Veränderung eines bestimmten Parameters (umweltbedingt oder systemisch) eine Veränderung im Lernprozess auslösen. Im Folgenden sollen die Hauptfaktoren, die CLI in Bereich der Phonologie auslösen, umrissen werden, ohne jedoch die einzelnen Faktoren als isoliert zu behandeln, sondern vielmehr in dem Bemühen, ihre gegenseitige Abhängigkeit und/oder Synergie darzustellen.

Das Erlernen einer LX, insbesondere wenn es sich nicht um die erste Fremdsprache handelt, wird wesentlich durch die im Laufe der Zeit gesammelten Vorerfahrungen geprägt. Unabhängig davon, ob diese auf der L1 oder einer früher erlernten Fremdsprache basieren, liegt stets ein

Erfahrungsfundament vor, das sowohl Lernstrategien als auch zuvor gemachte Lernerfahrungen umfasst. Dies umfasst nicht nur das Erlernen des Zielsprachlichen Systems, sondern auch die Lernerfahrungen in einem bestimmten Lernumfeld oder in mehreren Lernumgebungen. Dabei spielt die Art des Input, dessen Häufigkeit und Qualität eine wichtige Rolle; mit Qualität ist auch die *attitude* gemeint, mit der der Input vermittelt wurde und/oder mit der er aufgenommen und verarbeitet wird. Mit *attitude* ist hier die Einstellung (positiv oder negativ) gemeint, mit der die LX unterrichtet und gelernt wird, sowie die Gesamtheit der mit dem Lernen verbundenen Erfahrungen.

In dem Moment, in dem die Voraussetzungen für das Lernen gegeben sind (Wille, Notwendigkeit und ein geeigneter Ausgangskontext), stoßen mehrsprachige Lerner:innen auf eine der ersten Hürden: die Ähnlichkeit bzw. Unterschiedlichkeit zur Hintergrundsprache(n) (vgl. de Angelis 2007b). Dies ist zudem einer der Hauptfaktoren, die das CLI auslösen. Aus Studien (vgl. z. B. de Angelis 2007b) geht hervor, dass das CLI eher zwischen eng verwandten Sprachen auftritt. Die Sprachverwandtschaft wird in der Regel im genetischen Sinne verstanden, doch kann die enge Beziehung zwischen zwei Sprachen über rein genetische Aspekte hinausgehen. Dieser Faktor kann mit der Gemeinsamkeit bestimmter formaler Merkmale einhergehen oder nicht, oder es kann das Konzept der „wahrgenommenen sprachlichen Distanz“ (auf Englisch *perceived language distance*) ins Spiel kommen (vgl. de Angelis 2007b). Je enger die Verbindung ist, desto wahrscheinlicher ist das Auftreten von CLI im Sinne des *overall level* (Ringbom 2002, S. 2). Auf der Ebene, die Ringbom (ebenda) als *item level* bezeichnet, wirkt zudem das Phänomen der Prototypikalität (*prototypicality*), wie es von Kellerman (1983) theoretisiert wurde. Hierbei ist jedoch entscheidend, dass die Übertragung nicht linear verläuft: Während Lernende prototypische Bedeutungen tendenziell als übertragbar (*language-neutral*) wahrnehmen, stufen sie weniger prototypische oder idiomatische Bedeutungen als sprachspezifisch (*language-specific*) ein und sind daher vorsichtiger, diese in die LX zu übertragen. Kellerman weist jedoch nach, dass die tatsächliche Transferrate stark von der subjektiven Psychotypologie beeinflusst wird, was dazu führen kann, dass selbst prototypische Kernelemente bei wahrgenommener typologischer Distanz nicht übertragen werden.

Es ist nicht auszuschließen, dass eine Ähnlichkeit gleichzeitig mit einem Unterschied zwischen den konkurrierenden Sprachen einhergeht: Es handelt sich hierbei um relative Dimensionen, so dass zum Beispiel zwei oder mehr Sprachen genetisch voneinander entfernt sein können (und daher potenziell unähnlich sind), aber durch eine Ähnlichkeit verbunden sind, die auf der Wahrnehmung der Lerner:innen selbst beruht. In diesem Fall spricht man von *perceived similarity/proximity* und *perceived dissimilarity/distance*. Wenn CLI zwischen genetisch unterschiedlichen Sprachen stattfindet, ist es plausibel, dass Sprecher:innen das übertragene Material als ähnlich empfinden, ungeachtet der genetischen Distanz. Wenn die Sprachen hingegen (unabhängig davon, ob sie als ähnlich oder unterschiedlich wahrgenommen werden)

als entfernt empfunden werden, ist es wahrscheinlicher, dass das CLI behindert wird, was zu positiven oder negativen Effekten führen kann. Letztere manifestieren sich konkret in Lernschwierigkeiten. In diesem Zusammenhang führt die *Contrastive Analysis Hypothesis* (Lado, 1957) diese Schwierigkeiten auf strukturelle Unterschiede in den konkurrierenden Systemen zurück, die im Kopf der Sprecher:innen miteinander verglichen werden.

Auf phonetischer Ebene betreffen die Diskrepanzen, die aus dem Vergleich zwischen Sprachen hervorgehen, nicht nur die akustischen und artikulatorischen Merkmale der sprachlichen Kategorien, sondern auch die spezifischen Beziehungen, die zwischen diesen Kategorien innerhalb jeder Sprache bestehen. Ein unterschiedliches Gleichgewicht zwischen den *cues* in der Ausgangs- und der Zielsprache kann die Wahrnehmung der Zielsprache beeinflussen (vgl. (Bozena & Levy, 2014)). Aus diesem Vergleich ergibt sich gewöhnlich ein *mapping* im neuen Sprachsystem, das nicht nur von einem emergierenden System abhängt, welches das der L1 oder einer anderen, später erlernten Sprache sein kann (vgl. z.B. de Angelis 2007b, S.12), sondern auch von breiteren phonologischen Kriterien. Im Verlauf des *mapping*-Prozesses tritt das Phänomen der *selective attention* (selektive Aufmerksamkeit, Bozena und Levy 2014) in Erscheinung, die es ermöglicht, sich auf die akustischen und artikulatorischen Merkmale zu konzentrieren, die helfen könnten, möglichst viele Kontraste in der Zielsprache zu erkennen, auch wenn diese in der L1 nicht unbedingt relevant sind. Die selektive Aufmerksamkeit auf bestimmte phonetische Merkmale kann die Schwierigkeiten in der Wahrnehmung von Kontrasten erklären, die für die L1 keine Bedeutung haben, und somit das Verständnis der LX beeinflussen. Die Wahrnehmung der LX könnte jedoch verbessert werden, wenn nützliche Hinweise aus der L1 auch in der LX wiederverwendet werden können, auch wenn dies in anderen phonetischen Kontexten geschieht. Clements (2003) hat vorgeschlagen, dass Sprachen eine begrenzte Anzahl phonetischer Merkmale verwenden, um verschiedene Kontraste auszudrücken. Daher ist es für den Lerner:innen vorteilhafter, sich auf diese wiederverwendbaren Merkmale zu konzentrieren, anstatt auf L1-spezifische Kategorien.

Um die Komplexität der Wechselwirkungen zwischen dem genetischen Faktor und anderen Faktoren zu verstehen, werden hier drei mögliche Szenarien aufgeführt, in denen CLI auf der Grundlage der Ähnlichkeit zwischen konkurrierenden Sprachsystemen wahrscheinlich ausgelöst wird (vgl. de Angelis 2007b): 1) Sprecher:innen verfügen über Kenntnisse von Hintergrundsprachen, die genetisch ähnlich und unähnlich sind; 2) Sprecher:innen verfügen über Kenntnisse von Hintergrundsprachen derselben Familie, die jedoch nicht alle derselben Untergruppe angehören; 3) Sprecher:innen verfügen über Kenntnisse von Hintergrundsprachen, die derselben Familie und derselben Untergruppe angehören. In diesen drei Szenarien spielen somit Verhältnisse der sprachlichen Nähe und Distanz eine Rolle: Das Verständnis für das Überwiegen des einen oder des anderen Faktors und somit auch die Richtung des CLI kann nur durch die Analyse weiterer Faktoren gewonnen werden, wie beispielsweise das

Kompetenzniveau in jeder Sprache (z.B. *proficiency* in der L2, Ringbom, 1987; Ringbom, 2005; Möhle, 1989; Vogel, 1992; Ecke, 2001; Fouser, 2001; Cenoz, 2005; Gallardo del Puerto, 2007), die zeitliche Nähe der letzten Verwendung der Sprache (*recency of use*, vgl. z.B. Hammarberg et al., 1993; Hammarberg & Hammarberg, 2005; Hammarberg 2001) und ihr Status (*L2-status factor*, vgl., Dewaele, 1998; Kopečková et al., 2016; Williams & Hammarberg, 1998; Flynn et al., 2004), die Dauer des Aufenthalts an einem Ort, an dem eine bestimmte Sprache des Repertoires gesprochen wird (*Length Of Residence*, LOR), die Dauer der Exposition (*Length Of Exposure* oder auch *input quantity*), die Qualität des *Input* (*input quality*, das auch die Art des Lernkontextes einschließt), die Reihenfolge des Erwerbs, die Dominanz (siehe z.B. Puig-Mayenko et al., 2018) und, nicht zuletzt, die emotionale Bindung zu den erlernten Sprachen (*psychological attachment*, vgl. de Angelis 2007b; Hoffmann, 2001). Theoretisch wäre die Sprache, von der das CLI ausgehen sollte, diejenige mit den günstigsten Werten bei diesen Variablen: die höchste Kompetenz, eine sehr aktuelle Nutzung, die längste und qualitativ beste Expositionszeit sowie die positivste emotionale Bindung (z.B. eine Sprache, die mit angenehmen Erfahrungen und/oder besonders positiven Lernmotiven verbunden ist). Oftmals weist die im Repertoire vorhandene Sprache die meisten dieser Merkmale auf und ist die L1, die daher in der Regel zuerst untersucht wird. In einem realen Kontext zeigen diese Variablen jedoch nicht nur unterschiedliche Werte für alle Sprachen des Repertoires, sondern unterliegen auch zeitlichen Schwankungen. Der Zeitverlauf selbst, der oft mit einem Erosionsprozess der Sprachkenntnisse einhergeht, ist nicht immer ein Prädiktor für die Richtung des CLI. Ein Beispiel hierfür ist der von de Angelis (1999, in de Angelis 2007b) berichtete Fall, in dem der *main supplier* nicht die zuletzt erlernte Sprache war, sondern eine (fremde) Sprache, die über Jahrzehnte hinweg nicht verwendet wurde, zu der jedoch eine emotionale Bindung bestand (vgl. de Angelis 2007b).

Zusammenfassend lässt sich festhalten, dass die Phänomene, welche das CLI beeinflussen, durch eine enge Wechselwirkung vieler sprachinterner und -externer Faktoren gekennzeichnet sind. Dies erschwert insbesondere die präzise Vorhersage der Richtung und des Umfangs des CLI. Die in der Literatur untersuchten Szenarien und Hypothesen stellen jedoch eine wertvolle Grundlage für das Verständnis der Dynamik dar, die das Lernen und die Wahrnehmung in einem mehrsprachigen Kontext formt. Ein umfassendes Modell, welches sowohl strukturelle als auch psychologische Dimensionen des Sprachlernens berücksichtigt, kann dazu beitragen, die Komplexität des CLI zu erfassen und spezifische Transfermuster differenzierter zu erklären.

2.2.4 Theoretische Modelle des LX-Erwerbs

Die Forschung zum Erwerb einer Drittsprache greift auf eine Reihe von Modellen zurück, um die Mechanismen des *Cross-Linguistic Influence* (CLI) zu erklären. Um der spezifischen Fragestellung dieser Arbeit gerecht zu werden, ist eine Differenzierung notwendig, die auch die

in Abschnitt 2.2.1 diskutierte Entwicklung von der *Second Language Acquisition*- zur *Third or Additional Language Acquisition*-Forschung widerspiegelt:

Im ersten Teil (Abschnitt 2.2.4.1) stehen Modelle im Fokus, die ursprünglich im Kontext der *Second Language Acquisition* (SLA) für die phonologische Ebene entwickelt wurden, in der aktuellen L3-Forschung jedoch als zentraler theoretischer Ausgangspunkt dienen und auf den multilingualen Kontext übertragen werden. Dazu zählen das *Perceptual Assimilation Model* in seiner L2-Erweiterung (PAM-L2; Best & Tyler, 2007), das *Speech Learning Model* (SLM; Flege, 1995) bzw. seine Revision (SLM-r; Flege & Bohn, 2021) sowie das *Second Language Linguistic Perception Model* (L2LP; Escudero, 2005).

Im zweiten Teil (Abschnitt 2.2.4.2) werden hingegen übergreifende Modelle betrachtet, die spezifisch die Interaktion mehrerer Sprachen im Blick haben, wie das *Cumulative Enhancement Model* (CEM; Flynn et al., 2004), das *Typological Primacy Model* (TPM; Rothman, 2011), das *Linguistic Proximity Model* (LPM; Westergaard et al., 2017) oder die *L2-Status-Hypothese* (Hammarberg, 2001; Bardel & Falk, 2007, 2012). Auch wenn diese Ansätze primär für die Morphosyntax entwickelt wurden, sind sie für die vorliegende Arbeit von Relevanz, da sie die Debatte über die *Quelle* des Transfers (L1 vs. L2) theoretisch rahmen.

2.2.4.1 Spezifische Modelle aus den *Second Language Acquisition Studies*

Bevor die spezifischen Modelle im Detail betrachtet werden, ist eine Einschränkung vorzuschicken: Nicht alle der folgenden Ansätze modellieren beide Dimensionen – Wahrnehmung und Produktion – gleichermaßen explizit. Dennoch gibt es eine grundlegende Annahme, die als gemeinsamer Nenner dient: der *Perceptual Magnet Effect* (PME; Kuhl, 1991; Kuhl & Iverson, 1995). Dieser Effekt beschreibt die Filterung der eingehenden akustischen Information durch die Kategorien des bereits vorhandenen phonologischen Systems (z.B. der L1). Betrachtet man das phonologische System als ein eigenständiges “Ökosystem”, das durch sprachspezifische Regeln organisiert ist, wird verständlich, warum objektiv ähnliche Laute (wie der Schwa-Laut /ə/ in der Abbildung 1) trotz gleicher IPA-Transkription unterschiedlich wahrgenommen werden. Der PME erklärt, wie diese Nuancen durch den L1-Filter “assimiliert” werden. In der Abbildung 1 wird der Versuch unternommen, den PME schematisch darzustellen:

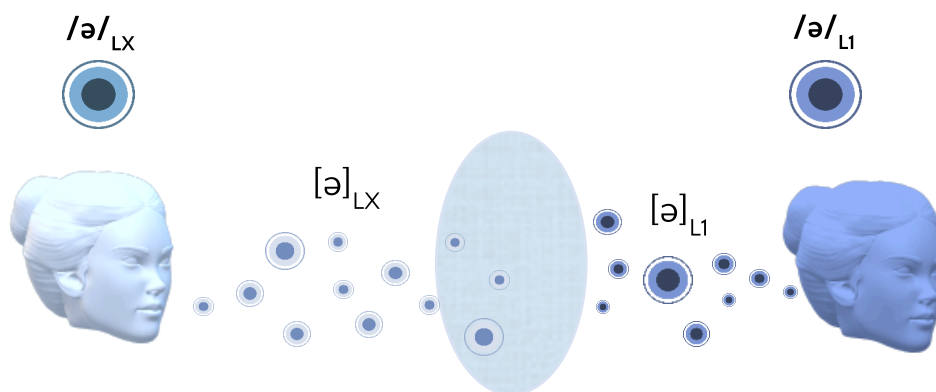


Abbildung 1: Schematische Darstellung des PME-Filterungsprozesses. Der Input (links), der von einer Person erzeugt wird und aus den einzelnen Realisierungen (auch *exemplars* genannt) des Phonems /ə/ der LX besteht, wird von einer hypothetischen zweiten Person durch den "Filter" ihrer eigenen L1 wahrgenommen. Eigene Abbildung.

Das *Perceptual Assimilation Model* in seiner L2-Erweiterung (PAM-L2; Best & Tyler, 2007) baut auf dieser Grundannahme auf, verzichtet jedoch auf abstrakte mentale Repräsentationen zugunsten einer direkten Wahrnehmung artikulatorischer Gesten. Best & Tyler skizzieren Szenarien der Assimilation, basierend auf der wahrgenommenen Ähnlichkeit zwischen L1- und LX-Lauten. Grundsätzlich werden drei Szenarien unterschieden: 1) Assimilation an eine L1-Kategorie (als "gutes", "akzeptables" oder "abweichendes" Exemplar); 2) Assimilation als nicht-kategorisierbarer Sprachlaut; 3) Nicht-Assimilation als Sprachlaut. Für die L3-Forschung besonders relevant ist die Vorhersagekraft bei Minimalpaaren: Werden beide Laute eines LX-Kontrasts als äquivalent zu einer *einzelnen* L1-Kategorie wahrgenommen (*Single Category Assimilation*), ist die Diskriminierung schwierig. Wird hingegen nur einer der Laute assimiliert (*Category-Goodness Difference*), ist die Unterscheidung oft leichter, da der abweichende Laut potenziell zur Bildung einer neuen Kategorie führen kann.

Das *Speech Learning Model* in seiner revidierten Form (SLM-r; Flege & Bohn, 2021) erweitert den Fokus auf die Interaktion von Wahrnehmung und Produktion über die gesamte Lebensspanne. Im Gegensatz zum PAM-L2 operiert das SLM-r mit phonetischen Kategorien als mentalen Repräsentationen im Langzeitgedächtnis. Eine zentrale Neuerung der Revision ist die Abkehr von der "Kritischen Periode": Die Fähigkeit, neue Kategorien zu bilden, bleibt lebenslang erhalten. Das Modell postuliert, dass die Bildung einer neuen Kategorie nur dann erfolgt, wenn Lerner:innen eine phonetische Distanz (Dissimilarität) zwischen dem LX-Laut und der nächstgelegenen L1-Kategorie wahrnehmen. Wird diese Distanz nicht erkannt, erfolgt eine Assimilation, die dauerhaft zu einer verzerrten Produktion führen kann. Das SLM-r betont zudem

die Rolle des Inputs und der *Category Precision*: Lerner:innen gewichten phonetische Merkmale (*feature weighting*) oft anders als L1-Sprecher:innen, was zu divergenten Kategorien führt.

Das *Second Language Linguistic Perception Model* (L2LP; Escudero, 2005; Escudero & Yazawa, 2024) hingegen integriert explizit lerntheoretische Ansätze (wie die Optimalitätstheorie) und formuliert klare Hypothesen für den Anfangszustand (*Initial State*). Basierend auf der *Full Copying Hypothesis*, wird angenommen, dass Lerner:innen zu Beginn eine vollständige Kopie ihres L1-Wahrnehmungssystems auf die LX übertragen. Lernen wird als Fehlerkorrektur verstanden: Wenn die L1-Kopie bei der Verarbeitung des LX-Inputs versagt (*mismatch*), wird das System angepasst. Escudero unterscheidet drei Lernszenarien (*New, Similar, Subset*), wobei der Erfolg davon abhängt, ob für die LX-Laute neue Kategorien geschaffen oder bestehende L1-Kategorien geteilt bzw. verschoben werden müssen. Dabei ist die Fehlerkorrektur (*error-driven learning*) zentral: Wenn die kopierte L1-Kategorie nicht mit dem LX-Input übereinstimmt (*mismatch*), wird das System angepasst, um eine neue Zielkategorie zu bilden (*New Scenario*). Ein Alleinstellungsmerkmal des L2LP ist die Annahme, dass Wahrnehmung und Produktion spiegelbildlich verknüpft sind.

2.2.4.2 Modelle aus den Third or Additional Language Acquisition Studies

Während die phonologischen Modelle primär die Mechanismen der Verarbeitung beschreiben, adressieren die hier diskutierten übergreifenden Modelle (CEM, TPM, L2-Status, LPM) eine zentrale Frage der Mehrsprachigkeitsforschung: Welche der bereits gelernten Sprachen (L1 oder L2) fungiert als primäre Quelle des Transfers für die weitere Sprache (die sogenannte L3 bzw. LX)?

Eine Position, die den Einfluss der L1 relativiert, vertritt das *Cumulative Enhancement Model* (CEM; Flynn et al., 2004). Es geht davon aus, dass Mehrsprachigkeit kumulativ wirkt: Jede zuvor gelernte Sprache kann den Erwerb der L3 bzw. LX positiv verstärken (*enhancement*), sofern die Strukturen kompatibel sind. Ein negativer Transfer wird theoretisch ausgeschlossen oder als irrelevant betrachtet. Die L1 hat hier keine privilegierte Rolle mehr; vielmehr scheint bei Erwachsenen das spezifische Wissen aus einer L2 oft leichter zugänglich zu sein.

Das *Typological Primacy Model* (TPM; Rothman, 2011) präzisiert diesen Gedankengang, indem es die Auswahl der Transferquelle nicht dem Zufall oder der Chronologie überlässt, sondern der Psychotypologie. Rothman argumentiert, dass der Initial State der L3 bzw. LX selektiv bestimmt wird: Das kognitive System wählt diejenige Sprache (L1 oder L2) als Basis, die als typologisch ähnlicher zur L3 bzw. LX wahrgenommen wird – unabhängig davon, ob diese Wahrnehmung objektiv korrekt ist. Diese "ökonomischste" Wahl bestimmt den Transfer von syntaktischen Eigenschaften en bloc.

Im Kontrast dazu betont die *L2 Status Factor Hypothesis* (Bardel & Falk, 2007, 2012) die Rolle des kognitiven Speichermodus. Basierend auf Paradis' (2004) Unterscheidung zwischen prozeduralem (L1) und deklarativem Gedächtnis (L2/L3 bzw. LX) argumentieren die Autorinnen, dass L2 und L3 bzw. LX kognitiv "verwandter" sind. Da beide Sprachen später gelernt und oft stärker über deklarative Mechanismen verarbeitet werden, neigen Lerner:innen dazu, bei der L3 bzw. LX-Produktion auf die L2 zurückzugreifen (*Foreign Language Effect*), da die L1 als "Nicht-Fremdsprache" unterdrückt wird.

Einen integrativeren Ansatz bietet schließlich das *Linguistic Proximity Model* (LPM; Westergaard et al., 2017). Es lehnt die Idee eines vollständigen Transfers einer einzigen Sprache (wie im TPM) ab und postuliert stattdessen ein eigenschaftsbasiertes Lernen (*property-by-property*). Transfer erfolgt lokal und selektiv immer dann, wenn eine spezifische linguistische Struktur in der L3 bzw. LX Ähnlichkeiten zu irgendeiner der vorherigen Sprachen aufweist. Im Gegensatz zum CEM erlaubt das LPM ausdrücklich auch negativen Transfer und erkennt an, dass sowohl L1 als auch L2 gleichzeitig Einfluss nehmen können (kombinierter CLI), abhängig von der strukturellen Nähe der jeweiligen Eigenschaften.

2.3 Wahrnehmung und Produktion von Vokalen im LX-Erlernen

2.3.1 Die *Consonant Vowel- Hypothesis*

Vokale, die den Schwerpunkt dieser Arbeit bilden, stellen ein universelles, aber hochgradig variables Merkmal menschlicher Sprachen dar. Während das durchschnittliche Inventar etwa sechs Vokale umfasst, reicht das Spektrum von minimalen Systemen wie im Yimas (Papua-Neuguinea) mit nur zwei Vokalen bis hin zu komplexen Systemen wie dem Deutschen mit 14 Monophthongen (vgl. Maddieson, 2013; Wiese, 2000). Nach Maddieson (2013) lassen sich Inventare nach ihrer Größe klassifizieren: Reduzierte Systeme umfassen 2 bis 4 Vokale, mittlere Systeme 5 bis 6 und große Systeme 7 bis 14. Dabei sind Sprachen mit mittleren Inventaren weltweit am häufigsten vertreten, während die extrem kleinen oder großen Systeme eine signifikante geografische Ungleichverteilung aufweisen.

Im Sprachfluss alternieren Vokale und Konsonanten, wobei sich diese beiden Klassen nicht nur durch ihre akustische Konfiguration unterscheiden – Vokale durch eine periodische Signalstruktur, Konsonanten primär durch aperiodische Rauschanteile –, sondern auch durch ihre funktionale Last. Diese Diskrepanz ist Gegenstand der kognitiven Linguistik und führte zur Formulierung der *Consonant-Vowel-Hypothesis* (CV-Hypothese). Sie postuliert eine funktionale Arbeitsteilung: Während Konsonanten primär lexikalische Informationen (die "Wurzel" der Bedeutung) tragen, vermitteln Vokale

vorrangig morphosyntaktische und prosodische Informationen (vgl. Boudelaa, 2015; Nespor et al., 2003).

Diese funktionale Unterscheidung ist für den Spracherwerb von zentraler Bedeutung. Die prosodischen und rhythmischen Eigenschaften der Vokale dienen Säuglingen als unverzichtbare Orientierungshilfe (*cues*), um das Silbeninventar zu segmentieren und syntaktische Grenzen – etwa die Relation zwischen Kopf und Komplement – zu identifizieren (vgl. Hochmann et al., 2011). Im Kontext des L1-Erwerbs wurde die CV-Hypothese dahingehend präzisiert, dass Kleinkinder ihre Aufmerksamkeit selektiv auf Vokale richten, um strukturelle Informationen aus dem Input zu extrahieren und grundlegende grammatische Beziehungen zu erlernen (vgl. Hochmann et al., 2011, S. 1446).

Vor diesem Hintergrund ist es entscheidend, die spezifische Rolle der Vokale auch für die Entwicklung des phonologischen Systems einer LX neu zu bewerten. Wenn Konsonanten das lexikalische Gerüst einer Sprache bilden, so sind es die Vokale, die maßgeblich zur phonologischen Feinabstimmung sowie zur Wahrnehmung von Prosodie und Akzent beitragen. Die Analyse, wie Lerner:innen Vokale in einer LX wahrnehmen und produzieren, gewährt daher tiefe Einblicke in die Mechanismen der phonologischen Verarbeitung und zeigt auf, wie suprasegmentale und strukturelle Informationen in frühen Lernphasen internalisiert werden.

2.3.2 Die Wahrnehmung und Produktion eines LX Vokalsystems: Forschungsstand und -richtungen

Nach Kuhl (1991) lässt sich beim Menschen eine andere Erwerbssequenz von Vokalen beobachten als bei Tieren, was die Wahrnehmung dieser Laute betrifft. Des Weiteren scheint das Erlernen von Vokalen sich in grundlegender Weise von dem der Konsonanten zu unterscheiden. Kuhl & Miller (1978) sowie Eimas et al. (1971) berichteten von bemerkenswerten Ergebnissen in prälinguistischen (Tests mit Säuglingen) bzw. nicht-linguistischen Kontexten (Tests mit Tieren), welche die Hypothese stützen, dass es sowohl für Menschen als auch für Tiere eine "natürliche Grenze" (Schwartz et al., 2005) zwischen den stimmhaften und stimmlosen Plosiven gibt. Im Falle der Vokale scheint es jedoch keine universellen "natürlichen Grenzen" zu geben, die eine eindeutige Unterscheidung der Vokal-Kategorien ermöglichen. Wie von Kuhl (1991) beschrieben, scheinen diese für Menschen wahrnehmbar zu sein, jedoch nicht für Affen. Gemäß Kuhl et al. (1992) entwickelt sich beim Menschen die Fähigkeit zur Unterscheidung von Vokal-Kategorien in einer Lernphase, die etwa sechs Monate nach der Geburt beginnt. In diesem Zeitraum erfolgt die Bildung der Prototypen, welche den Kern der finalen Kategorien im Erwachsenenalter bilden (vgl. Kuhl et al., 1992; Schwartz et al., 2005).

Einige Studien legen nahe, dass das Erlernen von Vokalen nicht für alle Vokallaute gleichmäßig erfolgt, sondern dass einige Vokale schneller und effizienter erlernt werden als andere. Diese Vokale fungieren als "Referenzpunkte" (vgl. Polka & Bohn, 2003) oder *perceptual magnets* (in Polka & Bohn 2003, vgl. auch Kuhl, 1991), wodurch der Erwerb des gesamten L1-Vokalsystems erleichtert wird. Wie von Polka und Bohn (2003) zitiert, die wiederum auf Medin & Barsalou (1987) verweisen, befinden sich diese Referenzpunkte an idealen Positionen in auffälligen Vokal-Dimensionen. Dadurch werden sie zu prototypischen Beispielen der Kategorie, zu der sie gehören (vgl. Polka & Bohn 2003, S. 221). Die höhere Salienz dieser Werte führt zu einem Phänomen, das als *perceptual asymmetry* bezeichnet wird. Dabei fällt die Unterscheidung zwischen Vokalen in einem Kontrast in eine Richtung leichter als in die andere. Diese Sichtweise unterscheidet sich von der symmetrischen, euklidischen Betrachtung des Vokalraums, die jedoch bereits seit den 1970er Jahren durch zahlreiche Studien gestützt wird (vgl. Rosch & Mervis, 1975; Tversky & Gati, 1978). In diesem Kontext sind die Ergebnisse von Polka & Werker (1994) sowie von Polka & Bohn (1996) von besonderem Interesse, da sie in ein plausibleres theoretisches Modell eingeordnet werden können. Die Ergebnisse beider Studien stützen die Hypothese der perzeptuellen Asymmetrie sowohl bei der Wahrnehmung von L1- als auch LX-Kontrasten. Im Jahr 1994 wurde jedoch noch die Hypothese aufgestellt, dass die Asymmetrie bei LX-Kontrasten stärker ausgeprägt sei als bei L1-Kontrasten. Die Ergebnisse der im Jahr 1996 durchgeführten Studie widerlegten jedoch diese Annahme. Im Rahmen der Studie wurden zwei Gruppen von Säuglingen getestet, die Englisch oder Deutsch als L1 erlernten. Der Test zur Unterscheidung der Kontraste /ɛ/-/æ/ (als "nicht-deutsch" definierter Kontrast) und /u/-/y/ (als "nicht-englisch" definierter Kontrast) wurde mit den beiden Gruppen durchgeführt, wobei die Säuglinge zwischen 6 und 8 Monaten bzw. 10 und 12 Monaten alt waren. Die Ergebnisse widersprachen den aufgestellten Hypothesen, indem sie eine einfachere Unterscheidung in den Richtungen /y/ → /u/ und /ɛ/ → /æ/ unabhängig von der sprachlichen Herkunft oder dem Alter der Säuglinge aufzeigten. Die Ergebnisse zeigten, dass die Vokale /u/ und /æ/, unabhängig davon, ob es sich um vertraute (L1) oder weniger vertraute Laute handelte, als Referenzvokale für beide Gruppen von Säuglingen fungierten. Dies widerlegt die sogenannte *Native Language Familiarity Hypothesis* und stellt sie in Frage (vgl. Polka und Bohn, 2003). Diese Ergebnisse lassen den Schluss zu, dass ein natürlicher, universeller Wahrnehmungsbias vorliegt. Auf der Grundlage dieser Phänomene war es möglich, die *Peripherality Hypothesis* (vgl. Polka & Bohn 2003) zu stützen, die postuliert, dass die markanteren Vokale diejenigen sind, die in peripheren Positionen im Vokalraum liegen. Periphere Vokale markieren tatsächlich die grundlegenden Grenzen des Vokalraums und weisen weniger akustische Überlappungen zwischen verschiedenen Sprecher:innen im Vergleich zu anderen Vokalen auf, wie in früheren Studien gezeigt wurde (vgl. Peterson & Barney, 1952). Diese letzte Studie zeigte, dass periphere Vokale in einem Identifikationstest weniger Verwechslungsraten aufwiesen als zentrale Vokale (vgl. Polka & Bohn 2003). Laut den Autor: innen entwickelt sich der Vokalraum durch eine schrittweise

Expansion: Von anfänglich zentraleren Realisierungen, die typisch für frühe kindliche Laute sind, bis hin zu einer ausgereiften Konfiguration, die durch deutlichere periphere Punkte gekennzeichnet ist, was mit den Angaben in Kuhl et al. (1992, vgl. Polka & Bohn, 2003) übereinstimmt.

Die Untersuchung dieses Phänomens wurde in einer späteren Revision von Polka und Bohn (2011) fortgesetzt, die darauf abzielte, die Anwendbarkeit des *Natural Referent Vowel Framework* (NRV) auf die vokalische Entwicklung von Säuglingen und Erwachsenen zu überprüfen, die Stimuli in einer L2 ausgesetzt sind. Im Wesentlichen würde diese Hypothese die größere wahrnehmungsspezifische Sensibilität für periphere Vokale erklären, die als Referenzpunkte für den Beginn des schrittweisen Erlernens des neuen vokalischen Systems dienen. Insbesondere die Ergebnisse einer Studie, die an Zuhörern von *Southern British English* (SBE) durchgeführt wurde, deren L1 Dänisch war, unterstützten die Hypothese, dass perceptuelle Asymmetrien sowohl bei Säuglingen als auch bei Erwachsenen vorhanden sind, auch in einer frühen Lern- oder Expositionsphase gegenüber einer L2, bei der noch keine bedeutende phonologische Sprachkenntnis vorausgesetzt werden kann. Die Studie analysierte die Wahrnehmung des Kontrasts /ɒ/-/ʌ/, der im SBE, aber nicht im Dänischen vorkommt. Der Kontrast stellt einen tiefen hinteren Vokal gegen einen zentralen hinteren Vokal gegenüber und verschiebt die Analyse in einen bisher unerforschten Bereich des Vokalraums. Die Zuhörer, die eine *Go/No-Go*-Diskriminierungsaufgabe erhielten, bei der die Stimuli in einem KVK-Kontext präsentiert wurden, zeigten tatsächlich eine Antwortgenauigkeit, die von der Reihenfolge der Stimuluspräsentation abhing: 50,5 % in der Richtung /ɒ/ → /ʌ/ und 69,2 % in die entgegengesetzte Richtung. Dies stützt die Hypothese, dass das wahrnehmungsspezifische Bias auch im Erwachsenenalter in Situationen existiert, in denen noch keine phonemische Wahrnehmung von Fremdlauten vorliegt (vgl. Polka & Bohn 2011, S. 469). Ähnliche Ergebnisse wurden in einer Stichprobe von englischen L1-Zuhörern erzielt, die den Kontrasten /u/-/y/ und /ʊ/-/y/ aus dem Deutschen ausgesetzt waren (Polka et al., 2005, vgl. Polka & Bohn, 2011): Die Engländer zeigten Wahrnehmungsasymmetrien in die erwartete Richtung (leichtere Diskriminierung in die Richtung /y/ → /u/ und /y/ → /ʊ/), während die Kontrollgruppe der Deutschen L1 keine Asymmetrie zeigte. Da es sich um eine Stichprobe von Erwachsenen handelte und nicht um Säuglinge wie in den Studien von Polka & Werker (1994) und Polka & Bohn (1996), war dies nicht überraschend: Bei den Zuhörern mit L1-Deutsch kann das Erlernen des vokalischen Systems als nahezu stabil betrachtet werden, daher sind L1-Stimuli in beide Richtungen gleich wahrnehmbar. In einem System, das noch als neu betrachtet werden kann, wie bei den Zuhörern mit L1 Englisch, ist die Asymmetrie deutlich und erwartet. Um den möglichen Effekt experimenteller Bedingungen zu testen, stellten Polka & Bohn (2011) fest, dass die *multiple-talker*-Bedingung (Stimuli, die von mehreren Sprecher:innen ausgesprochen wurden) keinen signifikanten Einfluss hatte. Hinsichtlich der Manipulation des *ISI* kann jedoch festgestellt werden, dass die Reduzierung auf 500 ms tatsächlich ein überraschendes Ergebnis

lieferte: Dadurch wurde das auditive Gedächtnis unterstützt, was die Belastung durch die phonologische Kodierung kompensierte (dies kann auf die Fremdheit des Stimulus im Vergleich zur *L1* zurückgeführt werden; vgl. Polka & Bohn, 2011). Dieses Ergebnis stützt die Hypothese, dass das wahrnehmungsspezifische Bias phonologisch ist und nicht von einer allgemeinen natürlichen akustischen Wahrnehmung abhängt. Aus diesen und weiteren Ergebnissen geht die allgemeine Plausibilität des *NRV*-Frameworks hervor, dessen Schlüsselpunkte in der folgenden Tabelle (Tabelle 1) zusammengefasst werden, die in Polka & Bohn (2011, S. 474 ff.) dargestellt sind, mit besonderem Fokus auf die Wahrnehmung durch Erwachsene, die im Mittelpunkt dieser Untersuchung stehen:

Tabelle 1: Übersichtliche schematische Darstellung der Grundpfeiler des *NRV*-Frameworks mit Fokus auf die Wahrnehmung bei Erwachsenen.

Grundpfeiler	Beschreibung
Natural Referent Vowels	• Periphere Vokale wie /i/, /a/, /u/ sind natürliche Referenzpunkte. • Sie ziehen die Aufmerksamkeit von Säuglingen auf sich und bieten Stabilität für das Lernen. • Eine universelle Prädisposition, die auf das Vorhandensein dieser Vokale in allen Sprachen basiert.
Rolle der sprachlichen Erfahrung	• Die sprachliche Erfahrung erweitert die anfängliche Prädisposition. Lerner:innen erstellen L1 vokalische Kategorien und optimieren die Wahrnehmung für die L1. • In reifen Hörer:innen: Direktionale Asymmetrien und Wahrnehmungsbiases nehmen für Nicht-L1-Vokale zu. • Direktionale Asymmetrien und Wahrnehmungsbiases verringern sich für L1-Vokale.
Überwindung und Wiederaufleben von Wahrnehmungs-Biases	• Reife Hörer:innen können die anfänglichen Biases für Vokale der L1 ignorieren.

	• Allerdings kann der Bias gegenüber natürlichen Vokalen in bestimmten Verarbeitungsaufgaben auch bei den L1-Vokalen wieder auftauchen.
Salienz und Stabilität der natürlichen Referenzvokale	• Die Salienz ergibt sich aus der Fokussierung der Formanten (Frequenzkonvergenz), wodurch die akustische Energie erhöht wird. • Maximale Fokussierung für Vokale wie [i], [a], [u] und [y]. • Vokal /y/: Wird zu einem fokalen Vokal aufgrund der Konvergenz von F2 und F3, auch wenn sie im traditionellen F1/F2-Raum nicht sichtbar ist.

Es ist von entscheidender Bedeutung, sich auf die Rolle des /y/ als fokalen Vokal zu fokussieren. Obwohl er nicht als peripher betrachtet wird (er nimmt nicht die Spitzenposition im Vokaltriangel ein), spielt dieser Vokal eine wesentliche Rolle innerhalb des NRV-Frameworks sowie darüber hinaus. Schwartz et al. (2005) sowie Polka & Bohn (2003) diskutieren die entscheidende Funktion des Vokals innerhalb der *Dispersion-Focalization Theory* (DFT), in der das Konzept der Fokalisierung (*focalization*), welches im NRV-Framework formuliert wurde, eine prominente Rolle einnimmt. In der DFT erfolgt die Beschreibung von Vokalen anhand der Werte der ersten vier Formanten (F1, F2, F3, F4). Dabei stellt F4 einen Standardwert für männliche Sprecher dar (3560 Hz) und die verbleibenden Formanten sind im maximal verfügbaren Vokalraum verteilt (vgl. Schwartz et al. 2005, S. 427). Diese Werte liefern eine dreidimensionale Beschreibung (F1-F2 und F3, wo relevant, vgl. Schwartz et al. 2005). Im Jahr 2005 wurde von Schwartz et al. (2005) eine dreidimensionale Beschreibung der Position im Vokalraum vorgestellt, welche in der DFT vereinfacht durch das Verhältnis F1-F2 dargestellt wird. Dabei stellt F2 den "tatsächlichen" Wert der zweiten Formanten dar, welcher sich aus einer Funktion von F2, F3 und F4 ergibt. Wie in Polka & Bohn (2003) dargelegt, lässt sich dieser Raum durch vier "natürliche" Eckpunkte beschreiben: Dies sind die Vokale /u/, /a/, /i/ und /y/. Letzterer Vokal weist einen minimalen Wert von F1 auf, ebenso wie /i/ und /u/, und eine starke Konvergenz (F2-F3). Dadurch wird er sowohl auf wahrnehmbarer als auch produktiver Ebene zu einem fokalen Vokal, wie in Polka & Bohn (2003) und Ménard et al. (2007) dargestellt. Im Rahmen des Reviews von Polka & Bohn (2003, vgl. Best et al., 1997), wird die "wahrnehmungsankernde" Rolle von /y/ beim Kontrast zwischen den norwegischen Vokalen

/ʊ/-/y/ deutlich: die Perzeption, die der Richtung /ʊ/ → /y/ folgt, erweist sich tatsächlich als effektiver. Dies wird jedoch beim Kontrast /i/-/y/ in Frage gestellt, bei dem /i/ eine peripherere Position einnimmt und dennoch, wie in der Studie von Best & Faber (2000) gezeigt, die Diskriminierung besser in der Richtung /i/ → /y/ ist. In der Studie von Best & Faber (2000) werden Säuglinge (3–5 Monate alt) mit L1 Englisch der Wahrnehmung des norwegischen Kontrasts /i/-/y/ unterzogen. Hier kann die DFT, insbesondere das Konzept der Fokalisierung, das Phänomen effektiver erklären als das Konzept der *peripherality*, das in Polka & Bohn (2003) betont wird: Wenn /y/ keinen L1-Laut darstellt (es ist also kein erlernter Prototyp, vgl. Schwartz et al. 2005) und keine traditionell periphere Position einnimmt (es ist kein Eckpunkt im Vokaltriangel), dann ist die wahrscheinlichste Erklärung, dass in der Wahrnehmung eher die wahrnehmungsaktive Salienz aufgrund seiner fokalen Natur eine dominierende Rolle gespielt hat. Diese Erkenntnisse zur fokalen Natur von /y/ sind für die vorliegende Arbeit von zentraler Bedeutung, da der Vokal /y:/ auch im hier untersuchten Sprachpaar eine Schlüsselrolle spielt und als potenzieller "Attraktor" im Wahrnehmungsraum fungieren könnte.

Die Studien von Polka und Bohn (2003; 2011; vgl. auch Polka et al., 2021) sowie von Schwartz et al. (2005) leisten einen wesentlichen Beitrag zum Verständnis der Wahrnehmung (und, wie von Ménard et al. 2007 betont, auch der Produktion) sprachlicher Laute. Sie geben wichtige Impulse für die Forschung zu den universellen Aspekten des Lernens der L1. Im Rahmen der genannten Studien wurden auch Tests zur Wahrnehmung von Lauten einer Fremdsprache (LX) an Erwachsenen ohne vertieftes Wissen dieser Sprache durchgeführt, um zu überprüfen, ob die beschriebenen Asymmetrien auch ohne signifikante Erfahrung in der LX auftreten. An dieser Stelle erscheint es sinnvoll, die Ausprägung und Entwicklung der Wahrnehmungsasymmetrien in Bezug auf das Sprachniveau der Fremdsprache zu untersuchen.

Die Anwendbarkeit des NRV-Frameworks, der *Peripherality Hypothesis* und der DFT im Kontext des Fremdsprachenerwerbs wurde bislang jedoch noch nicht systematisch erforscht. Es liegt die Vermutung nahe, dass die perzeptiven Grundprinzipien der L1 auch die Wahrnehmung und Produktion einer LX regulieren. Eine empirische Überprüfung dieser Annahme könnte aufzeigen, inwieweit diese Verarbeitungsmechanismen eine sprachübergreifende Gültigkeit besitzen, oder ob im Fremdsprachenerwerb abweichende Prozesse wirksam werden.

Im Folgenden werden empirische Studien diskutiert, die belegen, dass bestimmte Entwicklungsmuster – wie die Tendenz zur Zentralisierung oder die Expansion in die Peripherie – nicht auf spezifische Sprachpaare beschränkt sind, sondern in unterschiedlichen linguistischen Konstellationen wiederkehren. Diese Arbeiten liefern somit wichtige Hinweise darauf, dass es sich um übergreifende Phänomene der vokalischen Entwicklung in der Fremdsprache handelt. Es sei darauf hingewiesen, dass die Entwicklung des vokalischen Systems gemäß Polka und Bohn (2003) einem zentrifugalen Verlauf folgt, bis es sich in einer "prototypischen" Weise im

Hinblick auf die Zielsprache (in ihrer Studie die L1) konfiguriert. Dieser Verlauf manifestiert sich in den ersten sechs Lebensmonaten. Wie in Abschnitt 2.2.3 dargelegt, können die Beziehungen zwischen den verschiedenen einflussnehmenden Faktoren sehr komplex und schwer zu entwirren sein, insbesondere wenn neben den internen linguistischen Faktoren auch externe und kontextuelle Faktoren hinzukommen, die mit der Lernumgebung, der Aufnahme des Inputs oder auch nur mit den experimentellen Modalitäten zusammenhängen. Zur Vereinfachung der Analyse kann die Rolle der Sprachkompetenz, verstanden als Erfahrung in der LX, jedoch vorerst isoliert werden, um deren Einfluss auf die progressive Veränderung des vokalischen Systems in der LX zu untersuchen.

Die Studie von Lee & Cho (2020) belegt die entscheidende Rolle der sprachlichen Kompetenz bei der Wahrnehmung von LX-Vokallauten. Gegenstand der Untersuchung war die Vorhersagekraft eines auf der L1 basierenden Modells. Das Modell, welches die Wahrnehmung von L1-Lauten prognostizieren kann, wurde an eine Gruppe koreanischer Lerner des amerikanischen Englisch mit unterschiedlichen Kompetenzniveaus appliziert. Die genannten Probanden wurden, wie in derartigen Untersuchungen üblich, anhand der *Length of Residence* (LOR) gemessen. In Bezug auf die *Length of Residence* (LOR) wurden die Lerner in zwei Gruppen unterteilt: a) Lerner mit geringerer Erfahrung (4 Jahre Aufenthalt in den USA) und b) Lerner mit höherer Erfahrung (11,1 Jahre Aufenthalt). Des Weiteren verfügten die Versuchsteilnehmer über keine nennenswerten Erfahrungen mit dem *Standard Southern British English* (SSBE). Die Resultate demonstrierten, dass das Modell die Wahrnehmung der Vokale der Gruppe mit geringerem Erfahrungsschatz im amerikanischen Englisch sowie der SSBE-Vokale mit einer gewissen Genauigkeit vorhersagen konnte, im Vergleich zur Gruppe mit höherem Erfahrungsniveau.

Weitere Erkenntnisse liefert die jüngste Studie von Cho & Lee (2023), welche sich mit der Beziehung zwischen der Wahrnehmungs- und Produktionsdimension sowie ihrer Dynamik in Abhängigkeit von der sprachlichen Kompetenz befasst. Im Rahmen der Studie wurden Probanden mit koreanischer L1, die in einem immersiven L2-SSBE-Kontext lebten, anhand ihres Kompetenzniveaus unterschieden. In diesem Fall wurde die Erfahrung anhand des Alters bei der Ankunft sowie der LOR gemessen. Im Rahmen der Analyse wurden sowohl Vokale berücksichtigt, die in ihrer Struktur den Vokalen des Koreanischen ähneln (/i, ʌ, u/), als auch "neue" Vokale, die spezifisch für das SSBE sind (/eɪ, əʊ, ɑ, ɒ/). Die Ergebnisse der Studie lassen den Schluss zu, dass lediglich bei Individuen mit höherer Erfahrung ein signifikanter Zusammenhang zwischen Wahrnehmung und Produktion festgestellt werden konnte. Die Ergebnisse deuten darauf hin, dass die Werte der Wahrnehmungsgenauigkeit höher sind als die der Produktionsgenauigkeit. Dies wies auf eine Entwicklung der Kompetenz in die Richtung Perzeption → Produktion hin. Eine differenzierte Betrachtung der einzelnen Vokale ergab, dass bei den koreanisch-ähnlichen Vokalen die Gruppe mit höherer Erfahrung eine stärkere

Wahrnehmungs-Produktions-Verknüpfung aufwies als die Gruppe mit geringerer Erfahrung. Es ist jedoch bemerkenswert, dass bei der letztgenannten Gruppe lediglich der Vokal /i/ eine signifikante Korrelation zwischen Wahrnehmung und Produktion aufwies, während der Vokal /u/ allgemein hohe Werte für Wahrnehmungs- und Produktionsgenauigkeit zeigte. Für die "neuen" Vokale, die im koreanischen System keine direkte Entsprechung haben, konnte lediglich /ɒ/ (ein tendenziell peripherer Vokal) eine negative Korrelation zwischen Wahrnehmung und Produktion nachgewiesen werden. Dies bedeutet, dass eine hohe Wahrnehmungsgenauigkeit mit einer niedrigen Produktionsgenauigkeit einhergeht und umgekehrt. Die Daten legen insgesamt nahe, dass für die Vokale /i/, /u/, /ɒ/ stärkere Korrelationen zwischen Wahrnehmung und Produktion bestehen als für die anderen analysierten Vokale. Dies wird insbesondere bei den Probanden mit geringerer Erfahrung ersichtlich. Für den Vokal /i/ lagen die Wahrnehmungs- und Produktionsgenauigkeitswerte bei 82,2 % bzw. 84,4 %, für den Vokal /u/ bei 95,8 % bzw. 94,4 %. Demgegenüber betrugen die Werte für den zentralen Vokal /ʌ/ lediglich 84,4 % (Wahrnehmung) und 73,3 % (Produktion).

Die Studie von Oh (2019) zielte speziell darauf ab, die Auswirkungen der Kompetenz (Erfahrung) in LX auf die Realisierung koreanischer Vokale bei 20 Lerner:innen mit Mandarin-Chinesisch als L1 (im Durchschnitt 22,6 Jahre alt) mit unterschiedlichem Erfahrungsniveau zu untersuchen. Letzteres wurde anhand der Expositionszeit gemessen, wobei der am wenigsten erfahrene Lerner 6 Monate und der erfahrenste 2 Jahre Erfahrung in LX hatte. Die Ergebnisse bestätigten die Hypothese, dass mehr Erfahrung einer zielkonformen Umsetzung entspricht. Darüber hinaus gibt es relevante Überlegungen, die die Autorin auf die Rolle der L1 und insbesondere auf die Ähnlichkeits-Differenz-Beziehung zwischen konkurrierenden Phonemen zurückführt. Generell zeigen die Vokalsysteme der beiden Gruppen zwei unterschiedliche Muster, auf zwei unterschiedliche Kompetenzniveaus zurückgeführt werden können (vgl. Missaglia & Sendlmeier, 1999). Der Vokalraum der weniger erfahrenen Lerner:innen zeigt eine Konzentration der Vokale entlang der Peripherie des Dreiecks und eine deutliche Überlappung der Kategorien. Im Gegensatz dazu zeigt der Vokalraum der erfahreneren Lerner:innen eine klare Differenzierung der Vokalkategorien mit einer deutlichen Verschiebung der zentralen Vokale (/ɛ/, /ʌ/, „neu“ für chinesische Lerner:innen) in Richtung ihrer Zielposition. Der Autor stellte fest, dass die „ähnlichen“ Vokale, d.h. diejenigen, die im chinesischen System eine Entsprechung haben (/i/, /u/, /a/, alle peripher), allen Sprecher:innen die geringsten Schwierigkeiten bereiteten. Dieses Ergebnis unterstützt einerseits Trofimovich et al. (2001) und Baker et al. (2008), die behaupten, dass es die „neuen“ Laute (ohne Entsprechung) sind, die die meisten Schwierigkeiten verursachen; andererseits widerspricht es dem SLM (Flege, 1995, vgl. Oh, 2019). Auf den Bildern erscheinen sie tatsächlich als die Vokale, die die geringste Verschiebung im Vokalraum erfahren haben. Der Autor berichtet, dass der tiefe Vokal /ʊ/, der peripher, aber zentraler als /i/ und /u/ liegt, von beiden Sprechergruppen zielsprachlich realisiert wurde. Auch hier handelt es sich um einen „ähnlichen“ Vokal. Es ist jedoch wichtig anzumerken, dass dieses Ergebnis nicht mit der

früheren Studie von Kim (2013) übereinstimmt, in der festgestellt wurde, dass chinesisch-koreanische Lerner:innen nicht in der Lage waren, dieses Phonem innerhalb eines Jahres zu lernen. Als mögliche Ursache führt der Autor die unterschiedlichen experimentellen Kontexte an, die im Fall von Kim (2013) das Lesen von Pseudowörtern in einer Trägerphrase beinhalteten, während im Fall von Oh (2019) das Lesen von echten Wörtern in Isolation erfolgte.

Die in der Studie von Oh (2019, siehe auch Missaglia & Sendlmeier, 1999) beobachteten strukturellen Muster von Vokaldreiecken werden auch in der Studie von Lee & Rhee (2019) bestätigt, die mit Daten aus dem *Genie Speech Corpus* ("Genie SpeechCor") durchgeführt wurde, das Sprachdaten von koreanischen L1-Lerner:innen des Englischen als Fremdsprache enthält. Die Daten von 30 erwachsenen männlichen Sprechern wurden auf der Grundlage des Sprachniveaus in verschiedene Gruppen eingeteilt: Stufe 1-2, Stufe 3 und Stufe 4-5 (die höchste). Die Gruppe mit dem höheren Sprachniveau weist einen breiteren Vokalraum und somit eine größere spektrale Differenzierung auf, während die Gruppe mit dem niedrigeren Sprachniveau stärker zentralisiert ist. Nach Ansicht der Autor:innen könnte dies auf die unterschiedliche Dichte der beiden Vokalsysteme zurückzuführen sein: Das Englische hat ein viel reicheres Vokalinventar, was weniger erfahrene Sprecher dazu veranlassen würde, Phoneme in einem eher begrenzten und zentralisierten Bereich zu realisieren. Die Daten zeigen jedoch nur für den Kontrast /i/-/ɪ/ einen signifikanten Effekt der Spracherfahrung.

Die Studie von Mora (2021) ist hier von besonderem Interesse, da sie eine Längsschnittperspektive auf die Auswirkungen von Erfahrung auf die Genauigkeit von LX-Vokalisierung bietet. In diesem Fall konzentriert sich die Analyse auf die Produktion nur des englischen /ʌ/-/æ/-Kontrasts durch Lerner:innen des Englischen als Fremdsprache mit Spanisch/Katalanisch als L1. Die Lerner:innen, die sich alle auf einem fortgeschrittenen Sprachniveau befanden, nahmen an einem phonetischen Training mit hoher Variabilität (HVPT) teil, um dessen Wirksamkeit zu testen. HVPT ist ein Training, das darauf abzielt, die Wahrnehmung und Produktion von LX-Kontrasten zu verbessern, die als komplex zu erlernen gelten, indem die Proband:innen Ziellauten ausgesetzt werden, die von mehreren Sprecher:innen und in unterschiedlichen phonologischen Kontexten produziert werden (vgl. Mora 2021, S. 10). In diesem Sinne kann argumentiert werden, dass die Proband:innen dieser Studie im Laufe der Zeit immer mehr Erfahrung mit dem /ʌ/-/æ/-Kontrast gesammelt haben. Das letztendliche Ziel der Studie war es, die Sensitivität verschiedener akustischer Maße (Euklidischer Abstand, Pillai-Score, Mahalanobis-Abstand) für überlappende Vokalkategorien und deren Unterscheidbarkeit zu testen. Abgesehen von diesem sehr spezifischen Aspekt der Studie liegt der Schwerpunkt hier auf den Auswirkungen des HVPT auf die Vokalproduktion im Englischen. Das Paradigma wurde in vier Sitzungen durchgeführt, um die Phonemwahrnehmung und -produktion durch eine AX-Diskriminierungsaufgabe, eine Identifikationsaufgabe und eine sofortige Wortwiederholungsaufgabe (in dieser Reihenfolge, siehe ebd.) zu testen. Der

Fortschritt in der Produktion wurde mit Hilfe einer verzögerten Wortwiederholungsaufgabe (DWR) überprüft, die vor und nach dem Test durchgeführt wurde. Die Verteilung der normalisierten Formantenwerte (vgl. Mora 2021, S. 12) zeigt eine allgemein größere Überlappung der Realisierungen bei den Lerner:innen. In der Posttestbedingung scheinen die Kategorien jedoch stärker voneinander getrennt zu sein. Im Allgemeinen scheinen die Vokale der Lerner:innen weiter im Vokalraum verteilt zu sein und eine größere Varianz aufzuweisen als die der L1-Sprecher:innen.

Eine noch detailliertere Untersuchung dieses “Verhaltens” findet sich in der Studie von Oh & Shin, (2023), die eine Stichprobe koreanischer L1-Sprecherinnen analysiert, die Englisch in einem immersiven Kontext mit unterschiedlichem Grad an Englisch Erfahrung lernen. In diesem Fall wurde die Erfahrung anhand des Erwerbsalters (AOA) und der Aufenthaltsdauer (LOR) gemessen. Die Stichprobe, bestehend aus 45 Frauen (18–52 Jahre), wurde dem *Speech Accent Archive* Korpus entnommen (Weinberger, 2015) . Die Ergebnisse zeigten eine Korrelation zwischen vokalspektralen Merkmalen und LOR, was darauf hindeutet, dass ein immersiver Lernkontext für die Entwicklung der Genauigkeit der Vokalproduktion förderlich ist. Die in der Studie angegebenen Graphiken (siehe Oh & Shin, 2023, S. 13) veranschaulichen die Streuung der Realisierungen von mehr und weniger erfahrenen Probandinnen im Vergleich zu einer Kontrollgruppe mit L1 Englisch. Insbesondere bei erfahreneren Sprecherinnen ist eine Dissimilation von Kontrasten, eine geringere Streuung der Existenzbereiche der Vokale (weniger Variation in der Produktion einzelner Vokale) sowie das Vorhandensein spezifischer phonetischer Merkmale, die für dialektale Varianten der Zielsprache typisch sind, zu beobachten. Der letztgenannte Faktor trägt dazu bei, zu zeigen, dass eine größere Exposition gegenüber der Zielsprache auch zur Entwicklung „feinerer“ Merkmale des zielsprachlichen Vokalsystems geführt hat.

Aus den dargelegten Studien lassen sich einige Tendenzen ableiten. Zunächst begünstigt die Sprachkompetenz, die im Allgemeinen als verlängerte Exposition gegenüber dem Input in der LX verstanden wird, die Entwicklung des Vokalsystems in Richtung der Zielpositionen. Diesbezüglich ist festzuhalten, dass die Rolle des externen Inputs, sei es in Form eines immersiven Kontexts oder einer gezielten Erhöhung desselben durch ein spezifisches Paradigma, nicht unterschätzt werden darf. In der Tat lassen sich zwei Haupttendenzen feststellen: Vokalsysteme neigen entweder dazu, sich auf die Peripherie der Vokalraum-Ecken zu konzentrieren, oder es kommt zu einer verstärkten Zentralisierung (vgl. Missaglia & Sendlmeier, 1999). Wie in den Studien von Lee & Rhee (2019) sowie von Oh & Shin (2023) dargelegt, lässt sich eine positive Korrelation zwischen einer größeren Erfahrung bzw. Kompetenz in der Zielsprache und einem erweiterten Vokalraum, deutlicher definierten Ecken sowie einer Platzierung der zentralen Vokale in den zentraleren Zonen des Dreiecks beobachten. Dies geht mit einer klareren Unterscheidung der Kategorien und somit einer geringeren

Überlappung derselben einher. In den frühen Lernstadien zeigt sich eine stärkere Überlappung der Vokale sowie eine größere Streuung ihrer Realisierungen, was auf eine hohe Varianz hindeutet.

Die Analyse der Wahrnehmung und Produktion im sprachlichen Bereich erfordert einen multidimensionalen Ansatz, der die Wechselwirkungen zwischen externem Input, Sprachkompetenz und vorbestehenden vokalischen Konfigurationen berücksichtigt. Zusammenfassend bestätigen die diskutierten Studien, dass die Umstrukturierung des Vokalraums – sei es durch Zentralisierung oder Expansion in die Peripherie – ein sprachübergreifendes Phänomen darstellt, das maßgeblich von der Kompetenz und dem Input gesteuert wird. Dennoch bleiben Fragen offen, insbesondere wie sich diese Dynamiken in einem komplexeren, mehrsprachigen System verhalten, in dem die *Cross-Linguistic Influence* (CLI) durch die Konkurrenz mehrerer Referenzsprachen an Komplexität gewinnt.

Die vorliegende Arbeit knüpft genau hier an: Sie untersucht, ob und inwieweit sich diese anscheinend sprachübergreifenden Entwicklungstendenzen auch in der spezifischen Sprachkonstellation dieser Studie manifestieren. Ziel ist es, durch die Analyse einer homogenen Stichprobe die Wechselwirkung zwischen perceptiven Asymmetrien (im Sinne des NRV) und den Mechanismen des CLI bei der produktiven Neuordnung des Vokalraums empirisch zu beleuchten.

3. Untersuchte Vokalsysteme vis-à-vis und Forschungsstand zu Wahrnehmung und Produktion im DaF-Erlernen

3.1. Die Vokalsysteme des Italienischen und des Deutschen im Vergleich

Um die Schwierigkeiten beim Erwerb des deutschen Vokalsystems wissenschaftlich zu fundieren, muss die Ausgangslage der Lernenden (L1) nicht als isoliertes System, sondern in ihrer inhärenten kontrastiven Dynamik zur Zielsprache betrachtet werden.

Das italienische Standardsystem wird klassischerweise als Heptavokalsystem in betonten Positionen beschrieben (vgl. Grassi et al., 2003; Maturi, 2014). Nach Krämer (2009) besteht es aus sieben Monophthonge: /i/, /e/, /ɛ/, /a/, /o/, /ɔ/, /u/. Es ist zu konstatieren, dass alle Vokale des Inventars in betonter Position auftreten können, wohingegen in unbetonter Position das Inventar auf fünf Vokale reduziert wird (vgl. Mioni, 1993, Bertinetto & Loporcaro, 2005). Nachfolgend werden die distinktiven Merkmale der italienischen Vokale angeführt (Tabelle 2):

Tabelle 2: Distinktive Merkmale des Italienischen, adaptiert nach Heinz & Schmid (2021, S.47, Tabelle 11)

Merkmal/ Vokal	/i/	/e/	/ɛ/	/a/	/ɔ/	/o/	u/
hoch	+	-	-	-	-	-	+
tief	-	-	+	+	+	-	-
hinten	-	-	-	+	+	+	+
gerundet	-	-	-	-	+	+	+

Die Vokale unterscheiden sich hinsichtlich ihrer vokalischen Qualität, die durch zwei Hauptartikulationspositionen (vorn/hinten) und vier Öffnungsgrade (hoch, mittel-hoch, mittel-tief, tief) bestimmt wird. Üblicherweise (vgl. Calabrese, 1987, 1988) wird ein solches Vier-Stufen-System durch die Verwendung von zwei Höhenmerkmalen ([±hoch] und [±tief]) beschrieben, wobei die mittleren Vokale für beide Merkmale negativ spezifiziert sind (vgl. aber

Krämer, 2009, für die Unterscheidung nach dem Merkmal ATR; siehe auch Halle & Stevens, 1969).

Bezüglich des Merkmals der Rundung sind die vorderen Vokale ungerundet, während drei hintere Vokale gerundet sind. Der tiefe Vokal ist weder vorne noch hinten: Phonetisch betrachtet ist er ein zentraler bis vorderer Vokal, doch da er an keinen phonologischen Prozessen beteiligt ist, die die vorderen Vokale betreffen, wird er häufig als phonologisch hinterer Vokal betrachtet (vgl. Krämer, 2009, S. 50).

Hinsichtlich des Merkmals der Länge ist festzustellen, dass diese im Italienischen keinen phonematischen Status auf segmentaler Ebene besitzt. Die Variation der Länge der Vokale ist vielmehr kontextabhängig und resultiert aus prosodischen sowie lexikalischen Bedingungen (vgl. Bertinetto & Loporcaro, 2005; Krämer, 2009). Die Länge eines Vokals im Italienischen ist abhängig von seiner Position im Wort (vgl. Krämer, 2009, S. 51): Vokale im Auslaut sind stets kurz, unabhängig davon, ob sie betont oder unbetont sind¹ (ebenda). Lange Vokale treten in betonten offenen Silben auf, während Vokale in betonter offener Antepänultima als halblang klassifiziert werden können und kurze Vokale in allen anderen Positionen zu verorten sind (vgl. ebenda). Bertinetto und Loporcaro (2005) betonen die suprasegmentale Komponente der Vokalquantität, wonach Vokale am Ende einer Intonationsphrase eine prosodische Dehnung erfahren. Zudem ist festzustellen, dass am Wortende lediglich vier unbetonte Phoneme auftreten können, da das Phonem /u/ in dieser unbetonten Position niemals erscheint – mit seltenen Ausnahmen in lexikalischen Entlehnungen wie dem Wort *bantu* ['bantu], das an das typische italienische Akzentmuster angepasst wurde (vgl. Bertinetto & Loporcaro, 2005).

Obwohl das Phänomen der Reduktion im Italienischen keinen phonematischen Status hat, kann es im speziellen Fall der Vokale in unbetonter Position am Wortende auftreten. Wenn zwei Vokale aufeinander folgen und an der Wortgrenze stehen (-V##V-), kann der zweite Vokal manchmal die Merkmale eines Semivokals annehmen (z. B. *tanti accenti* ['ta:ntj atʃ:e:nti], "viele Akzente", vgl. Bertinetto & Loporcaro, 2005, S. 139) oder der vollständigen Elision unterliegen (*gli amici arrivano* [ʎ a'mi:tʃ a'r:i:vano], "die Freunde kommen", vgl. Bertinetto & Loporcaro, 2005, S. 139), oder sogar zu einem Diphthong führen. Die Resultate dieser Prozesse unterliegen einer signifikanten Variation, die auch auf die idiosynkratischen Merkmale der Sprecher: innen wie etwa die Sprechgeschwindigkeit, zurückzuführen ist. Das Einfügen eines glottalen Plosivs zwischen benachbarten Vokalen ist jedoch nicht üblich (vgl. ebenda).

Das vokalische System des Deutschen besteht hingegen nach Wiese (2000) aus den folgenden Monophthongen: /i:/, /ɪ/, /y:/, /ʏ/, /e:/, /ɛ/, (/ɐ:/), /ø:/, /œ/, /a:/, /a/, /o:/, /ɔ/, /u:/, /ʊ/, (/ə/). Laut Wiese (2000, S.10) repräsentieren das Vokal- und Konsonanteninventar des Deutschen zwei Systeme, die das phonemische Inventar der deutschen Sprache beschreiben. Es gibt

¹ Krämer (2009, S. 51) weist darauf hin, dass bei einigen Varietäten des Norditalienischen die betonten finalen Vokale lang sind.

jedoch Übergangsphänomene von Konsonanten zu Vokal (wie die Vokalisation von /r/), die sich von dieser Perspektive abwenden und daher besser in einer vereinheitlichten phonemischen Dimension beschrieben werden können (vgl. ebenda). Hier wird ausschließlich das vokalische System als getrennt vom konsonantischen System behandelt. Laut Wiese (2000, S. 20) können die distinktiven Merkmale der deutschen Vokale in der folgenden Matrix zusammengefasst werden, wobei der Autor auf ein «fairly uncontroversial set of features» zurückgreift (Tabelle 3):

Tabelle 3: Distinktive Merkmale der deutschen Vokale, adaptiert nach Wiese (2000, S. 20, Abbildung 6)

Merkmal/ Vokal	i:	ɪ	e:	ɛ:	ɛ	a:	ɑ	o:	ɔ	u:	ʊ	y:	ʏ	ø:	œ	ə
konsonantisch	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
hoch	+	+	-	-	-	-	-	-	-	+	+	+	+	-	-	-
tief	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-
vorne	+	+	+	+	+	-	-	-	-	-	-	+	+	+	+	-
hinten	-	-	-	-	-	-	-	+	+	+	+	-	-	-	-	-
gerundet	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	-
ATR	+	-	+	-	-	-	-	+	-	+	-	+	-	+	-	-
lang	+	-	+	+	-	+	-	+	-	+	-	+	-	+	-	-

Das Merkmal ATR weist auf eine artikulatorische Bewegung der Zungenwurzel in Richtung des vorderen Mundraums hin (vgl. ebenda und Halle & Stevens, 1969). Auf Grundlage dieser artikulatorischen Gestes können daher, laut Wiese (aber vgl. Wurzel, 1981), [+ATR] und [-ATR] Vokale unterschieden werden (symmetrisch zum Kontrast [+tense] vs. [-tense]). Diese sind respektive: /i:/, e:/, o:/, u:/, y:/, ø:/ und /ɪ, ɛ:/, ɛ, a:/, ɑ, ɔ, ʊ, ʏ, œ, ə/.

Ebenso bedeutsam ist das Merkmal der Länge, codiert mit [+/-lang], die meistens mit dem Merkmal der Gespanntheit korreliert. Beispiel hierfür ist der Kontrast /i:/ vs. /ɪ/ (jeweils [+ATR] und [-ATR]).

Das Merkmal [vorne], beschreibt die vordere Artikulation der Vokale /i:/, /ɪ/, /y:/, /ʏ/ (alle [+hoch]) und von /e:/, /ɛ:/, /ɛ/, /ø:/, /œ/ (alle [-hoch]). Der zentrale Vokal /ə/ ist für alle Merkmale mit deren negativem Wert spezifiziert.

Gemäß der klassischen Darstellung des vokalischen Systems (Abbildung 2), können das standarditalienische und standarddeutsche Vokalsystem wie folgt dargestellt werden²:

² Die Darstellung des deutschen Vokalinventars enthält das Phonem /ɐ/, das nicht in Wieses Inventar vorgesehen ist.

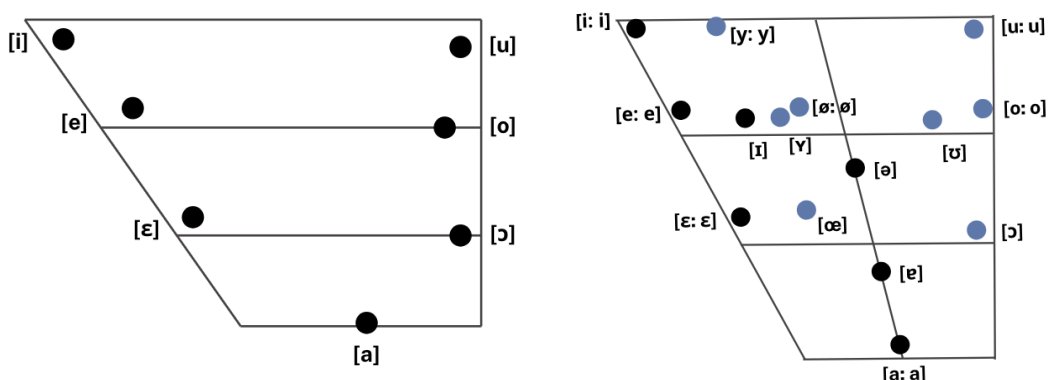


Abbildung 2: Vokaltrapeze des (betonten) standarditalienischen Vokalsystems in Anlehnung an Bertinetto (2010) und des standarddeutschen Vokalsystems (orale Monophthonge) adaptiert nach Kleiner & Knöbl (2015). Im deutschen Vokaltrapez sind gerundete Vokale in Blau gekennzeichnet. Eigene Darstellung.

Im Folgenden wird präzisiert, welcher Standardbegriff der Gegenüberstellung der beiden Vokalsysteme zugrunde liegt. Angesichts der ausgeprägten diatopischen Variabilität, die sowohl den italienischen als auch den deutschen Sprachraum kennzeichnet, wäre eine reduktionistische Definition unzureichend. Vielmehr gilt es, der Pluriarealität der Aussprachenormen Rechnung zu tragen und einen dynamischen Gebrauchsstandard zugrunde zu legen. Die vorliegende Arbeit untersucht nämlich *CLI*-Effekte anhand von Stichproben, die – obwohl sie unter kontrollierten Bedingungen erhoben wurden – die phonetische Realität und die tatsächliche artikulatorische Praxis des italienischen und deutschen Standards widerspiegeln.

Der Begriff des *Gebrauchsstandards* für das Deutsche reflektiert diejenige Varietät, die Sprecher:innen in formellen Gesprächskontexten verwenden, in denen die Standardsprache intendiert wird (vgl. Alber, 2022, S. 27; Lameli, 2004). Da keine allgemein anerkannte, abgeschlossene Definition von Standardsprache existiert (vgl. Alber 2022, S. 27), wird diese hier durch drei wesentliche Merkmale charakterisiert:

1. Sie ist primär als Gebrauchsnorm zu klassifizieren, da sie sich auf empirische Befunde stützt und somit die tatsächliche Sprechwirklichkeit abbildet;
2. sie umfasst überregional gebräuchliche Ausspracheformen, wobei auch national oder großregional übliche Varianten integriert sind;
3. ihre phonologische Repräsentation wird als grundsätzlich einheitlich betrachtet (vgl. Kleiner & Knöbl, 2015, S.31).

Im Gegensatz zur deutschen Situation ist die Standardisierung des Italienischen durch eine langwierige historische Entwicklung geprägt, die erst in der zweiten Hälfte des 20. Jahrhunderts eine relative Stabilisierung erfand. Bis heute sind italienische Sprecher:innen meist durch

spezifische regionale Merkmale gekennzeichnet, die – außer im Bereich der professionellen Sprecherziehung (z. B. Theaterpädagogik) – unmittelbar identifizierbar bleiben. Trotz der sprachpolitischen Zensurversuche nach der Einigung 1861 blieb Italien durch einen permanenten Kontakt zwischen Standard und Dialekt geprägt (vgl. Crocco, 2017, S. 91). Diese Dynamik förderte die Entstehung neuer Varietäten: Die ursprünglich scharfe Diglossie zwischen Dialekt und Hochsprache schwächte sich ab und wich einer differenzierten regionalen Variation (vgl. D'Achille, 2003; Berruto, 1987).

In der Forschung werden insbesondere das Florentinische, das Mailändische und das Römische als prominente Varianten identifiziert (vgl. Bertinetto & Loporcaro, 2005). Während das Florentinische aus historisch-literarischen Gründen das Fundament bildet, erlangten das Mailändische und das Römische aufgrund ihrer Rolle als wirtschaftliche und mediale Zentren eine dominante Stellung. Dennoch bleibt eine exakte Grenzziehung zwischen diesen Varietäten problematisch, da sich sprachliche Merkmale oft über Kilometer hinweg kontinuierlich verschieben („Verdunkelung der Grenzen“). Das traditionelle Modell des Standarditalienischen, welches eng an die florentinische Variante angelehnt ist, wird heute fast ausschließlich von einer kleinen Gruppe Gelehrter gesprochen (vgl. Bertinetto & Loporcaro, 2005). Modernere Ansätze wie die von Canepari (1999, 2007) postulierte „neutrale Aussprache“ weisen zwar eine größere Vielfalt auf, es bleibt jedoch unklar, inwieweit dieses Modell tatsächlich als allgemeiner Aussprachestandard fungiert (vgl. Crocco, 2014, S. 93).

Vor diesem Hintergrund wird in der vorliegenden Arbeit unter dem Begriff *Standarditalienisch* kein abstrakt-idealisiertes Modell verstanden, sondern der auf Italienisch sogenannte *Neo-standard* (vgl. Berruto, 2012). Dieser fungiert als ein Gebrauchsstandard gebildeter Sprecher:innen in formellen Kontexten und trägt der inhärenten regionalen Variation des italienischen Sprachraums explizit Rechnung.

Die akustische Referenzierung der standarddeutschen Vokalphoneme stützt sich in dieser Arbeit primär auf Sendlmeier & Seebode (2006), Pätzold & Simpson (1997) sowie – spezifisch für das Phonem /ə/ – auf Zimmerer & Reetz (2011). Bei der Interpretation dieser Daten ist jedoch eine erhebliche methodische Heterogenität zu berücksichtigen. Diese betrifft einerseits die Probandenzahl, die von kleineren Stichproben (n=24 bei Pätzold & Simpson, 1997) bis hin zu umfangreicheren Datensätzen (n=127 bei Sendlmeier & Seebode, 2006) reicht. Andererseits divergieren die experimentellen Settings: Während die Daten von Pätzold & Simpson (1997) auf kontrollierter, gelesener Sprache basieren (*PhonDat90-Datenbank*, Thon & Dommelen, 1992), legen Sendlmeier & Seebode (2006) sowie Zimmerer & Reetz (2011) Analysen spontansprachlicher Korpora vor (*Kiel Corpus for Spontaneous Speech*, IPDS-München, 1994). Ein wesentlicher statistischer Unterschied besteht zudem in der Datenaufbereitung: Während Sendlmeier & Seebode (2006) sowie Zimmerer & Reetz (2011) Mittelwerte für die ersten zwei Formanten (F1, F2) angeben, basieren die Ergebnisse von Pätzold & Simpson (1997)

ausschließlich auf Medianwerten, wobei hier zusätzlich der dritte Formant (F3) aufgeführt wird. Es ist darauf hinzuweisen, dass der Vokal /ɛ:/ in dem von Pätzold & Simpson verwendeten Datensatz nicht enthalten ist. Diese Abweichung ist darauf zurückzuführen, dass der Vokal in der deutschen Sprache einen fraglichen Status aufweist (vgl. Sendlmeier & Seebode, 2006, vgl. Wiese, 2000).

Basierend auf den oben dargelegten Quellen werden im Folgenden die akustischen Kennwerte der deutschen Standardvokale zusammenfassend dargestellt. Tabelle 4 bildet die Mittelwerte ab, die Sendlmeier & Seebode (2006) sowie – im Falle des Phonems /ə/ – Zimmerer & Reetz (2011) entnommen wurden. Die Werte für den Schwa-Laut sind dabei jeweils mit a) (Sendlmeier & Seebode, 2006) und b) (Zimmerer & Reetz, 2011) gekennzeichnet, um die jeweilige Provenienz zu verdeutlichen. In Tabelle 5 werden hingegen die Medianwerte nach Pätzold & Simpson (1997) aufgeführt.

Tabelle 4: Mittelwerte zu F1 und F2 der deutschen Vokale nach Sendlmeier & Seebode (2006) und Zimmerer & Reetz (2011)

Vokal	M (F1)	M (F2)	W (F1)	W (F2)
/a/	694	1372	836	1586
/a:/	737	1275	896	1517
/e:/	348	2126	434	2461
/ɛ/	489	1817	608	2040
/ɛ:/	482	1902	584	2166
/ɪ/	369	1902	433	2095
/i:/	263	2171	302	2533
/ɔ/	537	1074	605	1200
/o:/	383	841	440	889
/ʊ/	391	1010	442	1081
/u:/	310	854	345	956
/ʏ/	373	1543	426	1670
/y:/	302	1722	320	1810
/œ/	474	1477	564	1654
/ø:/	371	1501	440	1605
ə (a)	517	1447	572	1763
ə (b)	509	1712	510	1793

Tabelle 5: Medianwerte zu F1, F2 und F3 der deutschen Vokale nach Pätzold & Simpson (1997)

Vokal	M (F1)	M (F2)	M (F3)	W (F1)	W (F2)	W (F3)
/a/	608	1309	2466	751	1460	2841
/a:/	639	1225	2477	779	1347	2785
/e:/	372	1879	2486	431	2241	2871
/ɛ/	498	1639	2451	592	1944	2867
/ɪ/	343	1803	2483	391	2136	2867
/i:/	290	1986	2493	329	2316	2796
/ɔ/	506	1060	2415	573	1174	2825
/o:/	380	907	2415	438	953	2835
/ʊ/	382	1058	2363	450	1184	2749
/u:/	309	961	2366	350	1048	2760
/ʏ/	374	1431	2284	406	1612	2631
/y:/	310	1505	2205	342	1667	2585
/œ/	437	1504	2179	509	1767	2640
/ø:/	375	1458	2220	434	1646	2573
ə	370	1521	2368	420	1746	2811

Die Berücksichtigung der methodischen Heterogenität der Quellen ist insbesondere deshalb von Relevanz, da die Realisierung von Vokalphonemen maßgeblich von jeweiliger Erhebungsmethode beeinflusst wird. Die Spontaneität der gesprochenen Sprache wirkt sich signifikant auf die Beschaffenheit des Vokalraums (vgl. Kohler & Rodgers, 2001). In der vorliegenden Arbeit wurde die Aufgabenart (vom semi-freien Sprechen bis hin zur kontrollierten Wortliste) daher explizit als Faktor berücksichtigt.

Ein weiterer potenziell einflussreicher Faktor, der jedoch im Rahmen dieser Arbeit nicht operationalisiert wurde, ist die Sprechgeschwindigkeit. Untersuchungen von Siebenhaar & Hahn (2019) verdeutlichen deren Komplexität: Während das Verhältnis zwischen langen und kurzen Vokalen bei Geschwindigkeitsänderungen weitgehend stabil bleibt, führt eine Erhöhung des Tempos einerseits zur Reduktion weiträumiger Vokalsysteme und andererseits zur Expansion kleinerer Vokalräume. Trotz einer beobachteten Unabhängigkeit von diatopischen Faktoren betonen die Autoren die Notwendigkeit einer detaillierten Kontextualisierung der Daten unter Berücksichtigung sozialer und stilistischer Variablen (vgl. Siebenhaar & Hahn 2019, S. 882).

Analog zur Darstellung des Deutschen weist auch die Quellenlage für das Italienische eine signifikante methodische Varianz auf. Die Daten von Bertinetto & Loporcaro (2005, S. 136), die auf Erhebungen von Ferrero et al. (1978) zurückgreifen, basieren sulla Artikulation isolierter Wörter, realisiert von zehn Sprechern aus dem florentinischen Raum. Kritisch anzumerken ist hierbei, dass in der Originalquelle keine expliziten Angaben zum Geschlecht der Proband:innen vorliegen und die Darstellung ausschließlich auf Mittelwerten basiert, ohne die Standardabweichung (Standardfehler) auszuweisen.

Einen komplementären Ansatz bieten Heinz & Schmid (2021). Ihre Daten basieren auf der Analyse gelesener Sprache (ein Textauszug aus Giorgio Bassanis Roman *Il giardino dei Finzi-Contini*), wobei hier ausschließlich Vokale in betonter Position berücksichtigt wurden. Im Gegensatz zur eher homogenen florentinischen Gruppe bei Ferrero et al. (1978) spiegelt diese Quelle eine ausgeprägte diatopische Variation wider: Die Daten stammen von einer Sprecherin aus dem Kanton Tessin (mit lombardischem Substrat) und einem Sprecher aus der Toskana.

Im Folgenden werden die akustischen Kennwerte (F1, F2) für das Italienische in Tabelle 6 zusammengefasst. Hierbei ist festzuhalten, dass im Gegensatz zum Deutschen – basierend auf der aktuellen Literaturrecherche – für das Italienische ausschließlich Mittelwerte verfügbar sind. Die Werte von Heinz & Schmid (2021) wurden dabei als primäre Referenz ausgewählt, da sie einerseits den aktuellen Forschungsstand abbilden und andererseits eine geschlechtsspezifische Differenzierung ermöglichen.

Tabelle 6: Medianwerte zu F1 und F2 der deutschen Vokale nach Pätzold & Simpson (1997)

Vokal	M (F1)	M (F2)	W (F1)	W (F2)
/i/	328	2133	371	2290
/e/	406	1886	418	2054
/ɛ/	529	1814	567	1830
/a/	694	1467	785	1565
/ɔ/	602	1069	554	1220
/o/	448	1133	442	1250
/u/	379	1114	376	1200

Im Folgenden wird der Fokus auf die zwei untersuchten Phoneme /ə/ und /y:/ gelegt, indem sowohl ihr Status im deutschen und italienischen Sprachsystem als auch wesentliche Elemente für die vorliegende Argumentation hinsichtlich der Variation dieser Phoneme in verschiedenen Dimensionen dargelegt werden.

3.1.1 Der Status von /ə/ im Deutschen

Um die Komplexität der vokalischen Kontraste im Spracherwerb zu erfassen, ist zunächst der besondere Status des zentralen Vokals /ə/ zu betrachten. Dieser nimmt im deutschen Phonemsystem eine Sonderstellung ein: Aufgrund seiner Artikulation in der Mitte des Mundraums ist er der einzige Vokal, der keines der bei Wiese (2000) postulierten distinktiven Merkmale der Zungenhöhe aufweist. Diese theoretische „Unterspezifiziertheit“ ist eng mit der anhaltenden Diskussion über seinen phonematischen Status verknüpft. Während die Beschränkung auf unbetonte Silben e die Unmöglichkeit, als alleiniger Wortvokal zu fungieren, gegen eine Einordnung als Phonem sprechen, belegt die Existenz von (fast) Minimalpaaren wie *fehlend* vs. *elend* o *Gebein* vs. *geben* ([ə] vs. [ɛ] bzw. [e:]) seine distinktive Funktion (vgl. Wiese 2000, S. 11 und 16).

Für die vorliegende Untersuchung ist jedoch weniger die rein phonologische Statusfrage entscheidend, als vielmehr die Tatsache, dass diese strukturelle Besonderheit das Phonem /ə/ besonders anfällig für artikulatorische Variationen macht. Da /ə/ primär in unbetonten Präfixen (*ge-*, *be-*) und Suffixen (*-e*, *-en*) auftritt, fungiert er als zentraler Knotenpunkt für diatopische und diaphasische Reduktionsprozesse. Wie im Folgenden dargelegt wird, ist es gerade diese Instabilität des Schwa, die im Gebrauchsstandard zu einer ausgeprägten regionalen Vielfalt führt.

3.1.1.1 *Vokalreduktion und Elision im Gebrauchsstandard*

„Die zuvor beschriebene strukturelle Sonderstellung von /ə/ bildet die theoretische Basis für die weitreichenden phonologischen Abweichungen, die für den deutschen Gebrauchsstandard charakteristisch sind. Wie Barbour & Stevenson (1998) konstatieren, manifestiert sich hier eine ausgeprägte Tendenz zur Vokalreduktion, insbesondere bei Vokalen in unbetonter Position, die zugunsten des zentralen Schwa-Lautes abgeschwächt werden. Dieses Merkmal unterscheidet den Gebrauchsstandard deutlich von der präskriptiven Standardsprache und tritt regelmäßig bei Funktionswörtern wie Pronomen und Artikeln auf.

Hinsichtlich der konkreten Realisierung von /ə/ ist zu beobachten, dass dieses häufig zu [ɐ] oder silbischen Konsonanten wie /ŋ/ oder /m/ reduziert wird, sofern entsprechende phonetische Umgebungen vorliegen (z. B. *er* [ɐ], *den* [dŋ], *dem* [dɪm]; vgl. Barbour & Stevenson, 1998). Ein weiteres aktuelles Phänomen ist die Apokope in der ersten Person Singular der Imperativformen.

Die Realisierung von /ə/ in seinen in der Standardsprache vorgesehenen Positionen ist jedoch kein statisches Phänomen, sondern stark diatopisch, diastratisch (vgl. Hahn & Siebenhaar, 2019; Lange et al., 2024) sowie diaphasisch konditioniert (Kohler & Rodgers, 2001). Im Vergleich zu anderen Vokalen tritt die Elision bei /ə/ weitaus häufiger auf, wobei in reduzierten Kontexten selten artikulatorische Spuren des Vokals verbleiben (vgl. Helgason et al., 1996).

Ein signifikanter regionaler Schwerpunkt der Schwa-Synkope findet sich im bayerisch-österreichischen Raum, wo etwa die Realisierung des Suffixes <-en> maßgeblich durch den linken Konsonantenkontext determiniert wird (vgl. Schirmunski, 2010). Diese diatopische Variation ist eng mit dem Register verknüpft: Während die vollständige Aussprache von /ə/ in Suffixen wie <-en>, <-em> und <-el> das formale Standardregister der „deutschen Hochlautung“ (Krech et al. 2009) kennzeichnet, ist die Elision typisch für informelle, schnellere Sprache. Das *Duden Aussprachewörterbuch* (Kleiner & Knöbl, 2023) beschreibt diese Tendenz weniger präskriptiv als ältere Werke (z. B. Krech, 1982) und differenziert nach Artikulationsgenauigkeit. Bei hoher Präzision bleibt /ə/ fast vollständig erhalten. Bei mittlerer Präzision entfällt /ə/ vor allem nach Plosiven, Frikativen (außer [ʁ], [ʝ]) und Lateralen. Bei

niedriger Präzision, die oft mit hoher Sprechgeschwindigkeit korreliert (vgl. Hahn & Siebenhaar, 2019), ist die Elisionshäufigkeit so hoch, dass fast alle Positionen betroffen sind.

Besonders aufschlussreich für die vorliegende Arbeit sind die Befunde von Kohler & Rodgers (2001) zur generationenspezifischen Variation. Jüngere Sprecher:innen neigen signifikant stärker zur Elision als ältere Generationen. Im sächsischen Raum stieg die Elisionshäufigkeit bei Männern von 9 % auf 19 % und bei Frauen von 21 % auf 33 %. Im Bairischen ist dieser Kontrast noch drastischer: von 1 % auf 69 % bei Männern und von 3 % auf 64 % bei Frauen. Auch die diaphasische Variation – der Unterschied zwischen gelesener (44 % Auslassung) und spontaner Rede (64 % Auslassung) – unterstreicht die Dynamik dieses Lautes.

Zusammenfassend dienen diese Skizzierungen dazu, die Dimensionen der ständigen Veränderung von /ə/ in Zeit und Raum aufzuzeigen. Von zentraler Bedeutung ist dabei die Variation in den jüngeren Generationen, die in dieser Untersuchung als Kontrollgruppe für die Aussprache des Deutschen herangezogen werden.

3.1.2 Das Phonem /y:/ im Deutschen

Neben dem Schwa-Laut stellt der geschlossene gerundete Vorderzungenvokal /y:/ eine zentrale Komponente des deutschen Vokalsystems dar. Im Gegensatz zu /ə/ weist /y:/ eine stabile Merkmalskombination auf, die durch die Merkmale [+vorne] und [+rund] definiert ist. Diese phonologische Spezifikation ist entscheidend für die Aufrechterhaltung distinktiver Kontraste zu den ungerundeten Gegenstücken (wie /i:/) sowie zu den hinteren gerundeten Vokalen (wie /u:/).

Hinsichtlich der Realisierung im Gebrauchsstandard lassen sich jedoch spezifische Abweichungen von der präskriptiven Norm beobachten. Ein wesentliches Phänomen ist die Entrundung (Delabialisierung), bei der der Lippenrundungsgrad reduziert wird, was zu einer akustischen Annäherung an /i:/ führt. Barbour & Stevenson (1998) weisen darauf hin, dass diese Tendenz insbesondere in bestimmten regionalen Varietäten (wie dem Mitteldeutschen) verbreitet ist, jedoch auch als allgemeines Merkmal weniger betonter oder schnellerer Redeformate auftritt.

Ein weiterer Faktor ist die Variabilität der Vokaldauer und die damit korrelierende Qualität. Während im Standarddeutschen eine strikte Opposition zwischen dem langen, gespannten /y:/ und dem kurzen, ungespannten /Y/ besteht, zeigen Studien wie die von O'Brien & Smith (2010), dass im Gebrauchsstandard Prozesse des *Fronting* oder der Zentralisierung auftreten können. Diese Dynamiken führen dazu, dass der artikulatorische Zielpunkt nicht immer die extreme vordere Position erreicht, was die Identifikation der Kategorie für Lerner:innen erschweren kann.

Abschließend lässt sich feststellen, dass /y:/ trotz eines etablierten phonematischen Status eine erhebliche artikulatorische Variabilität aufweist, die primär diatopisch und diaphasisch bedingt ist.

3.1.3 Die Vokale [ə] und [ɪ] in regionalen Varietäten des Italienischen

Im Gegensatz zum deutschen Phonemsystem ist der zentrale Vokal /ə/ kein Bestandteil des italienischen Standards bzw. des *Neo-standards*, sondern stellt eine spezifisch regionale Realität dar. Das prägende Phänomen ist hierbei die Vokalreduktion in unbetonter Position, die primär für die Dialekte der zentralen und südlichen Makroregionen charakteristisch ist. Während im Norden nur vereinzelte Belege existieren – wie etwa im Pontremolese (vgl. Cavarani, 2015) –, verläuft die systematische südliche Grenze der Reduktion finaler unbetonter Vokale auf /ə/ entlang der Isoglosse Cetraro-Bisignano-Melissa (vgl. Savoia & Maiden, 1997; Trumper, 1997).

Es handelt sich hierbei um das sprachliche Gebiet der hochsüdlichen Dialekte, die sich von den extrem südlichen Dialekten vor allem durch den phonologischen Prozess der systematischen Reduktion unbetonter Vokale in finaler Position unterscheiden. Dieses Phänomen manifestiert sich in acht italienischen Regionen, darunter die gesamte Campania, Molise und Basilikata. Es endet an der Grenze zum Salento (südliches Apulien) und dem südlichen Kalabrien, wo Abschwächungen zu [ə] zwar auftreten, jedoch als unsystematische Erscheinung und nicht als phonologischer Prozess einzustufen sind (vgl. Ledgeway, 2016, S. 251).

Die Relevanz dieser regionalen Verteilung wird durch die Daten des *Istituto Nazionale di Statistica* (ISTAT, 2017) zur Sprachverwendung im Jahr 2015 untermauert: Etwa 32,2 % der Bevölkerung sprachen in der Familie sowohl Italienisch als auch Dialekt. In den für diese Arbeit relevanten Altersgruppen der 6- bis 24-Jährigen sowie der 25- bis 34-Jährigen sprachen jeweils 25,4 % und 32,3 % im familiären Umfeld beide Sprachformen. Im Freundeskreis nutzten 25,4 % der Jüngeren und 32,5 % der Älteren (25–34 Jahre) sowohl Italienisch als auch Dialekt. Die höchsten Zahlen der Dialektsprechenden fanden sich neben dem Nordosten (Trentino-Südtirol, Venetien) vor allem im Zentrum und Süden, mit Spitzenwerten in Kampanien, Basilikata, Kalabrien und Sizilien.

Obwohl die Mehrheit in formellen Kontexten ausschließlich Italienisch verwendet (77,5 %), weisen diese Daten darauf hin, dass bei einem beachtlichen Teil der Sprecher:innen – insbesondere in Regionen mit systematischer Vokalzentralisierung – bereits Kenntnisse von Sprachsystemen vorhanden sind, die deutlich vom Standard abweichen. Die Berücksichtigung dieser (passiven und eventuell auch aktiven) Dialektkompetenz bietet einen entscheidenden Ausgangspunkt für die Annahme, dass artikulatorische Substrate existieren, welche die Realisierung der deutschen Vokale im Lernprozess beeinflussen können.

Neben dem Schwa-Laut stellt auch der vordere hohe gerundete Vokal /y:/ im italienischen Kontext eine rein regionale bzw. dialektale Realität dar. Während gerundete Vorderzungenvokale dem Standarditalienischen völlig fremd sind, belegt Cavarani (2015) deren Präsenz in der Region Lunigiana im Nordwesten der Halbinsel. Insbesondere im Pontremolese, einer Variante an der Schnittstelle zwischen Ligurien, der Toskana und der Emilia, finden sich Phoneme wie /y/, /ø/ und [œ]. Der phonologische Status dieser Laute ist jedoch aufgrund des geringen funktionalen Kontrastwertes (insbesondere zwischen /y/ und /ø/) nicht eindeutig geklärt.

Laut Cavarani (2015) ist das Vorkommen gerundeter Vokale primär für die lombardischen Dialekte kennzeichnend. Dass sich das Pontremolese in dieser Hinsicht von anderen Dialekten der Lunigiana unterscheidet, ist somit höchstwahrscheinlich auf den spezifischen Einfluss dieser lombardischen Systeme zurückzuführen (vgl. Cavarani, 2015, S. 20). Einen besonderen Fall des Sprachkontakts stellt zudem die Situation in Südtirol dar. In diesem Gebiet der nördlichen Makroregion ist das Deutsche massiv präsent, was zu einer spezifischen Schichtung im Repertoire führt, die sowohl diatopische Varianten als auch spezifische Varianten des Standards umfasst (vgl. Ciccolone, 2016, S. 29).

Die Untersuchung dieser diatopischen Variationen verdeutlicht eine fundamentale Divergenz zwischen der soziolinguistischen und dialektologischen Realität in Deutschland und Italien. Die unterschiedlichen historischen und sozialen Hintergründe beider Sprachräume haben zu entsprechend differenzierten Resultaten hinsichtlich der gegenwärtigen Aussprache der jeweiligen Standard- und Regionalvarietäten geführt.

Die Einbeziehung dialektologischer Hintergründe und der ISTAT-Daten (2017) dient dem Ziel, das sprachliche Profil der Proband:innen differenzierter zu erfassen. Da die vorliegende Arbeit als systematischer Beitrag zum Kontakt zwischen den jeweiligen Standardsystemen des Deutschen und Italienischen konzipiert ist, liegt der primäre Fokus auf dem Vergleich dieser Normvarietäten.

Gleichwohl wurden die im Forschungsprozess gewonnenen Erkenntnisse über das dialektale Substrat der Teilnehmer:innen als ergänzender Kontextfaktor berücksichtigt. Diese explorative Herangehensweise ermöglicht es, potenzielle artikulatorische Synergien oder Interferenzen (etwa bei /ə/ oder /y:/) fundierter einzuordnen, ohne den eigentlichen Fokus der Untersuchung zu verschieben. Damit wird ein Aspekt dokumentiert, der zwar nicht im unmittelbaren Zentrum der Fragestellung steht, jedoch die Komplexität der Datenbasis verdeutlicht und wertvolle Anhaltspunkte für zukünftige, spezifischere Studien zum Einfluss der Dialektkompetenz bieten kann.

3.2 Artikulatorische Zielwerte und Transferhypothesen

Um die experimentelle Untersuchung vorzubereiten, ist es notwendig, von der rein abstrakten Ebene des Systemvergleichs zu einer empirisch fundierten Erwartungshaltung überzugehen. Der hier gewählte kontrastive Ansatz steht in der strukturalistischen Tradition (Lado, 1957), die den Vergleich phonologischer Systeme als Mittel zur Visualisierung von Symmetrien und Dysymmetrien (Missaglia & Sendlmeier, 1999, S. 75; vgl. Alber, 2022) auf der Basis artikulatorisch-akustischer Merkmalen nutzt. Obwohl dieser Ansatz primär auf einer abstrakten, phonologischen Ebene verbleibt, bietet er die entscheidenden Ausgangspunkte für die Formulierung von Hypothesen im experimentellen Bereich.

Wie Missaglia & Sendlmeier (1999) berichten, liefert Mioni (1976, hier zitiert nach der Übersetzung von Missaglia & Sendlmeier, 1999) einen fundierten Überblick über die Schwierigkeiten der Lerner sowohl auf der Wahrnehmungs- als auch auf der Produktionsebene:

Das deutsche Vokalsystem bereitet Italienern [in Produktion und Perzeption] eine Reihe von Schwierigkeiten: Es bedient sich des Merkmals Länge, das dem Italienischen fremd ist, da im Italienischen die Länge mit der Silbenstruktur verbunden ist; es macht weiterhin einen größeren Gebrauch von Merkmal Gespanntheit als das Italienische (z.B. gespanntes [i:] – ungespanntes zentralisiertes [ɪ] usw.), und es hat eine Reihe vorderer gerundeter Phoneme (/y:/, /ʏ/, /ø:/, /œ/), die im Italienischen völlig fehlen. Die ersten zwei Merkmale sind miteinander verknüpft; sie können mit der folgenden Formulierung zusammengefaßt [sic in Missaglia & Sendlmeier, 1999, A.D.L.] werden: Das Merkmal Länge ist mit dem der geringeren Öffnung und der geringeren Zentralisierung verbunden; das Merkmal Kürze mit größerer Öffnung und größerer Zentralisierung.“ (Originaltext in Mioni, 1976, S. 111-112, Übersetzung von F. M., Missaglia & Sendlmeier 1999, S. 74).

Die Schwierigkeiten betreffen dabei nicht nur die Monophthonge, sondern erstrecken sich auch auf die Diphthonge (Mioni, 1976; vgl. Missaglia & Sendlmeier, 1999; Maturi, 2014). Eine empirische Bestätigung dieser theoretischen Annahmen findet sich in der akustischen Visualisierung der Vokalrealisierungen. In Abbildung 3 (in Anlehnung an Missaglia & Sendlmeier, 1999) wird die spezifische Besetzung des Vokalraums durch italienische Lernerinnen (IM) im Vergleich zu deutschen L1 Sprecher:innen (DM) deutlich.

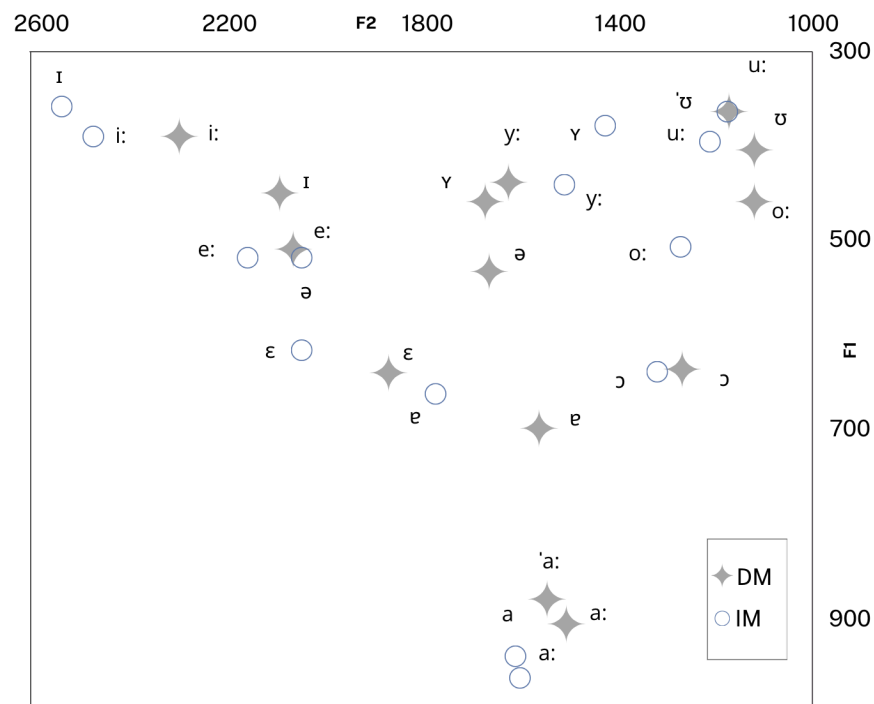


Abbildung 3: F2/F1-Grafik der deutschen Vokale, realisiert durch Lernende mit L1 Italienisch, in Anlehnung an Missaglia & Sendlmeier (1999, S. 92, Abb. 5). Die im Original grafisch uneindeutigen Symbole für den Wortakzent (angegeben als "I") wurden zur besseren Lesbarkeit und zur Angleichung an den IPA-Standard ersetzt (z. B. ['a] und ['u]). Eigene Darstellung.

Bei der Betrachtung der Daten (gelesene Sprache) zu den IM fällt insbesondere auf, dass der zentrale Bereich des Vokaldreiecks nicht besetzt ist: Das Phonem /ə/, das in der L1-deutschen Aussprache eine typisch zentrale Position einnimmt, wird erheblich nach vorne verschoben und nimmt nahezu die Position – und somit die Artikulation – eines deutschen /e:/ ein. Eine Verschiebung des Artikulationsortes ist ebenfalls bei den hinteren Vokalen auffällig: In diesem Fall erfolgt die Verschiebung in die entgegengesetzte Richtung, nämlich nach hinten, wo der italienische Vokalraum traditionell stärker besetzt ist.

Dort, wo sich Korrespondenzen herstellen lassen, entsprechen die visuellen Befunde von Missaglia & Sendlmeier (1999) den von Maturi (2014, S. 154) beschriebenen häufigsten phonetischen Assimilationsmustern in der italienischen Aussprache des Deutschen:

- [y] > [ju] oder [u]
- [ø] > [e]
- [œ] > [ɛ]
- [ə] > [e]
- [ʊ] > [i] oder [u]

Diese empirischen Anhaltspunkte dienen im weiteren Verlauf dieser Arbeit als Referenz, um die Realisierungen der untersuchten Probandengruppe vor dem Hintergrund etablierter CLI-Muster einzuordnen.

3.3 Standarddeutsch für Deutschlerner:innen

Angeichts der zuvor beschriebenen Variabilität zentraler deutscher Phoneme wie /ə/ und /y:/ stellt sich zwangsläufig die Frage, wie diese Variabilität im Kontext von Deutsch als Fremdsprache (DaF) behandelt wird. Die Natur des heutigen Standarddeutsch, seine phonetische Charakterisierung sowie seine Rezeption sind nicht nur zentrale Gegenstände der Variationslinguistik, sondern bilden die (oft problematische) Grundlage für die Aussprachevermittlung.

Die Suche nach einer allgemeingültigen Norm führt unweigerlich zu Ansätzen, die versuchen, die regionale Variabilität zugunsten einer überregionalen Standardisierung zu abstrahieren. So bezieht sich etwa Hollmach (1998) einen zentralen Stellenwert ein (vgl. Paschke 2009, S. 212). Demnach bezieht sich der geographische Kontext seiner Beschreibung auf das vereinigte Deutschland, wobei der Ausgangspunkt der sachlich-neutrale Sprachgebrauch ist, und die Allgemeingültigkeit basiert auf globalen und auditiven Akzeptanzbewertungen, die über regionale und soziale Eigenheiten hinausgehen (vgl. Paschke, 2009). In diesem Zusammenhang stellt sich die Frage, welches Standarddeutsch für Deutschlerner:innen als Referenz dient. Berücksichtigt man zudem die Schwierigkeit, klare Grenzen zwischen Gebrauchsstandard und Standardsprache zu ziehen – insbesondere vor dem Hintergrund der Massenkommunikation und der fließenden Übergänge innerhalb der diastratischen Dimension –, erscheint diese Fragestellung von besonderer Relevanz.

Eine Durchsicht der neuesten Literatur zur Ausspracheerziehung zeigt eine immer deutlicher werdende Überschreitung der Grenzen von Phonetik und Phonologie in den DaF-Bereich, was darauf hinweist, wie der Fortschritt in den Bereichen der experimentellen Phonetik und Phonologie, insbesondere im Bereich der Studien zu *Second und Third Language Acquisition*, positiv mit den Entwicklungen in der Sprachdidaktik einhergehen kann. Gegenstand dieser Untersuchungen im Rahmen dieses fachlichen Synkretismus sind die normabweichenden Ausspracheschwierigkeiten. Generell lassen sich zwei Hauptursachen für die Abweichung von der Norm identifizieren: 1) der Effekt des negativen Transfers (aus der L1 oder anderen Sprachen des Repertoires); 2) ein Imitationseffekt eines Modells in der Umgebung der Lerner:innen, das von der Norm abweicht. Weit entfernt von einer rein präskriptiven Haltung verfolgt dieser fachliche Synkretismus das Ziel, ein Orientierungsmodell für die Aussprache des Deutschen bereitzustellen, insbesondere für das vokalische System. Im Mittelpunkt steht dabei die Verständlichkeit als primäres, jedoch nicht hinreichendes Kriterium (vgl. Richter, 2011). Die Literatur der letzten dreißig Jahre zeigt jedoch eine wichtige Perspektivverschiebung, die durch

den Vergleich zwischen Diez (1999)³, Richter (2011) und Dietz (2015) exemplifiziert wird. Während ersterer, hauptsächlich auf die Didaktik des Deutschen für ungarische Lerner:innen fokussiert, offen erklärt, wie auch die folgenden zwei, die Notwendigkeit, die Deutschdidaktik mit phonetischen Elementen zu integrieren, ist ebenso explizit in ihrem «theoretischen Minimum» ihr Verzicht auf die «systematische Behandlung von dialektalen, soziolektalen, emotionalen oder umgangssprachlichen Varianten» (Dietz & Tronka 1998, S. 117). In Richter (2011) gewinnt der neue Begriff der *Interkulturellen Germanistik* an Bedeutung, der einerseits auf eine neue Aufmerksamkeit für Unterschiede sowie Gemeinsamkeiten zwischen Kulturen und Sprachsystemen hinweist; andererseits bleibt die Frage der Referenznorm der Zielsprache wenig vertieft – oder vielleicht als nicht entscheidend angesehen. Das Hauptziel in Richter (2011) ist daher die oben erwähnte Verständlichkeit. Dem entgegen setzt sich Dietz (2015) ab, der «die zu theoretisch (phonologisch-phonetisch) ausgerichtete oder ganz ausbleibende Vermittlung aussprachedidaktischen Wissens an zukünftige Lehrkräfte» betont (Dietz, 2015, S. 1). Er hebt hervor, dass «Verständlichkeit nicht das alleinige Kriterium» (2015) sei und unterstreicht die Bedeutung der Vermittlung von metalinguistischem Wissen an Lerner:innen, was unter anderem das Bewusstsein für phonetische und phonologische Phänomene umfasst (wie zum Beispiel Phänomene der Vokal-Elision und deren diastratische, diatopische und diaphasische Dimension). Die Entwicklung der metalinguistischen Dimension, zuerst bei den Lehrkräften und dann bei den Lerner:innen, fördert ein funktionales Verständnis für die pragmatische Nutzung und das Verständnis der deutschen Sprache, das vor allem bei der Verwendung der Zielsprache bei der Vermeidung von Verzerrungen der Ziel-Laute von Bedeutung ist. Die schrittweise Etablierung der Soziophonetik und Pragmatik im Bereich DaF kann jedoch noch als in einem frühen Entwicklungsstadium betrachtet werden: Es sind noch zahlreiche Experimente erforderlich, und es gibt viele Fortschritte zu erzielen. Aus der Literatur ergibt sich daher Folgendes: Die wissenschaftliche Auseinandersetzung mit der Normfrage weist eine deutliche Asymmetrie auf. Während die Problematik in der soziolinguistischen und dialektologischen Forschung intensiv diskutiert wird, bleibt die Definition der für Lerner:innen relevanten sprachlichen Norm in der Sprachvermittlungsforschung weitgehend unberücksichtigt.

3.4 Forschungsstand zur Wahrnehmung von deutschen Vokalen durch LX-Lerner:innen

Die jüngsten Studien zur Wahrnehmung deutscher Vokale bei erwachsenen Lerner:innen im formalen Kontext decken ein Spektrum ab, in dem Untersuchungen mit Englisch (Amerikanisch)

³ Es sei darauf hingewiesen, dass die zitierte Quelle online unter der Adresse https://opus.bibliothek.uni-augsburg.de/opus4/frontdoor/deliver/index/docId/67272/file/Dietz_Ausspracheschulung.pdf (letztlich abgerufen am 31.05.2025) abrufbar ist, wobei die Autoren als Dietz, G. und Tronka, K. aufgeführt sind. In der Bibliographie dieser Arbeit wird der Eintrag jedoch gemäß der Zitierweise in der offiziellen Publikationsliste von G. Dietz wiedergegeben, die unter folgender Adresse einsehbar ist: <https://www.uni-augsburg.de/de/fakultaet/philhist/professuren/germanistik/deutsch-als-zweit-fremdsprache-und-seine-didaktik/team/gunther-dietz/profil/> (letztlich abgerufen am 31.05.2025).

als Ausgangssprache überwiegen. In dieser Übersicht werden experimentelle Studien zu Deutsch als LX (L2/L3) betrachtet, wobei Arbeiten mit rein didaktischem Fokus ausgeschlossen bleiben.

In dieser Übersicht wird Deutsch als "Fremdsprache" (oder LX) eingestuft, da das Erlernen dieser Zielsprache formal und strukturiert (schulisch) erfolgt. Aktuelle Studien fokussieren sich hauptsächlich auf vier große Konstellationen von Ausgangssprachen (L1) beim Erlernen des Deutschen: 1) Germanische Sprachen (Englisch (Amerikanisch)-Deutsch), 2) Slawische Sprachen (Polnisch-Deutsch), 3) Türkische Sprachen (Türkisch-Deutsch) und 4) Romanische Sprachen (Italienisch-Deutsch). Die vorliegende Arbeit konzentriert sich besonders auf die vierte Gruppe. Entsprechend liegt ein besonderer Schwerpunkt auf den spezifischen Wahrnehmungsschwierigkeiten, die Lerner:innen mit L1 Italienisch beim Erwerb von Deutsch als Fremdsprache erfahren können.

3.4.1 Ausgangssprache Englisch

Studien zum sprachlichen Kontakt zwischen germanischen Sprachen wie Englisch und Deutsch haben gezeigt, dass sowohl Lerner:innen als auch ungeübte Hörer:innen Schwierigkeiten haben, das Merkmal [±gespannt] (z. B. /y:/ vs. /ʏ/, /i:/ vs. /ɪ/) sowie Kontraste zwischen gerundeten Vokalen (z. B. im Deutschen) zu unterscheiden. Die Relevanz dieser Forschungsthemen ergibt sich aus einem kontrastiven Ansatz, der die wechselseitige Beeinflussung der phonologischen Systeme von Englisch und Deutsch untersucht: Da das englische Vokalsystem keine vorderen gerundeten Monophthonge enthält, muss im Aufbau des interlingualen phonologischen Systems eine neue phonologische Kategorie geschaffen werden. Dieser Prozess sollte nicht nur aus der Perspektive einer „Sprechergruppe“ betrachtet werden, sondern auch die idiosynkratische Natur individueller Lernprozesse berücksichtigen, was unterschiedliche Assimilationsmuster zur Folge hat.

Mayr & Escudero (2010) untersuchten die Wahrnehmung von drei Kontrasten, die aus sechs gerundeten deutschen Vokalen (/y:/-/u:/, /ʏ/-/ɤ/, /ø:/-/œ/) bestehen. Diese Kontraste wurden in einzelnen Reizen im konstanten Kontext /bVt/ präsentiert. Obwohl die Studie eine erhebliche Uneinheitlichkeit in den von den Testpersonen verfolgten wahrgenommenen „Pfaden“ aufzeigte, kristallisierten sich zwei Haupttrends heraus. Einerseits ordnen Hörer:innen die LX-Kontraste einer einzigen L1-Vokal-Kategorie zu. Andererseits neigen sie dazu, multiple Assoziationen zu bilden. Die festgestellten perzeptuellen Muster scheinen zudem die Qualität der Diskriminationsleistung von LX-Lauten gut vorherzusagen. Dies stützt die Annahme des L2LP-Modells (Escudero, 2009), wonach interlinguistische perzeptuelle Muster die Lernentwicklung der LX bestimmen (vgl. Mayr & Escudero, 2010).

Die gerundeten vorderen Vokale des Deutschen, charakterisiert durch die Merkmale [+gerundet, +vorne], gelten im Vergleich zum englischen Vokalsystem als „neue Phoneme“. Basierend darauf stellte Schultheiss (2008) die Hypothese auf, dass diese Phoneme (/y:/, /ʏ/, /ø:/ und

/œ/) von Lerner:innen entweder ihren ungerundeten vorderen englischen Entsprechungen oder ihren gerundeten hinteren Gegenstücken zugeordnet werden. Zur Überprüfung dieser Hypothese wurde eine Studie mit 118 Universitätsstudenten durchgeführt, die in vier Gruppen unterteilt wurden, um den Einfluss von Deutschkenntnissen und immersiver Erfahrung zu untersuchen: Anfänger ohne Deutschkenntnisse, Studenten mit Sekundarstufen-Deutschkenntnissen nach dem 3. Semester (ohne immersive Erfahrung), Studenten mit Sekundarstufen-Deutschkenntnissen nach dem 5. Semester (ohne immersive Erfahrung) sowie Studenten mit Sekundarstufen-Deutschkenntnissen nach dem 5. Semester, die zusätzlich mindestens 12 Monate immersive Erfahrung gesammelt hatten.

Die Analyse konzentrierte sich auf die vorderen Vokale (/i:/, /ɪ/, /y:/, /ʏ/, /e:/, /ø:/, ɛ/, /ɛ:/, /œ/), zentralen Vokale (/a/, /a:/) und hinteren Vokale (/ɑ:/, /ʊ/, ɔ:/, /ɔ/) des Deutschen, wobei ein besonderes Augenmerk auf die vorderen gerundeten Vokale gelegt wurde. Der verwendete Test umfasste Pseudowörter und reale Wörter mit der Struktur /b/C/ə/ (stimmhafter bilabialer Plosiv, gefolgt von einem Konsonanten /n/, /r/, /ʃ/ oder /t/, und /ə/).

Entgegen den Erwartungen zeigte sich bei den Ergebnissen zur Vokalzuordnung, dass alle gerundeten deutschen Vokale, mit Ausnahme von /y:/, mehrheitlich den englischen Phonemen /ʊ/ oder /ə/ zugeordnet wurden. Die Phoneme /y:/ und /ʏ/ wurden zwar tendenziell ihren englischen Entsprechungen /u/ und /ʊ/ zugewiesen, jedoch ordneten interessanterweise 7,6 % der Hörer das ungespannte /ʏ/ der englischen Entsprechung /ʊ/ zu. Dies deutet darauf hin, dass das Merkmal vorne/hinten stärker berücksichtigt wurde als das Merkmal [±gerundet]. Da diese Zuordnung jedoch nicht für das gespannte /y:/ beobachtet wurde, lässt sich zudem eine gewisse Relevanz des Merkmals gespannt/ungespannt feststellen.

In Bezug auf den Einfluss von Erfahrung auf die Wahrnehmung zeigte sich eine gewisse Variabilität, die vom wahrgenommenen Vokal und teils vom phonemischen Kontext abhing. Generell wiesen Hörer:innen mit größerer Erfahrung eine höhere Konsistenz bei der Auswahl des entsprechenden englischen Vokals auf, was insbesondere für die „neuen Phoneme“ /ʏ/, /œ/ und /ø:/ galt. Die Gruppen 1 und 2 sowie die Gruppen 3 und 4 zeigten dabei ein ähnliches Wahrnehmungsverhalten. Im Gegensatz dazu hatte die An- oder Abwesenheit von Vokallänge keinen Einfluss auf die Wahrnehmung (vgl. Schultheiss, 2008, S. 83–84).

Das Merkmal [±gespannt] wird als besonders relevant in der Studie von Polka (1995) beschrieben, die die Relevanz dieses Merkmals für die Wahrnehmung deutscher Vokale untersuchte. Zehn Zuhörer:innen ohne Deutschkenntnisse wurden den Kontrasten /y:-u/ und /ʏ-ʊ/ in einer AXB-Diskriminationsaufgabe ausgesetzt. Trotz fehlender Vorerfahrung zeigten die meisten Zuhörer:innen eine hohe Genauigkeit bei der Diskriminierung der ungespannten Vokale. Bei den gespannten Vokalen erreichten *alle* Teilnehmer:innen hohe Akkuratheitswerte, was die Annahme stützt, dass die *category goodness* für Vokale mit diesem Merkmal stärker ausgeprägt ist.

Die Studie von Strange et al. (2009) beleuchtet den signifikanten Zusammenhang zwischen der Präzision bei der Unterscheidung von LX-Lauten und deren Verwendungskontext. Die Untersuchung konzentrierte sich auf die Wahrnehmung von Vokalen der norddeutschen Standardsprache und des Pariser Französisch. Daran nahmen elf Hörer:innen teil, die weder Deutsch noch Französisch beherrschten und auch keine andere Sprache mit gerundeten vorderen Vokalen kannten. Analysiert wurde jedoch ausschließlich die Wahrnehmung der deutschen Vokale. Die Analyse der Muster der Wahrnehmungsassimilation und der *category goodness rating* ergab, dass es den Proband:innen schwerer fiel, auf die phonetischen Informationen des spezifischen Kontextes zuzugreifen, wenn die Vokale in einem Satzkontext präsentiert wurden. Im Satzkontext wurden die norddeutschen Standardvokale /u:/ und /y:/ beide als gleichermaßen gute Exemplare (*equally good exemplars*) der amerikanischen Englisch-Vokale /u:/ bewertet. Im isolierten Kontext hingegen wurden die hinteren gerundeten Vokale der norddeutschen Standardsprache als bessere Exemplare für die hinteren gerundeten Vokale des amerikanischen Englisch angesehen als die vorderen Vokale der norddeutschen Standardsprache.

Die Studie von Smith & Baker (2011) untersuchte die Rolle der erlernten deutschen Varietät (Österreichisches Deutsch vs. Norddeutsche Standardsprache) beim Erlernen der Vokalkategorien /i:/- /y:/- /u:/, wobei sie sich auf die unterschiedliche Salienz des Merkmals [±gerundet] konzentrierte. Hintergrund dieser Untersuchung ist die Beobachtung, dass im Österreichischen Deutsch das Merkmal [±gerundet] aufgrund von Entrundungsphänomenen weniger stark ausgeprägt ist. Dies reduziert die akustischen Hinweise zur Unterscheidung von /i:/ und den gerundeten Vokalen (/y:/, /u:/), könnte jedoch gleichzeitig den akustischen Gegensatz zwischen /y:/ und /u:/ verstärken. Die Methode umfasste sowohl Wahrnehmungs- als auch Produktionsdaten, wobei der vorliegende Abschnitt die Wahrnehmungsdaten präsentiert. Die Experimentalgruppe bestand aus 10 Lerner:innen, die mindestens 16 Monate in einem immersiven Kontext in Süddeutschland (Bayern) und Österreich verbracht hatten. Da die Mehrheit der Teilnehmenden angab, die meiste Zeit in Österreich verbracht zu haben, wurde Österreichisches Deutsch als ihre Zielvarietät betrachtet. Die Ergebnisse und deren Interpretation zeigten, dass die Lerner:innen des Österreichischen Deutschen tendenziell präzisere Leistungen erzielten als die Lerner:innen der Norddeutschen Standardsprache. Für diese Ergebnisse wurden zwei mögliche Erklärungen aus den soziolinguistischen Daten abgeleitet: Erstens könnte die Diskriminationsfähigkeit durch die höhere Quantität und Qualität des Inputs gefördert worden sein, da die Lerner:innen des Österreichischen Deutschen zusätzlich zum Österreichischen Deutsch mehreren süddeutschen Varietäten ausgesetzt waren (nach Bradlow & Pisoni, 1999). Zweitens wurde vermutet, dass eine größere Distanz zum lokalen Modell zugunsten des Hochdeutschen die Tendenz zu einer präziseren Leistung in der Experimentalgruppe begünstigte, da die Mehrheit der Lerner:innen des Österreichischen Deutschen im soziolinguistischen Fragebogen angab, nicht wie die Einheimischen sprechen zu

wollen, im Gegensatz zur Kontrollgruppe der Norddeutschen Standardsprache, die neutral antwortete.

Ein weiterer wichtiger Faktor, der das Deutsche vom Englischen unterscheidet, ist das Verhältnis (*weighting*) zwischen Vokalquantität und -qualität. Wie Predeck (2022)⁴ berichtet, stützt sich die L1-Wahrnehmung der langen und kurzen deutschen Vokale in erster Linie auf das Merkmal der Dauer und erst in zweiter Linie auf das der Qualität. Mittels eines *Identification-* und *Rating Task* wird die Natur des *weightning* bei 11 ZuHörer:innen ohne Erfahrung im Deutschen und mit Amerikanisch-Englisch als L1 untersucht.⁵ Die Ergebnisse der Untersuchung legen nahe, dass für die untersuchte Gruppe das Merkmal der Dauer nicht entscheidend für die Diskriminierung der langen und kurzen deutschen Vokale ist. Vielmehr scheint es, dass die Zuhörer:innen stärker auf die spektralen Merkmale der Vokale angewiesen sind. Dies legt nahe, dass die Gewichtung in Abhängigkeit von der Ausgangssprache unterschiedlich konfiguriert sein kann. Im speziellen Fall des amerikanischen Englisch, in dem das Merkmal der Länge keinen phonemischen Status hat, erlangt die Vokalqualität eine höhere Relevanz im Vergleich zu Zuhörer:innen mit L1 Deutsch (vgl. Predeck, 2022).

Für die Hypothese einer intrinsischen Verbindung zwischen Wahrnehmung und Produktion sprechen die Ergebnisse von Stratton (2025), die sich auf die Wahrnehmung der Kontraste /y:/-/u:/, /ʊ/-/ʏ/, /ɪ/-/y:/ durch amerikanisch-englische L1-Hörer:innen konzentrieren, die entweder explizites (n = 16) oder implizites (n = 13) Training erhalten hatten. Die Ergebnisse liefern Evidenz für das Potenzial des expliziten phonetischen Trainings im Vergleich zum impliziten Training oder dessen vollständigem Fehlen. Die Teilnehmer:innen, Student:innen des Deutschen als Fremdsprache, die einem expliziten Training unterzogen wurden, zeigten eine Verbesserung von 19 % in der Diskrimination neuer Laute sowie eine Steigerung ihrer Leistung in der perzeptiven Kategorisierung um 14 % (vgl. Stratton, 2025).

3.4.2 Ausgangssprache Polnisch

Die Kategorie der Studien zur Vokalwahrnehmung des Deutschen in der Sprachkonstellation Polnisch-Deutsch wird durch die Studien Nimz & Khattab (2015) und Nimz (2022) repräsentiert. Der rote Faden, der diese Studien verbindet, wird durch das Interesse an der Wahrnehmungssensibilität (und bei Nimz, 2022 auch an der Produktionssensibilität) für die Vokallänge der deutschen Vokale dargestellt, die im Polnischen nicht phonologisch ist.

In Nimz & Khattab (2015) geht dieses Interesse mit einem Fokus auf die Rolle der Rechtschreibung und damit an der Interaktion zwischen phonologischem Lernen und lexikalischen Repräsentationen einher, mit der Frage, ob letztere dazu beitragen, phonologisches Wissen zu festigen. Die hier vorgestellte Perzeptionsstudie wird in einem größeren Kontext, der

⁴ In der Studie von Predeck (2022) nahmen L1-amerikanisch-englische Hörer die deutschen Vokale als „different native categories“ wahr, mit Ausnahme der Paare /ɛ:/-/ɛ/ und /u:/-/ʊ/.

⁵ Predeck (2022, S. 59) berichtet, dass zwei der Teilnehmer:innen Erfahrung mit anderen Sprachen neben Englisch hatten.

auch ein Produktionsexperiment umfasst, in Nimz (2022) beschrieben. In beiden Studien finden wir also eine Population von 20 Sprecher:innen mit vergleichbarem Sprachniveau (fortgeschritten, zwischen C1 und C2 nach dem Gemeinsamen Europäischen Referenzrahmen für Sprachen, GER). Ausgangspunkt ist wiederum der kontrastive Ansatz, der die polnische Sprache, deren phonologisches System nicht nach der Länge unterscheidet, auf der einen Seite und die deutsche Sprache, die dieses Merkmal eindeutig phonematisch nutzt, auf der anderen Seite stellt. Die Schwierigkeiten der Lerner:innen wurden in einem Diskriminationsexperiment mit Pseudowörtern und Stimuli, die immer im gleichen phonologischen Kontext standen (aus artikulationsökonomischen Gründen bilabial), und in einem Bewertungstest mit echten Wörtern aus dem deutschen Lexikon getestet. Die Ergebnisse zeigten, dass die Lerner:innen tatsächlich größere Schwierigkeiten bei der Diskrimination der Vokallänge als bei der Unterscheidung der Qualität (gespannt/ungespannt) aufwiesen. Diese Überprüfung wurde durch eine Manipulation der Stimuli ermöglicht, die der in Sendlmeier (1981) verwendeten Methode ähnelt. Ein Vergleich zwischen den Aufgaben zeigte, dass die Diskrimination der Vokallänge im *Judgement Task* genauer war. Allerdings konnte die statistische Analyse keinen signifikanten Vorteil der orthographischen Komponente gegenüber den phonologischen Fähigkeiten nachweisen. Daher wurde als mögliche Ursache der metalinguistische Hintergrund der Teilnehmer:innen in Betracht gezogen: Da die meisten von ihnen keine metalinguistischen Kenntnisse über die Rolle der Vokallänge im deutschen Vokalsystem besaßen, ist es wahrscheinlich, dass das lexikalische Wissen nicht hinreichend gefestigt war, um die Entwicklung phonologischer Kompetenzen zu unterstützen.

In dem Werk von Nimz (2022), das im Vergleich zu Nimz & Khattab (2015) detailliertere Informationen zu dieser Wahrnehmungsstudie liefert, wird neben den bereits erwähnten Ergebnissen eine interessante Beobachtung gemacht. Einerseits wurde die Hypothese bestätigt, dass polnische Lerner:innen Schwierigkeiten bei der Wahrnehmung der Vokallänge im Deutschen haben. Andererseits zeigte sich jedoch, dass auch die Kontrollgruppe mit L1-Deutsch erhebliche Schwierigkeiten hatte, insbesondere bei den mittleren Vokalen: Während die Genauigkeit bei der Wahrnehmung der Länge tiefer Vokale bei 94 % lag, erreichte sie bei mittleren Vokalen lediglich 69 %. Der unterschiedliche Einfluss der Vokallänge auf tiefe im Vergleich zu mittleren Vokalen wird zudem in Sendlmeier (1981) bestätigt. Hinsichtlich der Wahrnehmung prototypischer Vokale (ohne Manipulation von Länge und Qualität) wurden die größten Schwierigkeiten beim Kontrast /a:/-/a/ festgestellt (Genauigkeit von 59 %). Dieses Ergebnis lässt sich gemäß Flege (1995) als Effekt der *Equivalence Classification* interpretieren, wonach LX-Vokale auf der Grundlage einer perzeptuellen Ähnlichkeit den L1-Vokalen zugeordnet werden. In ähnlicher Weise könnte gemäß Best und Tyler (2007) von einer *Single Category Assimilation* gesprochen werden, wobei beide Ansätze die Wahrnehmungsschwierigkeiten durch ein Ähnlichkeitsverhältnis erklären. Im Gegensatz dazu betrachtet Escudero (2009) /a:/ als ein neues Phonem, dem ein Merkmal (die Länge)

zugeordnet wird, die zuvor nicht im phonologischen System der L1 kodiert war. Obwohl Escudero die Möglichkeit einräumt, dass unter diesen Bedingungen eine neue Kategorie gebildet werden kann, kommt die Autorin zu dem Schluss, dass die Ergebnisse die Vorhersagen des Modells nicht bestätigen (vgl. Nimz, 2022, S. 112).

3.4.3 Ausgangssprache Türkisch⁶

Die Konstellation mit Türkisch als Ausgangssprache ist durch eine reichhaltigere Literatur vertreten als die für Polnisch. Die meisten Beiträge liefert sicherlich Nimz (u.a. Nimz, 2011; Nimz, 2015), die auch eine doppelte Perspektive auf Produktion und Wahrnehmung einnimmt (Nimz, 2011). Beschränkt sich die Darstellung hier auf die Untersuchung der Wahrnehmung, so geht es Nimz auch bei dieser Sprachkombination vor allem um die Wahrnehmung der Qualität und Quantität der deutschen Vokale. Wie im Falle des Polnischen zeigt der kontrastive Ansatz auch im Türkischen, dass das Merkmal Länge im Türkischen nicht die gleiche Bedeutung hat wie im Deutschen. Im Falle des Türkischen ist sein Status sogar eher umstritten: In der Regel hat sie keinen phonologischen Status, bei arabischen und persischen Lehnwörtern ist dieser eher oszillierend (im Arabischen hat die Länge phonologischen Status, im Persischen manchmal nicht, vgl. Nimz, 2011; Nimz, 2015). Da türkische Sprecher:innen den Merkmalen Dorsalität, Höhe und Rundung phonologischen Status zuschreiben und daher lernen müssen, auch das Merkmal Länge perzeptiv zu integrieren, wird vermutet, dass dies ihre größte Schwierigkeit darstellt. In Nimz (2011) werden die Paare /i:/-/ɪ/, /a:/-/a/ und /u:/-/ʊ/ analysiert, wobei, wiederum in Anlehnung an Sedlmeier (1981), die entsprechenden Manipulationen von Länge und Qualität angewendet werden, um die Sensitivität für beide Merkmalen zu testen. Die Teilnehmer sind 22 Deutschlerner:innen, die drei Jahre lang intensiven Deutschunterricht erhalten haben⁷. Die Ergebnisse zeigen, dass die größten Schwierigkeiten tatsächlich bei der Unterscheidung der Stimmqualität liegen, insbesondere beim Kontrast /u:/-/ʊ/ in der Bedingung, in der die Länge manipuliert wurde.

3.4.4 Ausgangssprache Italienisch

Bisher wurde die Wahrnehmung deutscher Vokale durch italienische Lernerinnen lediglich in einer einzigen Studie (Missaglia, 2000) untersucht. Die zentralen Herausforderungen beim Erwerb deutscher Vokale durch italienische L1-Sprecher:innen ergeben sich primär aus zwei Gründen. Zum einen ist das italienische Vokalsystem deutlich kleiner als das deutsche: Das

⁶ In der vorliegenden Übersicht werden die Ergebnisse der Studie von Darcy & Krüger (2012) bewusst ausgeklammert, da diese auf einer Stichprobe türkischsprachiger Kinder basieren, die bereits zwischen dem zweiten und vierten Lebensjahr in einem frühen immersiven Kontext dem Deutschen ausgesetzt waren. Obwohl in der vorliegenden Dissertation das Erwerbsalter nicht primär als Ausschlusskriterium für den zielsprachlichen Erfolg gewertet wird, weichen die Lernbedingungen dieser Probanden signifikant von dem hier untersuchten formalen Fremdspracherwerb im Erwachsenenalter ab. Zudem lässt sich bei Darcy & Krüger (2012) ein Migrationshintergrund der Teilnehmer:innen aufgrund der sozialen Rahmenbedingungen des untersuchten Kontexts – wenngleich nicht explizit thematisiert – nicht ausschließen, was die Vergleichbarkeit mit den übrigen LX-Studien einschränkt. Dennoch bildete Darcy & Krüger (2012) eine wertvolle Referenz für die methodologische Gestaltung der eigenen Wahrnehmungsstudie im folgenden Kapitel.

⁷ Es wird nicht weiter spezifiziert, welches Kompetenzniveau zum Zeitpunkt der Messung vorlag.

Italienische weist 7 Vokale in betonter und 5 in unbetonter Position auf, während das Deutsche 15 bzw. 9 Vokale besitzt (vgl. Missaglia, 2000). Zum anderen liegt dies am Status von Vokallänge und -qualität, da die Vokallänge im Italienischen auf segmentaler Ebene keinen phonematischen Status hat und nicht zur Unterscheidung von Wörtern dient (vgl. Missaglia, 2007). Vielmehr werden Vokale im Italienischen hauptsächlich durch ihre spektrale Qualität unterschieden. Interessanterweise besitzt das Merkmal Länge jedoch auf suprasegmentaler Ebene einen phonologischen Status im Italienischen (vgl. Missaglia, 2007, S. 243). Dies hat zur Folge, dass die Wahrnehmung von L1-Italienisch sehr sensibel auf Längenunterschiede reagiert, was einen potenziellen positiven Transfer ermöglicht (vgl. Missaglia, 2007, S. 309). Experimentelle Ergebnisse belegen entsprechend, dass sich italienische Hörer:innen bei der Unterscheidung deutscher Vokale eher auf die Länge als auf die Qualität verlassen (vgl. Missaglia, 2000; Missaglia, 2007). Da die Qualität deutscher Vokale allerdings durch eine komplexere Kombination akustischer Merkmale bestimmt wird (wie z. B. [±hoch], [±vorne], [±gespannt], [±gerundet]) als im weniger komplexen italienischen System, müssen L1-italienische Sprecher:innen spezielle Wahrnehmungsmuster entwickeln. Ziel dabei ist es, die ihnen zur Verfügung stehenden akustischen Hinweise optimal für die Diskrimination deutscher Vokale zu nutzen. Diese Unterschiede in den Vokalsystemen sind die Ursache für die Entwicklung von Wahrnehmungskategorien und -strategien, die die Grundlage für mögliche Wahrnehmungsdiskrepanzen legt.

Die Studie von Missaglia (2000) setzt genau an diesem Punkt an. Um die Beziehung zwischen Vokallänge und -qualität bei der Wahrnehmung deutscher Vokale durch italienische Lerner:innen zu untersuchen, wurden 45 erwachsene Deutschlernerinnen einem Kategorisierungstest mit manipulierten Stimuli unterzogen. In der Studie werden keine weiteren Angaben zur Versuchsstichprobe gemacht. Konkret wurden 56 betonte Vokale in deutschen Minimalpaaren manipuliert, indem die vollen Perioden in der stationärsten Zone des gesamten langen Vokals eliminiert und die Perioden der kurzen Vokale verdoppelt wurden. Anschließend wurden Minimalpaare mit Vokalen gebildet, die der deutschen Orthophonie widersprechen, d.h. mit gespannten (geschlossenen, aber kurzen) und ungespannten (offenen, aber langen) Vokalen. An dieser Stelle wird lediglich eine Auswahl an Ergebnissen kommentiert, die für die vorliegende Arbeit als besonders relevant erachtet werden. Die langen, gespannten, geschlossenen, nicht zentralisierten Vokale in extremen artikulatorischen Positionen, die einer Verkürzung unterzogen wurden, wurden als die entsprechenden ungespannten, offenen, zentrale Vokale wahrgenommen. Generell zeigte sich in dieser Studie, dass mittlere Vokale die größten Unterschiede in der Wahrnehmung zwischen den Lerner:innen und der Kontrollgruppe mit L1 Deutsch aufwiesen. Der Schwa-Laut wurde in der Studie von Missaglia (2000) nicht berücksichtigt. Was genauer die Wahrnehmung vom manipulierten Vokal [ɛ:] angeht, wurde er sowohl von den italienischen Sprecherinnen als auch bei der Kontrollgruppe der L1 deutsche Sprecher:innen als [ɛ] wahrgenommen, während [e] als [ɪ]. Bei gerundeten palatalen Vokalen

eine Veränderung der Quantität in den meisten Fällen führte dazu, dass die Hörerinnen den jeweils anderen Vokal desselben Oppositionspaars wahrnahmen. Das bedeutet, dass beispielsweise eine Verkürzung eines langen /y:/ überwiegend als dessen kurze Variante /Y/ interpretiert wurde, anstatt zu einer Wahrnehmung als ein völlig anderer Vokal zu führen. Dies legt nahe, dass bei diesen Vokalen die Länge als primäres Unterscheidungsmerkmal innerhalb des jeweiligen Kontrasts fungiert. Dies ist laut Missaglia (Missaglia, 2000, S. 309) auf die ausgeprägte Sensibilität der Versuchspersonen gegenüber der Variation der Vokallänge als akustischem Hinweisreiz für die Diskrimination der Stimuli zurückzuführen.

Zusammenfassend zeigt sich, dass die Wahrnehmung deutscher Vokale bei LX-Lerner:innen stark von der Ausgangssprache geprägt ist, nämlich durch das jeweilige sprachspezifische Wahrnehmungsmuster, das die akustischen Merkmale (wie Länge und Qualität) zur Diskriminierung deutscher Vokale nutzt. Die Untersuchung der Wahrnehmung des gerundeten vorderen Vokals /y:/ wurde in den angeführten Sprachkonstellationen berücksichtigt, während zur Wahrnehmung des zentralen mittleren Vokals /ə/ keinerlei Studien gefunden wurden – ein wesentlicher Ausgangspunkt für die im folgenden Kapitel präsentierte Studie.

3.5 Forschungsstand zur Produktion von deutschen Vokalen durch LX-Lerner:innen

Die Anzahl der berücksichtigten Quellen in diesem Bereich ist, ähnlich wie bei der Übersicht über neuere Studien zur Wahrnehmung deutscher Vokale, gering. Auswahl und Terminologie richten sich nach den gleichen Prinzipien, die im Einleitungsabschnitt zur Wahrnehmung definiert wurden. Die Studien untersuchen die Produktion deutscher Vokale im Kontext von Sprecher:innen aus verschiedenen Sprachfamilien. Konkret wurden Sprachkonstellationen aus germanischen (Englisch/Amerikanisch), slawischen (Polnisch), türkischen (Türkisch) und romanischen Sprachen (Französisch und Italienisch) untersucht. Demnach umfassen die untersuchten Sprachkonstellationen Englisch (Amerikanisch) – Deutsch, Polnisch – Deutsch, Türkisch – Deutsch, Französisch – Deutsch und Italienisch – Deutsch.

3.5.1 Ausgangssprache Englisch

Die Studie von O'Brien & Smith aus dem Jahr 2010 verfolgt einen kontrastiven Ansatz und untersucht den zielsprachlichen Erwerb des Merkmals [+gerundet] der deutschen Vokale /y:/ und /u:/.

Im Fokus dieser Untersuchung steht insbesondere die Rolle der Erstsprache (L1-Variante: Nordamerikanisches Englisch) bei der Realisierung des im Deutschen ähnlichen Phonems /u:/ und des neuen Phonems /y:/. Die zugrundeliegende Hypothese ist die eines Transfers (positiv oder negativ), basierend auf spezifischem Vorwissen über die untersuchten Vokale: Einerseits wird beleuchtet, wie Probandinnen aus der *North Central Region* – denen literaturmäßig eine geringe Tendenz zur Frontalisierung hinterer Vokale zugeschrieben wird – /u:/ aussprechen;

andererseits ist [y] eine allophone Variante des englischen Phonems /u/ in westkanadischen und nördlichen Binnendialekten. Diese Annahmen führen zu differenzierten Hypothesen bezüglich des Einflusses der Kontaktsprache (CLI).

An der Studie nahmen 72 Universitätsstudentinnen der Mittelstufe teil, die zwischen einem und 14 Jahren Deutsch lernten und eine hohe Lernmotivation angaben. Die Aufgabe bestand darin, Zielwörter (Pseudowörter) in einem deutschen und einem englischen Trägersatz vorzulesen. Die anschließende Analyse der Variation der Formantenwerte F1, F2, F3 der Vokale in den Pseudowörtern lieferte folgende Ergebnisse. Obwohl die Realisierung von /u:/, das dem englischen /u/ ähnelt, von der Realisierung von /y:/ unterschieden wird, ist sie nicht das Ergebnis eines direkten Transfers der L1-Kategorie. Die Unterscheidung zwischen den beiden Phonemen bleibt zwar erhalten, entspricht jedoch nicht der zielsprachlichen Norm, sondern ist eher ungerundet. Im Wesentlichen realisieren die Lernenden /u:/ aufgrund von Hyperkorrektur (vgl. Flege et al., 2003; O'Brien & Smith, 2010) mit höheren F2- und F3-Werten, was auf eine frontlere und weniger gerundete Realisierung hindeutet und somit eine *Interlanguage*-spezifische Aussprache darstellt. Im Gegensatz dazu liegt die Realisierung des englischen Phonems /u/ deutlich näher am deutschen Ziel. Die Realisierung des Phonems /y:/, das dem englischen System phonologisch fremd ist, unterscheidet sich klar sowohl vom englischen /u/ als auch vom deutschen /u:/. Interessanterweise assimilieren Sprecherinnen des westkanadischen Englisch, die eine Tendenz zur Frontalisierung von /u/ zum Allophon [y] zeigen, das Phonem /y:/ nicht zu /i:/. Dennoch handelt es sich auch hier um eine nicht zielsprachliche Realisierung, die typisch für eine *Interlanguage* ist.

Der Ansatz von Smith & Baker (2011) untersucht die Rolle der erlernten Varietät, anstatt sich primär auf die Ausgangssprache zu konzentrieren. Die Studie verglich die Produktion der deutschen Vokale /i:/- /y:/- /u:/ durch Deutschlernende des österreichischen Deutsch mit einer Gruppe von Lernenden des Norddeutschen. Ein relevanter Unterschied zwischen den beiden Varietäten ist die Rundung dieser Phoneme, wobei die norddeutsche Standardsprache eine stärker ausgeprägte Rundung aufweist. Die Analyse der Formantenwerte (F1, F2, F3) ergab, dass das Phonem /i/ von beiden Lernergruppen ähnlich realisiert wurde. Beim Phonem /y/ wurde hingegen ein signifikanter Unterschied festgestellt; die Realisierung beider Gruppen lag leicht unter der Produktion der jeweiligen Kontrollgruppen (österreichisches Deutsch bzw. norddeutsche Standardsprache). Für das Phonem /u/ zeigte die Gruppe der Deutschlernenden der österreichischen Varietät eine zielähnlichere Realisierung.

Diese Ergebnisse untermauern die bereits für die Wahrnehmung berichteten Befunde und weisen die Gruppe der Deutschlernenden der österreichischen Varietät als die genaueste Gruppe sowohl auf der Wahrnehmungs- als auch auf der Produktionsebene aus. Eine mögliche Erklärung für diese höhere produktive Genauigkeit sehen die Autorinnen in der Vielfalt des deutschen Inputs, dem die Lernenden nach eigenen Angaben ausgesetzt waren. Es wird vermutet, dass diese Vielfalt die Fähigkeit der Lernenden geschärft hat, potenziell komplexe

Kontraste wie /y:/-/u:/ wahrzunehmen und zu produzieren, was zur Entwicklung einer "neutralen" Realisierung von Kontrasten führen könnte, die eine effektive Kommunikation unabhängig von der gesprochenen Varietät ermöglicht. Diese Hypothese wird durch die im Fragebogen erhobenen Verhaltensdaten gestützt: Die Lernenden des österreichischen Deutsch zeigten eine beträchtliche Bereitschaft, vom lokalen Sprachmodell abzuweichen und eine Aussprache anzustreben, die näher an der Norm liegt (vgl. Smith & Baker, 2011, S. 129-130).

3.5.2 Ausgangssprache Polnisch

Die Studie von Nimz & Khattab (2020) untersucht als erste den Einfluss der Orthographie auf die Produktion deutscher Vokale durch Polnische L1-Sprecher:innen. Dabei lag ein Hauptaugenmerk auf der Frage, ob orthographische Markierungen einen potenziellen Vorteil bei der Realisierung der Vokallänge bieten, da die Vokallänge im Polnischen nicht phonologisch ist.

Für die Untersuchung wurden die Vokalkontraste /i:/-/ɪ/, /a:/-/a/ und /u:/-/ʊ/ herangezogen. Die Stichprobe bestand aus 18 Deutschlerner:innen mit mittlerem bis fortgeschrittenem Kompetenzniveau. Diese Experimentalgruppe stammt aus einem gesteuerten DaF-Unterricht (Sprache und Kultur), der keinen besonders intensiven phonetischen Schwerpunkt hatte. In einer Bildbenennungsaufgabe, bei der die Hälfte der Wörter orthographisch markiert war und die andere Hälfte nicht, konnte nachgewiesen werden, dass die orthographische Markierung die korrekte Aussprache der deutschen Langvokale *tatsächlich* unterstützt. Konkret waren die Langvokale in der unmarkierten Bedingung 12 % länger als die Kurzvokale, während dieser Unterschied in der markierten Bedingung 36 % betrug. Obwohl ein Vergleich mit einer Deutsch-L1-Kontrollgruppe zeigte, dass die Realisierung der polnischen Lerner:innen nicht zielähnlich war, reichten die Unterschiede zwischen den beiden experimentellen Bedingungen aus, um zu dem Schluss zu kommen, dass die markierte Bedingung eine Realisierung begünstigt, die dem L1-Gegenstück ähnlicher ist.

Die Studie von Lipinska (2015) untersuchte das Phänomen der Kreuzsprachigen Interferenz (CLI) bei mehrsprachigen Sprecher:innen. Im Fokus stand dabei die Frage, ob Personen, die drei Sprachen aktiv beherrschen – nämlich Polnisch (L1), Englisch (L2) und Deutsch (L3) in dieser Reihenfolge – in der Lage sind, L3-Phonemrealisierungen von ähnlichen Phonemen in L1 und L2 zu unterscheiden. Die 11 Teilnehmer:innen waren Studierende der englischen Philologie mit fortgeschrittenen Kenntnissen sowohl im Englischen als auch im Deutschen. Ein wichtiger Unterschied zu früheren Studien ist der Umstand, dass diese Gruppe formalen Unterricht in englischer und deutscher Aussprache erhalten hatte (Lipinska, 2015, S. 85). Analysiert wurden polnische Vokale, die aufgrund ihrer Ähnlichkeit potenzielle Quellen für CLI darstellen, da sie den Vokalen /u:/ und /ɔ:/ im Englischen sowie dem deutschen Phonem /œ/ ähneln. Konkret handelte es sich um /u/, /ɔ/ und /ɛ/ im Polnischen. Die Untersuchung umfasste eine Satzerschließungsaufgabe in allen drei Sprachen, bei der Vokale in Pseudowörtern im Kontext /bVt/ präsentiert und anschließend phonetisch transkribiert wurden. Die Analyse der F1- und

F2-Werte zeigte eine signifikante Überlappung der Vokalkategorien über alle drei Sprachen hinweg, was auf einen deutlichen Einfluss der L1 auf L2 und L3 hindeutet. Insbesondere wurde der mittlere Vokal /œ/ häufig ähnlich den polnischen und englischen Vokalen /u/ bzw. /u:/ realisiert. Trotz der spezifischen Ausspracheausbildung und des hohen Sprachkompetenzniveaus folgert die Autorin abschließend, dass die Ähnlichkeit der Phoneme in den drei Sprachen die Bildung einer neuen L3-Kategorie – wie im *Speech Learning Model* (SLM) von Flege (1995) postuliert – erschwert.

Der vorliegende Überblick über die polnisch-deutsche Sprachkombination schließt mit der Studie von Nimz (2022), die bereits im Abschnitt über die Wahrnehmung erwähnt wurde. Ziel der Studie war es, die Rolle der Orthographie bei der Realisierung der Länge im Deutschen zu untersuchen. Zu diesem Zweck wurde eine Bildbenennungsaufgabe durchgeführt. Die Ergebnisse basieren auf 12 Proband:innen mit polnischer L1, die Deutsch als L2 auf mittlerem bis fortgeschrittenem Niveau erlernen. Hinsichtlich der Realisierung der Vokalqualität wurde festgestellt, dass die Realisierungen der Lerner:innen gemäß der Hypothese nicht als gleichwertig mit denen der Kontrollgruppe angesehen werden können. Dies ist darauf zurückzuführen, dass deutschen lange Vokale kürzer, während kurze Vokale lang ausgesprochen werden. Es konnte jedoch festgestellt werden, dass die Lerner:innen für das Unterscheidungsmerkmal der Länge sensibel sind, was auf eine perzeptuelle Modellierung der *Interlanguage* hindeutet (vgl. Nimz, 2022, S. 136). Hinsichtlich der Qualität wurde bei der Analyse von F1 und F2 eine zentralere Realisierung von /ɛ/ als sein deutsches Pendant festgestellt. Dies lässt auf einen CLI-Effekt seitens der L1 schließen, die diesen Vokal in einer zentraleren und offeneren Position realisiert als das Deutsche. Hinsichtlich der Realisierung von /e:/ vs. /ɛ/ wurde ein signifikanter Unterschied zu den Realisierungen der L1-Kontrollgruppe festgestellt, da die Realisierungen der Lerner:innen zwischen dem einen und dem anderen Vokal liegen. Insbesondere für /e:/ wurde eine größere Variation in der Formantenbewegung festgestellt (Werte gemessen an den beiden Extremen, 25 % und 75 %), was auf eine Diphthongierung hindeutet. Eine mögliche Erklärung könnte in einer Verbindung zwischen wahrnehmungsbezogenen, orthographischen und produktiven Faktoren liegen. Da das Graphem <e> im Polnischen normalerweise als /ɛ/ ausgesprochen wird, neigen die Lerner:innen instinktiv zu einer solchen Realisierung, die jedoch im Gegensatz zum wahrgenommenen Ziel steht. In dem Bestreben, eine Aussprache zu erreichen, die dem wahrgenommenen Ziel näher kommt, wird die Aussprache von einem polnischen [ɛ] zu einer geschlosseneren Variante in der Nähe des polnischen [i] umgestaltet. Bei /a:/ überschneiden sich die Lerner:innen mit der Aussprache von /a/, wodurch die statistisch signifikanten Unterschiede fast unhörbar werden. In Bezug auf die Realisierung des deutschen /o:/ zeigen sich Überschneidungen zwischen den Realisierungen der Lerner:innen und denen von /ɔ/. Die Autorin schließt jedoch einen Interferenzeffekt auf der phonetisch-phonologischen Ebene von L1 aus, da das deutsche /o:/-Phonem eine größere Ähnlichkeit mit dem polnischen /u/-Phonem aufweist (vgl. Hentschel

1986; Nimz, 2022, S. 145). Es wird daher vermutet, dass es sich um eine orthographische Interferenz handelt, bei der das Graphem <o> in der Regel mit dem Phonem /ɔ/ assoziiert wird.

3.5.3 Ausgangssprache Türkisch

Da es nur wenige Studien zur Realisierung deutscher Vokale durch erwachsene Lerner:innen mit der Sprachkombination Türkisch–Deutsch gibt, stützt sich dieser Abschnitt hauptsächlich auf die akustische Analyse aus der bereits im Kontext der Perzeption erwähnten Studie von Nimz (2011). Die akustische Untersuchung (Nimz, 2011) konzentrierte sich auf die Vokalrealisierungen von 8 der 22 Proband:innen des Perzeptionsexperiments. Hierbei wurden die deutschen Vokale /i:/, /ɪ/, /a:/, /a/, /u:/, /ʊ/ hinsichtlich der phonologisch relevanten Merkmale Länge und Qualität analysiert. Während im Türkischen der Status der Vokallänge unsicher ist, besitzen beide Merkmale im Deutschen phonologischen Status. Aus den Angaben zur Stichprobe geht hervor, dass die Teilnehmer:innen des Intensivlernpfads des Deutschen keinen spezifischen phonetischen Unterricht erhalten hatten. Die Proband:innen absolvierten eine Bildbenennungsaufgabe, bei der sie einsilbige Wörter aus ihrem Übungsmaterial benennen sollten. Die Analyse der Vokaldauer zeigte dabei eine Tendenz zur Angleichung der deutschen Vokallängen an die der L1 (Türkisch), was auf einen deutlichen Einfluss der L1 (CLI) hindeutet. Hinsichtlich der Vokalqualität wurden signifikante Abweichungen in den F1- und F2-Werten zwischen den Lernerrealisierungen und denen der deutschen Kontrollgruppe festgestellt, ausgenommen bei /i:/. Darüber hinaus zeigte sich eine Schwierigkeit bei der zielsprachlichen Unterscheidung der Merkmale gespannt/ungespannt: Deutsche ungespannte Vokale, die zentral im Vokalsystem liegen, wurden aufgrund des L1-Einflusses von den Lerner:innen dezentral realisiert.

3.5.4 Romanische Ausgangssprachen: Französisch und Italienisch

Die Makrokategorie der Lernenden mit einer romanischen L1 wird primär durch die Konstellationen Französisch–Deutsch und Italienisch–Deutsch repräsentiert. Im Vergleich zu germanischen oder slawischen Ausgangssprachen sind diese Kombinationen in der Literatur unterrepräsentiert, da der Forschungsschwerpunkt häufig auf der Konsonantenproduktion liegt.

3.5.4.1 Die Konstellation Französisch–Deutsch

In Bezug auf die deutsch–französische Konstellation ist insbesondere die Studie von Zimmerer & Trouvain (2015) von Relevanz. Aus einer kontrastiven Perspektive zeigen sich gewisse Gemeinsamkeiten und Unterschiede zwischen dem phonologischen System des Französischen und des Deutschen, was zu differenzierten Transferhypothesen führt. So sind im Französischen, wie auch im Deutschen, Phoneme wie /y/ und /ø/ durch das Merkmal der Rundung gekennzeichnet. Darüber hinaus ist festzustellen, dass zum phonologischen Inventar des Französischen, wie auch des Deutschen, der mittlere Vokal /ə/ gehört.

Hinsichtlich des Reichtums des Inventars kann das Französische als eine romanische Sprache betrachtet werden, die mit 11 Vokalen dem Deutschen (mit ca. 15 Vokalen) relativ nahe kommt. Diese Merkmale reichen bereits aus, um das Französische von anderen romanischen Sprachen wie dem Italienischen (mit nur 7 Vokalen) zu unterscheiden. Trotz dieser Ähnlichkeiten gibt es jedoch signifikante Unterschiede, die zu Ausspracheschwierigkeiten auf Seiten der Lerner:innen und folglich zu Verständnisproblemen bei Hörer:innen mit deutscher L1 führen können. Ein wesentlicher Unterschied besteht im Status der Vokallänge im Französischen, die im Gegensatz zum Deutschen nicht phonematisch ist. Darüber hinaus existieren, obwohl französische Phoneme wie /y/ dem deutschen /y:/ akustisch sehr ähnlich sind, feine artikulatorische Merkmale, die sie voneinander unterscheiden – wie z. B. eine frontalere Realisierung im Französischen, die es bei der Aussprache durch Lerner:innen potenziell ununterscheidbar vom deutschen ungespannten /y/ macht. Um diese möglichen Ausspracheschwierigkeiten empirisch zu testen, wurde ein Forced-Choice-Entscheidungs experiment durchgeführt. Die Teilnehmer:innen des Experiments waren zwei Gruppen von „Beurteiler:innen“: Die erste bestand aus drei phonetisch geschulten Hörer:innen, die zweite aus sechs Universitätsstudenten, die alle Deutsch als ihre L1 hatten. Im Rahmen der Untersuchung zur Rolle der Kompetenz wurden Minimalpaare verwendet, die von insgesamt 56 Französischlerner:innen ausgesprochen wurden. Diese Lerner:innen waren in zwei Gruppen nach ihrem Kompetenzniveau unterteilt: 22 Fortgeschrittene (Niveau B2–C1) und 34 Anfänger (Niveau A1–B2). Die Ergebnisse der Untersuchung bestätigten das Vorhandensein systematischer Schwierigkeiten bei der Produktion von Vokalen. Es zeigte sich, dass diese Schwierigkeiten umgekehrt proportional zum Kompetenzniveau verliefen. Hinsichtlich der Quantität verursachten längere Vokale am wenigsten Probleme, wohingegen kürzere und gerundete Vokale die größten Schwierigkeiten auslösten. Diese Fehlermuster traten zwar auf höheren Kompetenzniveaus in geringerem Maße auf, blieben jedoch in ihren Proportionen unverändert. Die Autoren weisen darauf hin, dass aufgrund der limitierten Anzahl an Items keine allgemeingültigen Generalisierungen gezogen werden können. Sie vermuten jedoch eine starke Beeinflussung durch die L1 aufgrund der feinen artikulatorischen Differenzen zwischen dem französischen Vokal /y/ und den ähnlichen deutschen Vokalen /y:/ und /ʏ/. Die Untersuchung der Produktion deutscher Vokale durch französische Lerner:innen kann daher als ergebnisoffen bezeichnet werden. Dies lässt Raum für weitere Untersuchungen, sowohl im Hinblick auf die Realisierung des Merkmals der Dauer als auch allgemeiner an der Schnittstelle zur Wahrnehmungsdimension (vgl. Zimmerer & Trouvain, 2015, S. 1723).

3.5.4.2 Die Konstellation Italienisch-Deutsch

Die zweite untersuchte romanische Sprache, das Italienische, weist im Vergleich zum Französischen deutlich ausgeprägtere Unterschiede zum deutschen Vokalsystem auf. Die Schwierigkeiten, die Lerner:innen mit L1 Italienisch auf der produktiven Ebene erfahren, resultieren primär aus den signifikanten Divergenzen zwischen den italienischen

Wahrnehmungsmustern und den konfigurativen Unterschieden der beiden Inventare. Neben den bereits erörterten perzeptiven Aspekten sind hierbei die unterschiedlichen artikulatorischen Gewohnheiten von zentraler Bedeutung. Das Italienische ist durch einen weniger komplexen Vokalraum charakterisiert, der tendenziell größere artikulatorische Bewegungen erfordert, um Vokalphoneme voneinander zu differenzieren. Trifft dieses System auf das wesentlich komplexere deutsche System, in dem die distinktiven artikulatorischen Bewegungen notwendigerweise präziser und kleiner ausfallen müssen, entstehen erhebliche Ausspracheschwierigkeiten.

Die Etablierung eines erweiterten Vokalinventars, dessen Realisierung eine angemessene Balance von Länge, Spannung, Rundung sowie Anteriorität/Posteriorität erfordert (vgl. Missaglia 2007, S. 244), stellt für die Lerner:innen eine beträchtliche Herausforderung dar. Auch die für unbetonte Positionen im Standarddeutschen typische Vokalreduktion, die zu zentralisierten, zum /ə/ tendierenden Realisierungen führt, wird im Standarditalienischen nicht systematisch angewendet (vgl. ebenda). Eine Ausnahme bilden bestimmte Phänomene der verketteten (gelesenen) Sprache (*connected speech*), wie sie bei Missaglia & Sendlmeier (1999) zu beobachten sind. Dort unterliegen Vokale, die durch Wortgrenzen getrennt sind, häufig einer Fusion (Sandhi-Phänomene). Dies führt zu einem Vokal, der im akustischen Raum eine Mittelposition zwischen den beiden Ausgangsvokalen einnimmt (z. B. „tramontana#e#il“, wo [a]#[e] → [ɛ], vgl. Missaglia & Sendlmeier 1999, S. 83). Dieses Phänomen ist auf das Prinzip der artikulatorischen Ökonomie zurückzuführen, beeinflusst durch Sprechgeschwindigkeit und idiosynkratische Merkmale der Sprecher:innen (vgl. ebenda). Mögliche Ausnahmen von diesen Schwierigkeiten könnten Fälle sein, in denen eine regionale Variante des Italienischen im Repertoire der Sprecher:innen vorhanden ist, die eine Reduktion oder Realisierung von Lauten ähnlich dem /ə/ vorsieht – wie etwa im mittleren Süditalienisch im Auslaut (z. B. [e]# → [ə]). Eine ähnliche Situation ergibt sich bei der Realisierung von /y:/, das zwar im Standarditalienischen fehlt, aber in einigen norditalienischen Varianten vorkommt.

Die einzigen experimentellen Beiträge zur Vokalproduktion von Lerner:innen sind Missaglia & Sendlmeier (1999) und Missaglia (2000), die bereits im Abschnitt zur Wahrnehmung erwähnt wurden. Beide beziehen sich auf die gleiche Produktionsstudie, so dass im Folgenden auf die erstgenannte Studie verwiesen wird, die eine ausführlichere Beschreibung der Studie enthält. Diese Studie ist besonders relevant, da sie die einzige experimentelle Vergleichsstudie zur Vokalproduktion darstellt, jedoch keine individuellen Daten zu Formantwerten für einzelne Vokale liefert. Nichtsdestotrotz liefert diese Studie wichtige Erkenntnisse durch die Darstellung der Voklräume der Lerner:innen und deren Beschreibung in einem kontrastiven Schlüssel.

Die Studie konzentriert sich auf die Realisierung von Vokalen bei einer Stichprobe von 10 Studentinnen mit vergleichbarer Herkunft (Lombardei) und unterschiedlichen Kompetenzniveaus im Deutschen (von Anfängerinnen bis Fortgeschrittenen). Als Kontrollgruppe

wurde eine Gruppe von 5 Universitätsdozent:innen⁸ mit L1 Deutsch gebildet, die an der Fakultät für Fremdsprachen der Katholischen Universität Mailand tätig sind. Die Aufgabe bestand darin, die Erzählung „Nordwind und Sonne“ zuerst auf Italienisch und dann auf Deutsch zweimal vorzulesen. Aus der Analyse der F1- und F2-Werte wurden die Vokalräume auf kartesischen Ebenen dargestellt, die Unterschiede zwischen den Gruppen und zwischen den einzelnen Sprecher:innen zeigen. In der Studie von Missaglia & Sendlmeier (1999) ist jedoch nur die Kennzeichnung der Eckvokale (/i:/, /a:/, /u:/) in den F2/F1-Diagrammen zu finden.

Der Gruppenvergleich offenbart einen wesentlichen Unterschied: Die Vokalräume der Lerner:innen erscheinen als klare, scharf umrissene Vokaldreiecke, während die der Kontrollgruppe stumpf und trapezförmig wirken. Während bei den Lerner:innen die Streubereiche der fünf gespannten Vokale /i, e, a, o, u/ laut Datenbeschreibung klar erkennbar sind, sind die Grenzbereiche in der Kontrollgruppe nicht klar definiert und scheinen in einem Kontinuum ineinander überzugehen (vgl. Missaglia & Sendlmeier, 1999, S. 88). In Bezug auf das artikulatorische Verhalten der Lerner:innen wurden zwei Haupttendenzen identifiziert:

- **“Leere“ Vokalräume:** Eine starke Konzentration in peripheren Bereichen (/i, a, u/) bei gleichzeitigem Mangel an Realisierungen im zentralen Bereich. Dies wird als Effekt eines unzureichenden Vokalinventars in der Interlanguage interpretiert, was zu begrenzter artikulatorischer Variabilität führt (vgl. ebd., S. 87).
- **“Volle“ Vokalräume mit Dezentralisierung:** Diese Räume weisen zwar eine Präsenz von Vokalen der Qualität „E“ auf (erklärbar durch die Häufigkeit des Graphems <e>), zeigen jedoch eine Tendenz zur dezentralen Realisierung von /ə/ und /ɐ/.

Die Autoren schlussfolgern, dass dieses Verhalten weniger auf ein „artikulatorisches Defizit“ zurückzuführen ist, sondern auf den Magnet-Effekt (Kuhl, 1991; Kuhl & Iversen, 1995) der L1 auf die Phonologie des Deutschen. Dies äußert sich in der Dezentralisierung vorderer Phoneme (/i:/-/ɪ/, /ɛ/, /a:/-/ɑ/), der Zentralisierung hinterer Phoneme (/u:/-/ʊ/, /o:/) sowie der Tendenz, Realisierungen in den äußeren Zonen zu konzentrieren, wo sich die nativen Phoneme /i, a, u/ befinden (vgl. Missaglia & Sendlmeier, 1999, S. 91).

3.6 Abschließende Anmerkungen und übergeordnete Forschungsfrage

Die Durchsicht der Literatur zur Wahrnehmung und Produktion von Vokalen in verschiedenen Sprachkonstellationen verdeutlicht eine starke Dominanz der Kombination (Amerikanisches) Englisch-Deutsch, während andere Sprachpaare vergleichsweise unterrepräsentiert bleiben. Besonders deutlich wird dieses Forschungsdefizit in der Kategorie der romanischen Sprachen,

⁸ Die Kontrollgruppe besteht aus 4 Frauen und 1 Mann. Im Text von Missaglia & Sendlmeier (1999) wird nicht genauer beschrieben, ob für den Vergleich lediglich die Daten bezogen auf die 4 Frauen in Betracht gezogen wurden.

wobei für die Zwecke der vorliegenden Arbeit insbesondere das Italienische von zentralem Interesse ist.

Zwar liegen für die italienisch-deutsche Konstellation erste wegweisende psycholinguistische (Missaglia, 2000) und akustisch-phonetische Untersuchungen (Missaglia & Sendlmeier, 1999) vor, jedoch bleiben wesentliche Aspekte der Interaktion zwischen rein linguistischen Parametern und den soziolinguistischen Faktoren, die die untersuchte Population charakterisieren, bislang unerforscht. Insbesondere die Rolle des regionalen Backgrounds (dialektales Substrat) wurde in diesem Zusammenhang noch nicht systematisch beleuchtet. Darüber hinaus fehlt für die italienisch-deutsche Kombination ein integrativer Ansatz, der sowohl die Wahrnehmung als auch die Produktion deutscher Vokale bei derselben Probandengruppe vergleichend untersucht.

Aus diesen theoretischen und empirischen Desideraten ergibt sich die zentrale Forschungsfrage, die der vorliegenden Arbeit zugrunde liegt:

Wie manifestiert sich Cross-Linguistic Influence (CLI) in der Wahrnehmung und Produktion deutscher Vokale bei erwachsenen italienischen Lernern:innen an der Schnittstelle zwischen sprachenstrukturellen Faktoren und soziolinguistischer Variabilität?

3.7 Forschungsdesiderat und Zielsetzung der Studie

Die vorliegende Studie zielt darauf ab, eine signifikante Forschungslücke hinsichtlich der Wahrnehmung und Produktion von Vokalen bei erwachsenen Lerner:innen des Deutschen als Fremdsprache zu schließen. Die Untersuchung konzentriert sich auf eine Sprachkonstellation, deren Einzelsysteme sich sowohl in ihrer quantitativen als auch in ihrer qualitativen phonologischen Struktur deutlich voneinander unterscheiden. Während das Italienische nach der Taxonomie von Maddieson (2013) über ein mittelgroßes Vokalsystem mit einer durchschnittlichen Ausdehnung im Vokalraum verfügt, weist das Deutsche ein großes und dichtes Vokalinventar auf (vgl. ebd.).

Qualitativ betrachtet zeichnet sich das deutsche System durch eine Vielzahl von Phonemen aus, die dem italienischen Standardausgangssystem völlig fremd sind. In der vorliegenden Arbeit liegt der Fokus explizit auf dem vorderen hohen gerundeten Vokal /y:/ und dem zentralen mittleren Vokal /ə/. Diese beiden Phoneme des Standarddeutschen gelten aufgrund ihrer intrinsischen und phonotaktischen Merkmalen als potenziell schwierig zu erlernen (vgl. Maturi, 2014).

Insbesondere das Phonem /y:/ weist mit dem Merkmal der Rundung bei Vorderzungenvokalen ein Merkmal auf, das im Standarditalienischen nicht existiert. Es muss daher als "neues" Phonem betrachtet werden, das von den Lerner:innen die Bildung einer völlig neuen Kategorie erfordert

(vgl. Flege, 1995). In ähnlicher Weise verlangt das Phonem /ə/ eine artikulatorische und auditive Sensibilität im zentralen Bereich des Vokalraums, der im Italienischen üblicherweise unbesetzt bleibt (vgl. Flege, 1995).

Auf dieser theoretischen Basis soll untersucht werden, inwiefern die Perzeption und Produktion dieser beiden Phoneme nicht allein durch streng linguistische Faktoren (primär den Einfluss der L1) bedingt sind, sondern in Interaktion mit soziolinguistischen Variablen stehen. Von besonderem Interesse sind dabei die sprachliche Kompetenz und Dominanz sowie der Umfang des individuellen sprachlichen Repertoires (einschließlich regionaler Varietäten).

Schließlich verfolgt die Arbeit das Ziel, die komplexe Beziehung zwischen Wahrnehmung und Produktion empirisch zu beleuchten. Dieser Beitrag strebt an, einen ersten umfassenden Überblick über die Charakteristika der *Interlanguage* Italienisch-Deutsch (DaF) zu liefern, wobei der Schwerpunkt auf Assimilationsphänomenen hinsichtlich des Schwa-Lauts und des Phonems /y:/ liegt. Die daraus generierten Erkenntnisse sollen über den rein theoretischen Rahmen hinausgehen und potenziell in didaktischen Kontexten (z. B. durch ein gezieltes Phonetiktraining) Anwendung finden können.

4. Eine psycholinguistische Studie zur Wahrnehmung deutscher Vokale

4.1 Forschungsfragen und Hypothesen

Die zentrale Forschungsfrage dieser Studie betrifft die Wahrnehmung der Phoneme /ə/ und /y:/ durch erwachsene Lerner:innen des Deutschen als Fremdsprache. Insbesondere wird die Fähigkeit der Hörerinnen und Hörer getestet, die deutschen Phoneme /ə/ und /y:/ zu diskriminieren und darüber hinaus von perzeptiv ähnlichen deutschen Phonemen zu unterscheiden (Maturi, 2014; Missaglia, 2007), mit denen sie kontrastiert werden. Dabei geht es um die Diskriminierung von /ə/ gegenüber den ähnlichen Phonemen /ɛ/ und /e:/ sowie von /y:/ gegenüber den ähnlichen Phonemen /i:/ und /u:/.

Diese Forschungsfrage basiert auf der Annahme, dass LX-Phoneme, die für die Lerner:innen neue Kategorien darstellen, schwer zu erlernen sind (Escudero, 2005). Diese Hypothese repräsentiert das typische Lernszenario, in dem Lerner:innen, deren L1 (Italienisch) ein eher mittelgroßes Vokalsystem aufweist mit einem relativ reichen und dichten Vokalsystem (wie dem Deutschen) konfrontiert werden. Die Herausforderung besteht im Wesentlichen darin, neue Lautkategorien zu erlernen indem die Effekte der *phonemic equation* (Escudero, 2005) schrittweise überwunden werden. Auf einer abstrakten Ebene werden die kontrastierenden LX-Phoneme zu einer einzigen L1-Lautkategorie verbunden. Dieses Szenario ist typisch für die erste Lernphase (*initial state* in Escudero, 2005, vgl. auch Best & Tyler, 2007; Flege, 1995), in der die Lerner:innen zum ersten Mal mit der neuen Sprache konfrontiert werden. Das Szenario verändert sich im Laufe des Lernprozesses (*learning task* in Escudero, 2005), was im Szenario des Erlernens einer neuen Kategorie besondere Schwierigkeiten mit sich bringt: Das Streben nach einer zielsprachlichen (optimalen, vgl. Escudero, 2005) Wahrnehmung erfordert die Bewältigung von Wahrnehmungsdiskrepanzen und die Verschiebung der Wahrnehmungsgrenzen zwischen L1- und L2-Kategorien (vgl. Escudero, 2005, S. 159). Nach Escudero (vgl. ebd.) erfordern diese Aufgaben eine umfassende Exposition mit reichhaltigem L2-Input. Im Folgenden werden die möglichen Schwierigkeiten beim Erlernen der Phoneme /ə/ und /y:/ und ihrer Unterscheidung von Phonemen, die potenziell mit ihnen assimilierbar sind, sowie die möglichen Gründe dafür näher erläutert.

Im Falle von /ə/ stellt das Fehlen jeglicher Merkmale aufgrund der zentralen Position im deutschen Vokaltrapez (vgl. Wiese, 2000; Browman & Goldstein, 1992) eine potentielle Wahrnehmungsschwierigkeit dar, da die Hörer:innen mit italienischer L1 dem fremden Phonem potentiell ein Merkmal zuordnen müssen, um den Laut kategorisieren zu können (vgl. Hayes-Harb

& Barrios, 2021).

Im Falle von /y:/ stellt sich die Frage nach der Gewichtung der Merkmale anders. Bei /y:/, das u.a. durch die Merkmale [+gerundet] und [+vorne] gekennzeichnet ist, besteht die Wahrnehmungsschwierigkeit für Deutschlerner:innen darin, diese zwei Merkmale, die in ihrer L1 nicht kombiniert vorkommen, perzeptiv als einige Einheit zu erkennen.

Die Rechtschreibung könnte eine Rolle unter den Faktoren spielen, die die Wahrnehmung von LX-Lauten in der Lernphase beeinflussen. In jüngerer Zeit wurde insbesondere die unterstützende Funktion von Orthographie einer eingehenden Untersuchung unterzogen. Eine Übersicht von Hayes-Harb & Barrios (2021) zeigt: Die Forschung zur Rolle der Rechtschreibung in der LX-Wahrnehmung ist bislang uneinheitlich. Es gibt jedoch Hinweise auf einen negativen Transfer (vgl. Bassetti & Atkinson, 2015), insbesondere wenn Lerner:innen mit einem transparenten orthografischen System (z. B. Italienisch) auf ein weniger transparentes System treffen (z. B. Englisch oder Deutsch; siehe (Erdener & Burnham, 2005).

Lerner:innen mit solchen transparenten L1-Systemen tendieren demnach dazu, sich stärker auf Orthographie zu verlassen als Lerner:innen mit weniger transparenten L1-Systemen (vgl. Erdner & Burnham, 2005). In vorliegendem Fall kann daher nicht ausgeschlossen werden, dass die italienischsprachige Lerngruppe die vertraute Zuordnung des Phonems /e/ oder /ɛ/ mit dem Graphem <e>⁹ auch auf das Deutsche überträgt. Dies könnte Auswirkungen auf das Erlernen von /ə/ haben. Die Annahme zu /y:/ kann ähnlich formuliert werden: die Nutzung des deutschen Graphems <ü>, was ähnlich zu dem italienischen Graphem <u> ist, könnte zu unterschiedlichen Ergebnissen bei der Entwicklung der Kategorie führen. Tatsache ist es, dass im Italienischen das Graphem <u> immer das Phonem /u/ repräsentiert. Maturi (2014) berichtet über die Ausspracheschwierigkeiten deutscher Laute, die für Lerner:innen mit italienischer L1 als schwieriger bekannt sind, dass das finale <e> der deutschen Orthografie, das in der Aussprache als [ə] realisiert werden sollte (so in Maturi, 2014 S. 154), zu [e] wird (ebenda). Diese Schwierigkeiten könnten je nach dem Sprachkompetenzniveau im Deutschen sowie dem Einfluss früherer Sprachkenntnisse unterschiedliche Ausprägungen annehmen, die jedoch in Maturi (2014) nicht berücksichtigt werden.

Im Hinblick auf die spezifischen experimentellen Bedingungen der vorliegenden Studie werden folgende Hypothesen aufgestellt.

H1) Es wird die übergeordnete Hypothese aufgestellt, dass die Wahrnehmung von /ə/ und /y:/ bei den Lerner:innen aufgrund des Einflusses der L1 – im Sinne von *magnet effect* (Kuhl, 1991, Kuhl & Iverson, 1995), *equivalence classification* (Flege, 1995), *phonemic equation* (Escudero, 2009) – schlechter (d.h. mit niedrigerer Akkuratheit und höheren Reaktionszeiten) ausfällt als bei der

⁹ Nach der Standardaussprache wird aber das Graphem <e> in unbetontem Wortauslaut immer als /e/ ausgesprochen.

Kontrollgruppe von Hörer:innen mit Deutsch als L1. Zugleich wird erwartet, dass die Kontrollgruppe mit Italienisch als L1 und ohne Deutschkenntnisse aufgrund fehlender Vorerfahrungen im Deutschen schlechter abschneidet als die Experimentalgruppe der Lerner:innen.

H1a) Bezüglich der Wahrnehmung von /ə/ wird die Hypothese aufgestellt, dass das Phonem mit akustisch ähnlichen Phonemen des Deutschen verwechselt wird, was auf einen zugrunde liegenden Einfluss der L1 zurückzuführen ist. Dieser Einfluss (negativer Transfer) bewirkt, dass die Wahrnehmung die Phoneme der Fremdsprache (LX) in Richtung der vorderen Bereiche des Vokalraums filtert. Konkreter wird angenommen, dass die deutschen Kontraste /ə/-/e:/ und /ə/-/ɛ/ nicht zielsprachlich bzw. mit niedriger Akkuratheit und hohen Reaktionszeiten wahrgenommen werden. Außerdem ist zu bedenken, dass in der L1 die Kontraste zwischen tiefen und hohen L1-Frontvokalen /ɛ/ und /e/ eine geringe funktionelle Belastung haben, was die Unterscheidung von Minimalpaaren im Deutschen weiter erschweren könnte.

H1b) Bezüglich der Wahrnehmung von /y:/ wird die Hypothese aufgestellt, dass das Phonem /y:/ mit deutschen Phonemen assoziiert, die bestimmte Merkmale mit ihm teilen, nämlich /i:/ [+vorne] und /u:/ [+gerundet]. Da im Standarditalienischen kein Phonem existiert, das beide Merkmale – Vorderartikulation und Rundung – wie /y:/ kombiniert, ist anzunehmen, dass auf der Wahrnehmungsebene ein nicht zielsprachlicher Merkmalsausgleich stattfindet, der auf einen zugrunde liegenden Einfluss (negativen Transfer) der L1 zurückzuführen ist. Konkreter wird angenommen, dass die deutschen Kontraste /y:/-/i:/ und /y:/-/u:/ nicht zielsprachlich bzw. mit niedriger Akkuratheit und hohen Reaktionszeiten wahrgenommen werden.

H2) Bezüglich des Einflusses außerlinguistischer Faktoren auf die Wahrnehmung von /ə/ und /y:/ wird die übergeordnete Hypothese aufgestellt, dass diese maßgeblich durch eben jene Faktoren beeinflusst wird.

H2a) Bezüglich der Wahrnehmung von /ə/ wird die Hypothese aufgestellt, dass Vorerfahrungen mit dem Englischen – gemessen an der sprachlichen Dominanz – eine Grundlage für einen positiven Transfer bilden können. Das Phonem /ə/ wird dabei als theoretisch bereits „bekanntes“ Phonem aus dem Englischen betrachtet, da Englisch in italienischen Schulen (ab dem 6. Lebensjahr) als erste Fremdsprache eingeführt wird. Darüber hinaus besteht insbesondere bei der jungen Bevölkerung ein relativ gleichmäßiger Zugang zu englischem Input. Es sei jedoch angemerkt, dass die vorliegende Hypothese unter Berücksichtigung mehrerer kontextueller Faktoren mit Vorsicht zu interpretieren ist. Der schulische Englischunterricht in Italien zeichnet sich traditionell durch einen geringen Fokus auf die Vermittlung der Aussprache sowie durch eine stärkere Ausrichtung auf Grammatik und kommunikative Kompetenzen aus (vgl. Gauci, 2008). Zudem wird er häufig von nicht-L1 Lehrkräften (d.h. Lehrkräfte, bei denen das Deutsche nicht die L1 ist) erteilt und findet in einem gesellschaftlichen Kontext statt, in dem der alltägliche Gebrauch des

Englischen lange Zeit eingeschränkt war (vgl. Leonidou, 2020). Diese Rahmenbedingungen könnten sich negativ auf die Qualität des phonetischen Inputs auswirken und die Ausbildung einer präzisen phonologischen Kategorie – wie der des /ə/ – erschweren.

H2b) Hinsichtlich sowohl des Phonems /ə/ als auch des Phonems /y:/ wird die Hypothese aufgestellt, dass die Zielsprachlichkeit ihrer Wahrnehmung vom Grad der sprachlichen Dominanz sowie vom Kompetenzniveau im Deutschen (gemäß GER) beeinflusst wird. Konkret wird angenommen, dass mit höherer Dominanz und höherem GER-Niveau eine zielsprachlichere Wahrnehmung (d.h. mit höherer Akkuratheit und niedrigeren Reaktionszeiten) einhergeht.

H2c) Es wird angenommen, dass Vorkenntnisse in einem Dialekt der L1 (Italienisch), in dessen phonologischen System der Vokal /ə/ oder /y:/ enthalten ist, die Grundlagen für einen positiven Transfer bilden können. Konkreter wird erwartet, dass der positive Transfer zu höherer Akkuratheit und niedrigeren Reaktionszeiten führt.

H2d) Es wird angenommen, dass umfangreiche sprachliche Kenntnisse (im Sinne der Anzahl der im Laufe des Lebens erlernten Sprachen) das Erlernen einer Kategorie wie /ə/ und /y:/, die im standardsprachlichen L1-System nicht vorhanden ist, begünstigen können. Diese Hypothese stützt sich auf eine Argumentation, die in der Diskussion der Studie von Smith & Baker (2011) dargelegt wird. Dort wird angenommen, dass eine größere Bandbreite an sprachlichen Vorerfahrungen das Sprachenlernen positiv beeinflussen kann. Für die vorliegende Untersuchung wird erwartet, dass sich dieser Effekt in einer höheren Akkuratheit und in kürzeren Reaktionszeiten bei den Teilnehmer:innen mit einem breiteren sprachlichen Repertoire niederschlägt.

Die Hypothesen zur perceptiven Assimilation anhand des Hintergrunds (L1 Italienisch kontrastiert mit der LX Deutsch) sind in der folgenden Abbildung zusammengefasst (Abbildung 4):

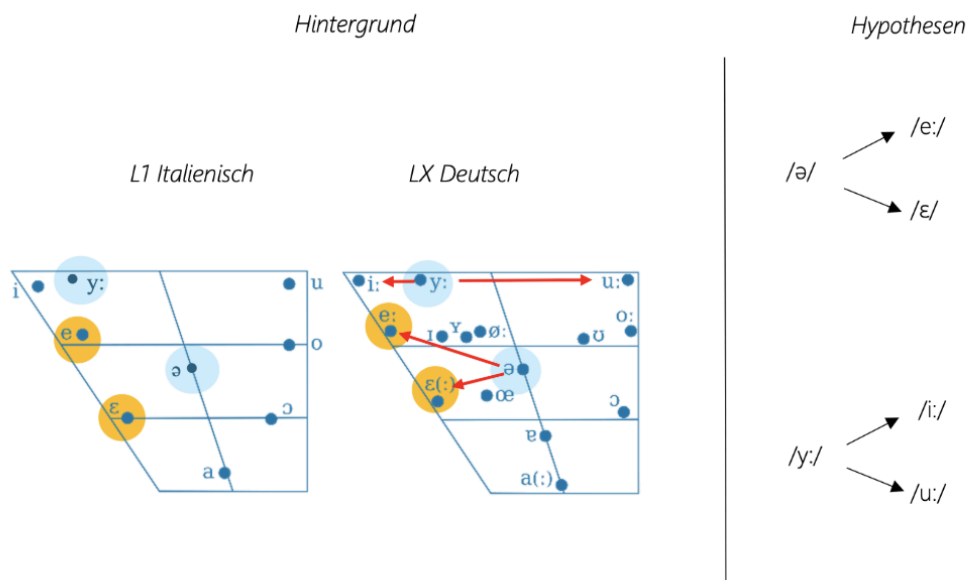


Abbildung 4: Darstellung der Hypothesen anhand der kontrastierten Systeme des Standarditalienischen (vgl. z.B. Maturi, 2014) und des Standarddeutschen (vgl. z.B. Wiese 2000). Farbige Markierungen zeigen neue Zielvokale (blau) und artikulatorische N%öhe (gelb); rote Pfeile deuten auf die hypothesierten perzeptiven Assimilationsmustern. Eigene Darstellung.

4.2 Methodisches Vorgehen

Zur Beantwortung der Forschungsfragen wurde ein psycholinguistisches Experiment zur Diskrimination von Vokalkategorien (im Englischen *Oddity-discrimination task* genannt) durchgeführt. Konkret handelte es sich um ein *Oddity-Discrimination*-Experiment, das einem AXB-Experiment¹⁰ ähnelt, bei dem jedoch die Konfiguration der Stimuluswiedergabe dem Muster XAA/AXA/AAX folgt. In diesem Schema gibt es eine Sequenz von drei akustischen Stimuli, von denen zwei identisch und einer unterschiedlich sind, der erkannt werden muss. Die akustischen Reize enthalten deutsche Vokale, und das Ziel des Versuchsmodells ist es, die Fähigkeit der Zuhörer:innen zu beurteilen, zwischen verschiedenen deutschen Vokalkategorien zu unterscheiden. Diese Fähigkeit wird in Form von Akkuratheit und Reaktionszeiten gemessen. Die Sequenzen wurden ebenfalls in zwei Formen präsentiert: als Sequenzen mit jeweils unterschiedlichen Stimuli an unterschiedlichen Positionen (XAA/AXA/AAX) oder als Sequenzen mit identischen Stimuli (AAA, XXX).

¹⁰ Das AXB-Paradigma ist ein Wahrnehmungsexperiment, bei dem den Versuchspersonen drei Stimuli in der festen Reihenfolge A–X–B präsentiert werden. Der mittlere Stimulus X ist der Zielreiz, und die Aufgabe besteht darin zu entscheiden, ob X dem ersten (A) oder dem dritten (B) Stimulus ähnlicher ist.

Die gewählte Abweichung vom klassischen AXB-Paradigma erscheint im Hinblick auf die Zielsetzung der vorliegenden Untersuchung besonders geeignet. Intendiert ist eine erweiterte Analyse des Wahrnehmungsverhaltens der Lerner:innen, die über die bloße Zuordnung des Zielreizes X zu einer der beiden Vokalkategorien (A oder B) hinausgeht. Ziel ist es, einerseits zu erfassen, ob die Versuchspersonen überhaupt in der Lage sind, eine phonetische Differenzierung zwischen den dargebotenen Lauten vorzunehmen (wobei pro Sequenz nur zwei Vokalkategorien getestet werden), und andererseits zu untersuchen, ob sich mögliche Wahrnehmungsschwierigkeiten eher in Form einer nivellierten Wahrnehmung, d.h. der identischen Repräsentation akustisch unterschiedlicher Stimuli, oder in Form einer Fehldiskriminierung, d.h. der fehlerhaften Identifikation des abweichenden Elements, manifestieren.. Darüber hinaus werden durch das Vorhandensein des Stimulus X in allen möglichen Positionen Verzerrungen vermieden, die durch die feste Position des Stimulus X im AXB-Modell entstehen können. Schließlich verringert das Fehlen von zwei Referenzkategorien ("A" und "B") in derselben Sequenz die kognitive Belastung des Arbeitsgedächtnisses, wodurch das Design flexibler und für eine Population mit sehr unterschiedlichen Sprachkenntnissen geeignet wird.

Für die Gruppe der Lerner:innen wurde das GERS-Kompetenzniveau auf der Grundlage des Ergebnisses des zuletzt absolvierten Sprachtests ermittelt. Zusätzlich zur Durchführung des Experiments wird am Ende des Experiments ein Fragebogen ausgefüllt, um soziolinguistische Informationen zu sammeln (Marian et al. 2007, siehe Appendix).

Im Allgemeinen wurde die gesamte Sitzung im einsprachigen Modus (Deutsch) für Sprecher ab B1 durchgeführt, da sie über ausreichende Kompetenzen verfügten, um die Inhalte angemessen zu verstehen. Diese Wahl wurde auf der Grundlage der Tatsache getroffen, dass auf dem B1-Niveau davon ausgegangen wird, dass die Sprecher: innen bereits ein mittleres Kompetenzniveau erreicht hat, das es ihnen ermöglichen, einfache Gespräche zu führen und Anweisungen zu verstehen. Der einsprachige Modus wurde gewählt, um so genau wie möglich ein klassisches didaktisches Umfeld nachzubilden, in dem die Sprecher: innen die deutsche Sprache möglicherweise gelernt haben. Es wurde nicht angenommen, dass diese Wahl für die Lerner:innen mit niedrigeren Kompetenzniveaus nachteilig gewesen sein könnte, da bei ihnen die Interferenz des Italienischen noch massiv ist.

Die Teilnehmer:innen erhielten eine Aufwandsentschädigung für ihre Teilnahme am Experiment.

4.2.1 Versuchspersonen: Beschreibung der drei Stichproben

4.2.1.1 *Experimentalgruppe: Hörer:innen mit L1 Italienisch und Deutschkenntnis*

Die Experimentalgruppe umfasst eine Stichprobe von 36 (Durchschnittsalter 21,99 Jahre, SD: 3,09) gesunden Lerner:innen von Deutsch als Fremdsprache, die alle an der Universität Verona studieren.

Alle Teilnehmer:innen abgesehen von 3 Frauen gaben „Italienisch“ als L1. Die 3 Teilnehmerinnen gaben an, „Dialekt“ als L1 zu haben. Insbesondere gab eine von ihnen (GT_PER) an, dass sie den Dialekt der Region Venetien als L1 hatte und dass sie das regionale Italienisch erst später gelernt hat. Die beiden anderen Teilnehmerinnen (DR_PER, VA_PER) gaben an, den Dialekt (des mittleren südlichen Gebiets der Region Kampanien bzw. des äußersten südlichen Gebiets der Region Sizilien) gleichzeitig mit dem Regionalitalienischen erlernt zu haben, was man als funktionale Zweisprachigkeit bezeichnen könnte.

Was die Kenntnis eines Dialekts der eigenen L1 betrifft, so gaben insgesamt 26 Teilnehmer:innen an, Dialektkenntnisse zu haben (ob aktiv oder nur passiv). Die folgenden Gebiete sind vertreten: das Gallo-italische Gebiet (1 Frau aus der Lombardei), das Gebiet Veneto (20 Teilnehmer:innen, die sich auf die Regionen Trentino-Südtirol und Veneto verteilen, wobei die Mehrheit der beherrschten Varietäten aus Veneto stammt), das mittlere südliche Gebiet (2 Proband:innen aus der Region Kampanien) und das äußerste südliche Gebiet (1 Probandin aus der Region Sizilien).

Alle Teilnehmer:innen gaben an, dass sie Deutsch in einem formalen (schulischen) Kontext gelernt haben. Im Vergleich zu allen anderen Fällen heben sich die Proband:innen BB_PER, LG_PER und SLL_PER dadurch ab, dass sie angaben, Deutsch als Autodidakt mit Hilfe von didaktischen Hilfsmitteln, aber nicht auf natürliche Weise gelernt zu haben.

Das durchschnittliche Lernalter für Deutsch, das von allen Teilnehmer:innen immer als Fremdsprache und nie als zweite L1 eingestuft wurde, betrug 11,2 Jahre.

Hinsichtlich der Deutschkenntnisse weisen die Teilnehmer:innen eine Bandbreite auf, die alle Niveaustufen bis C1 abdeckt (A1, A2, B1, B2, C1). Die Mehrheit der Stichprobe weist ein B1-Niveau auf, was im Allgemeinen ein mittleres Niveau anzeigt, vergleichbar mit den Kenntnissen, die erforderlich sind, um sich aus den häufigsten Situationen befreien zu können. Gemäß dem GERS ist ein Lerner:innen auf B1-Niveau insbesondere in der Lage 1) die Hauptpunkte klarer, standardsprachlicher Äußerungen zu vertrauten Themen verstehen, die regelmäßig in Arbeit, Schule, Freizeit usw. vorkommen; 2) die meisten Situationen bewältigen, die auf Reisen in einem Gebiet auftreten können, in dem die Sprache gesprochen wird, 3) einfache, zusammenhängende Texte zu vertrauten oder persönlichen Interessengebieten verfassen, 4) Erfahrungen und Ereignisse, Träume, Hoffnungen und Ambitionen beschreiben und kurz Gründe und Erklärungen für Meinungen und Pläne geben (vgl. Europarat, 2010).

Die Sprachkompetenz im Deutschen wurde auch in Bezug auf die Dominanz gemessen, wobei die Dominanz anhand der Angaben aus dem soziolinguistischen Fragebogen ermittelt wurde. Die Teilnehmer:innen wurden dazu aufgefordert, die Sprachen, die sie kennen, auf der Grundlage ihrer Dominanz anzugeben und somit jeder Sprache eine Position ab 1 (maximale Dominanz) zuzuweisen. Die Mehrheit der Teilnehmer:innen gab eine Sprachdominanz zwischen 3

(“drittbekannteste”) und 4 (“viertbekannteste”) an: Dies lässt darauf schließen, dass das Deutsche keine dominante Rolle im Sprachrepertoire der Teilnehmer:innen einnimmt.

In Bezug auf die Erfahrung in einem immersiven Kontext, gemessen durch Frage 8 des ersten Abschnitts des LEAP-Fragebogens, hatten 12 Proband:innen noch nie Kontakt mit deutschem L1-Input in einem immersiven Kontext, 15 Proband:innen für einige Wochen bis maximal zwei Monate, 5 Probanden für 6 Monate, während eine Minderheit (3 Proband:innen) seit 11–12 Monaten in Kontakt war.

In Bezug auf die Englischkenntnisse, die anhand der ersten Frage im ersten Abschnitt des LEAP-Fragebogens als Dominanz gemessen wurden, zeigt sich, dass die Mehrheit der Teilnehmer:innen, sowohl männliche als auch weibliche, Englisch als dominante L2 in ihrem sprachlichen Repertoire angeben.

Was den Umfang des Sprachrepertoires der Versuchspersonen angeht, wies die Mehrheit der Lerner:innen (22 Personen) 5 Sprachen im Sprachrepertoire auf. Die wenigsten gaben nur 3 oder 6 Sprachen im Sprachrepertoire an (jeweils 3 und 4 Personen).

4.2.1.2 Kontrollgruppe 1: Hörer:innen mit L1 Italienisch ohne Deutschkenntnis

In die Analyse wurden insgesamt 27 Personen (Durchschnittsalter: 21,39, SD: 2,42) einbezogen. Es handelte sich um gesunde Erwachsene, die an der Universität Verona studieren.

Alle Teilnehmer:innen gaben Italienisch als ihre L1 an mit Ausnahme einer Frau, die einen Dialekt (sizilianischen) als ihre L1 angab.

Hinsichtlich der Dialektkenntnisse zeigte sich, dass die meisten Teilnehmer:innen (17 Proband:innen) über (passive und/oder aktive) Kenntnisse eines Dialekts verfügten. Im Einzelnen handelt es sich um 17 Teilnehmer:innen. Die folgenden Gebiete sind vertreten: das Gallo-Italische Gebiet (6 Proband:innen), das Gebiet Veneto (5 Personen), das Toskanische Gebiet (1 Frau), das Mittlere Gebiet (1 Mann), das Südliche Mittelgebiet (6 Personen), das Südliche Extremgebiet (4 Frauen), das Sardische (1 Frau).

In Bezug auf die Kompetenz in Englisch, gemessen an der sprachlichen Dominanz, gab die Mehrheit der Teilnehmer:innen (20, 12 Frauen und 8 Männer) an, dass sie Englisch als dominanteste Zweitsprache in ihrem Sprachrepertoire haben, gefolgt von insgesamt 6 Personen (4 Frauen und 2 Männer), die einen Dominanzgrad von 3 angaben, und einer deutlichen Minderheit von 1 Person (1 Frau), die angab, dass sie Englisch als vierte Sprache in ihrem Repertoire habe.

4.2.1.3 Kontrollgruppe 2: Hörer:innen mit L1 Deutsch

In die Analyse wurden 25 Proband:innen einbezogen (Durchschnittsalter 24,64 Jahre, SD: 4,97). Alle waren gesunde Erwachsene. Die Mehrheit der Stichprobe bestand aus Student:innen, wobei eine Person eine Ausnahme darstellte.

Hinsichtlich der Kenntnis eines Dialekts der L1 wurde in Bezug auf die Herangehensweise der Teilnehmer:innen bei der Angabe solcher Fähigkeiten ein Unterschied zu den beiden anderen Gruppen festgestellt. Die meisten Teilnehmer:innen zögerten erheblich, ihre (auch nur passive) Kompetenz in einer Dialektvarietät des Deutschen zu erklären, wohingegen in den beiden anderen Gruppen die einzige anfängliche Schwierigkeit darin bestand, zu klären, dass der Dialekt als eine eigene Sprache zu anerkennen ist. Im Allgemeinen wurde die Angabe von Dialektkompetenzen in den anderen beiden Gruppen als eine Möglichkeit erlebt, die eigene sprachliche Vielfalt zu zeigen. In der Gruppe der L1 Deutsch-Sprecher:innen, die über Kenntnisse in einem Dialekt der L1 verfügten, konnte diese positive Wahrnehmung dagegen nicht beobachtet werden. Insgesamt gaben 10 Personen allgemeine Vorkenntnisse in einem Dialekt der L1 an. Die in der Gruppe der Teilnehmer:innen mit Vorkenntnissen in einem Dialekt vertretenen Bundesländer sind: Mecklenburg-Vorpommern (1 Frau), Nordrhein-Westfalen (1 Frau), Rheinland-Pfalz (1 Mann), Hessen (6 Proband:innen) und Baden-Württemberg (1 Frau). Da die meisten Teilnehmer:innen keine weiteren Angaben zu ihrer Herkunft machten, wird an dieser Stelle eine Klassifizierung auf geographischer Basis, bezogen auf die Bundesländer, vorgeschlagen.

Hinsichtlich der Englischkenntnisse in Bezug auf die Dominanz gab die Mehrheit der Teilnehmer:innen (21 Proband:innen) an, Englisch als die zweite dominante Sprache in ihrem Repertoire zu haben, während eine Minderheit (4 Frauen) es als die dritte dominante Sprache einstufte.

4.2.2 Aufbau des Experiments und Stimuli

Das Experiment wurde mit der Software *PsychoPy*® (Version 2021.2.3, Peirce et al., 2019) programmiert. Das experimentelle Design orientiert sich an der in Darcy & Krüger (2012) angewandten Methodik, in der eine *Oddity Discrimination-Task* für eine experimentelle Gruppe türkischsprachiger Kinder, die Deutsch als zweite Sprache ("L2", so in Darcy & Krüger 2012) erlernen, eingesetzt wurde. Das Design, das auf dem Prinzip der *Gamification* basiert, wurde auch für eine erwachsene Zielgruppe mit unterschiedlichen Deutschkenntnissen wie bei der vorliegenden Studie als geeignet erachtet. Die Benutzeroberfläche von Darcy & Krüger (2012)

wurde gemäß den von Carbotti (2015) definierten Richtlinien für die Gestaltung digitaler Inhalte¹¹ grafisch modifiziert. Das Hauptziel bestand darin, die grafische Benutzeroberfläche so benutzerfreundlich und ansprechend wie möglich zu gestalten, wobei die aktuellen grafischen Standards berücksichtigt wurden, an die die Zielgruppe dieses Experiments potenziell gewöhnt ist. Gemäß den Empfehlungen von Carbotti (2015) wurde bei der Programmierung des Experiments besonderes Augenmerk auf die Einbindung eines "Erzähler"-Charakters gelegt, der visuell durch die Anweisungen führt, jedoch ohne akustische Unterstützung auskommt. In Anlehnung an das Modell von Darcy & Krüger (2012) werden die Charaktere des Experiments durch Roboter dargestellt: ein farblich von den anderen unterscheidbarer Eingangs-Erzähler, um eine spätere Wiedererkennung zu vermeiden, sowie drei weitere Roboter – zwei identische weibliche und ein männlicher, der zwischen ihnen positioniert ist. Die experimentelle Benutzeroberfläche blieb während des gesamten Experiments konstant (vgl. Carbotti, 2015, S. 162) und war durch eine überwiegende Statik sowie eine einfache Wahl der Schriftarten gekennzeichnet, um potenzielle Ablenkungseffekte zu minimieren. Die einzige bewusste Abweichung von diesem Prinzip bestand in der Integration einer Mundbewegung der Roboter, die mit der Wiedergabe des entsprechenden akustischen Stimulus synchronisiert wurde. Laut Carbotti (2015) kann die Verwendung von Animationen zur Hervorhebung interaktiver Bereiche die Erfahrung mit digitalen Inhalten verbessern. In diesem Fall ermöglichte die animierte Mundbewegung den Teilnehmer:innen eine mühelose Identifikation der akustischen Stimulusquelle. Zudem wurde dadurch das weitere Element der Anthropomorphisierung des sprechenden Charakters integriert, wodurch die Interaktion näher an eine reale Kommunikationserfahrung herangeführt wurde.

Nachfolgend werden zwei Beispiel-Layouts dargestellt (Abbildung 5):

¹¹ Es sei an dieser Stelle darauf hingewiesen, dass sich die Richtlinien von Carbotti (2015) auf die Gestaltung von Inhalten für Kinder beziehen. Obwohl sich das vorliegende Experiment an Erwachsene richtete, erschien es dennoch sinnvoll, diesen grafischen Strategien zu folgen, da sie als geeignet für Lerner:innen aller Kompetenzstufen gelten und somit eine grundlegende Zugänglichkeit für jede Zielgruppe gewährleisten. Diese Entscheidung steht im Einklang mit dem Prinzip des Universal Design for Learning, das eine möglichst barrierefreie und zugängliche Gestaltung von Lernmaterialien für alle Nutzergruppen unabhängig von Alter oder Kompetenzniveau anstrebt (CAST, 2024. Universal Design for Learning Guidelines Version 3.0. Siehe <https://udlguidelines.cast.org>, letztlich abgerufen am 15.05.2025)

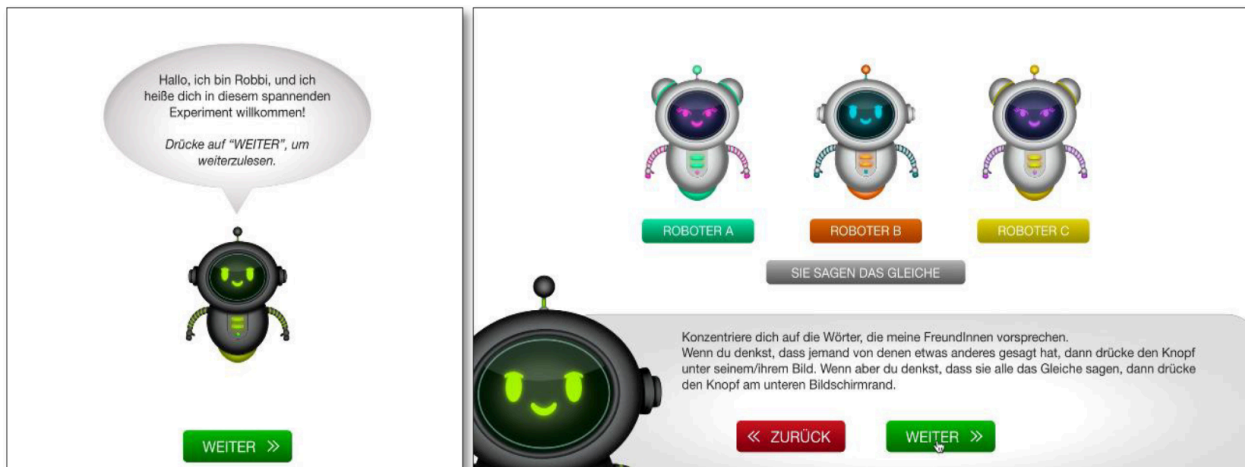


Abbildung 5: Beispiele für das Layout des Experiments. Links: Der Erzähler-Charakter stellt sich den Teilnehmer:innen vor und erklärt seine Funktion. Rechts: der Erzähler gibt einen Ausblick auf das Experiment-Layout und erläutert den Ablauf. Eigene Darstellung.

Die gewählte Silbenstruktur (KVKV) der Stimuli folgt dem für die deutsche Sprache typischen trochäischen Akzentmuster und wurde gewählt, weil sie die unmarkierte Form der deutschen Wörter repräsentiert (vgl. Wiese, 2000). Pseudowörter wurden genutzt, um lexikalische Einflüsse zu minimieren. Die KVKV-Sequenz ist auch im Italienischen vorherrschend, das durch das Vorhandensein einer offenen Endsilbe gekennzeichnet ist. Die Zielphoneme /ə/ und /y:/ wurden in die Pseudowörter gemäß den Regeln der deutschen Phonotaktik eingebettet. Dabei erscheint /ə/ typischerweise – wenn auch nicht ausschließlich – in wortfinaler unbetonter Position, während /y:/ bevorzugt in wortinitialer oder -medialer betonter Position vorkommt. Für /y:/ wurde ausschließlich die wortinterne Position verwendet. Der phonologische Kontext, in dem die beiden Phoneme realisiert wurden, wurde über das gesamte Experiment hinweg konstant gehalten: Für /ə/ wurde der Kontext #{/ko:b}/# und für /y:/ der Kontext #{/z/}{po:}/# gewählt. Die Pseudowörter wurden aufgebaut, indem die Zielphoneme an ihren jeweiligen Positionen gemäß den in den Hypothesen formulierten Assoziationsmustern eingefügt wurden: In den Sequenzen wurde daher /ə/ entweder mit /e:/ oder mit /ɛ/ in Kontrast gesetzt (H1a) und /y:/ mit dem Phonem /u:/ oder /i:/ (H2a). Zu den Assoziationsmustern wurden zwei Kontrollkontraste hinzugefügt, einer für jedes Zielphonem (/ə/-/i:/ und /y:/-/e:/). Die nachfolgende Darstellung zeigt die Assoziationsmuster gemäß den experimentellen Hypothesen mit den jeweils hinzugefügten Kontrollkontrasten (vgl. Abbildung 6):

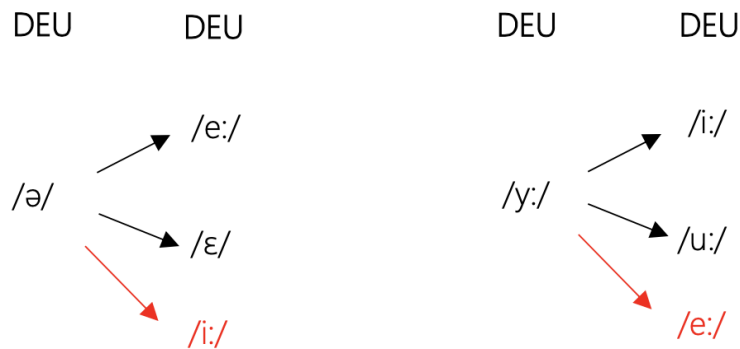


Abbildung 6: Darstellung der Assimilationsmuster (schwarz) nach den Hypothesen und der Kontrollkontraste (rot). Eigene Darstellung.

Sequenzen von Stimuli (im Folgenden als "Triplets" oder "Abfolgen" genannt) wurden entweder dadurch gebildet, dass jedes Mal ein Pseudowort mit einem anderen Phonem als die beiden anderen in der Sequenz platziert wurde (daher "Kontraste" genannt) oder indem dasselbe Pseudowort dreimal wiederholt wurde (daher "gleiche Abfolgen" genannt). Die mit dem Phonem /ə/ kontrastierende Vokale wurden in einem als "Block /ə/" bezeichneten Block von Stimuli zusammengefasst, und die mit dem Phonem /y:/ kontrastierende Vokale in einem als "Block /y:/" bezeichneten Block. Für jeden Kontrasttyp mit den beiden untersuchten Zielphonemen (/ə-e:/, /ə-ɛ/, /ə-i:/; /y:-u:/, /y:-i:/, /y:-e:/) wurden vier Sequenzmuster realisiert: zwei, in denen das Zielphonem die Position X einnimmt (AAX, AXA), und zwei, in denen das Zielphonem die Positionen A besetzt, während an der Position X ein anderes Phonem realisiert wird (ebenfalls AAX, AXA). Diese Kontraste werden im Folgenden als "primäre Kontraste" bezeichnet. Dasselbe Verfahren wurde auf die sogenannten "sekundären Kontraste" (/e:-ɛ/, /e:-i:/, /ɛ:-i:/, /i:-u:/, /e:-i:/, /e:-u:/) angewendet, bei denen die Zielphoneme nicht enthalten waren.

Insgesamt wurden so 24 Kontraste pro Block gebildet. Zu diesen wurden 24 gleiche Abfolgen pro Block hinzugefügt, 4 für jeden beteiligten Vokal. Der Stimulusdatensatz bestand insgesamt aus 96 Sequenzen (48 pro Block). Die Stimuli wurden dann innerhalb jedes Blocks kontrolliert randomisiert. Für die Durchführung des Experiments wurden zwei Versionen entwickelt. In der ersten Version wurden die Stimulusblöcke in der Reihenfolge Block /ə/ – Block /y:/, und in der zweiten Version in der umgekehrten Reihenfolge, also Block /y:/ – Block /ə/, vorgegeben. Beide Versionen wurden abwechselnd durchgeführt.

Zur konkreten Umsetzung der akustischen Reize wurden diese von deutschen L1-Sprecher:innen innerhalb eines Trägersatzes nach dem Modell von Darcy & Krüger (2012) aufgezeichnet: "Ich habe ein(e) _____ gesehen". Es sei darauf hingewiesen, dass in Darcy & Krüger (2012) der unbestimmte Artikel "ein" unabhängig vom nachfolgenden Stimulus immer gleich gehalten wurde;

in diesem Experiment wurde eine größere Flexibilität gewählt, die von der Intuition der Sprecher:innen abhing, die den Stimulus aussprechen würden. Dies begünstigte eine möglichst spontane Aussprache des gesamten Satzes. Für die vorliegende Studie wurden 2 Sprecher:innen (1 Frau und 1 Mann) rekrutiert, die zum Personal des Instituts für Germanistik der Universität Marburg gehören und sich freiwillig zur Aufnahme ihrer Stimme meldeten. Die Anweisung für die Sprecher:innen lautete, die Sätze spontan und mit neutraler Intonation auszusprechen. Zudem wurden die Sprecher:innen angewiesen, die Pseudowörter konsequent auf der vorletzten Silbe zu betonen. Vor der eigentlichen Aufnahme der Trägersätze wurde den Sprecher:innen ausreichend Zeit eingeräumt, um die Zielwörter zu trainieren und eine möglichst natürliche Artikulation sicherzustellen. Besonderes Augenmerk galt dabei jenen Pseudowörtern, die eine für das Standarddeutsche ungewöhnliche Phonotaktik aufweisen. Dies betrifft insbesondere das Pseudowort “kob[ɛ]”, bei dem der mitteltiefe Vorderzungenvokal in wortfinaler Position auftritt, sowie das Pseudowort “kob[e:]”, bei dem der finale Vokal explizit als Vollvokal und nicht als das in diesem Kontext übliche Schwa zu realisieren war. Da die Sprecher:innen über eine linguistische Ausbildung verfügten, wurden die kontrastierenden Phoneme zu ihrer Unterstützung bereits in IPA-Transkription in die Zielwörter integriert (z. B. *kob[ɛ]*). Demnach lautete die Zielaussprache der Stimuli wie folgt: [ˈkoːbə], [ˈkoːbe], [ˈkoːbɛ], [ˈkoːbi], [ˈzyːpo], [ˈzuːpo], [ˈziːpo], [ˈzeːpo]. Die Aufnahmen erfolgten in einer schallisolierten Kabine, in welcher die Sprecher:innen die Sätze anhand von Folien lasen, die sie eigenständig umblättern. Die Aufnahmen wurden unter Verwendung der Software *Audacity*® (Version 3.4.2, 2023) durchgeführt. Im Anschluss wurden die Aufnahmen einer gemeinsamen Evaluierung durch die Versuchsleiterin sowie zwei Betreuerinnen dieser Dissertation unterzogen. Keine der zwei Stimmen wiesen markierte regionale Färbungen auf. Um eine dritte, ebenso geeignete weibliche Stimme zu generieren, wurde die ursprünglich aufgezeichnete weibliche Stimme mithilfe der Software *Audacity*® modifiziert, um eine jüngere weibliche Sprechstimme zu imitieren. Im Anschluss wurden die einzelnen Stimuli mit der Software *Praat* (Version 6.4.22, Boersma & Weenink, 2024) aus dem Satzkontext isoliert und mit *Audacity*® hinsichtlich Lautstärke (normalisiert auf -20,0 dB) und Dauer (0,71 Sek.) angepasst.

4.2.3 Prozedere

4.2.3.1 Testung der Sprachkenntnisse vor der Experimentphase

Teilnehmer:innen der Gruppe der Lerner:innen, die keine Selbstauskunft über ihr Sprachniveau auf der Grundlage des letzten Sprachtests (nicht älter als 2/3 Jahre) geben konnten oder die im Besitz eines datierten Zertifikats waren, wurden einem Einstufungstest des Goethe-Instituts unterzogen. Der Test wurde nach der Unterzeichnung der Einverständniserklärung durchgeführt.

Der Test ist von kurzer Dauer und enthält eine Reihe von Aufgaben, mit denen das schriftliche und mündliche Verständnis der deutschen Sprache überprüft werden soll. Er folgt einem ähnlichen Modell wie die offiziellen Goethe-Sprachtests. Die am Ende des Tests angegebene

GER-Niveaustufe ist ein Richtwert und dient nur zur Orientierung. Das galt auch für diejenigen, die ihre GER-Niveaustufe auf der Grundlage ihrer Zeugnisse angegeben haben.

Die Teilnehmer:innen, für die der Test bestimmt war, wurden angewiesen, den Deutschunterricht zur Vorbereitung auf die Testsitzung nicht wieder aufzunehmen. Außerdem wurde der Name des Tests vor der Versuchssitzung nicht bekannt gegeben. Vor Beginn des Einstufungstests, der auf demselben Medium und mit denselben Kopfhörern wie das Experiment durchgeführt wurde, wurden die Teilnehmer:innen angewiesen, die Fragen auf der Grundlage ihrer tatsächlichen und aktuellen Deutschkenntnisse zu beantworten, also keine Fragen zu beantworten, bei denen sie sich der Antwort überhaupt nicht sicher waren, um so die Möglichkeit zu minimieren, dass ihr Sprachniveau zufällig berechnet wurde. Darüber hinaus gab es keine Unterstützung bei der Durchführung des Tests. Während der Durchführung des Tests nahm die für die Durchführung des Experiments verantwortliche Person in angemessenem Abstand hinter dem/der Teilnehmer/-in Platz, um eine bessere Konzentration zu ermöglichen. Nach Abschluss des Einstufungstests wurden das Kompetenzniveau offengelegt und die zugrunde liegenden Bewertungsmaßstäbe erläutert. Auf Anfrage wurde ein Bericht über die korrekten, fehlerhaften und fehlenden Antworten erstellt, der auf dem am Ende des Tests erstellten Abschlussbericht basierte. Es wird anekdotisch berichtet, dass die Teilnehmer:innen nach der Besprechung der erzielten Ergebnisse häufig ein größeres Bewusstsein für ihre eigenen Kompetenzen zum Ausdruck brachten und so die Bedeutung einer möglichst wahrheitsgetreuen Erfassung der Kompetenzen im Versuchskontext noch besser verstanden.

4.2.3.2 Die Experimentphase

Das Experiment wurde erst nach erfolgter Klärung der Modalitäten und Ziele sowie nach Unterzeichnung der Einverständniserklärung durchgeführt. Die Ziele des Experiments wurden in allgemeiner, aber dennoch klarer Form erläutert, ohne dass die Teilnehmer:innen dadurch in irgendeiner Weise beeinflusst werden konnten. Keine(r) von ihnen wusste also, auf welche Ebene der Sprachanalyse sich die Untersuchung konzentrierte. Letztere fand in Präsenz in einer einzigen Sitzung statt. Die Rekrutierung der Gruppe der Lerner:innen sowie der Kontrollgruppe ohne Deutschkenntnisse erfolgte im Rahmen eines Forschungsaufenthalts der Versuchsleiterin an der Universität Verona, wobei die Rekrutierung und Durchführung der Experimente direkt von der Versuchsleiterin geleitet wurde.

Für die Durchführung des Experiments wurde ein ruhiger Raum gewählt, in dem sich neben der Versuchsperson ausschließlich die Versuchsleiterin befand, um mögliche Ablenkungen zu vermeiden. Das Experiment wurde auf demselben Gerät durchgeführt – einem Laptop, der mit professionellen Kopfhörern ausgestattet war. Bereits zu Beginn der Studie zeigte sich eine allgemeine Anspannung bei den Teilnehmer:innen, die offenbar unter Leistungsdruck standen. Um dem entgegenzuwirken, wurde den Teilnehmer:innen mehrfach versichert, dass es sich keinesfalls

um einen Test der Deutschkompetenzen mit positiver oder negativer Bewertung handelte. Stattdessen wurden die Teilnehmer:innen dazu ermutigt, sich mit ihren aktuellen sprachlichen Fähigkeiten wohlfühlen. Die Versuchspersonen wurden gebeten, eine bequeme Sitzposition einzunehmen, da das Experiment etwa 20 Minuten dauern würde. Darüber hinaus bestand die Möglichkeit, zwischen der Nutzung der Maus oder des Trackpads des Laptops zu wählen, je nachdem, welche Methode die Teilnehmer:innen als komfortabler empfanden. Während der experimentellen Sitzung wurde von der Versuchsleiterin in angemessenem Abstand hinter den Teilnehmer:innen Platz genommen, um eine ungestörte Konzentration zu ermöglichen. Gleichzeitig wurde so die Überwachung des ordnungsgemäßen Ablaufs des Experiments sichergestellt. Der allgemeine Aufbau des Experiments war durch folgende Sequenzen gekennzeichnet: 1) Einleitung mit Instruktionen, 2) kurze Testsimulation, 3) erster Stimulusblock, 4) Pause, 5) zweiter Stimulusblock, 6) Ende des Experiments. Die Gesamtdauer der Sitzung betrug ca. 20 Minuten, abhängig von der Länge der Pause, die der Teilnehmer:innen nach eigenem Ermessen einlegen konnte. Vor Beginn des Experiments wurden die Teilnehmer:innen angewiesen, den gespielten Stimuli große Aufmerksamkeit zu schenken, wobei sie sich nicht auf den Inhalt (als Pseudowörter), sondern auf ihre "Form", d. h. die Art und Weise ihrer Aussprache und etwaige Unterschiede, konzentrieren sollten. Damit die Teilnehmer:innen so spontan wie möglich antworten, aber gleichzeitig ihre Antworten nicht erzwingen, wurden sie einerseits angewiesen, so schnell und spontan wie möglich zu antworten, aber, wenn sie völlig unsicher waren, welche Option sie wählen sollten, nichts anzuklicken und zu warten, bis das Experiment zur nächsten Sequenz überging. Dieses Kriterium stimmt mit den Überlegungen von Isbell (2016) überein, dass die Erfassung der Wahrnehmung mittels erzwungener Wahlmöglichkeiten Zufall oder Raten als störende Elemente induzieren kann (Isbell, 2016, S. 62). Die in Phase 2 verwendeten Stimuli (kurze Testsimulation mit insgesamt 5 Triplets) enthielten andere Kontrastmuster als die für das Experiment vorgesehenen, jedoch an denselben Positionen (Wortinneres und Wortende) und boten sowohl Kontraste als auch gleiche Abfolgen. Ziel war es, die Teilnehmer:innen an die Art der Aufgabe zu gewöhnen, ohne jedoch zu viele Anhaltspunkte zu geben. Die ersten Dreiergruppen bestanden aus echten Wörtern (z. B. *Hase, Hose, Hase; Biene, Bohne, Biene*) und anschließend aus Kontrasten zwischen Pseudowörtern mit einem anderen Silbenkontext als dem für die Versuchsphase gewählten (z. B. *Gabo, Gebo, Gabo*). Diese Wahl war durch den Wunsch motiviert, die Teilnehmer:innen allmählich daran zu gewöhnen, sich von den lexikalischen Einflüssen zu lösen und sich nur auf den akustischen Aspekt der Stimuli zu konzentrieren. Während der Versuchsphase wurden die Stimuli automatisch abgespielt; erst am Ende der Wiedergabe hatten die Teilnehmer:innen maximal 4 Sekunden Zeit, ihre Antwort zu wählen.

4.2.3.3 Die Erfassung der Daten zu Sprachkompetenz, -dominanz und Sprachrepertoire

Die Erfassung außersprachlicher Daten erfolgte in zwei Momenten der gesamten Experimentsitzung. Erst wurde das GER-Niveau der Proband:innen entweder nach dem Ergebnis des letzten

erworbenen Sprachzertifikat oder durch einen kurzen Online-Einstufungstest erfasst. War das Zertifikat älter als zwei/drei Jahre oder lag kein Zertifikat vor, auf das sich die Selbstauskunft stützen konnte, wurden die Teilnehmer:innen vor der Durchführung des Experiments dem Online-Einstufungstest des Goethe-Instituts unterzogen, um ihr aktuelles GER-Kenntnisniveau zu ermitteln¹².. Für den Einsatz des Goethe-Einstufungstests gibt es mehrere Gründe. Der Goethe-Einstufungstest bietet eine abwechslungsreiche Reihe von Kurztests zur Beurteilung der Lese- und Hörverstehenskompetenz auf der Grundlage von Kriterien, die mit denen von Texten der Amtssprache vergleichbar sind. Er kann außerdem in aller Ruhe ausgefüllt werden, ohne dass die Testteilnehmer unter Druck gesetzt werden.

Zusätzlich zur Durchführung des Experiments wurde am Ende des Experiments der Fragebogen *Language Experience And Proficiency Questionnaire* (LEAP-Q, Marian et al., 2007) ausgefüllt, um außerlinguistische Informationen zu sammeln (Marian et al., 2007, siehe Appendix). Der LEAP-Q ist ein Fragebogen, der für die Erfassung außerlinguistischer Daten in gesunden erwachsenen Populationen mit zweisprachigen oder mehrsprachigen Kompetenzen konzipiert wurde. Diese zeichnen sich potenziell durch verschiedene sprachliche und kulturelle Hintergründe aus und verfügen über mindestens eine Abiturschulung (vgl. Marian et al. 2007, S. 962). Ein wesentlicher Vorteil dieses Erhebungsinstruments besteht darin, dass es in zahlreichen übersetzten Fassungen frei zur Verfügung steht, darunter auch in italienischer und deutscher Sprache. Die Teilnehmer:innen, die über keine ausreichenden Deutschkenntnisse für das klare Verständnis der Inhalte verfügten, haben die italienische Fassung ausgefüllt, während alle anderen die deutsche Version erhalten haben. Der Fragebogen umfasst einen ersten generellen Teil, der der Erfassung von Angaben zum Hintergrund, zum Sprachrepertoire, zum sprachlichen Verhalten, zu eventuellen Erfahrungen in Deutschland, zum Schulungsniveau und zur eventuellen Anweisung von Behinderungen gewidmet ist. Der zweite Teil des LEAP-Q besteht aus einer maximalen Anzahl von 5 Karten (eine für jede bekannte Sprache). Mithilfe dieser Karten ist eine Vertiefung der spezifischen Sprachkenntnisse und des Sprachverhaltens zu jeder angegebenen Sprache möglich. Die Beantwortung des Fragebogens durch eine Person, die zwei Sprachen beherrscht, dauert durchschnittlich 15 Minuten (vgl. Marian et al. 2007). Die Bearbeitung einer zusätzlichen Karte erfordert etwa 5 Minuten (ebd.). Die Teilnahme erfolgte entweder vor Ort oder im häuslichen Umfeld, je nach Bedarf der Teilnehmer:innen. In beiden Fällen stand die Unterstützung der Versuchsleiterin zur Verfügung. Insbesondere bei der Interpretation spezifischer Termini wurde auf die Expertise der Versuchsleiterin zurückgegriffen. Die Teilnehmer:innen wurden während des Ausfüllens des Fragebogens gebeten, etwaige Dialektkenntnisse zu spezifizieren, indem sie den Ort angaben, an dem die entsprechende dialektale Varietät erworben wurde. Wurde die dialektale Varietät gleichzeitig mit dem regionalen Italienisch erworben, wurde diese als "2L1" klassifiziert.

¹²Der Goethe-Einstufungstest ist unter <https://www.goethe.de/de/spr/kur/tsd.html> abrufbar (letztlich abgerufen am: 25.02.2025).

4.2.4.Datenanalyse

Zur Beantwortung der Forschungsfragen wurden zwei abhängige Variablen gemessen: Akkuratheit und Reaktionszeit. Der Datensatz umfasste die von *PsychoPy*® automatisch erstellten Outputdateien, in die die relevanten außerlinguistischen Informationen aus den ausgefüllten Fragebögen integriert wurden. Vor der Durchführung der Analysen wurden die Daten anonymisiert und die Namen der Teilnehmer:innen durch alphanumerische Codes ersetzt. In die Modelle der statistischen Analysen wurden die Items (Sequenzen) sowie die Versuchspersonen als Zufallsfaktoren aufgenommen. Als feste Faktoren gingen die Gruppen (Lerner:innen, Kontrollgruppe mit L1 Italienisch, Kontrollgruppe mit L1 Deutsch), die Blöcke (Block /ə/ und Block /y:/), die Art der Sequenz (Kontrast- vs. identische Abfolge), die spezifischen Kontraste, deren Richtung sowie verschiedene außerlinguistische Bedingungen ein (GER-Niveau im Deutschen, Dominanz im Deutschen und Englischen sowie die Breite des sprachlichen Hintergrundrepertoires).

Bei einer Untergruppe von Teilnehmer:innen, die angaben, über Dialektkompetenz zu verfügen, die potenziell (im Sinne eines positiven Transfers) die Wahrnehmung von /ə/ und /y:/ beeinflusst, wurde auch der Einfluss von Dialekthintergrundwissen getestet, indem die Variable bezogen auf die Dialektkenntnisse als fester Faktor aufgenommen wurde. Die statistische Analyse dieser Untergruppe wurde auf einem rein deskriptiven Niveau durchgeführt und wird im Folgenden als "explorative Studie" bezeichnet.

Für die statistischen Methoden wurde die Software *Jamovi* (Version 2.4.11.0, 2025) verwendet. Für die deskriptiven Statistiken wurden durchschnittliche Prozent- und SE-Werte (Standardfehler) für die Akkuratheit sowie Durchschnittswerte und SE für die Reaktionszeiten¹³ herangezogen, die ausschließlich aus den korrekten Antworten berechnet wurden. Darüber hinaus wurden alle Werte, die als zu hoch angesehen wurden und die statistischen Analysen möglicherweise beeinträchtigen könnten, ausgeschlossen, wobei ein maximaler Schwellenwert von 3 Sekunden festgelegt wurde (alle Daten, die länger als 3 Sekunden waren, wurden daher ausgeschlossen). Bei den inferentiellen Analysen wurde das Signifikanzkriterium durch den p-Wert ausgedrückt ($p < 0,05 = *$, $p < 0,01 = **$, $p < 0,001 = ***$).

Im Fall der Akkuratheit wurden generalisierte gemischte Modelle (GMM) und Post-Hoc-Tests mit Bonferroni-Korrektur verwendet, um die Interaktion zwischen Akkuratheit und den strukturellen und außerlinguistischen Variablen zu analysieren. In der vorliegenden Analyse wurde eine a-posteriori akustische Analyse der Stimuli durchgeführt, da initial keine Notwendigkeit für eine tiefgehende Analyse gesehen wurde. Diese Analyse erfolgte, um eine vertiefte Untersuchung der beobachteten Besonderheiten in der Variation der Akkuratheitswerte in Bezug auf die Richtung der primären Kontraste sowie die Art der sekundären Kontraste zu ermöglichen. Dabei kamen ein

¹³ Hier werden die Reaktionszeiten in Sekunden mit zwei Dezimalstellen angegeben. Daher wurden, wo möglich, Aufrundungen vorgenommen.

gemischtes lineares Modell und ein Post-Hoc-Test mit Bonferroni-Korrektur zum Einsatz. Für diese tiefgehende Analyse wurden die Stimuli isoliert, annotiert und durch ein *Praat* (Boersma & Weenink, 2025) Skript extrahiert.

In Bezug auf die Analyse der Reaktionszeiten wurden lineare gemischte Modelle sowie Post-Hoc-Tests unter Verwendung der Bonferroni-Korrektur auf die logarithmischen Werte der Reaktionszeiten mit einer Dauer von höchstens drei Sekunden angewendet.

Im Hinblick auf die Analyse von taskbedingten Variablen, wie etwa die Stimulusposition innerhalb der Triplets, das Geschlecht der Stimme der abgespielten Stimuli sowie die Experimentphase, fand die Anwendung von Generalisierten Gemischten Modellen mit Post-Hoc-Tests und einer Bonferroni-Korrektur statt. Die Auswahl dieser Modelle erfolgte unter Berücksichtigung der nicht-normalen Datenverteilung, welche durch einen Shapiro-Wilk-Test überprüft wurde.

4.3 Ergebnisse

4.3.1 Akkuratheit

Die deskriptive statistische Analyse der Durchschnittswerte der Akkuratheit wurde zunächst mit dem Vergleich zwischen den Gruppen begonnen. Die Lerner:innen sind als "Lernende", die Kontrollgruppe mit L1 Italienisch ohne Deutschkenntnisse als "KO_ITA" und die Kontrollgruppe mit L1 Deutsch als "KO_DEU" gekennzeichnet. Die folgende Grafik (Abbildung 7) zeigt die Ergebnisse:

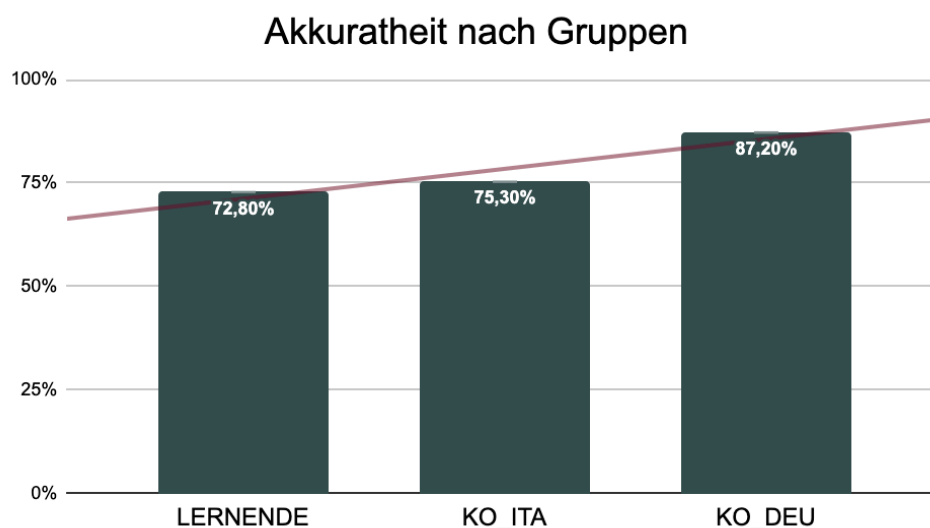


Abbildung 7: Durchschnittliche Akkuratheitswerte der Probandengruppen mit Standardfehlern (SE) und Trendverlauf.

Der beobachtete Trend zeigt die niedrigsten durchschnittlichen Akkuratheitswerte für die Gruppe

der Lerner:innen (72,80 %, SE: 0,75 %), gefolgt von der Kontrollgruppe ohne Deutschkenntnisse (75,30 %, SE: 0,84 %) und schließlich, mit der höchsten durchschnittlichen Akkuratheit (87,20 %, SE: 0,68 %), die Kontrollgruppe mit L1 Deutsch. Um den Effekt der unabhängigen Variable *Gruppe* auf die abhängige Variable *Akkuratheit* abzuleiten, wurde ein verallgemeinertes gemischtes Modell angewendet.

Folgendes Modell wurde angewendet:

$$lme4::glmer\ Akkuratheit \sim 1 + Gruppe + (1 | Items) + (1 | Versuchsperson).$$

Als abhängige Variable wurde die Akkuratheit angegeben, während die Variable *Gruppe* als Faktor berücksichtigt wurde. Als Zufallsfaktoren wurden die Variable *Items* und *Versuchsperson* angegeben.

Die Ergebnisse zeigten einen signifikanten Effekt ($p < .001$) des Faktors *Gruppe*. Der Post-Hoc-Test ergab einen signifikanten Unterschied zwischen der Gruppe der Lerner:innen und der Kontrollgruppe mit deutscher L1 ($p < .001$) und zwischen der Kontrollgruppe ohne Deutschkenntnisse ($p < .001$), jedoch keinen signifikanten Unterschied zwischen der Gruppe der Lerner:innen und der Kontrollgruppe ohne Deutschkenntnisse ($p = 0,9$).

Im Detail wurden die durchschnittlichen Akkuratheitswerte innerhalb jeder Gruppe analysiert, wobei zwischen den Werten in den beiden verschiedenen Stimulusblöcken unterschieden wurde, die in der folgenden Grafik (Abbildung 8) mit “/ə/” und “/y:/” gekennzeichnet sind. Diese zwei Blöcke umfassen die (primären und sekundären) Kontraste und die gleichen Abfolgen in Bezug auf die Wahrnehmung von /ə/ und /y:/.

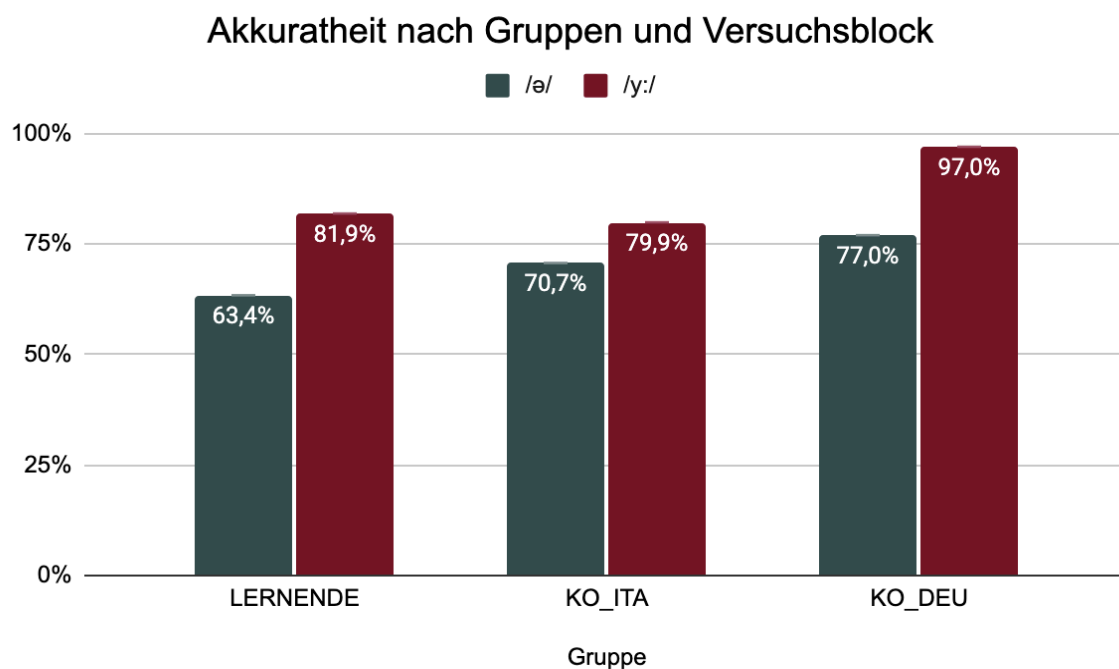


Abbildung 8: Durchschnittliche Akkuratheit nach Probandengruppen und Versuchsblock. Die Fehlerbalken geben den Standardfehler (SE) an.

Auf deskriptiver Ebene sind deutlich niedrigere mittlere Akkuratheitswerte bei den Sequenzen zur Wahrnehmung von /ə/ und ähnlichen Phonemen zu beobachten als bei denen zu /y:/ und ähnlichen Phonemen. Für den Versuchsblock zu /ə/ (“/ə/” in der Graphik) ist die gleiche Tendenz wie in der vorangegangenen Grafik (Abbildung 5) zu beobachten, mit den niedrigsten durchschnittlichen Akkuratheitswerten für die Gruppe der Lerner:innen (63,4 %, SE 8,1 %), gefolgt von denen der Kontrollgruppe ohne Deutschkenntnisse (70,7 %, SE 8,7 %) und den höchsten Werten der Kontrollgruppe mit Deutsch L1 (77,0 %, SE 8,4 %). Es zeigt sich also eine generelle Schwierigkeit bei der Wahrnehmung der Stimuli für diesen Versuchsblock, die für jede Gruppe unterschiedlich ausgeprägt ist. Ein anderes Bild ergibt sich hingegen bei der Wahrnehmung des Versuchsblocks zu /y:/ (“/y:/” in der Graphik), bei dem der Trend hier nicht demselben Verlauf folgt, sondern höhere Durchschnittswerte für die Gruppe der Lerner:innen (81,9%, SE 6,4 %) im Vergleich zur Kontrollgruppe ohne Deutschkenntnisse (79,9%, SE 7,7 %) und sehr hohe Werte für die Kontrollgruppe mit deutscher L1 (97%, SE 3,4 %) aufweist.

Um den Effekt des Faktors *Block* auf die Akkuratheit und die Interaktion zwischen dem Faktor *Gruppe* und *Block* zu berechnen, wurde folgendes Modell angewendet:

lme4::glmer Akkuratheit ~ 1 + Gruppe + Block + Gruppe:Block + (1 | Versuchsperson) + (1 | Items).

Als abhängige Variable wurde die Akkuratheit angegeben. Als Faktoren wurden die Variablen *Gruppe* und *Block* berücksichtigt. Als Zufallsfaktoren wurden sowohl die Variable *Versuchsperson* als auch die Variable *Items* angegeben.

Die Ergebnisse zeigten hohe Signifikanz ($p < .001$) für alle drei festen Effekte. Insbesondere der für den Faktor *Block* durchgeführten Post-Hoc-Test zeigte eine hohe Signifikanz ($p < .001$) beim Vergleich zwischen Block /ə/ und Block /y:/ und bestätigte damit, was bereits auf der deskriptiven Ebene ersichtlich war. Darüber hinaus ergab der Post-Hoc-Test für die Interaktion *Gruppe*Block* Folgendes: 1) für alle drei Gruppen gibt es einen hochsignifikanten Unterschied in den mittleren Akkuratheitswerten zwischen den Versuchen in Block /ə/ und Block /y:/; 2) in Bezug auf Block /ə/ sind die mittleren Akkuratheitswerte der Lerner:innen-Gruppe signifikant niedriger ($p < .001$) als die der Kontrollgruppe mit Deutsch L1, aber nicht signifikant anders als die der Kontrollgruppe ohne Erfahrung ($p = 0,439$), während die Kontrollgruppe ohne Deutschkenntnisse Durchschnittswerte zeigte, die nicht signifikant niedriger waren als die der Kontrollgruppe mit Deutsch L1 ($p = 0,939$); 3) In Bezug auf Block /y:/ sind die mittleren Akkuratheitswerte der Lerner:innen-Gruppe signifikant niedriger ($p < .001$) als die der Kontrollgruppe mit Deutsch L1, aber nicht signifikant verschieden von denen der Kontrollgruppe ohne Deutschkenntnisse ($p = 1,000$), während es einen signifikanten Unterschied ($p < .001$) in den mittleren Akkuratheitswerten zwischen den beiden Kontrollgruppen zugunsten der Kontrollgruppe mit Deutsch L1 gibt.

Was die Variation der Akkuratheit in Abhängigkeit von der Art der Abfolge (Kontraste oder gleiche Abfolgen) betrifft, so ist im Allgemeinen die gleiche Tendenz wie zuvor zu beobachten, wobei die Gruppe der Lerner:innen in beiden Bedingungen die niedrigsten Werte aufweist, gefolgt von der Kontrollgruppe ohne Deutschkenntnisse und der Kontrollgruppe mit Deutsch L1. Auch für die gleichen Abfolgen werden niedrigere Werte beobachtet als für die Kontraste: Für die Lerner:innen ergeben sich durchschnittliche Prozentwerte von 70,5 % (SE 7,6 %) bzw. 75,2 % (SE 7,2 %), für die Kontrollgruppe ohne Deutschkenntnisse 74,5 % (SE 8,4 %) bzw. 76,2 % (SE 8,2 %), für die Kontrollgruppe mit Deutsch L1 85,8 % (SE 7 %) bzw. 88,5 % (SE 6,4 %).

Um den Effekt der Variable zu testen, wurde folgendes Modell angewendet:

$$\text{lme4::glmer Akkuratheit} \sim 1 + \text{Gruppe} + \text{gleich_ungleich_Bedingung} + \text{Block} + \\ \text{Gruppe:gleich_ungleich_Bedingung} + \text{Gruppe:Block} + \text{gleich_ungleich_Bedingung:Block} \\ + \text{Gruppe:gleich_ungleich_Bedingung:Block} + (1 | \text{Items}) + (1 | \text{Versuchsperson})$$

Als abhängige Variable wurde die Akkuratheit angegeben, während als Faktoren die Variablen *Gruppe*, *gleich_ungleich_Bedingung* und *Block* berücksichtigt wurden. Als Zufallsfaktoren wurden die Variablen *Items* und *Versuchsperson* angegeben.

Unter den Ergebnissen, einschließlich der Ergebnisse des Post-Hoc-Tests für die Variable *gleich_ungleich_Bedingung*, ist ein besonders signifikanter Effekt ($p < .001$) der Variable

gleich_ungleich_Bedingung festzustellen, nicht aber der Interaktion zwischen dieser und der Variable *Gruppe* ($p=0,398$).

4.3.1.1 Ergebnisse zur *perzeptiven Assimilation* von /ə/ zu /e:/ oder /ɛ/

In diesem Unterkapitel konzentrieren wir uns auf Ergebnisse, die sich auf Assimilationsmuster und damit auf die Wahrnehmung von Stimuli in H1 beziehen. Die Analyse umfasst daher einen Vergleich der Bedingung, in der die Stimuli in Form von Kontrasten (mit /ə/ in der odd- und nicht-odd-Bedingung) erscheinen, und der Bedingung, in der sie in Form von gleichen Abfolgen erscheinen. Darüber hinaus werden die Ergebnisse für die einzelnen beteiligten und für die Hypothese relevanten Kontraste angegeben. Darüber hinaus werden am Ende des Abschnitts die Ergebnisse zur Variation der Mittelwerte der Akkuratheit und des Dominanzniveaus im Englischen berichtet. Beginnend mit dem Vergleich der beiden Arten von Stimuli Abfolgen (Kontraste/gleiche Abfolgen), zeigt das folgende Balkendiagramm (Abbildung 9) die durchschnittlichen prozentualen Werte von Akkuratheit und den Standardfehler (SE) für diese beiden Bedingungen in allen drei Gruppen:

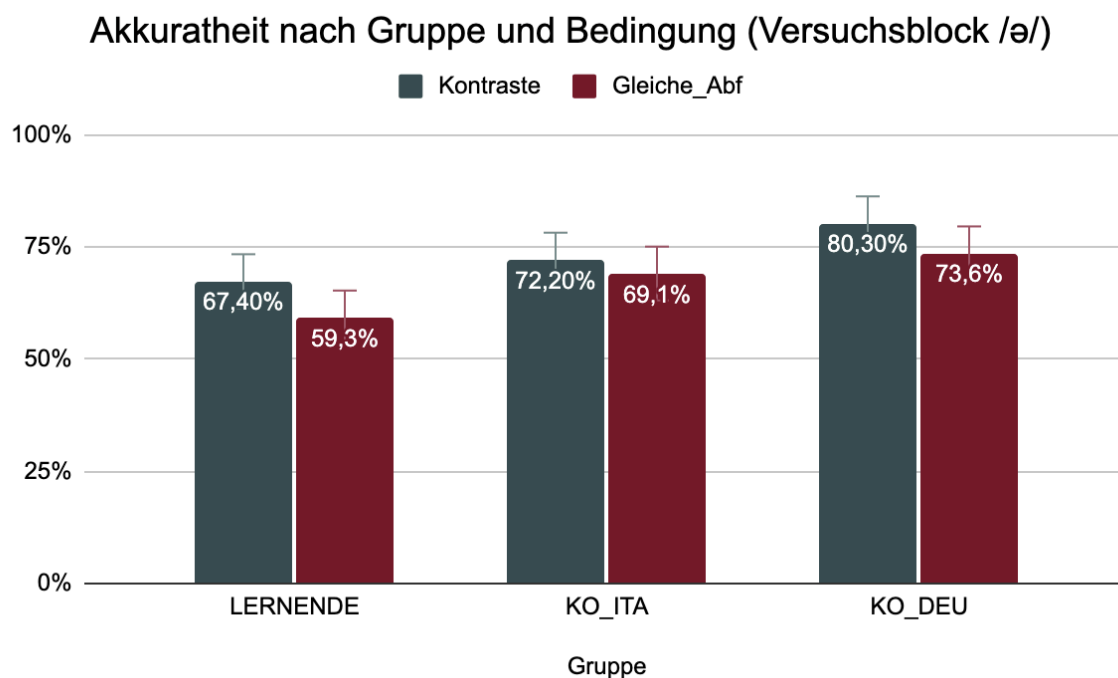


Abbildung 9: Durchschnittliche Akkuratheit im Versuchsblock /ə/ für Kontraste und gleiche Abfolgen pro Probandengruppe. Die Fehlerbalken geben den Standardfehler (SE) an.

Die beiden dargestellten Bedingungen umfassen alle Versuche, in denen /ə/ vorkommt. Die Ergebnisse zeigen generell niedrigere Werte für die Gruppe der Lerner:innen. Insbesondere in dieser Gruppe ist der Unterschied in der Akkuratheit in den beiden Bedingungen relativ gering

(Kontraste: 67,4%, SE 1,6% ; gleiche Abfolgen: 59,3%, SE 1,7%). Das Verhältnis zwischen den beiden Bedingungen ist bei der Kontrollgruppe ohne Deutschkenntnisse ähnlich, wo die höchsten Werte der ersten Bedingung vorbehalten sind (Kontraste: 72,2%, SE 1,76%; gleiche Abfolgen: 69,1%, SE 1,85%). Bei der Kontrollgruppe der Teilnehmer:innen mit deutscher L1 ist der Unterschied zwischen den beiden Bedingungen ebenfalls gering (Kontraste: 80,3%, SE 1,63%; gleiche Abfolgen: 73,6%, SE 1,83%). Um den Effekt der unabhängigen Variable Kontrast/gleiche Abfolgen (*gleich_ungleich_Bedingung*) zwischen den Gruppen zu testen, wurde folgendes Modell angewendet:

$$\text{lme4::glmer Akkuratheit} \sim 1 + \text{Gruppe} + \text{gleich_ungleich_Bedingung} + \text{Gruppe} \\ \text{gleich_ungleich_Bedingung} + (1 \mid \text{Item}) + (1 \mid \text{Versuchsperson}).$$

Als abhängige Variable wurde die Akkuratheit angegeben, während als Faktoren die Variablen *Gruppe* und *gleich_ungleich_Bedingung* berücksichtigt wurden. Als Zufallsfaktoren wurden die Variablen *Item* und *Versuchsperson* angegeben.

Die Ergebnisse zeigen einen hochsignifikanten Effekt des Faktors *Gruppe* ($p < .001$). Für die Variable *gleich_ungleich_Bedingung* sowie für die Interaktion *Gruppe*gleich_ungleich_Bedingung* konnte kein signifikanter Effekt festgestellt werden (jeweils $p = 0,066$ und $p = 0,055$). Bei dem Post-Hoc Test ergibt sich einen hochsignifikanten Unterschied für den Vergleich *KO_DEU* – *Lernende* ($p < .001$), für die Vergleiche *KO_DEU* – *KO_ITA* und *KO_ITA* – *Lernende* konnte kein signifikanter Effekt festgestellt werden (jeweils $p = 0,182$ und $p = 0,131$).

Im Folgenden wird die Kontrastbedingung analysiert. Unter der Variable *Kontrast* fallen sowohl Primärkontraste sowie Sekundärkontraste. Da die Primärkontraste ($/ə-e:/$, $/ə-ε/$ und der Kontrollkontrast $/ə-i:/$) die Forschungsfragen unmittelbar adressieren, wird im Folgenden ausschließlich die Verteilung der Akkuratheitswerte für diese Kontraste in den drei Gruppen dargestellt (Abbildung 10).

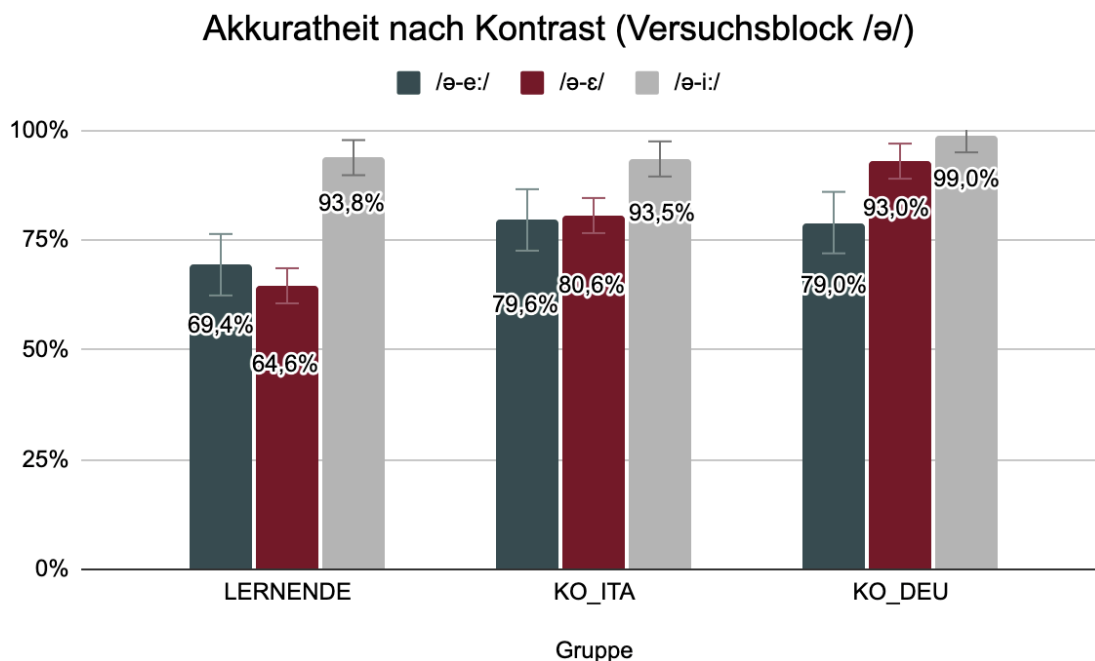


Abbildung 10: Durchschnittliche Akkuratheit in der Perzeption der Kontraste /ə-ɛ/, /ə-e:/ und /ə-i:/ nach Probandengruppe. Die Fehlerbalken stellen den Standardfehler (SE) dar.

Generell finden sich die niedrigsten Werte in der Gruppe der Lerner:innen, während die höchsten Werte in der Kontrollgruppe mit deutscher L1 zu finden sind. Der Kontrast mit den höchsten mittleren Prozentwerten ist der Kontrollkontrast /ə-i:/. Was die Verteilung der Werte in den Kontrasten /ə-e:/ und /ə-ɛ:/ betrifft, so sind sie je nach Gruppe unterschiedlich gestaltet. Die Gruppe der Lerner:innen zeigt besonders niedrige Werte für den /ə-ɛ/-Kontrast (64,6%, SE 4,0%), gefolgt von denen für den /ə-e/-Kontrast (69,4%, SE 3,85%). Bei der Kontrollgruppe ohne Deutschkenntnisse weisen /ə-e:/ und /ə-ɛ/ eine entgegengesetzte Verteilung zu der der Lerner:innen auf: Der /ə-ɛ/-Kontrast weist die höchsten Werte auf (80,6%, SE 3,83%), gefolgt von denen für den /ə-e/-Kontrast (79,6%, SE 3,89%). Bei der Kontrollgruppe mit deutscher L1 finden sich die höchsten durchschnittlichen Prozentwerte für /ə-ɛ/ und /ə-i:/ (jeweils 93%, SE 2,56%; 99%, SE 1,01%). Für /ə-e:/ liegen die Werte jedoch deutlich niedriger und näher an denen der Kontrollgruppe ohne Deutschkenntnisse (79%, SE 4,09%).

Bei der Wahrnehmung der sekundären Kontraste (/ɛ:-ɛ/, /ɛ-i:/ und /e:-i:/) ergeben sich für den Kontrast /ɛ-i:/ in allen drei Gruppen besonders hohe Prozentangaben, wobei die Anteile zwischen den Gruppen dem oben bereits beobachteten Trend folgen: Für die Gruppe der Lerner:innen sind die Werte anteilig am niedrigsten (94,4%, SE 3,8%), während für die Kontrollgruppe ohne

Deutschkenntnisse und die Gruppe mit deutscher L1 die Werte durchweg höher sind (98,1%, SE 2,6%, bzw. 99%, SE 2%). Der zweifellos problematischste Kontrast ist für alle drei Gruppen /e:-i:/, der für die Lerner:innen einer Akkuratheit von 16,3% (SE 0,9%) entspricht, für die Kontrollgruppe ohne Deutschkenntnisse sinken die Werte noch weiter auf 10,8% (SE 1,4%), während für die Kontrollgruppe mit deutscher L1 die Werte mit 30% (SE 2,7%) am höchsten sind. Da in diesem Fall der Vokal /ə/ nicht den Vokal mit der niedrigsten Salienz darstellt und sich diese Vokale alle in peripheren Bereichen befinden, die theoretisch ein geringeres reziprokes Salienzverhältnis aufweisen als das zwischen ihnen und einem zentralen Vokal, wurde es als angemessener erachtet, sich auf die intrinsischen qualitativen Merkmale der beteiligten Vokale zu konzentrieren. Die Anwendung des oben erwähnten linearen gemischten Modells und insbesondere der Post-Hoc-Test haben ergeben, dass für den Vergleich der Vokale /e:-/i:/ kein signifikanter Unterschied ($p=1$) zwischen ihren F2-Werten besteht.

Um den Effekt der unabhängigen Variable *Kontrast* und *Gruppe* auf die abhängige Variable *Akkuratheit* zu testen, wurde folgendes Modell angewendet:

$$lme4::glmer \text{ Akkuratheit} \sim 1 + \text{Gruppe} + \text{Kontrast} + \text{Gruppe}:\text{Kontrast} + (1 | \text{Items}) + (1 | \text{Versuchsperson}).$$

Als abhängige Variable wurde die Akkuratheit angegeben, während als Faktoren die Variable *Gruppe* und die Variable *Kontrast* berücksichtigt wurden. Als Zufallsfaktoren wurden die Variablen *Items* und *Versuchsperson* angegeben.

Die Ergebnisse zeigten einen hochsignifikanten Effekt für den Faktor *Kontrast* und *Gruppe* ($p<.001$); für ihre Interaktion konnte ein leicht signifikanter Effekt festgestellt werden ($p=0,042$). Aus dem Post-Hoc Test haben sich hochsignifikante Unterschiede in der Akkuratheit zwischen dem Kontrast /e:-i:/ und allen anderen Vokalkontrasten (/e:-ε/, /ə-ε/, /ə-i/, /ε-i/). Paarweise Vergleiche bestätigten, dass die Akkuratheit bei /e:-i:/ im Vergleich zu /e:-ε/, /ə-e:/, /ə-ε/, /ə-i:/ und /ε-i:/ signifikant abfällt ($p<.001$). Darüber hinaus unterscheiden sich die Kontraste mit Schwa-Lauten untereinander: Die Akkuratheit bei /ə-i:/ ist signifikant höher als bei /ə-ε/ ($p= 0,028$), bei /ə-e:/ ergibt sich als signifikant niedriger als bei /ə-i:/ ($p=0,013$). Zwischen den Kontrasten /ə-ε/ und /ə-e:/ besteht kein signifikanter Unterschied hinsichtlich der Akkuratheit ($p = 1,000$), wobei die Akkuratheit beim Kontrast /ə-e:/ leicht niedriger ausfällt.

Um ein umfassenderes Verständnis des Wahrnehmungsverhaltens der Teilnehmer:innen gegenüber Primärkontrasten zu erhalten, wurden die Akkuratheitswerte auch in Bezug auf die Richtung analysiert, in der die Phoneme innerhalb des Kontrasts präsentiert wurden. Insbesondere wurde die Analyse auf eine Teilmenge von Kontrasten mit /ə/ beschränkt, bei denen sich das kontrastive Element in der Anfangs- oder Endposition befand, so dass eine Richtung entsprechend dem Übergang, dem die Abfolge folgt, definiert werden konnte. Die Überlegung, die hinter dieser

Analyse steht, ist, dass die Akkuratheit höhere Werte in Richtungen zu den Randzonen des Vokalraums aufweisen kann, Zonen, die von Vokalen (insbesondere /i:/) mit hoher Salienz besetzt sind. Die folgende (Abbildung 11) zeigt die Verteilung und SE der mittleren Prozentwerte der Akkuratheit zwischen den Gruppen und Richtungen der Kontraste.

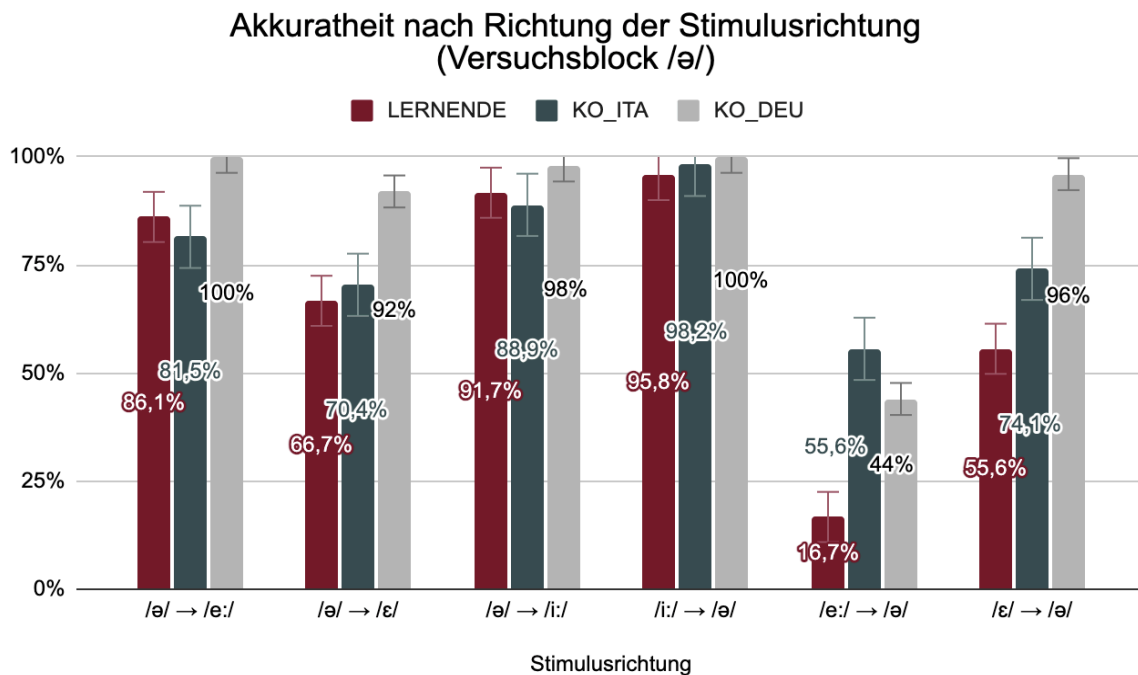


Abbildung 11: Durchschnittliche Akkuratheit in Abhängigkeit von der Stimulusrichtung im Versuchsblock /ə/. Die Fehlerbalken geben den Standardfehler (SE) an.

Es ist interessant zu beobachten, dass die Kontraste in den Richtungen /ə/ → /i:/ und /i:/ → /ə/ in allen drei Gruppen hohe und ziemlich vergleichbare Werte aufweisen, ebenso wie die Richtungen /ə/ → /e:/ und /ə/ → /ɛ/, wobei letztere die niedrigsten Werte in allen drei Gruppen aufweist. Besonders hohe Akkuratheitswerte sind in den Richtungen vom zentralen mittleren (/ə/) zum vorderen hohen Artikulationsort (/i:/) bzw. vom zentralen mittleren zum vorderen hohen (/e:/) und vorderen tiefen Artikulationsort (/ɛ/). Besonders niedrige Werte finden sich in den Richtungen /e:/ → /ə/ und /ɛ/ → /ə/ in beiden Gruppen mit italienischer L1 (Richtung /e:/ → /ə/ a) Lerner:innen: 16,7%, SE 6,3%; b) Kontrollgruppe ohne Deutschkenntnisse: 55,6%, SE 9,75%), während in der Kontrollgruppe mit deutscher L1 nur die Richtung /e:/ → /ə/ außergewöhnlich niedrige Werte aufweist (44%, SE 10,13%). Um den Effekt der Variable *Richtung* zu testen, wurde folgendes Modell angewendet:

$$lme4::glmer \text{ Akkuratheit} \sim 1 + \text{Gruppe} + \text{Richtung} + \text{Gruppe:Richtung} + (1 | \text{Items}) + (1 | \text{Versuchsperson}).$$

Als abhängige Variable wurde die Akkuratheit angegeben, während als Faktoren die Variable *Gruppe* und die Variable *Richtung* berücksichtigt wurden. Als Zufallsfaktoren wurden die Variablen *Items* und *Versuchsperson* angegeben.

Die Ergebnisse zeigen einen stark signifikanten Effekt des Faktors *Richtung* ($p < .001$). Für die Variable *Gruppe* und die Interaktion *Gruppe***Richtung* konnte kein signifikanter Effekt festgestellt werden (jeweils $p = 0,185$ und $p = 0,140$). Dem Post-Hoc-Test nach ergibt sich einen signifikanten Unterschied bei dem Vergleich $/e:/ \rightarrow /ə/ - /ə/ \rightarrow /ɛ/$ ($p < .001$) und $/e:/ \rightarrow /ə/ - /ɛ/ \rightarrow /ə/$ ($p < .001$). Im ersten Fall ist die Akkuratheit bei $/e:/ \rightarrow /ə/$ signifikant niedriger als bei $/ə/ \rightarrow /ɛ/$; im zweiten Fall erweist sie sich als niedriger bei $/e:/ \rightarrow /ə/$ als bei $/ɛ/ \rightarrow /ə/$.

Um zu verstehen, ob diese Ergebnisse eher durch die intrinsische Qualität der akustischen Stimuli bestimmt sind oder ob die Richtung an sich der Faktor ist, der am meisten für die Wahrnehmungsschwierigkeiten für die in der Grafik angegebenen Richtungen verantwortlich ist, wurde eine akustische Analyse der Stimuli durchgeführt, indem ein lineares gemischtes Modell auf die Werte der abhängigen Variable *F2* (zweiter Formant) angewendet wurde. Der zweite Formant gilt nämlich als besonders aussagekräftig für die horizontale Positionierung (vorne/hinten) von Lauten im Vokalraum. Folgendes Modell wurde angewendet:

$$\text{Imer } F2 - 1 + \text{Vokal} + \text{Geschlecht} + \text{Vokal:Geschlecht} + (1 | \text{Sprecher:in}).$$

Als abhängige Variable wurde die Variable *F2* angegeben, während als Faktoren die Variablen *Geschlecht* und *Vokal* berücksichtigt wurden. Als Zufallsfaktor wurde die Variable *Sprecher:in* angegeben.

Die Ergebnisse zeigen einen hochsignifikanten Effekt des Faktors *Vokal* ($p < .001$), einen nicht-signifikanten Effekt ($p = 0,217$) des Faktors *Geschlecht* und einen hochsignifikanten Effekt ihrer Interaktion ($p < .001$). Für den Faktor *Vokal* und die Interaktion mit dem Faktor *Geschlecht* wurde ein Post-Hoc-Test durchgeführt, aus dem sich hochsignifikante ($p < .001$) Unterschiede zwischen $/ə/ - /e:/$, $/e:/ - /ə/$ und $/i:/ - /ə/$ ergaben, was darauf hindeutet, dass sich die Phoneme sowohl bei Frauen- als auch bei Männerstimmen voneinander qualitativ gut unterscheiden.

4.3.1.2 Ergebnisse zur perzeptiven Assimilation von */y:/* zu */u:/* oder */i:/*

In diesem Unterkapitel wird der Schwerpunkt auf die Ergebnisse gelegt, die sich auf die Assimilationsmuster und damit auf die Wahrnehmung der Stimuli in H2 beziehen. Es handelt sich also um einen Vergleich zwischen der Bedingung, in der die Stimuli als Kontraste (mit */y:/* in der odd- und nicht-odd-Bedingung) erscheinen, und der Bedingung, in der sie als gleiche Abfolgen erscheinen. Darüber hinaus werden die Ergebnisse für jeden der beteiligten und für die Hypothese relevanten Kontraste berichtet.

Beginnend mit dem Vergleich der beiden Arten von Reizsequenzen (Kontraste/gleiche Abfolgen) zeigt das folgende Balkendiagramm (Abbildung 12) die durchschnittlichen Prozentwerte der Akkuratheit und SE für diese beiden Bedingungen in allen drei Gruppen.

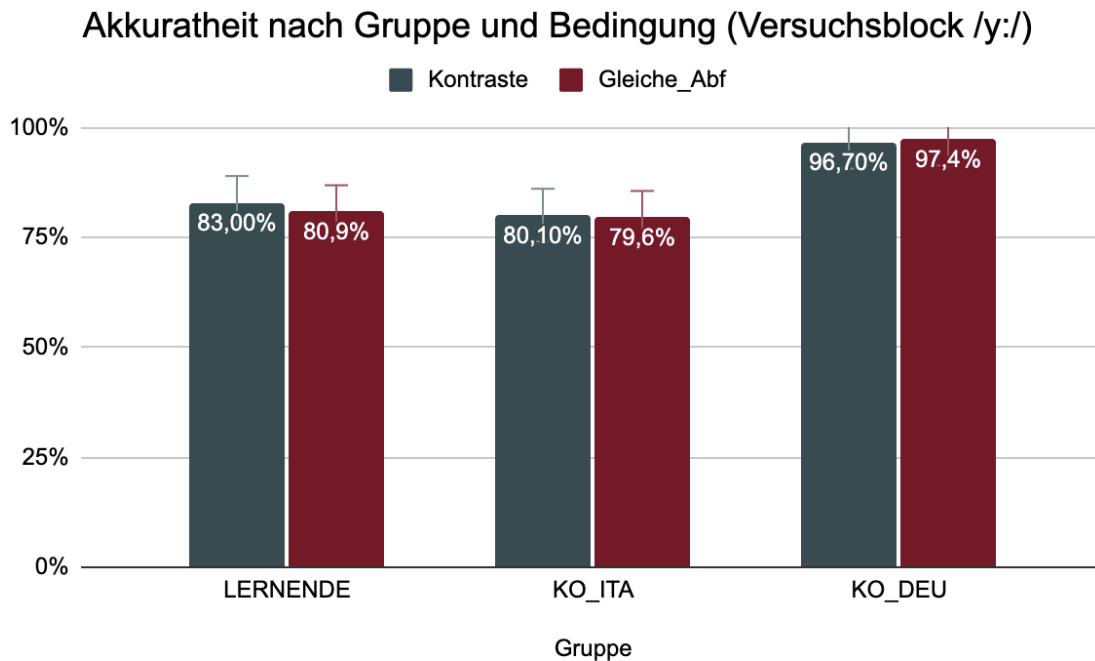


Abbildung 12: Durchschnittliche Akkuratheit im Versuchsblock /y:/ für Kontraste und gleiche Abfolgen pro Probandengruppe. Die Fehlerbalken geben den Standardfehler (SE) an.

Die in der Grafik dargestellten Bedingungen umfassen alle Versuche, bei denen /y:/ vorkommt. Im Allgemeinen sind die Verteilungsmuster der Akkuratheit vergleichbar mit denen, die in Bezug auf Block /ə/ beobachtet wurden. Im Gegensatz zu Block /ə/ unterscheiden sich die Lerner:innen und die Kontrollgruppe ohne Deutschkenntnisse jedoch nur minimal in den Kontrasten (83 %, SE 1,27 % bzw. 80,10 %, SE 1,57 %). Bei den gleichen Abfolgen weist die Gruppe der Lerner:innen höhere Werte für Akkuratheit (80,9%, SE 1,31%) auf als die Kontrollgruppe ohne Deutschkenntnisse (79,6%, SE 1,55%). Die absolut höchsten Ergebnisse sind in der Kontrollgruppe mit Deutsch L1 zu finden, in der die gleichen Abfolgen etwas höhere Werte aufweisen (97,4%, SE 0,63%) als die gleichen Kontraste (96,7%, SE 0,73%).

Um den Effekt der Variable *gleich_ungleich_Bedingung* zu testen, wurde folgendes Modell angewendet:

*lme4::glmer Akkuratheit ~ 1 + Gruppe + gleich_ungleich_Bedingung +
Gruppe:gleich_ungleich_Bedingung + (1 | Items) + (1 | Versuchsperson).*

Als abhängige Variable wurde die Akkuratheit angegeben, während als Faktoren die Variable *Gruppe* und die Variable *gleich_ungleich_Bedingung* berücksichtigt wurden. Als Zufallsfaktoren wurden die Variablen *Items* und *Versuchsperson* angegeben.

Das Modell lieferte nur für den Faktor *Gruppe* hochsignifikante Effekte ($p < .001$). Insbesondere spiegelte der Post-Hoc-Test das zuvor beobachtete Muster zwischen den Gruppen wider, demzufolge es einen hochsignifikanten Unterschied ($p < .001$) zwischen den Gruppen mit italienischer L1 und der Gruppe mit deutscher L1 gibt.

Im Folgenden wird die Kontrastbedingung analysiert. Unter der Variable *Kontrast* fallen sowohl Primärkontraste sowie Sekundärkontraste. Da die Primärkontraste (/y:-i:/, /y:-u:/, und dem Kontrollkontrast /y:-e:/) die Forschungsfragen unmittelbar adressieren, wird im Folgenden ausschließlich die Verteilung der Akkuratheitswerte für diese Kontraste in den drei Gruppen dargestellt (Abbildung 13).

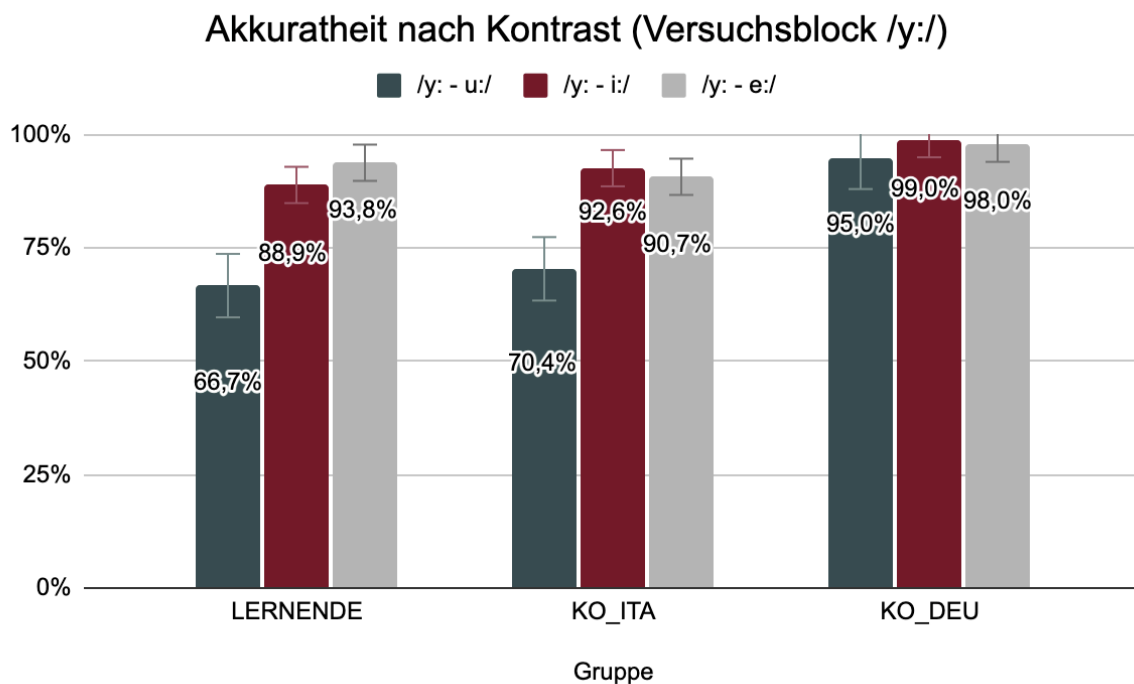


Abbildung 13: Durchschnittliche Akkuratheit in der Perception der Kontraste /y:-u:/, /y:-i:/ und /y:-e:/ nach Probandengruppe. Die Fehlerbalken geben den Standardfehler (SE) an.

Generell zeigt sich ein deutlicher Unterschied zwischen den beiden Gruppen mit italienischer L1 und der Kontrollgruppe mit deutscher L1, da in letzterer die Werte für alle Kontrastarten durchweg höher sind. Die Verteilung der Mittelwerte der Akkuratheit zeigt in der Lerner:innen-Gruppe und der Kontrollgruppe ohne Deutsch-Erfahrung ein ähnliches Muster: In beiden Fällen ist der Kontrast

/y:-u:/ durch besonders niedrige Werte gekennzeichnet, mit einem leichten Vorteil für die Kontrollgruppe (Lerner:innen: 66,7%, SE 3,94%; Kontrollgruppe mit L1 Italienisch: 70,4%, SE 4,41%).

Um den Effekt der unabhängigen Variable *Kontrast* und *Gruppe* zu testen, wurde folgendes Modell angewendet:

$$lme4::glmer \text{ Akkuratheit} \sim 1 + \text{Gruppe} + \text{Kontrast} + \text{Gruppe:Kontrast} + (1 | \text{Items}) + (1 | \text{Versuchsperson}).$$

Als abhängige Variable wurde die Akkuratheit angegeben, während als Faktoren die Variable *Gruppe* und die Variable *Kontrast* berücksichtigt wurden. Als Zufallsfaktoren wurden die Variablen *Items* und *Versuchsperson* angegeben.

Die Ergebnisse ergaben einen hochsignifikanten Effekt ($p < .001$) für die Variablen *Gruppe* und *Kontrast*. Insbesondere der für beide durchgeführte Post-Hoc-Tests ergab signifikante Effekte beim Vergleich zwischen der Gruppe der Lerner:innen und der Kontrollgruppe mit der deutschen L1 ($p < .001$) und zwischen letzterer und der Kontrollgruppe mit der italienischen L1 ($p < .001$), womit sich das bereits zuvor beobachtete Muster wiederholt. Was die Kontraste betrifft, so ergab der Test hochsignifikante p -Werte ($p < .001$) für den Vergleich zwischen /e:-i:/ und /e:-u:/, /u:-i:/, /y:-e:/ und /y:-i:/. Ein hochsignifikanter Effekt ($p < .001$) hat sich auch bei dem Vergleich /u:-i:/ - /y:-u:/ ergeben. Bei den Vergleichen /y:-e:/ - /y:-u:/ und /e:-u:/ - /y:-u:/ hat sich der Effekt als leicht signifikant ergeben (jeweils $p = 0,030$ und $p = 0,036$).

Hinsichtlich der Richtung des Kontrastes zeigt die Kontrollgruppe mit italienischer L1 die niedrigsten Akkuratheitswerte, mit 63% (SE 6,63%) für die Richtung /u:/ → /y:/, 77,8% (SE 5,71%) für die Richtung /y:/ → /u:/ und 88,9% (SE 6,16%) für die anderen Richtungen. Generell zeigt die Richtung /u:/ → /y:/ im Vergleich zu allen anderen Richtungen in beiden Gruppen mit italienischer L1 die niedrigsten durchschnittlichen Akkuratheitswerte (Lerner:innen: 66,7%, SE 5,59%; Kontrollgruppe: 63%, SE 6,63%). Im Gegensatz dazu zeigt die Kontrollgruppe mit deutscher L1 höhere Werte (98%, SE 2,00%).

Um den Effekt der Richtungsvariable zu testen, wurde folgendes Modell angewendet:

$$lme4::glmer \text{ Akkuratheit} \sim 1 + \text{Gruppe} + \text{Richtung} + \text{Gruppe:Richtung} + (1 | \text{Items}) + (1 | \text{Versuchsperson}).$$

Als abhängige Variable wurde die Akkuratheit angegeben, während als Faktoren die Variablen *Gruppe* und *Richtung* berücksichtigt wurden. Als Zufallsfaktoren wurden die Variablen *Items* und *Versuchsperson* angegeben.

Die Ergebnisse ergaben keine signifikanten Effekte für einen der drei Faktoren ($p = 0,987$, $p = 0,822$ bzw. $p = 0,726$). Ein lineares gemischtes Modell, das mit den akustischen Inputs von Block /y:/

durchgeführt wurde und dieselben Bedingungen erfüllt wie das von Block /ə/, zeigt einen hochsignifikanten Effekt des Faktors Vokal ($p < 0,001$), im Gegensatz zum Faktor Geschlecht ($p = 0,322$) und deren Interaktion ($p = 0,687$). Der Post-Hoc-Test zu diesem Faktor zeigt keine Ergebnisse, die darauf hindeuten, dass die Realisierungen [u:] und [y:] nicht ausreichend voneinander unterschieden sind; im Gegenteil, ihr Vergleich zeigt einen hochsignifikanten Unterschied ($p < 0,001$).

Was die sekundären Kontraste betrifft, so finden sich im Allgemeinen für alle Gruppen sehr hohe Akkuratheitswerte (von 90% aufwärts): Die einzige Ausnahme, die sich auf die beiden Gruppen mit italienischem L1 beschränkt, ist der Kontrast /e:-i:/: Die Gruppe der Lerner:innen zeigt hier eine durchschnittliche Akkuratheit von 61,8% (SE 8,13%), während die Kontrollgruppe 38% (SE 9,39%) aufweist. Im Lichte der Daten des zuvor beschriebenen statistischen Modells und des entsprechenden Post-Hoc-Tests ergibt sich tatsächlich keine Signifikanz im Vergleich der Realisierungen von /e:/ und /i:/ ($p = 1,000$), unabhängig vom Geschlecht ($p = 0,322$).

4.3.1.3 *Übergreifende Analyse: Akkuratheit und Faktoren der Sprachkompetenz, Sprachdominanz und des Sprachrepertoires*

In diesem Abschnitt wird über die Ergebnisse der Analysen zur Variation der Akkuratheitswerte in Bezug auf die Variablen der Sprachkompetenz, Sprachdominanz und des Repertoires berichtet. Insbesondere wird der Fokus auf die Deutschkompetenz gelegt, kodiert als *GER-Niveau*, die Dominanz des Deutschen und Englischen (kodiert jeweils als *Deutsch_LEAP(1)* und *Englisch_LEAP(1)*, Werte von 1 bis 5 gemäß der ersten Frage des LEAP-Fragebogens) sowie den Umfang des sprachlichen Repertoires (kodiert als *Sprachrepertoire*). Der Fokus der Analyse liegt auf der Zielgruppe der Lerner:innen, da diese die zentrale Untersuchungskategorie der Forschungsfragen darstellt.

Hinsichtlich der Deutschkompetenz, kodiert nach GER, werden auf der deskriptiven Ebene die höchsten Akkuratheitswerte für die Niveaus A1, B2 und C1 gefunden (jeweils 76,8%, SE 21,6%; 80,9%, SE 23,2%; 74,1%, SE 15,8%).

Was die Dominanz im Deutschen betrifft, so sind die niedrigsten Akkuratheitswerte auf den Sprachdominanzstufen 3 und 4 zu finden, nämlich 73,4% (SE 1,16%) bzw. 69,3% (SE 1,31 %), aber die höchsten Werte auf der niedrigsten Dominanzstufe (Stufe 5, 78%, SE 1,73%).

Was die Kompetenz im Englischen angeht, liegen die höchsten Werte an Akkuratheit bei Dominanzniveau 4 (84,4%, SE 3,72%) und 3 (74,1%, SE 1,35%).

Was die Breite des sprachlichen Repertoires (Anzahl der Sprachen im Sprachrepertoire) betrifft, finden sich auf deskriptiver Ebene die höchsten Akkuratheitswerte an beiden Extremen (Stufe 3: 79,2%, SE 2,39%; Stufe 6: 76,8%, SE 2,15%), während die Akkuratheitswerte in der Mitte ähnlich sind (Stufe 4: 71,9%, SE 1,73%; Stufe 5: 71,5%, SE 0,09%).

Bei dem folgenden Modell

lme4::glmer Akkuratheit ~ 1 + Deutsch_LEAP(1) + Deutsch_Niveau_Deutsch_GER + Englisch_LEAP(1) + Sprachrepertoire + Block + Englisch_LEAP(1) : Block + (1 | Items) + (1 | Versuchsperson)

wurde die Akkuratheit als abhängige Variable angegeben. Als Faktoren wurden die Variablen *Deutsch_LEAP(1)*, *Deutsch_Niveau_Deutsch_GER*, *Englisch_LEAP(1)*, *Sprachrepertoire*, *Block* und *Englisch_LEAP(1) *Block* berücksichtigt, während die Variablen *Items* und *Versuchsperson* als Zufallsfaktoren angegeben wurden. Das Modell ergab keinen signifikanten Effekt des Faktors *Deutsch_LEAP(1)* ($p=0,438$), *Niveau_Deutsch_GER* ($p=0,830$), *Englisch_LEAP(1)* ($p=0,270$), jedoch einen leicht signifikanten Effekt der Interaktion *Englisch_LEAP(1) *Block* ($p=0,025$).

4.3.1.4 Akkuratheit und Task-bedingte Variablen

Zum Abschluss der Analysen werden die Ergebnisse der Analysen für die aufgabenabhängigen Variablen, d.h. das Geschlecht der Stimmen, die die Stimuli während des Experiments vortrugen (*Stimme*), die Position des Stimulus innerhalb der Sequenz (*Position*) und die experimentelle Phase (*Exp_Phase*) vorgestellt.

Die erste Analyse wurde durchgeführt, um zu überprüfen, ob die Qualität der Stimmen in irgendeiner Weise mit den durchschnittlichen Akkuratheitswerten korreliert und somit die Antworten konditioniert.

Die zweite Analyse, die mit der ersten insofern zusammenhängt, als jedes Item einer anderen Position des Stimulus entspricht, ist durch die Hypothese motiviert, dass insbesondere die Anfangsposition des abweichenden Stimulus besonders lange im Arbeitsgedächtnis gehalten werden muss und daher schwieriger mit den nachfolgenden Stimuli zu vergleichen ist. Für diese Analyse wurden nur die Sequenzen mit Kontrasten in Betracht gezogen.

Die dritte Analyse wurde mit dem Wissen durchgeführt, dass der Ablauf des Experiments zwei verschiedene Auswirkungen auf die Akkuratheit haben könnte: Einerseits könnte sie aufgrund eines Lerneffekts zunehmen, andererseits aufgrund von Ermüdung abnehmen.

Was den Einfluss des Geschlechts der Stimmen (erste Analyse) betrifft, so werden zwar auf deskriptiver Ebene tendenziell niedrigere Werte für die Versuche mit weiblicher Stimme beobachtet, aber das gleiche Modell, das mit der gleichen Struktur und der gleichen Anzahl von Beobachtungen angewendet wurde, ergab identische Ergebnisse für die Variable Gruppe und ähnliche Ergebnisse für die beiden anderen festen Faktoren (*Stimme*: $p=0,419$; *Gruppe*Stimme*: $p=0,831$).

Was den Einfluss der Position des Stimulus innerhalb der Sequenzen (zweite Analyse) betrifft, so ist auf deskriptiver Ebene festzustellen, dass die Kontrollgruppe mit deutscher L1 im Allgemeinen die

höchsten Werte in allen drei Bedingungen auf vergleichbarem Niveau aufweist, gefolgt von der Kontrollgruppe mit der italienischen L1 und der Gruppe der Lerner:innen. Bei den Gruppen mit L1 sind die Akkuratheitswerte über die drei Bedingungen hinweg ähnlich verteilt: höhere Akkuratheitswerte in der mittleren Position und niedrigere Akkuratheitswerte in der letzten Position.

Folgendes Modell wurde angewendet:

$$lme4::glmer \text{ Akkuratheit} \sim 1 + \text{Gruppe} + \text{Position} + \text{Gruppe:Position} + (1 | \text{Items}) + (1 | \text{Versuchsperson}).$$

Als abhängige Variable wurde die Akkuratheit angegeben, während die Variablen *Gruppe* und *Position* als Faktoren berücksichtigt wurden. Als Zufallsfaktoren wurden die Variablen *Items* und *Versuchsperson* angegeben.

Während der Faktor *Gruppe* beim Vergleich der Kontrollgruppe mit deutscher L1 und der Gruppe der Lerner:innen einen hochsignifikanten Effekt ($p < .001$) ergab, waren die beiden anderen festen Faktoren nicht signifikant (*Position*: $p = 0,277$, *Gruppe*Position*: $p = 0,251$).

Was die dritte Analyse anbelangt, so ist auf deskriptiver Ebene der einzige eindeutig beobachtbare Trend derjenige, dass die Werte von Gruppe zu Gruppe entsprechend dem Verhältnis Lerner:innen<Kontr_ITA<Kontr_DEU steigen.

Bei dem folgenden Modell

$$lme4::glmer \text{ Akkuratheit} \sim 1 + \text{Gruppe} + \text{Exp_Phase} + \text{Gruppe:Exp_Phase} + (1 | \text{Items}) + (1 | \text{Versuchsperson})$$

wurde die Akkuratheit als abhängige Variable angegeben, während die Variablen *Gruppe* und *Exp_Phase* als Faktoren berücksichtigt wurden. Die Variablen *Items* und *Versuchsperson* wurden als Zufallsfaktoren berücksichtigt.

Das Modell und der entsprechende Post-Hoc-Test ergaben einen signifikanten Effekt für den Faktor *Gruppe* ($p < .001$), und zwar nur für den Vergleich zwischen der Gruppe mit deutscher L1 und der Gruppe der Lerner:innen ($p < .001$).

4.3.2. Reaktionszeiten

Unter den Reaktionszeitanalysen zeigte sich lediglich ein Effekt als statistisch signifikant, der im Folgenden zusammenfassend dargestellt wird.

Angesichts der Tatsache, dass es keine signifikanten Unterschiede zwischen den Gruppen gab, wurden die Daten für Block /ə/ und Block /y:/ zusammen betrachtet. Die folgende Abbildung

(Abbildung 14) zeigt die Boxplots, die die Verteilung der Reaktionszeiten (Sek) in den beiden Blöcken darstellen.

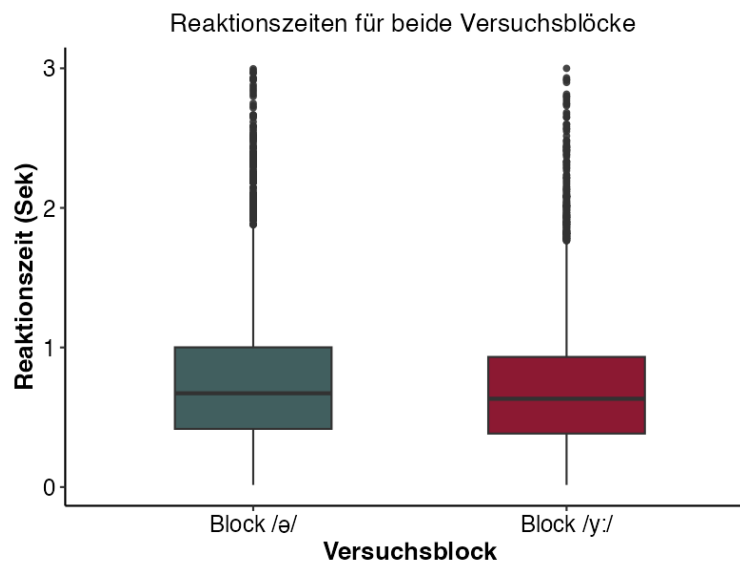


Abbildung 14: Verteilung der Reaktionszeiten in allen Gruppen (Lernende, KO_ITA, KO_DEU) für beide Blöcke.

Die Mittelwerte für Block /ə/ sind höher (0,79 Sek., SE 0.06 Sek.) als die für Block /y:/ (0,7 Sek., SE 0.05 Sek.).

Bei dem folgenden Modell

$$\text{lmer log_RT} \sim 1 + \text{Gruppe} + \text{Block} + \text{Gruppe:Block} + (1 | \text{Items}) + (1 | \text{Versuchsperson})$$

wurden die Reaktionszeiten (log_RT) als abhängige Variable angegeben. Die Variablen *Gruppe* und *Block* wurden als Faktoren berücksichtigt, während die Variablen *Items* und *Versuchsperson* als Zufallsfaktoren angegeben wurden.

Das Modell zeigt einen hochsignifikanten Effekt des fixen Faktors *Block* ($p < .001$).

Auf eine detaillierte Darstellung der übrigen Ergebnisse wird verzichtet, da sich keine weiteren signifikanten Unterschiede zeigten und die Reaktionszeiten nur eingeschränkt interpretierbar sind.

4.3.3. Eine Explorative Studie zum Einfluss der dialektalen Hintergrundkenntnis auf die Wahrnehmung des Deutschen: Akkuratheit und Reaktionszeiten

Um den möglichen positiven Transfer von Vorkenntnissen in einem Dialekt der L1 auf die Wahrnehmung des Deutschen zu testen, wurde eine explorative Studie an zwei kleinen Teilstichproben durchgeführt, die aus Lerner:innen und Proband:innen der Kontrollgruppe mit L1 Italienisch bestanden, die angaben, über Kompetenzen in einem Dialekt des Italienischen zu

verfügen, dessen phonologisches System die Phoneme /ə/ oder /y:/ enthalten. Für diese Auswertung wurden die Teilnehmer:innen aufgrund ihrer Selbstauskünfte im soziolinguistischen Fragebogen einem der nach Vignuzzi (2010) gezogenen Dialektgebiete zugeordnet. Lerner:innen und Probanden ohne Deutschkenntnisse wurden gemeinsam betrachtet, da sowohl bei der Akkuratheit als auch bei den Reaktionszeiten keine signifikanten Unterschiede zwischen den beiden Gruppen vorlagen. Die Mittelwerte von Akkuratheit und Reaktionszeiten wurden für alle Versuche berechnet, die /ə/ oder /y:/ enthielten, die sowohl in Kontraste als auch in gleichen Abfolgen vorkommen. Die Mittelwerte der Akkuratheit und der Reaktionszeiten für die Versuchsgruppen wurden mit den Mittelwerten aller anderen Teilnehmer:innen an der Hauptstudie mit L1 Italienisch verglichen.

4.3.3.1 *Wahrnehmung von /ə/*

Die experimentelle Teilstichprobe, die auf die Wahrnehmung von /ə/ getestet wurde, besteht aus 2 Teilnehmer:innen (1 weiblich und 1 männlich), die dem zentral-süditalienischen Makrobereich – genauer gesagt dem mittleren Gebiet – zuzuordnen sind. Die Teilnehmerin gehört zur Gruppe der Lerner:innen, während der Teilnehmer zur Kontrollgruppe ohne Deutschkenntnisse gehört. Die Teilnehmerin gab eine maximale Dominanz (Niveau 1) für Regionalitalienisch und Dialekt an. Der Sprecher schrieb seinem Dialekt ein mittleres Maß an Dominanz zu (3, nach Englisch). Die Kontrollgruppe bestand aus 63 Teilnehmer:innen mit L1 Italienisch, die entweder keine spezifische Dialektkompetenz aufwiesen oder angaben, über keinerlei Dialektkompetenz zu verfügen.

In Bezug auf die Akkuratheit zeigt die deskriptive Statistik, dass der durchschnittliche Akkuratheitswert der beiden Teilnehmer:innen höher ist (94,1%, SE 16,9%) als der der Kontrollgruppe (75,9%, SE 5,4%).

Bei den Reaktionszeiten ist der Durchschnitt der Gruppe der Lerner:innen (0,53 Sek., SE 0,27 Sek.) niedriger als der der Kontrollgruppe (0,74 Sek., SE 0,07 Sek.).

4.3.3.2 *Wahrnehmung von /y:/*

Die experimentelle Untergruppe, die auf die Wahrnehmung von /y:/ getestet wurde, bestand aus 8 Teilnehmer:innen (5 Frauen und 3 Männer), die dem Gebiet Veneto und dem Gallo-italischen Dialektraum zuzuordnen waren (vgl. Vignuzzi, 2010). Insbesondere gaben die Teilnehmer:innen an, aus den Regionen Trentino-Südtirol und Lombardei zu stammen. 2 der Teilnehmer:innen gaben eine hohe Dialektdominanz an (Stufe 2), 3 eine mittlere Dominanz (Stufe 3) und 2 eine niedrige Dominanz (Stufe 4 und 5). Die Kontrollgruppe bestand aus 42 Personen mit L1 Italienisch, die entweder keine spezifische Dialektkompetenz aufwiesen oder angaben, über keinerlei Dialektkompetenz zu verfügen.

In Bezug auf die Akkuratheit zeigen die deskriptiven Statistiken, dass die Durchschnittswerte in der Gruppe der Lerner:innen etwas höher sind (91%, SE 10,1%) als in der Kontrollgruppe (83,6%, SE 5,7%).

Hinsichtlich der Reaktionszeiten waren die Mittelwerte der Gruppe der Lerner:innen etwas höher (0,71 Sek., SE 0,13 Sek.) als die der Kontrollgruppe (0,69 Sek., SE 0,07 Sek.).

4.4 Diskussion

Die Ergebnisse zur Akkuratheit haben gezeigt, dass von den Phonemen /ə/ und /y:/ das erste den Deutschlerner:innen besondere Wahrnehmungsschwierigkeiten bereitet.

Die L1 hat einen deutlichen Effekt auf die Wahrnehmung von /ə/, was auf einen negativen Transfer hinweist. Im Gegensatz dazu war die perzeptuelle Leistung der Deutschlerner:innen bei der Wahrnehmung des Phonems /y:/ trotz seiner Abwesenheit im L1-System erfolgreicher. Dieses Gesamtergebnis kann prinzipiell als mit den Beobachtungen von Missaglia (2000) übereinstimmend betrachtet werden. In ihrer Studie, die darauf abzielte, die Sensitivität italienischer Deutschlerner:innen gegenüber vokalischer Qualität und Quantität in akustisch manipulierten Minimalpaaren zu untersuchen, stellte die Autorin fest, dass gerundete palatale Vokale häufig als das jeweils andere Element des Minimalpaars wahrgenommen wurden. Anders hingegen verhielt es sich bei den mittleren Vokalen (wobei der Schwa-Laut nicht in die Analyse einbezogen wurde): Hier führten verzerrte Wahrnehmungen der Laute zu einer Zuordnung der Stimuli zu völlig anderen Vokalen (wie z. B. [e], das als [i] wahrgenommen wurde), wobei zudem eine deutliche Diskrepanz zur Wahrnehmung der Kontrollgruppe mit L1 Deutsch festgestellt wurde.

Durch die vorliegende Studie konnten mögliche Ursachen für diese zentrale Diskrepanz in der perzeptiven Genauigkeit zwischen den beiden Stimulusblöcken im Detail untersucht werden. Zudem wurde durch die Einbeziehung des Schwa-Lauts in die Analyse das durch frühere Studien gezeichnete Gesamtbild weiter vervollständigt. Aufgrund des zentralen Artikulationsortes, durch den sich das Phonem /ə/ auszeichnet und der im italienischen Standardsystem kein direktes Gegenstück hat, bereitet dieses Phonem besondere perzeptive Schwierigkeiten, insbesondere, wenn es mit den ähnlichen Lauten des Deutschen /e:/ und /ɛ/ kontrastiert wird. Im Gegensatz dazu wird das Phonem /y:/ im Allgemeinen unter allen Bedingungen besser erkannt. Obwohl es mit keinem direkten Gegenstück im italienischen Lautsystem verglichen werden kann, ist es plausibel anzunehmen, dass sein Artikulationsort und damit seine höhere Salienz die Bildung der neuen phonologischen Kategorie begünstigt.

Die generelle Hypothese (H1) besagt, dass aufgrund des Einflusses der L1 eine fehlerhafte perzeptive Leistung der Deutschlerner:innen zu erwarten sei. Weiterhin wurde angenommen, dass die experimentalGruppe (die Lerner:innen) besser abschneiden würde im Sinne von höherer Akkuratheit und niedrigen Reaktionszeiten als die Italienischen L1-Hörer:innen ohne

Deutschkenntnisse, aber schlechter als die Gruppe von deutschen L1-Hörer:innen. Diese Hypothese kann nur zum Teil bestätigt werden: Gemäß den Erwartungen hat die experimentelle Stichprobe der Lerner:innen schlechter abgeschnitten (niedrigere Akkuratheitswerte) als die deutsche Kontrollgruppe, aber gleichzeitig auch schlechter als die Kontrollgruppe ohne Deutschkenntnisse. Dies ist wahrscheinlich auf mehrere Gründe zurückzuführen. Aus der einen Seite, haben die phonologischen Lautkategorien im Gehirn der Lerner:innen noch keinen endgültigen oder zumindest einen ungefähr festen Zustand erreicht. Daraus resultiert eine noch instabile mentale Repräsentation, die ein typisches Merkmal der *Interlanguage* ist (vgl. Escudero 2005). In diesem Zusammenhang ist die Dominanz der L1 und ein unzureichender Input der Zielsprache Deutsch ein möglicher Grund für das schlechtere Abschneiden der Experimentalgruppe im Vergleich zur deutschen Kontrollgruppe, die den Erwartungen entsprechend die höchsten Akkuratheitswerte aufweist.

Aufbauend auf den Überlegungen von Best & Tyler (2007, S. 32) zur Wahrnehmung von Phonemen in einer "L2" (wie in Best & Tyler, 2007 bezeichnet) durch sogenannte *naïve listeners*, lassen sich einige Überlegungen anstellen. Es ist möglich, dass die perzeptive Leistung der Kontrollgruppe ohne Deutschkenntnisse akkurater war als die durch die Hypothesen erwarteten Ergebnisse, da die Wahrnehmung der Stimuli ausschließlich auf phonetischer und nicht auf phonologischer Ebene erfolgte. Diese Beobachtung deckt sich mit den Erkenntnissen von Polka (1995), die zeigte, dass Hörer:innen ohne Vorerfahrung bei deutschen Vokalkontrasten (wie /y:-u/) oft hohe Akkuratheitswerte erzielen. Dies lässt sich darauf zurückführen, dass diese Probanden die Stimuli als rein akustische Ereignisse verarbeiten und die Unterscheidung anhand der sogenannten *category goodness* treffen, anstatt sie in ein noch instabiles L2-Phonemsystem einzuordnen. Allerdings muss berücksichtigt werden, dass dies nur eine von mehreren möglichen Erklärungen darstellt. Wie in Best & Tyler (2007) nämlich festgestellt, sind Sprachkenntnisse nur einer von vielen Faktoren, die eine Experimentalgruppe von der jeweiligen Kontrollgruppe unterscheiden können. Diese Erkenntnis unterstreicht die Komplexität der Analyse multilingualer Stichproben, wie sie auch in der vorliegenden Studie vorliegen. Im Falle der vorliegenden Untersuchung gaben die Teilnehmer:innen aus beiden Gruppen (Hörer:innen mit L1 Italienisch) an, sehr unterschiedliche sprachliche und erfahrungsbezogene Hintergründe im Umgang mit dem Fremdsprachenlernen zu haben. Der zunehmende Grad an Mehrsprachigkeit in der heutigen Gesellschaft macht es herausfordernd, Variablen wie sprachlichen Hintergrund, Exposition und Lernhaltungen streng zu kontrollieren (vgl. z. B. Hulstijn, 2011; Grosjean, 2010). Gleichzeitig gewinnt die ökologische Validität von Studien an Bedeutung, indem sie realistische und alltagsnahe Sprachlernsituationen berücksichtigt (vgl. Hulstijn, 2011). Somit trägt die vorliegende Untersuchung trotz (oder gerade wegen) dieser Komplexität dazu bei, ein möglichst realistisches Bild der Wahrnehmungen in multilingualen Hörer:innen zu zeichnen. Diese Erkenntnisse unterstreichen die Notwendigkeit, künftig stärker individuenzentrierte Ansätze (vgl. z. B. Mayr & Escudero, 2010; Wrembel & Amaro,

2018) zu verfolgen, um die komplexen inter- und intraindividuellen Variationen zu erfassen – als Grundlage für eine spätere Übertragung auf breitere Stichproben.

Was die Reaktionszeiten angeht, haben sie sich als weniger aussagekräftig erwiesen; trotzdem muss die Hypothese zu den gegenseitigen Verhältnisse in Bezug auf diese abhängige Variable abgelehnt werden. Anhand der Ergebnisse hat sich die experimentelle Gruppe als der schnellste bei der Diskrimination der Kategorien erwiesen. Hingegen wies die Kontrollgruppe der deutschen L1 Hörer:innen die höchsten Reaktionszeiten auf. In diesem Zusammenhang wurde ein Müdigkeitseffekt oder ein Lerneffekt anhand der nicht signifikanten Effekte für die Variable der Experimentphase ausgeschlossen.

Nachfolgend wird näher auf die einzelnen Bedingungen eingegangen.

Hinsichtlich der Lautwahrnehmung der drei verschiedenen Gruppen erwies sich die Unterscheidung von /ə/ von den anderen Lautkategorien (/e:/ und /ɛ/) als besonders schwierig und war durch signifikant niedrigere Akkuratheitswerte und – auf deskriptiver Ebene – höhere Reaktionszeiten im Vergleich zur Unterscheidung von /y:/ und den anderen Lautkategorien (/u:/ und /i:/) gekennzeichnet. Diese Ergebnisse stehen im Einklang mit den im Wahrnehmungsexperiment von Missaglia (2000) erhobenen Ergebnissen.

Die Hypothesen H1 a) und H1 b), die sich auf die erwarteten perzeptiven Assimilationsmuster beziehen (vgl. Maturi, 2014), besagen, dass /ə/ und /y:/ von der Gruppe der Lerner:innen mit den entsprechenden deutschen Phonemen (jeweils /e:/, /ɛ/ und /u:/ /i:/) assimiliert werden. Dies ist auf ein allgemeineres, für das Erlernen einer neuen Sprache typisches Phänomen zurückzuführen, das in der Literatur mit verschiedenen Begriffen wie *magnet effect* (Kuhl, 1991; Kuhl & Iverson, 1995), *equivalence classification* (vgl. Flege, 1995) oder *phonemic equation* (Escudero, 2005) bezeichnet wird und die beiden in Kontakt tretenden Sprachen charakterisiert.

Im Falle von /ə/ kann die Hypothese (H1a) zur perzeptiven Assimilation von /ə/ zu /e:/ oder /ɛ/ bestätigt werden, und die Ergebnisse liefern genauere Hinweise auf die Assimilationsmuster: Beide Kontraste, /ə-ɛ/ und /ə-e:/, erwiesen sich als vergleichbar schwer zu unterscheiden. Neue Evidenz zum Einfluss der Richtung des Kontrastes auf die Akkuratheit ergab die vorliegende Studie: Wenn der Kontrast die Richtung von einer in der "Peripherie" des Vokaltrapezes liegenden Position zu einer in der Mitte liegenden Position verfolgt (wie bei /e:/→/ə/), wird der Kontrast mit deutlich größerer Schwierigkeit (niedriger Akkuratheit) wahrgenommen. Diese Befunde liefern zusätzliche Unterstützung für die Theorie des *Natural Referent Vowel* von Polka & Bohn (2011).

Im Falle von /y:/ kann die Hypothese (H1b) zur perzeptiven Assimilation von /y:/ zu /u:/ oder /i:/ ebenfalls bestätigt werden und es wurden neue, genauere Belege für perzeptuelle Assimilationsmuster gefunden. Besondere Schwierigkeiten, die in ähnlichem Ausmaß sowohl die

experimentelle Gruppe der Lerner:innen als auch die Kontrollgruppe mit L1 Italienisch betreffen, liegen in der Wahrnehmung des Kontrastes /y:-u:/, bei der Genauigkeit und Reaktionszeit korrelieren. Mit besonderem Augenmerk auf die Gruppe der Lerner:innen kann davon ausgegangen werden, dass das Merkmal der Rundung durch die Wahrnehmungsmuster der L1 wahrgenommen und daher mit hinteren Positionen im Vokalraum – und somit mit dem deutschen /u:/ wie auch mit dem italienischen /u/ assoziiert wird. Die Tatsache, dass der Kontrast /y:-i:/ deutlich besser erkannt wird als /y:-u:/, legt folgende Interpretation nahe: /y:/ wird höchstwahrscheinlich als prinzipiell hinterer Vokal wahrgenommen, da das Merkmal [+gerundet] im Italienischen den hinteren Positionen im Vokalraum zugeordnet wird.

Die Untersuchung des Einflusses der Variablen der Sprachkompetenz, Sprachdominanz und des Sprachrepertoires (H2) zeigte, dass weder die Kompetenz im Deutschen noch die Dominanz im Deutschen (H2b), die durch einen soziolinguistischen Fragebogen erfasst wurde, Einfluss auf die Wahrnehmungsleistung ausübten. Diese Ergebnisse stimmen mit den früheren Befunden von Gallardo del Puerto (2007) überein, der keinen Unterschied in der Identifikation englischer Phoneme zwischen spanisch-baskischen Zweisprachigen feststellte, die sich in ihrem Gebrauch der beiden Sprachen unterschieden, und damit die geringe Aussagekraft der Dominanz in seiner Studie hervorhob. Die Tatsache, dass einerseits keine substantiellen Unterschiede zwischen den beiden Gruppen mit italienischer L1, andererseits aber erhebliche Diskrepanzen in der Akkuratheit zwischen den beiden Gruppen und der Kontrollgruppe mit deutscher L1 auftauchen, deutet in der Tat auf eine gewisse Rolle der Sprachkompetenz hin, die durch die Ergebnisse zur Akkuratheit stark erfasst wird. Der Befund, dass zwischen den beiden Gruppen mit italienischer L1 keine signifikanten Unterschiede festgestellt wurden, kann auf das Zusammenspiel verschiedener Faktoren zurückgeführt werden. Die Mehrheit der Lerner:innen befindet sich auf einem mittleren Sprachniveau (B1), wobei das Erlernen des Deutschen überwiegend in einem nicht-immersiven Kontext stattfand und die Dominanz des Deutschen bei den meisten Lerner:innen gering ist. Unter diesen Bedingungen könnte die Wahrnehmung sowohl durch phonologisches Wissen als auch durch phonetische Strategien beeinflusst werden. Da die vorliegende Studie keine Kontrolle über Qualität und Quantität des sprachlichen Inputs ermöglichte, ist dies ein wesentlicher Aspekt, der bei der Interpretation der Ergebnisse berücksichtigt werden muss. Diese Einschränkung verdeutlicht die Komplexität der Analyse mehrsprachiger Stichproben und weist darauf hin, dass künftige Untersuchungen den Einfluss von Inputvariablen gezielter erfassen sollten. Mögliche Einflüsse weiterer Sprachen, die von den Teilnehmer:innen beider Gruppen als Teil ihres multilingualen Sprachrepertoires angegeben wurden, können nicht ausgeschlossen werden. Die vorliegende Interpretation versteht sich als vorläufiger Ansatz, der potenzielle Einflussfaktoren aufzeigt und als Grundlage für weiterführende Forschungsarbeiten dienen kann.

Der Einfluss der Dominanz im Englischen auf die Wahrnehmung von /ə/ (H2a) war als Haupteffekt nicht signifikant. Die Interaktion zwischen dieser Variable und dem Faktor Block erwies sich jedoch als signifikant. Dieses Ergebnis verweist auf die inhärente auditive Komplexität zwischen den Aufgabenblöcken für /ə/-/y:/, wobei der zweite Block systematisch eine höhere Akkuratheit begünstigt. Das suggeriert, dass Teilnehmende mit einer ausgeprägten Sprachdominanz (Stufe 2) aufgrund stabilerer phonetischer Kategorien eine relativ homogene Leistung über beide Komplexitätsstufen hinweg aufweisen, während sich Individuen mit geringerer Dominanz (Stufe 4) als hochgradig anfällig für die strukturelle Aufgabenkomplexität zeigen.

Der Einfluss des Umfangs des Sprachrepertoires (H2d) – entgegen der Annahme von Smith & Baker (2011) – hat sich als nicht signifikant erwiesen. In der vorliegenden Studie wurde Englisch als potenzielle *supplier*-Sprache für die Wahrnehmung des Deutschen betrachtet, unter der Annahme, dass intersystemische Gemeinsamkeiten (wie die hohe Breite des vokalischen Repertoires und das Vorhandensein des Schwa-Lauts in beiden Sprachen) eine unterstützende Rolle für Deutschlerner:innen spielen könnten. Dass sich dies nicht über den Faktor der Englisch-Dominanz bzw. der Nutzungsfrequenz dieser Sprache zeigte, kann mehrere plausible Gründe haben. Es ist möglich, dass der nicht signifikante Einfluss des Englischen auf ein für die italienische Realität charakteristisches sprachliches Verhalten zurückzuführen ist (vgl. Leonidou, 2020). Das Erlernen der englischen Sprache konzentriert sich eher auf das Erlernen der Grammatik (vor allem im Sinne von Morphologie und Syntax), und die Aussprache des Englischen bleibt in der Didaktik oft ein Merkmal von geringer Bedeutung, vor allem wenn ein rein kommunikativer oder aufgabenorientierter Ansatz verfolgt wird (siehe Gauci, 2008). Dieser Faktor in Verbindung mit den unbestreitbaren intersystemischen Unterschieden zwischen dem italienischen und dem englischen System könnte dazu geführt haben, dass die Aussprache trotz der angegebenen hohen Dominanz im Sprachrepertoire unzureichend war.

Mit Blick auf den Einfluss der Kompetenz in L1-Dialektvarietäten des Italienischen auf die Wahrnehmung von /ə/ und /y:/ (H2c) konnte die Studie erste Hinweise liefern und ist in dieser Hinsicht eher explorativ. Sie stellt einen ersten Schritt zur Untersuchung von *Cross Linguistic Influence* (CLI) mit Blick auf die Kompetenz in L1-Dialektvarietäten dar. Smith & Baker (2011) haben Belege zur Unterstützung einer möglichen erhöhten Wahrnehmungssensitivität erbracht: Sie zeigten, wie die Vorkenntnis einer nicht-standardisierten Varietät des Deutschen eine größere Sensitivität für Vokalkontraste des Standarddeutschen begünstigt und als Ausgangspunkt für die Entwicklung eines bewussten Sprachverhaltens dient. Dabei wurde im Fall von Smith und Baker (2011) beobachtet, dass regionale Merkmale des Deutschen zugunsten standardsprachlicher Merkmale zurückgedrängt werden. Im Fall der vorliegenden explorativen Studie zeigen die deskriptiven Daten höhere Akkuratheit in der Wahrnehmung von /ə/ und /y:/ bei der Versuchsgruppe mit relevanten Vorkenntnissen in einem Dialekt der L1 als in der Vergleichsgruppe

ohne dialektale Kompetenz bzw. ohne relevante dialektale Kompetenz. Besonders deutlich ist der Unterschied in der Akkuratheit bei der Wahrnehmung von /ə/. Was die Reaktionszeiten betrifft, so scheinen deren niedrige Werte einen deskriptiven Zusammenhang mit den hohen Werten der Akkuratheit bei der Wahrnehmung von /ə/ herzustellen. Dies sind erste Belege in einem noch weitgehend unerforschten Zweig der CLI-Studien und gleichzeitig ein Anreiz, mit größeren Stichproben und einer systematischeren Kontrolle der Dialektfähigkeiten fortzufahren.

5. Eine phonetische Studie zur Produktion deutscher Vokale

5.1 Forschungsfragen und Hypothesen

Die zentrale Fragestellung dieser Studie fokussiert auf den Prozess der Produktion der deutschen Vokale /ə/ und /y:/ durch erwachsene Lerner:innen des Deutschen als Fremdsprache, die keine Hör- Seh- und Sprechstörungen aufweisen. Die Fragestellung wird sowohl aus einer intralinguistischen als auch interlinguistischen Perspektive behandelt.

Die intralinguistische Sichtweise zielt auf die Ergründung der Natur der erlernten Kategorien innerhalb des *Interlanguage*-Systems ab. Im Fokus der Untersuchung steht dabei die Frage, inwiefern die Phoneme /ə/ und /y:/ von artikulatorisch benachbarten Phonemen unter Berücksichtigung möglicher sprachspezifischer Merkmalgewichtungsstrategien auf produktiver Ebene getrennt sind, d. h. ob das *Interlanguage*-System eine eigene phonologische Struktur erreicht hat. In diesem Zusammenhang wird der Frage nachgegangen, inwiefern die Aussprache der untersuchten Phoneme und der benachbarten Phoneme in Abhängigkeit von strukturellen Faktoren sowie außerlinguistischen Einflussfaktoren wie Sprachkompetenz, -dominanz und dem Umfang des Sprachrepertoires variiert

Im Hinblick auf die strukturellen Faktoren wird der Frage nachgegangen, ob die Position im Wort bei der Aussprache von /ə/ eine Rolle spielt. Bei der Aussprache von /y:/ ist der phonologische Kontext – sowohl der linke als auch der rechte – von besonderem Interesse.

Aus außerlinguistischer Perspektive stellt sich zudem die Frage, ob die Sprachkompetenz – verstanden als GER-Kompetenzniveau (vgl. Europarat, 2010) sowie als Sprachdominanz – eine Rolle bei der Konfiguration der Kategorien innerhalb des *Interlanguage*-Systems spielt und ob letztere durch diese Faktoren variieren.

Im Zentrum der interlinguistischen Perspektive stehen zwei Fragen. Erstens wird untersucht, inwieweit die Realisierungen der Deutschlerner:innen als zielsprachlich zu betrachten sind. Zweitens wird der Einfluss des L1-Phonologischen Systems auf das *Interlanguage*-System beleuchtet. Dabei wird der Effekt der *phonetic equation* (Escudero, 2005) bzw. des Magnet-Effekts (Kuhl, 1991; Kuhl & Iverson, 1995) oder – nach Flege (1995) – der *equivalence classification* näher betrachtet.

Anhand dieser Forschungsfragen und Perspektiven werden folgende Hypothesen formuliert:

H1) Die untersuchten Phoneme /ə/ und /y:/ werden mit höher Wahrscheinlichkeit ähnlich wie Phoneme realisiert, die diesen artikulatorisch ähnlich sind bzw. mit denen sie gewisse distinktive

Merkmale teilen. Es wird daher ein grundsätzlicher Einfluss der L1 (CLI) auf die Entwicklung des *Interlanguage*-Systems hypothesiert, der mit einer nicht-zielsprachlichen Realisierung der deutschen Phoneme zusammengeht.

H1a) Es wird die Hypothese aufgestellt, dass /ə/ ähnlich wie die deutschen /ɛ/ oder /e:/ ausgesprochen wird, da sie qualitativ am nächsten liegen (vgl. Maturi, 2014). Die qualitative Nähe wird im Wesentlichen u.A. durch den Einfluss der Orthografie in L1 bestimmt, wonach das Graphem <e> entweder als /ɛ/ oder als /e/ realisiert werden kann.

H1b) Bei der Aussprache von /y:/ könnte das Zusammenspiel der Merkmale [+vorn] und [+gerundet] zu unterschiedlichen Ergebnissen führen, je nachdem, wie diese Merkmale gewichtet werden. Aus artikulatorisch-ökonomischer Sicht erscheint eine vordere Aussprache, d. h. eine Ähnlichkeit zu /i:/ oder /ɪ/, am wahrscheinlichsten. Es kann jedoch nicht ausgeschlossen werden, dass eine überwiegende Sensibilität zum Merkmal [+gerundet] zu hinteren Realisierungen, d. h. zu /u:/ und /ʊ/, führen könnte (vgl. Maturi, 2014). Die Hypothese, dass /y:/ wie /ʏ/ ausgesprochen wird, würde auf eine der Zielsprache naheliegende Realisierung hindeuten, wobei das Merkmal [+gerundet] mit dem Merkmal [+vorn] kombiniert werden. In diesem Fall wäre die Kombination beider Merkmale grundsätzlich adäquat, jedoch würde das gespannte, lange Phonem [y:] dann mit dem kurzen, ungespannten Phonem [Y] zusammenfallen, weil sie hinsichtlich der Länge nicht unterschieden werden. Diese Hypothese fußt auf dem Einfluss der L1 (Italienisch), wobei die Länge keinen phonematischen Status aufweist. Aus diesem Grund sind solche Phoneme als schwer realisierbar betrachtet (vgl. Zimmerer & Trouvain, 2015).

H2) Es wird die Hypothese aufgestellt, dass die Realisierung von /ə/ und /y:/ von streng linguistischen Faktoren abhängt, weil solche Variation mit der Entwicklung des *Interlanguage*-Systems in Richtung der Zielsprache zusammenhängt.

H2a) postuliert, dass die Artikulation von /ə/ in Abhängigkeit von Position im Wort (in präfixale, Wortstamm, präfinale und finale Position) variiert. Es wird genauer erwartet, dass die F1-Werte in finaler Position höher sind als in präfinaler Position (vgl. Flemming & Johnson, 2007; Lass, 2009 für das Englische). Aufgrund der Abwesenheit von vergleichbaren Daten aus der Literatur können allerdings genauere Hypothesen nicht formuliert werden (vgl. Christen et al. 2010, aber siehe Hove, 2020).

H2b) postuliert, dass die Realisierung von /y:/ vom phonologischen Erscheinungskontext (sowohl linken als auch rechten) beeinflusst wird. Insbesondere werden vordere und zwar labiale, labiodentale und/oder alveolare Artikulationsorte der benachbarten Konsonanten als fördernde Artikulationsorte für eine vorderere und gerundetere Artikulation von /y:/ angesehen (vgl. O'Brien & Smith, 2010).

H3) Es wird die Hypothese aufgestellt, dass die Realisierung von /ə/ und /y:/ von Faktoren wie Sprachkompetenz und Umfang des Sprachrepertoires beeinflusst wird, und zwar werden folgende spezifische Hypothesen formuliert:

H3a) postuliert, dass die Realisierung von /ə/ bei höherem Grad an Dominanz des Englischen, dessen phonologisches System ebenfalls das /ə/ enthält, einen zielsprachlichen bzw. zentraleren Artikulationsort einnimmt. Da das Englische im italienischen Schulsystem in der Regel ab dem sechsten Lebensjahr als erste Fremdsprache unterrichtet wird, wird es als geteilte Vorkenntnis betrachtet.

H3b) postuliert, dass die Realisierung von /ə/ und /y:/ bei einem höherem Deutschniveau, gemessen an GER-Niveau und Sprachdominanz, zielsprachlicher ist.

H3c) postuliert, dass im Falle einer möglichen Kompetenz in einem Dialekt der L1, der eines der beiden Phoneme enthält, kann dieses Hintergrundwissen bei der Aussprache der beiden deutschen Phoneme von Vorteil sein.

H3d) In Anbetracht der Erkenntnis (O'Brien & Smith, 2010), dass ein relativ großes Sprachrepertoire eine höhere Sensibilität für akustische Unterschiede in fremden Lauten und damit auch eine zielsprachliche Aussprache fördert, wird postuliert, dass bei einem relativ großen Sprachrepertoire die Aussprache von /ə/ und /y:/ zielsprachlicher ist.

H4) Es wird die Hypothese aufgestellt, dass die Artikulation von /ə/ und /y:/ von der jeweiligen Taskbedingung abhängt. Dieser Einfluss ist auf Koartikulationsphänomene zurückzuführen, die die spontane, wenig konditionierte Sprache charakterisieren. Es wird davon ausgegangen, dass Realisierungen bei konditionierteren experimentellen Bedingungen zielsprachlich artikuliert werden, während dies bei wenig konditionierten experimentellen Bedingungen nicht der Fall ist.

5.2 Methodisches Vorgehen

Für das Produktionsexperiment wurde ein kombinierter Aufgabenansatz verwendet, der aus einer Wegbeschreibung, einer Aufgabe "Nacherzählung" mithilfe von Abbildungen und einer Elizitationsaufgabe (Vorlesen von einzelnen Wörtern) bestand. Die Kombination verschiedener Aufgaben zielte darauf ab, die Sprachproduktion unter verschiedenen, zunehmend konditionierten Sprechbedingungen zu analysieren. Die Aufgaben wurden daher in folgender Reihenfolge durchgeführt: 1) die Wegbeschreibung, 2) die Nacherzählung, 3) Vorlesen. Für keine der experimentellen Aufgaben wurden Zeitbeschränkungen festgelegt.

Für diese Studie wurden dieselben Räumlichkeiten wie für das Perzeptionsexperiment verwendet.

An dieser Studie nahm eine Untergruppe von Lerner:innen teil, deren Deutschkenntnisse zwischen B1 und C1 lagen. Alle Teilnehmer:innen nahmen an der Studie teil, nachdem sie eine Einverständniserklärung unterschrieben hatten. Alle Teilnehmer:innen wurden mit einer

Aufwandsentschädigung vergütet. Für diejenigen, die sowohl am Perzeptions- als auch am Produktionsexperiment teilnehmen wollten, wurde zuerst das Produktionsexperiment durchgeführt. Da die genauen Ziele der Studie nicht bekannt gegeben wurden, wurde es vorgezogen, mit allgemeineren Sprechbedingungen zu beginnen und dann zu immer stärker konditionierten experimentellen Bedingungen wie denen des Perzeptionsexperiments überzugehen. Gemäß den von Flege, (2022) formulierten Richtlinien für die Durchführung von Produktionsstudien wurde der Sprachmodus während des gesamten Experiments auf Deutsch gehalten. Darüber hinaus erfolgte die Rekrutierung der Versuchspersonen über soziale Netzwerke (vgl. Flege 2022). Um einen dyadischen Dialogkontext für die Wegbeschreibung zu konstruieren, war es notwendig, zusätzliche Personen für die Teilnahme an diesem ersten Teil des Gesamtexperiments zu rekrutieren. Bei der Zusammenstellung der Paare wurde in erster Linie darauf geachtet, dass die beteiligten Personen über ein vergleichbares Sprachniveau im Deutschen verfügten. Das Niveau im Deutschen wurde nach den im Kapitel 4. erläuterten Kriterien bestimmt. Außerdem wurde darauf geachtet, möglichst Personen zu rekrutieren, die sich bereits kannten. Dies ist in den meisten Fällen gelungen. In den Fällen, in denen dies nicht möglich war, wurden Paare mit Personen gebildet, die in den ersten Kennenlerngesprächen mit der Versuchsleiterin ähnliche Persönlichkeitsmerkmale gezeigt hatten und als potenziell förderlich für die Schaffung eines angenehmen Gesprächskontextes angesehen wurden. Besonderes Augenmerk wurde auf die Neigung zu einem produktiven Gespräch, insbesondere mit einer unbekannten Person, sowie auf die Begeisterung für das Erlernen der deutschen Sprache gelegt. Wenn möglich, wurde darauf geachtet, dass sich die Teilnehmer bereits vor dem Experiment kannten, mit dem spezifischen Ziel, eine größere Vertrautheit mit dem/der zukünftigen Experimentalpartner/-in zu entwickeln. Diese Vorgehensweise hat sich in fast allen Fällen als erfolgreich erwiesen.

Für die Sprachaufnahmen wurden folgende Geräte verwendet, die von der Philipps-Universität Marburg zur Verfügung gestellt wurden: ein Handy Recorder H4nPro ZOOM, ein Mikrofon Sennheiser MKE 2-60 "Lavalier" und ein Thinkpad PC, der auch für das Perzeptionsexperiment verwendet wurde. Das Aufnahmeformat war WAV 44.1, 1kHz/16bit; die Aufnahmelautstärke war auf 100 eingestellt.

Was das experimentelle Design betrifft, so wurde für die Erhebung der Daten zu Sprachkompetenz, -dominanz und Sprachrepertoire derselbe Fragebogen (Marian et al., 2007) wie für das Perzeptionsexperiment und unter denselben experimentellen Bedingungen verwendet.

5.2.1 Versuchspersonen: Beschreibung der drei Stichproben

Im folgenden Abschnitt werden die drei Stichproben (Deutschlerner:innen, Kontrollgruppe mit L1 Italienisch und Kontrollgruppe mit L1 Deutsch) beschrieben.

5.2.1.1 Die Experimentalgruppe: Sprecher:innen mit L1 Italienisch und Deutschkenntnis

Die Daten, die den Analysen unterzogen wurden, basieren auf eine Stichprobe von 30 Versuchspersonen (22 Frauen, Durchschnittsalter: 23,5 Jahre, SD: 4.41; 8 Männer, Durchschnittsalter: 21,5 Jahre, SD: 3,63). Bei dieser Gruppe geht es um erwachsene Lerner:innen des Deutschen als Fremdsprache.

Alle Teilnehmer:innen haben das Italienische als L1 angegeben, wobei 4 Frauen angegeben haben, einen Dialekt des Italienischen als L1 erworben zu haben: GT_PROD erklärt, einen Dialekt aus dem venetischen Dialektgebiet als L1 zu besitzen, DR_PROD einen Dialekt aus dem südlichen Dialektgebiet (Neapel), MM_PROD das Friulanisch und EG_PROD einen Dialekt aus dem mittleren Dialektgebiet (Macerata).

Was die Dialektkompetenz betrifft, so gaben lediglich zwei Probandinnen – DR_PROD und OP_PROD – an, (passive oder sogar aktive) Kenntnisse eines Dialekts zu besitzen, dessen phonologisches System ein für die vorliegende Analyse relevantes Phonem enthält. Während die Fähigkeiten sowie der Dialekt der ersten Sprecherin bereits oben beschrieben wurden, gab die zweite an, einen Dialekt des südlichen Dialektraums zu beherrschen, genauer gesagt aus den Abruzzen.

Was das durchschnittliche Lernalter für das Deutsche angeht, liegt dies bei den Frauen bei 11.6 Jahren, während bei den Männern bei 13.8 Jahren.

Hinsichtlich der Kompetenzen im Deutschen laut des GER-Niveaus weist die Mehrheit der Proband:innen ein C1 Niveau auf, gefolgt von einer Subgruppe von 9 Proband:innen (6 Frauen, 3 Männer), die ein B1 Niveau aufweisen. Die Minderheit (6 Proband:innen, 3 Frauen, 3 Männer) weist ein B2 Niveau auf. Laut dem GER, ist das C1 Niveau als fortgeschritten zu betrachten, und zwar ist die Mehrheit der Proband:innen auf produktiver Ebene theoretisch in der Lage, klare, detaillierte Beschreibungen und Präsentationen zu komplexen Themen zu geben, dabei Unterthemen integrieren, bestimmte Punkte ausarbeiten und mit einem passenden Schluss abzurunden (vgl. Europarat 2020, S. 62).

Die Kompetenz im Deutschen wurde auch hinsichtlich der Dominanz kodiert. Bei der Mehrheit der Proband:innen (14 Proband:innen, 8 Frauen, 6 Männer) ist das Deutsche die drittbekannteste Sprache im Sprachrepertoire. Bei 8 Frauen und 1 Mann ist das Deutsche die viertbekannteste Sprache und bei der Minderheit die zweit- oder fünftbekannteste (jeweils 3 Frauen, 1 Mann; 2 Frauen).

Was die Englischkenntnisse betrifft, die hinsichtlich des Dominanzniveaus kodiert wurden, ist das Englische bei der Mehrheit der Proband: innen (17 Proband:innen, 11 Frauen, 6 Männer) die zweitbekannteste Sprache im Sprachrepertoire.

Betrachtet man das gesamte Sprachrepertoire der Proband: innen, ergeben sich alle als mehrsprachig. Die Mehrheit (17 Proband: innen: 15 Frauen, 2 Männer) haben 5 Sprachen in ihrem Sprachrepertoire angegeben, während die Minderheit entweder 3 (2 Frauen) oder 6 (1 Frau, 1 Mann).

5.2.1.2 Kontrollgruppe 1: Sprecher:innen mit L1 Italienisch

Die Kontrollgruppe mit italienischer L1, die im Rahmen dieser Studie als Referenz für die Aussprache des Italienischen betrachtet wird, besteht aus 20 Sprechern:innen (13 Frauen, 7 Männer). Die Daten bezogen auf diese Stichprobe stammen aus den *Corpora dei Lessici dell'Italiano Parlato e Scritto*¹⁴ (CLIPS, Albano Leoni, 2006). Im Einzelnen stammen die im Rahmen des vorliegenden Projektes analysierten Daten aus den *Corpora del Parlato*. Das Korpus besteht aus 5 unterschiedlichen Subkorpora (*Rundfunk und Fernsehen, dialogisch, gelesen, telefonisch, ortophonisch*), die jeweils Aufnahmen (im .wav-Format) und Transkripte (in .txt-Format) von Sprecher:innen aus 15 Städten Italiens enthalten. Für die vorliegende Studie wurden Aufnahmen aus den Korpora für dialogische und gelesene Sprache ausgewählt. Um eine mit der Versuchsgruppe qualitativ möglichst vergleichbare Kontrollgruppe zu bilden, wurden Aufnahmen ausgewählt, die die Sprache der in der Versuchsgruppe vertretenen Dialektgebiete repräsentieren. So enthält die Kontrollgruppe mehrheitlich Daten aus den Städten Venedig (venetisches Gebiet), gefolgt von Aufnahmen von Sprecher:innen aus den Städten Mailand und Bergamo (gallo-italisches Dialektgebiet), Neapel und Bari (südliches Dialektgebiet) und schließlich Catania (äußerster süditalienischer Bereich). Die Kontrollgruppe mit italienischer L1, die im Rahmen der vorliegenden Studie als Referenz für die Aussprache des Italienischen betrachtet wird, besteht aus 20 Sprechern:innen (13 Frauen, 7 Männer). Um das Ziel der Vergleichbarkeit der Daten weiter zu verfolgen, wurden außerdem Aufnahmen aus drei verschiedenen Aufnahmebedingungen ausgewählt: Wegbeschreibung, Vorlesen von Sätzen und Vorlesen von Wörtern. Dies waren die Bedingungen, die als am ehesten vergleichbar mit den Aufgaben "Wegbeschreibung", "Nacherzählung" und "Vorlesen" angesehen wurden. Eine der Stärken des CLIPS-Korpus ist neben der verfügbaren diatopischen Vielfalt seine große Homogenität, zumindest was die Schlüsselvariablen betrifft: Altersgruppe (18–30 Jahre), sozioökonomischer Status (mittlere/obere soziale Schicht), Bildungsniveau (Hochschulabsolvent:innen oder Student:innen), Wohnort in Groß- und Mittelzentren. Bezüglich der letzten Variable, sind alle Sprecher:innen entweder in der untersuchten Stadt oder in der am nächsten gelegenen Provinz geboren und zur Zeit der Erhebung lebten sie dort. Diese Charakteristika lassen eine Vergleichbarkeit der Sprecher:innen des CLIPS-Korpus mit den Sprecher:innen der experimentellen Stichprobe zu.

¹⁴ "Korpora und Lexika des gesprochenen und geschriebenen Italienischen" (Übersetzung von ADL).

5.2.1.3 Kontrollgruppe 2: Sprecher:innen mit L1 Deutsch

Die Kontrollgruppe der deutschen L1-Sprecher:innen besteht aus 10 Versuchspersonen (9 Frauen, Durchschnittsalter: 24,6 Jahre, SD: 4,26; 1 Mann, Alter: 23 Jahre). Auch diese Gruppe umfasst gesunde, mehrsprachige Erwachsene. Sie stellt grundsätzlich eine Teilgruppe der Kontrollgruppe dar, die ausschließlich am Perzeptionsexperiment teilgenommen hat und in Kapitel 4 beschrieben wurde. Zusätzlich wurden zwei weitere Personen eingeschlossen, die an der ersten (kooperativen) Aufgabe mitgewirkt haben.

Die Mehrheit der Versuchspersonen stammt aus dem Bundesland Hessen; eine Frau kommt aus Baden-Württemberg, der Mann aus Niedersachsen. Nur die Teilnehmerin aus Baden-Württemberg gab an, Vorkenntnisse in einem Dialekt ihres Bundeslandes zu haben; dennoch wies ihre Aussprache keine auffälligen regionalen Merkmale auf.

5.2.2 Aufbau des Experiments und Stimuli

Das Design folgt einem "Trichter"-Prinzip, das darauf abzielt, im größeren Kontext der Fragestellungen, die dieser Studie zugrunde liegen, auch die Variation der Vokalrealisierungen unter verschiedenen experimentellen Bedingungen zu erfassen. Frühere Studien haben nämlich gezeigt, dass die Realisierung von Vokalen von der Art der Aufgabe abhängt (vgl. z.B. Siebenhaar & Hahn 2019).

5.2.2.1 Die Aufgabe "Wegbeschreibung"

Die Wegbeschreibung wurde von der Versuchsleiterin geplant und durchgeführt und aus zwei Hauptgründen ausgewählt: 1) Diese stellt ein äußerst effektiver Aufgabentyp für die Analyse gesprochener Sprache (die Art von Sprache, die in der Literatur als *vernacular* oder *spoken language* bezeichnet wird) dar, die die Merkmale Spontaneität, unorganisierte Produktion, Verwendung von Füllwörtern, Vorhandensein von Fehlanfängen, Wiederholungen, Slangausdrücken und phonetischen Reduktionen umfasst (siehe Berríos et al., 2023); 2) Dank ihrer einfachen Struktur kann sie an die spezifischen Forschungszwecke angepasst werden (vgl. Berríos et al. 2023). Im konkreten Fall dieser Studie wurde die traditionelle Struktur und Art der Durchführung nach Brown et al. (1983) befolgt, die eine dyadische Gesprächsstruktur und die Definition von zwei Rollen beinhaltet: ein Richtungsgeber (*direction giver*) und ein Richtungsfolger (*direction follower*). Bei dieser Aufgabe erhalten die zwei Gesprächspartner:innen jeweils eine gedruckte Karte. Die Karten zeigen denselben Aufbau, unterscheiden sich jedoch darin, dass nur die Karte des sogenannten Richtungsgebers eine Route enthält. Diese Route soll der Richtungsgeber nun allein durch verbale Anweisungen so genau beschreiben, dass der Richtungsfolger sie auf seiner – identisch aufgebauten, aber leeren – Karte nachvollziehen kann. Der Richtungsfolger muss den Anweisungen folgen und die angegebene Route nachverfolgen und kann um eine Wiederholung der Anweisungen oder um allgemeine Erläuterungen bitten, so dass eine dyadische Kommunikation mit natürlichem Charakter entsteht (siehe Berríos et al., 2023).

In der für dieses Experiment konstruierten Wegbeschreibung wurden, ähnlich wie in der Umsetzung von Filipi (2014), auch kardinale Bezüge integriert, um eine zusätzliche Orientierungshilfe zu bieten. Ein stärker inhaltlich orientiertes Element der geplanten Variante der klassischen Wegbeschreibung bestand darin, ein Element des Geschichtenerzählens in den Inhalt der Karte zu implementieren: Die Route verband mehrere Szenarien zu einer kohärenten Geschichte, deren Protagonist ein fantastisches Wesen war, das – im Sinne der Kartenaufgabe – sein Zuhause erreichen sollte. Die erzählten Episoden orientierten sich an der Erzählstruktur nach Propp (1976): Ausgangszustand, Störung, Bewältigung durch den Helden und positive Auflösung. Angesichts der Heterogenität der Gruppe von Proband:innen, insbesondere in Bezug auf das Niveau der Deutschkenntnisse, an die sich die Untersuchung richten sollte, wurde eine Versuchsstruktur, die an ein allgemein bekanntes Erzählmuster erinnert, als ein Element angesehen, das eine engagiertere Herangehensweise an die Versuchssituation begünstigt. Das Vorhandensein von Erzählszenen und nicht von einzelnen Objekten wurde als ein Element betrachtet, eine reichhaltigere sprachliche Beschreibung begünstigte und in jedem Fall die Verwendung von gebräuchlichen Wörtern implizierte, die potenziell zum Lexikon von Lerner:innen mit einem Kompetenzniveau zwischen B1 und C1 gemäß der GER-Skala gehören. Die Szenen und damit auch die darin enthaltenen Elemente und die für ihre Umsetzung verwendeten Farben wurden gezielt nach dem Vorkommen der Phoneme /ə/ und /y:/ oder ihnen potenziell äquivalenter Phoneme ausgewählt, um die Verwendung dieser Wörter und damit die Umsetzung dieser Phoneme zu fördern. Um die Durchführung der Wegbeschreibung zu unterstützen, wurden, wie bei Wegbeschreibungen üblich, Wörter in die Szenen eingefügt, die in der Beschreibung als zu verwendende Pflichtwörter angegeben waren. Die beiden Versionen der Wegbeschreibung, die des Richtungsgebers und die des Richtungsfolgers, sind der Appendix aufgeführt.

5.2.2.2 Die Aufgabe „Nacherzählung“

Die zweite Aufgabe bestand darin, zwei kurze Bildergeschichten zu erzählen. Die Geschichten wurden auf der Grundlage von Illustrationen aus der Sammlung *Sortieren und erzählen: Bildergeschichten für 3- 8-Jährige - Jahreszeiten*. (2016) formuliert. Sie wurden in der Anzahl von 2 ausgewählt, um ein möglichst ausgewogenes Vorkommen und damit die Wiederverwendung von Wörtern mit /ə/ und /y:/ zu gewährleisten. Die ausformulierten Texte finden sich in der Appendix. Die Geschichten wurden mit Hilfe einer studentischen Hilfskraft mit L1 Deutsch formuliert, um die Natürlichkeit der Erzählung zu gewährleisten. Besonderes Augenmerk wurde auf die Auswahl von Bildern gelegt, die die Verwendung von Wörtern mit den relevanten Phonemen erfordern und gleichzeitig potenziell im Lexikon der Lerner:innen vorhanden sind. Um die produktive Leistung zu fördern, wurde wie bei der Wegbeschreibung auch bei dieser Aufgabe besonders auf den deskriptiven Charakter geachtet. Ebenfalls wie bei der Wegbeschreibung wurden die aus den Illustrationen konstruierten Geschichten nach der klassischen Erzählstruktur von Propp (1976) aufgebaut.

Die Geschichten wurden dann von einem Mitarbeiter des Deutschen Sprachatlas vorgelesen, der darauf trainiert worden war, die Geschichten mit einer modulierten Intonation zu lesen, die den erzählten Ereignissen entsprach und sich an den Kurzgeschichten orientierte, die Kindern typischerweise erzählt werden. Diese Wahl wurde getroffen, damit sich die Teilnehmer:innen den Inhalt besser merken konnten. Das Vorlesen der beiden Geschichten wurde im schalltoten Raum des Phonetiklabors des Fachbereichs Germanistik der Universität Marburg mit der Software *Audacity*® (Version 3.4.2, 2023) aufgenommen. Für die Erzählung wurde dieselbe männliche Stimme gewählt, die auch für die männlichen Stimmreize im Wahrnehmungsexperiment verwendet wurde. Diese war besonders klar, ohne ausgeprägten regionalen Akzent. Die Aufnahmen wurden auf dem Computer abgespeichert, der für die Durchführung des Experiments verwendet wurde und als Wiedergabemedium diente. Die Proband:innen durften die Geschichten nur einmal anhören und sich während der Wiedergabe Notizen machen, die sich auf das Notieren von Schlüsselwörtern auf einem Blatt Papier beschränkten. Am Ende der Wiedergabe der ersten Geschichte wurden ihnen die Illustrationen der gerade gehörten Geschichte in zufälliger Reihenfolge ausgehändigt: Die Proband:innen wurden nun gebeten, zunächst die Illustrationen in die ihrer Meinung nach richtige Reihenfolge zu bringen und dann mit eigenen Worten zu erzählen, was sie bei der Wiedergabe der Geschichte verstanden hatten.

5.2.2.3 Die Aufgabe "Vorlesen"

Im Rahmen der abschließenden Übung widmeten sich die Teilnehmer:innen der Aufgabe des Lesens von Wörtern. Die Präsentation der Wörter erfolgte mittels einer PowerPoint-Präsentation, wobei jede Folie einem Wort entsprach. Die präsentierten Wörter umfassten entweder Teile von Minimalpaaren, die /ə/ oder /y:/ enthielten oder Distraktoren. Die Silbenstruktur folgte u.a. dem trochäischen Betonungsmuster mit den Silbenstrukturen KV(K)KV, KV und KVK, das als prosodische Grundstruktur des Deutschen gilt (vgl. Wiese, 2000). Die insgesamt 20 Items wurden in einer zufällig generierten Reihenfolge präsentiert. Die Reihenfolge der Wörter, die während der Aufgabe präsentiert wurden, war wie folgt: *Lüge, Tier, Pappe, Kohle, Kiste, Stock, Mütze, Küste, Rose, Lotte, Kühle, Tür, Cola, Papa, Tasche, Stück, Liege, Putze, Rosa, Lotto*. Die Teilnehmer:innen hatten die volle Kontrolle über die PowerPoint-Präsentation und konnten die Folien frei durchblättern.

Die Proband:innen wurden gebeten, die Wörter der Präsentation laut zu lesen, und zwar so, wie es für sie spontan am angemessensten erschien. Die Wörter mussten einmal dem Ablauf der Präsentation folgend bis zur letzten Folie gelesen werden und dann noch einmal in umgekehrter Reihenfolge. Diese Vorgehensweise wurde gewählt, um eine hinreichend repräsentative Anzahl von Exemplaren zu erhalten und gleichzeitig die Möglichkeit eines Lerneffekts der gerade gelesenen Sequenz zu verringern.¹⁵ Die Teilnehmer:innen durften die auf dem Bildschirm

¹⁵ Die ursprüngliche Aufgabe bestand darin, die Wortfolge einmal zu lesen. Während der ersten Versuchssitzung zeigte die erste Teilnehmerin jedoch ein offensichtliches Anspannungsverhalten und wiederholte die Elemente der vorgegebenen Reihenfolge mehrfach, um sicherzustellen, dass ihre Aussprache den von ihr definierten Erwartungen an

dargestellten Begriffe mehr als einmal aussprechen, wenn sie dies für die Korrektur einer ihrer Meinung nach ungeeigneten Aussprache benötigten.

5.2.3 Prozedere

Sowohl die Experimentalgruppe als auch die Kontrollgruppe mit deutscher L1 nahmen an den gleichen Aufgaben unter den gleichen Versuchsbedingungen teil, mit dem einzigen Unterschied, dass die Kontrollgruppe an dem Experiment im Fachbereich Germanistik der Philipps-Universität Marburg teilnahm. Die Proband:innen betraten den Raum, in dem das Experiment stattfinden sollte, und nahmen nebeneinander an einem Schreibtisch Platz, an dem die Versuchsleiterin saß. Nach der Erklärung des Versuchsablaufs und der Unterzeichnung der Einverständniserklärung begann die Durchführung des ersten Experiments, bei dem zwei Stimmen aufgezeichnet wurden. Der Versuchsperson, die an der Aufgabe als Richtungsfolgender teilnahm, wurde ein Stift ausgehändigt, und beide Versuchspersonen wurden angewiesen, ihre eigene Karte zu betrachten, ohne auf die Karte der anderen Person zu schauen. Die Versuchsleiterin ließ die Versuchspersonen das Experiment selbständig durchführen und griff nur in Ausnahmefällen ein, wenn eine Versuchsperson sichtlich Schwierigkeiten mit der Aufgabe hatte. In solchen Fällen ermutigte die Versuchsleiterin einen der beiden Teilnehmer:innen, die/den andere/-n zu unterstützen. Nach Abschluss des ersten Experiments durfte die zusätzlich beteiligte Versuchsperson den Raum verlassen, und es wurde mit der nächsten Aufgabe fortgefahren.

Bei der Aufgabe "Nacherzählung" wurden zu jeder Geschichte die entsprechenden Aufnahmen abgespielt und jeweils nach dem Hören die entsprechenden Bilder ausgehändigt.

Nach der Vorleseaufgabe hatte die Versuchsperson die Möglichkeit, eine kurze Pause einzulegen, bevor die Experimentalsitzung weiter mit dem Perzeptionsexperiment fortsetzte.

Die Audioaufnahmen wurden mit einem mobilen Handheld-Recorder des Typs ZOOM H4n Pro durchgeführt. Die Digitalisierung des akustischen Signals erfolgte im WAV-Format mit einer Abtastrate (*sampling rate*) von 44,1 kHz und einer Quantisierung von 16 Bit. Der Aufnahmepegel (*recording level*) des Recorders wurde für alle Sitzungen konstant auf den Maximalwert 100 fixiert, während die Wiedergabelautstärke (*playback volume*) auf 73 eingestellt war. Als elektroakustischer Wandler kam ein Lavaliermikrofon der Marke Sennheiser (Kapselmodul MKE 2-60 in Kombination mit dem Speiseadapter K6) zum Einsatz. Das Lavaliermikrofon wurde bei den Aufgaben "Nacherzählung" und "Vorlesen" angewendet.

Bei der Erhebung der Daten zu Sprachkompetenz, -dominanz und Sprachrepertoire wurden sowohl die Versuchspersonen, die an beiden Experimenten teilgenommen hatten, als auch diejenigen, die nur an der Wegbeschreibung teilgenommen hatten, gebeten, den

eine "korrekte Aussprache des Deutschen" entsprach. Ab der zweiten Sitzung wurde daher die Methode angewandt, die Sequenz zweimal zu lesen, einmal in die eine und einmal in die umgekehrte Reihenfolge, um die Teilnehmer:innen in der Wiederholung ihrer Leistung zu unterstützen und so möglicherweise den Leistungsstress zu verringern.

soziolinguistischen Fragebogen auszufüllen. Dabei wurden die gleichen Kriterien berücksichtigt, wie sie in Kapitel 4 für das Perzeptionsexperiment beschrieben wurden.

5.2.4 Datenanalyse

Insgesamt wurden 99 Aufnahmen mit einer Gesamtdauer von 8 Stunden, 44 Minuten und 53 Sekunden für die Gruppe der Lerner:innen, 34 Aufnahmen mit einer Gesamtdauer von 1 Stunde, 46 Minuten und 4 Sekunden für die Kontrollgruppe der L1 Deutsch Sprecher:innen und 33 Aufnahmen mit einer Gesamtdauer von 2 Stunden, 30 Minuten und 35 Sekunden für die Kontrollgruppe der L1 Italienisch Sprecher:innen annotiert und analysiert.

Dazu wurden die Aufnahmen im .wav-Format in *Praat* (Version 6.4.22, Boersma & Weenink, 2024) geladen und annotiert. Da die dyadischen Dialoge für die Wegbeschreibung in Stereo aufgenommen wurden, wurden sie für die Annotation in Mono umgewandelt. Die Aufnahmen für die Nacherzählung und die Vorleseaufgabe mussten nicht konvertiert werden, da sie bereits als Mono-Dateien gespeichert waren.

Die Aufnahmen der Lerner:innen-Gruppe wurden von der Versuchsleiterin manuell annotiert, da zum Zeitpunkt der Analyse keine geeignete Software zur automatisierten Annotation von akzentuiertem Deutsch zur Verfügung stand. Die Annotation wurde ausschließlich von der Versuchsleiterin durchgeführt, da ihre L1 Italienisch ist und sie über entsprechendes Vorwissen bezüglich Ausspracheschwierigkeiten bei italienischen L1-Sprecher:innen verfügte. Für die Annotation der Aufnahmen der Kontrollgruppe mit L1 Deutsch wurde das Tool WebMAUS (Version 4.0, Schiel, 1999, 2015; Kisler et al., 2017) verwendet. Die von WebMAUS durchgeführte Annotation wurde anschließend von der Versuchsleiterin und einer studentischen Hilfskraft mit L1 Deutsch manuell überprüft. Alle Aufnahmen für die Kontrollgruppe der Sprecher:innen mit L1 Deutsch wurden in einer ersten Phase von der Versuchsleiterin annotiert, in einer zweiten Phase von einer studentischen Hilfskraft mit L1 Deutsch kontrolliert und in einer dritten Phase nochmals von der Versuchsleiterin kontrolliert. Dieses Vorgehen erwies sich als besonders hilfreich, da sich die Sprachproduktion in der L1 Deutsch als besonders koartikuliert erwies und Schwierigkeiten bei der Segmentierung bereiten konnte. Die Annotation der Sprachproduktion der Kontrollgruppe mit L1 Italienisch wurde sowohl von der Versuchsleiterin als auch von der studentischen Hilfskraft durchgeführt, da letztere über ausreichende Italienischkenntnisse verfügte. Dabei wurden die Aufnahmen zwischen den beiden Annotatorinnen aufgeteilt und anschließend von der Versuchsleiterin kontrolliert. Die Annotationen wurden immer unter denselben Annotationsregeln vorgenommen.

Im Allgemeinen wurde für die Annotation von Frauenstimmen der *view range* in *Praat* auf 5500 Hz und für Männerstimmen auf 5000 Hz eingestellt. In Fällen, in denen das Formantenmuster nicht klar erkennbar war, wurde die Formantenvisualisierung und Modulation des *dynamic range* verwendet. Bei der Annotation wurde auf folgende Regel geachtet: 1) Exemplare mit knarrender

Stimme wurden nicht annotiert, 2) die rechte Vokalgrenze wurde vor jedem Echo gesetzt, 3) die Anfangs- und Endvokalgrenzen wurden durch Beobachtung sowohl des Spektrogramms als auch des Oszillogramms bestimmt, wo die Wellenform begann, die typischen Konnotationen des identifizierten Vokalbeispiels zu verlieren, wurde die rechte Grenze gesetzt.

Die vokalische Annotation basiert auf IPA-Symbolen, die die phonologisch erwartete Aussprache kodifizieren, anstelle der tatsächlich akustisch realisierten Varianten. Referenz für die erwartete Aussprache war das *Große Wörterbuch der deutschen Aussprache* (Krech et al., 2009). Bei der Annotation des Vorkommens von /ə/ wurde dem IPA-Symbol eine Notation hinzugefügt, um die Position des Phonems innerhalb des Wortes zu identifizieren. Bei der Annotation von /y:/ wurden die Phoneme, die dem linken und rechten Kontext entsprechen, neben das Phonem gesetzt. Wörter, die die interessierenden Phoneme enthalten, wurden vollständig annotiert. Zusätzliche Vokalvorkommen wurden hinzugefügt, um eine möglichst ausgewogene Stichprobe zu erhalten, die nicht nur die relevanten Phoneme, sondern auch die Kernvokale im Allgemeinen umfasst. In Verfolgung des primären Ziels, die Forschungsfragen zu beantworten, wurde außerdem versucht, ein möglichst umfassendes Bild der allgemeinen Konfiguration des Vokalraums der drei untersuchten Sprechergruppen zu erhalten.

Für die Beantwortung der Fragestellungen wurde die Qualität der Realisierungen akustisch gemessen. Insbesondere wurden die Werte der ersten drei Formanten analysiert. Die ersten beiden Formanten gelten als die relevantesten für die Beschreibung der Vokalqualität, während der dritte Formant als relevant für die Analyse der Lippenprojektion in der Vokalartikulation und damit ihrer Rundung angesehen wird. Dies wurde in die Analyse der Qualität von /y:/ integriert.

Am Ende der Annotation wurden die Formantendaten für alle identifizierten Proben jedes Teilnehmers mit Hilfe eines *Praat*-Skripts extrahiert. Durch das Script wurden die F1-, F2- und F3 Werte in 3 verschiedenen Portionen des annotierten Vokals extrahiert. Der Code berechnet Start-, Mittel- und Endpunkt (*time_left*, *time_middle*, *time_right*) eines Vokalsegments, um an diesen drei Positionen Formantwerte zu extrahieren. Für die vorliegende Studie wurden ausschließlich die Formantwerte für die zentrale Portion des annotierten vokalischen Segments (*time_middle*) in Betracht gezogen. Diese Auswahl basiert auf der Annahme, dass das Zentrum des Vokals als der stabilste Abschnitt gilt und daher die zuverlässigsten qualitativen Informationen über den Vokal liefert (vgl. z.B. O'Brien & Smith, 2010).

Die erste zu analysierende Gruppe war die Versuchsgruppe. Das Verfahren zur Vorbereitung dieses Datensatzes, für den die deskriptiven und inferenzstatistischen Analysen durchgeführt wurden, legte das Verfahren fest, das auf die Kontrollproben angewendet werden sollte.

Was die rein linguistischen Informationen betrifft, so wurde die Variable der Position des Vorkommens von /ə/ innerhalb des Wortes und der linke und rechte phonologische Kontext des Vorkommens von /y:/ kodiert. Die Variable, die sich auf die Position von /ə/ innerhalb des Wortes

bezog, wurde durch die Kontexte *präfixal*, *Wortstamm*, *präfinal* und *final* charakterisiert, die jeweils die Position innerhalb der präfixalen Partikel (wie in <be->, <ge->), die Position innerhalb des Wortes (seltener Fall, aber vertreten durch Fälle wie das erste Auftreten von /ə/ im Wort *andere*) und die absolute Wortendstellung, typisch z. B. für das flektierende Verbalmorphem der ersten Person Singular <-e> (wie in *sage*, *würde*, *liege* usw.) oder das nominale Morphem Singular <-e> (wie bei *Pappe*, *Lüge* usw.).

Bei der Variable für den linken und rechten phonologischen Kontext in Bezug auf das Auftreten von /y:/ wurden die Kontexte analysiert, für die es in beiden Populationen (männlich und weiblich) eine Mindestanzahl von 10 Beobachtungen gab. So wurden für die linke Kontextvariable für beide Populationen die folgenden Artikulationsorte definiert: *labiodental*, *alveolar*, *uvular*, *velar* und *glottal*; für den rechten Kontext wurden die folgenden Ebenen definiert: *bilabial*, *alveolar*, *uvular* und *velar*. Für die Intergruppe-Analyse wurden die geteilten Kontexte *labiodental*, *alveolar*, *uvular*, *velar* und *glottal* analysiert.

Bei den Variablen der Sprachkompetenz, -dominanz und des Sprachrepertoires wurden die Datensätze von den Teilnehmer:innen um Angaben zur Deutschkompetenz (GER-Niveau), zur Dominanz des Deutschen und Englischen sowie zur Breite des sprachlichen Repertoires ergänzt. Diese Variablen wurden anhand der Skala von 1 (höchstes Niveau) bis 5 (niedrigstes Niveau) des soziolinguistischen LEAP-Fragebogens kodiert. Für eine explorative Analyse, wurde zudem die Variable für die Kenntnisse in einem Dialekt der L1 (Italienisch) kodiert.

Nach dieser Integration mit den interessierenden Variablen wurde die Verteilung der Daten durch Anwendung eines Shapiro-Wilk-Tests mit der *Jamovi*-Software (Version 2.4.11.0, 2025) überprüft. Für beide Populationen (männlich und weiblich) ergab der Test eine nicht vollkommene Normalverteilung der Daten, wobei der Test für einige Vokale $p < .001$ ergab. Da die Möglichkeit bestand, dass dies auf das Vorhandensein von Ausreißern zurückzuführen sein könnte, wurde der Datensatz in mehreren Phasen bereinigt. Zuvor wurden die Stellen aussortiert, an denen das *Praat*-Skript nicht in der Lage war, die drei Formanten zu messen. Die Löschung wurde der Imputation fehlender Daten vorgezogen, um ausschließlich mit Realdaten arbeiten zu können. Die in die Analyse einbezogenen Vokale wurden nach ihrer Relevanz für die Forschungsfragen ausgewählt, wobei das Vorhandensein von /ə/, /ɛ/, /e:/, /y:/, /ʏ/, /i:/, /ɪ/, /u:/ und /ʊ/ als unerlässlich angesehen wurde. Zu diesen wurden /a/, /a:/ und /o:/ hinzugefügt, um die analysierten Realisierungen innerhalb des Vokalraums besser kontextualisieren zu können.

Auf diese Weise war es nicht nur möglich, die der vorliegenden Studie zugrunde liegenden Forschungsfragen zu beantworten, sondern auch die Realisierungen der Kontraste /ɛ-e:/, /y:-Y/, /u:-ʊ/ zu analysieren und somit ein vollständigeres Bild der lautlichen Repräsentationen der Lerner:innen zu erhalten. Nach der Festlegung des Vokalinventars, auf das sich die Analyse beziehen sollte, wurden die Vokalformantendaten einem Bereinigungsprozess durch Identifizierung

und Entfernung von Ausreißer unterzogen, um die Zuverlässigkeit der Analyse zu gewährleisten. Ausreißer wurden mit Hilfe der Methode des Interquartilsbereichs (IQR) definiert, die üblicherweise zur Identifizierung von Extremwerten in nicht-normalen Verteilungen verwendet wird. Für jeden Vokal wurden die Werte von F1, F2 und F3 entsprechend dem ersten Quartil (Q1) und dem dritten Quartil (Q3) aufgeteilt. Anschließend wurden die untere und obere Grenze nach der folgenden Formel berechnet:

$$\text{Lower bound} = Q1 - 1.5 \times IQR$$

$$\text{Upper bound} = Q3 + 1.5 \times IQR$$

wobei der Interquartilsbereich definiert ist als:

$$IQR = Q3 - Q1$$

Alle Werte außerhalb dieser Bereiche wurden als Ausreißer betrachtet und aus der Analyse entfernt. Der IQR-basierte Ansatz wurde gewählt, weil er ein robustes Kriterium für den Umgang mit nicht-normalverteilten Daten bietet und den Einfluss von Extremwerten vermeidet, die die nachfolgenden statistischen Analysen beeinträchtigen könnten.

Nach der Streichung der Ausreißer wurde ein erneuter Shapiro-Wilk-Normalitätstest durchgeführt, der für beide Populationen eine Verteilung ergab, die einer Normalverteilung weitaus ähnlicher war als die Ergebnisse des ersten Tests. Da jedoch immer noch einzelne Fälle mit p-Werten kleiner als .001 vorhanden waren, wurde die Anwendung nachfolgender Analysemethoden präferiert, die für die Bearbeitung nicht normalverteilter Daten ausreichend robust sind.

Aus diesem Grund wurden deskriptive statistische Analysen auf der Grundlage des Medians der Daten durchgeführt, wobei der IQR für jeden Vokal ermittelt wurde.

Die akustische Darstellung der Vokalkontraste in diesem Kapitel folgt einer differenzierten Strategie zur Sichtbarmachung der erhobenen Daten. Als primäres beschreibendes Instrument dienen Streuungsellypsen im F2/F1- oder F3/F2-Vokalraum, die eine visuelle Einordnung der kategorialen Besetzung sowie der relativen Überlappungsbereiche der Phoneme ermöglichen. Die deskriptive Visualisierung für alle weiteren Untersuchungsparameter erfolgte ausschließlich mittels Boxplots. Dies betrifft die Darstellung von Positionseffekten bei /ə/, den Einfluss des phonologischen Kontextes bei /y:/ sowie die Variationen in Abhängigkeit von der Aufgabenart und den Variablen der Sprachkompetenz, -dominanz und Sprachrepertoire im Rahmen der übergreifenden Analyse.

Im Rahmen der vergleichenden deskriptiven Analyse wurden ausschließlich die Daten der weiblichen Versuchspersonen herangezogen. Ein Einbezug der männlichen Probanden wurde als nicht zielführend erachtet, da die männliche Datengruppe in der deutschsprachigen Kontrollgruppe

lediglich durch einen Teilnehmer vertreten war. Diese Einschränkung auf die weiblichen Probandengruppen gewährleistet zudem die methodische Konsistenz mit dem Vorgehen in der Inferenzstatistik.

In Bezug auf die Inferenzstatistik wurde die Analyse der Vokalrealisierung durch die Lerner:innen, sowohl bei männlichen als auch bei weiblichen Lerner:innen durchgeführt.

Der Verzicht auf eine Normalisierung der akustischen Daten begründet sich primär in der Zielsetzung, die interindividuelle Variation – insbesondere soziolinguistischer Art – in ihrer ursprünglichen Form abzubilden. Da die Untersuchung ausschließlich geschlechtshomogene Gruppen umfasst, wird die anatomisch bedingte Varianz innerhalb der Stichproben als kontrollierbar erachtet, was einen direkten Vergleich der Rohdaten rechtfertigt. Ergänzend dazu stützt sich dieses methodische Vorgehen auf drei wesentliche Aspekte: Erstens wurde in vergleichbaren Studien (vgl. Missaglia & Sendlmeier, 1999; O’Brien & Smith, 2010) bei geschlechtshomogenen Stichproben ebenfalls auf eine Normalisierung verzichtet. Zweitens werden alle Analysen konsequent geschlechtergetrennt durchgeführt. Drittens ergab eine explorative Voranalyse, dass eine Normalisierung der Daten keine substantiellen Veränderungen der inferenzstatistischen Tendenzen bewirkte, weshalb die Darstellung der gereinigten, geschlechtergetrennten als deskriptiv adäquater bewertet wurde.

Zur Analyse der Variation der F1- und F2-Werte in Bezug auf die Position von /ə/ im Wort, wurden vier Positionen in Betracht gezogen: die präfixale, die Position im Wortstamm, die präfinale und die absolute finale Position. Obwohl der Autorin bewusst ist, dass /ə/ in der Regel nicht im Wortstamm vorkommt (vgl. Wiese, 2000), hat die Studie Anlass zur Erhebung zahlreicher Exemplare für die Ausnahmeformen *andere* und *Neeles* – letzteres die Genitivform eines Eigennamens, der für eine der zwei Kurzgeschichten für die Aufgabe “Nacherzählung” verwendet wurde – gegeben. Nachfolgend werden Beispiele für alle Positionen aufgeführt (Tabelle 7). Die Beispiele wurden direkt aus dem Datensatz (also auch in flektierter Form) entnommen.

Tabelle 7: Beispiele für alle kodierten Positionen im Datensatz. Die /ə/-Position im Wort ist blau markiert

Position	Beispiele
präfixal	<i>be</i> ginnen, ange <i>fe</i> ngen, ge <i>ge</i> gangen
Wortstamm	<i>andere</i> (einziger Fall)
präfinal	Tü <i>re</i> n, ander <i>en</i> , zeich <i>ne</i> t
final	Büsch <i>e</i> , nehme <i>e</i> , grü <i>ne</i>

Für die Inferenzstatistik wurden lineare gemischte Modelle (LMM) verwendet. Für die explorative Studie zur Untersuchung des möglichen Einflusses von Dialektkenntnissen auf die Realisierung des Schwa wurde die Analyse – wie auch im Fall der Perzeptionsstudie – auf einer rein deskriptiven Ebene gehalten, da nur 2 Proband:innen beteiligt waren.

Für die Datenanalyse wurde *Jamovi* (Version 2.4.11.0, 2025) für die deskriptive Statistik und die Normalitätstestbedingung verwendet, *RStudio* (R Version 4.4.1, 2024-06-14 ucrt) für die Anwendung der IQR-basierten Methode zur Eliminierung von Ausreißern durch die Verwendung des Pakets *dplyr* (Version 1.1.4) und *Jamovi* wiederum für die Inferenzstatistik. Die Erstellung der Diagramme wurde mittels eines Codes in der Rj Editor-Erweiterung von *Jamovi* generiert.

5.3 Ergebnisse

5.3.1 Allgemeiner kontrastiver Übersicht über die Vokalsysteme

Die Analyse der Werte der ersten drei Formanten ermöglichte die Visualisierung des Vokalraums der drei Gruppen – Lerner:innen (kodiert als *Lernende*), Kontrollgruppe mit L1 Deutsch (kodiert als *Ko_DEU*) und Kontrollgruppe mit L1 Italienisch (kodiert als *Ko_ITA*) – in einem F2/F1-kartesischen Raum. In der Folge werden die Vokalräume der weiblichen und männlichen Lerner:innen durch die Medianwerte dargestellt (Abbildung 15 und 16). Den durch die Medianwerte definierten gesamten Vokalräumen werden jeweils Abbildungen gegenübergestellt, die die Realisierungen von /ə/ und /y:/ hervorheben, um die Streuung der Daten zu verdeutlichen.

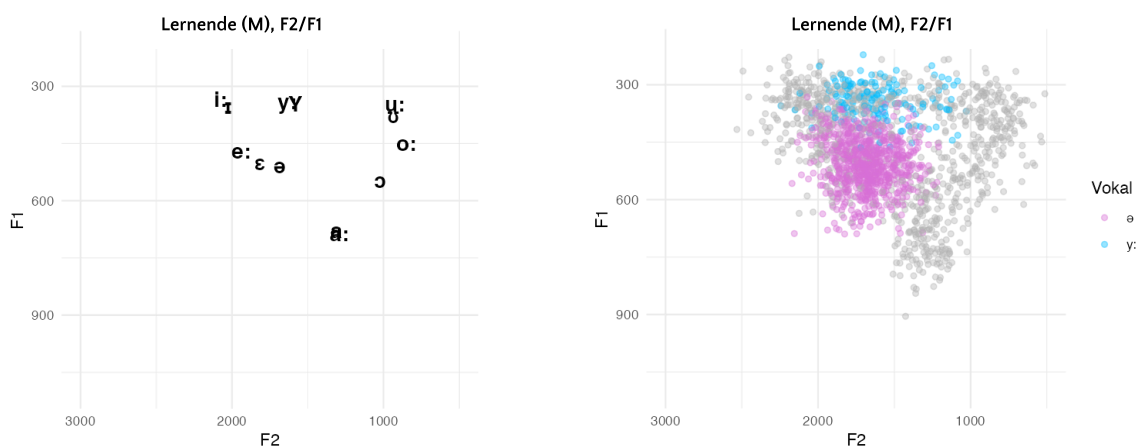


Abbildung 15: Links: F2/F1-Grafik des Vokalraums der Lerner (Männer, "M") durch Angabe der Medianwerte. Rechts: F2/F1-Grafik des Vokalraums der Lerner ("M") durch die Angabe der Einzelwerte ausgewählter Vokale mit Hervorhebung der Realisierungen für /ə/ (pink) und /y:/ (hellblau).

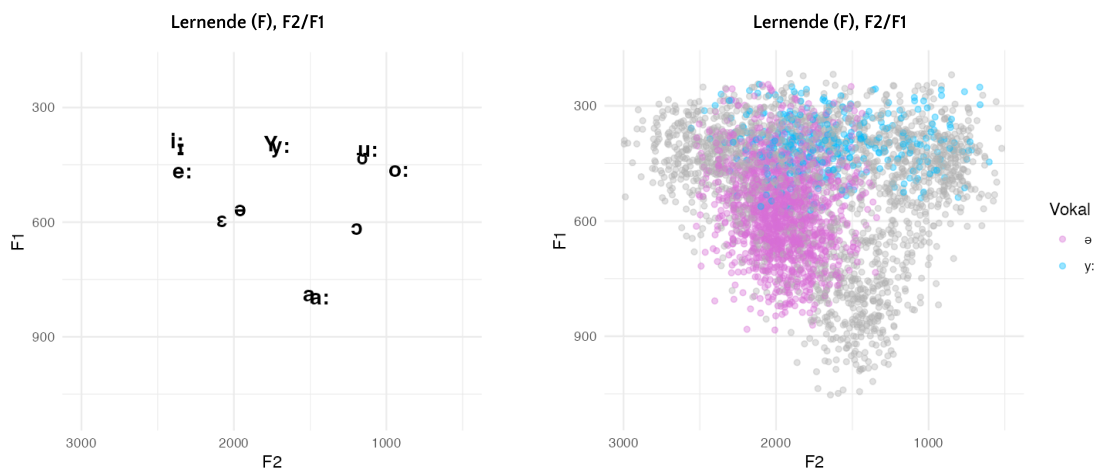


Abbildung 16: Links: F2/F1-Grafik der Realisierung ausgewählter Vokale des Deutschen bei den Lernerinnen (Frauen, "F") durch Angabe der Medianwerte. Rechts: F2/F1-Grafik des Vokalraums der Lernerinnen ("F") durch die Angabe der Einzelwerte ausgewählter Vokale mit Hervorhebung der Realisierungen für /ə/ (pink) und /y:/ (hellblau).

Die abgebildeten Realisierungen beziehen sich auf alle drei Experimentaufgaben. Bei der Analyse der Daten ist es von entscheidender Bedeutung, neben den personenbezogenen Differenzen in der Anzahl an produzierten Items, insbesondere bei den ersten beiden Aufgaben, auch die unterschiedliche Anzahl an Proband:innen pro Gruppe zu berücksichtigen. Diese Faktoren resultieren in einer insgesamt divergierenden Konfiguration hinsichtlich der Anzahl und Dichte der Realisierungen in den beiden Vokalräumen. Aus rein deskriptiver Perspektive lässt sich zudem ableiten, dass die Realisierungen von /ə/ insgesamt stark dezentralisiert sind und sich mit denen von /ε/ und /e:/ überlappen. Die erhöhte Konzentration von /ə/ im Vergleich zu den anderen Phonemen ist auf dessen hohe Häufigkeit im deutschen Lexikon zurückzuführen, was in den vorliegenden Abbildungen ersichtlich wird.

Hinsichtlich des /y:/ sind die Realisierungen sowohl bei Frauen als auch bei Männern über den höheren Arealen des Vokaltrapezes verstreut, insbesondere in Richtung der hinteren Artikulationsorte, wobei diese Tendenz bei den Frauen stärker ausgeprägt ist als bei den Männern.

In den nachfolgenden Grafiken (Abbildung 17) werden die Vokalräume der Lernerinnen mit denen der Sprecher:innen mit L1 Deutsch durch Angabe der Medianwerte in Kontrast gesetzt, indem die Medianwerte der Vokale im selben Vokalraum gegenübergestellt werden.

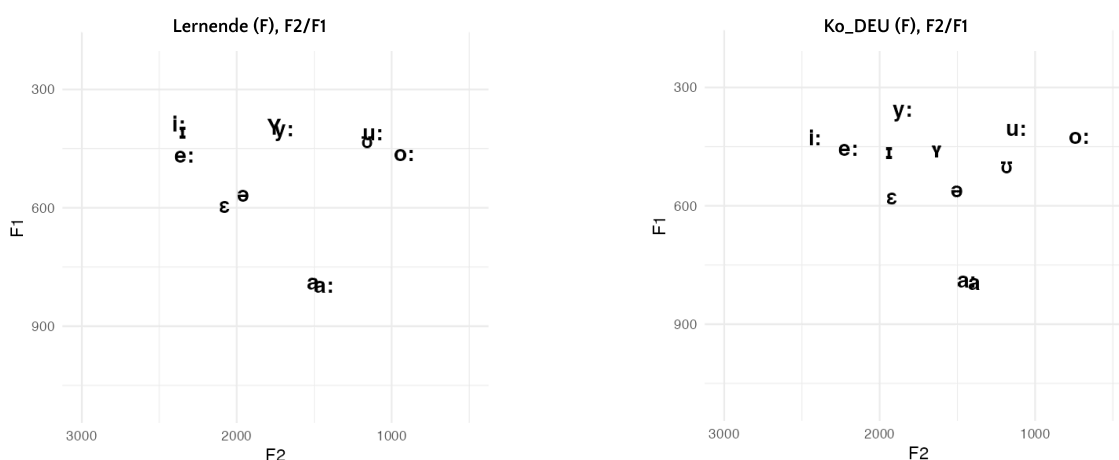


Abbildung 17: Vergleichende F2/F1-Grafik der Realisierung ausgewählter deutscher Vokale durch Angabe der Medianwerte für die Stichprobe der Frauen ("F"). Links: Lernerinnen. Rechts: Sprecherinnen mit L1 Deutsch.

Die Grafiken der Vokalräume zeigen die deutliche Tendenz der Lernerinnen, die deutschen Vokale in peripherischen Positionen des Vokalraums zu verorten. Als Ergebnis davon werden zentrale Artikulationsorte nicht besetzt und dadurch wird die Anlehnung an die Konfiguration des italienischen Vokalsystems deutlich.

In den folgenden F2/F1-Graphiken (Abbildung 18) wird das System der deutschen Vokale ausgesprochen durch die Lernerinnen mit den Realisierungen der italienischen Vokale ausgesprochen von L1-Sprecherinnen des Italienischen in Kontrast gesetzt.

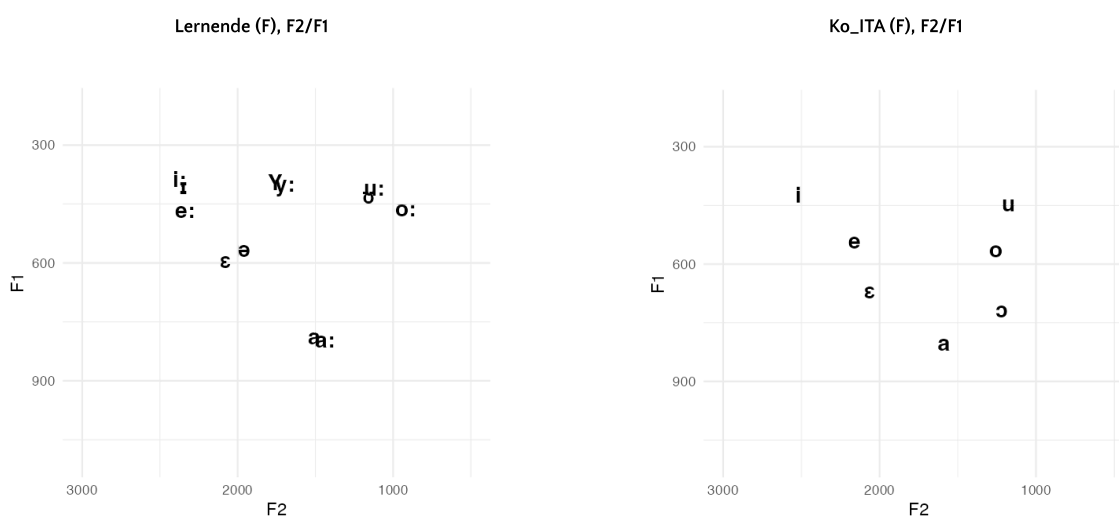


Abbildung 18: Vergleichende F2/F1-Grafik der Realisierung ausgewählter deutscher Vokale und der Vokale des Italienischen durch Angabe der Medianwerte für die Stichprobe der Frauen ("F"). Links: Lernerinnen. Rechts: Sprecherinnen aus dem CLIPS-Korpus mit L1 Italienisch.

Im Vokalraum der Lernerinnen fällt besonders auf, dass sich vordere und hintere Vokalcluster bilden, was auf einen Magnet-Effekt der L1-Kategorien in Richtung der Peripherie des Vokalraums hindeutet.

In dem Appendix befinden sich die Medianwerte zu den ersten drei Formanten und die entsprechenden IQRs für alle untersuchten Vokalen und Gruppen.

In den folgenden Abschnitten werden die Ergebnisse der Deskriptiv- und Inferenzstatistik präsentiert.

5.3.2 Ergebnisse zur Realisierung von /ə/

In diesem Abschnitt werden die Ergebnisse in Bezug auf die Realisierung von /ə/ präsentiert. Es wird die potenzielle Unschärfe der Trennung zwischen den deutschen Vokalkategorien /ə/-/e:/-/ɛ/, der Einfluss der Position von /ə/ innerhalb des Wortes und der Dominanz im Englischen auf die Realisierung dieses Phonems berücksichtigt.

In diesem Zusammenhang werden die Ergebnisse für die experimentelle Gruppe mit denen für die Kontrollgruppe mit deutscher L1 verglichen, um die Zielsprachlichkeit der Realisierungen zu überprüfen. Sie werden außerdem mit den Realisierungen der italienischen Phoneme /e/ und /ɛ/ verglichen, um eventuelle Kategorienüberlappungen und somit *Cross Linguistic Influence* (CLI)-Phänomene zu identifizieren. Was die Angaben zu Sprachkompetenz und -dominanz betrifft, wird auf die mögliche Aussagekraft der Variablen der Kompetenz im Deutschen gemessen an GER-Niveau und der Sprachdominanz eingegangen.

5.3.2.1 Realisierung von /ə/ in Kontrast zu /e:/und /ɛ/

In diesem Abschnitt werden die Ergebnisse zur Realisierung von /ə/ durch die Lerner:innen im Vergleich zu den Realisierungen der anderen zwei naheliegenden Phoneme präsentiert. Dazu wurden die Werte zu den ersten zwei Formanten (F1, F2, in Hz) in Betracht gezogen. Die Werte zu des ersten Formants werden in der Regel hinsichtlich des Offenheitsgrads (tief/hoch) für aussagekräftig gehalten, während die Werte des zweiten Formant werden in der Regel hinsichtlich des Artikulationsorts (vorne/hinter) für aussagekräftig gehalten. In der folgenden Abbildung (Abbildung 19) werden die Dispersionsellipsen¹⁶ (95%- Konfidenzniveau) zu den Kategorien /ə/, /e:/ und /ɛ/ dargestellt

¹⁶ Die Ellipsen basieren auf der t-Verteilung, die insbesondere bei kleineren Stichproben als robuster gilt, jedoch auch bei größeren Stichproben geeignete Konfidenzbereiche liefert.

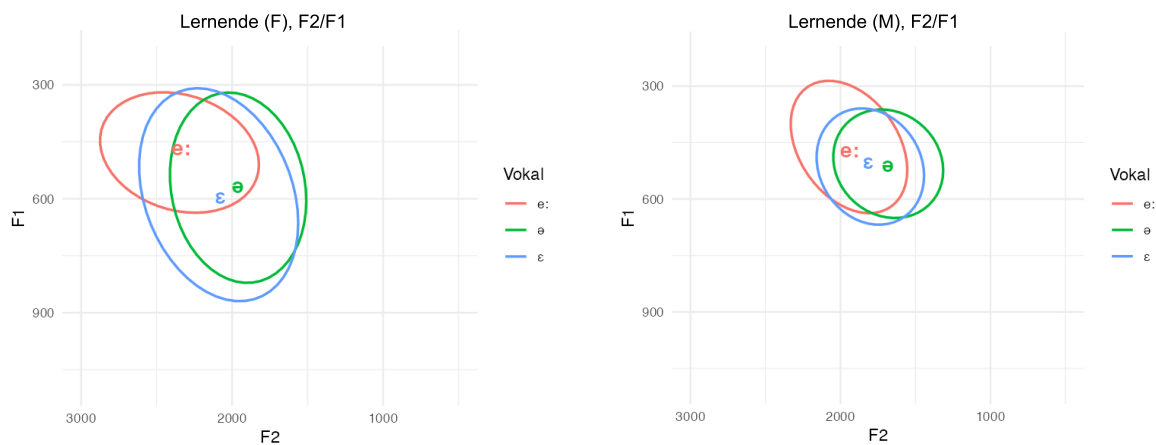


Abbildung 19: Dispersionsellipsen (95% Konfidenzniveau) und Medianwerte für /ə/, /e:/ und /ɛ/ für die Lerner:innen. Links: Frauen ("F"). Rechts: Männer ("M").

Auf deskriptiver Ebene zeigen die Dispersionsellipsen, dass sich sowohl bei den weiblichen als auch bei den männlichen Sprechern die Realisierungen von /ə/ mit denen von /e:/ und /ɛ/ überlappen, wobei die fast vollständige Konvergenz mit den Realisierungen von /ɛ/ besonders auffällig ist. Parallel kann man feststellen, dass sich die Artikulation von /e:/ bei den Frauen eher horizontal in Richtung der zentralen Artikulationsorte verstreut, während dies bei den Männern nicht der Fall ist und sich eher quer in die Richtung der tieferen Artikulationsorten (Richtung /ɛ/) verstreut.

Das statistische Modell

$$\begin{aligned} \text{Imer } F1 &= 1 + \text{Vokal} + (1 | \text{Versuchsperson}); \\ \text{Imer } F2 &= 1 + \text{Vokal} + (1 | \text{Versuchsperson}) \end{aligned}$$

wurde sowohl bei den Frauen als auch bei den Männern auf die Werte der F1 und F2 durchgeführt. Als abhängige Variable wurden jeweils F1 und F2 angegeben, während als Faktor die Variable *Vokal* und als Zufallsfaktor die Variable *Versuchsperson* angegeben wurde¹⁷.

Beginnend mit der Analyse der F2-Werte sowohl bei den Frauen als auch bei den Männern hat sich der Effekt für den Faktor *Vokal* als signifikant erwiesen ($p < .001$). Der Post-Hoc Test hat nämlich bei den beiden Gruppen dieselben Ergebnisse geliefert, nämlich eine hohe Signifikanz ($p < .001$) bei den Vergleichen /e:/- /ə/, /e:/- /ɛ/ und /ə/- /ɛ/.

¹⁷ Es wurde systematisch geprüft, ob für je getrennten Analysen neben den *Random Intercepts* auch *Random Slopes* für den Faktor *Vokal* (*Vokal* | *Versuchsperson*) implementiert werden können (vgl. Barr et al., 2013). Stichprobenartige Testungen zeigten jedoch, dass dieser maximale Ansatz die mathematischen Kapazitäten der vorliegenden Datenstruktur überschreitet: Bei den Modellen für F1 (Frauen) traten schwerwiegende Konvergenzprobleme auf, während die Modelle für F1 und F2 (Männer) jeweils einen singulären Fit aufwiesen. Lediglich das Modell für F2 (Frauen) konvergierte stabil. Im Sinne der Modellsparsamkeit und zur Gewährleistung einer einheitlichen, mathematisch stabilen und unverzerrten Parameterschätzung für alle Teilgruppen wurde daher für alle Analysen das rein strukturierte Random-Intercept-Modell beibehalten.

Fortsetzend mit der Analyse der F1-Werte, liefert das statistische Modell zum Teil vergleichbare Ergebnisse für Frauen und Männer. Der Effekt für den Faktor *Vokal* hat sich auch bei dieser Variable als signifikant erwiesen ($p < .001$). Laut des Post-Hoc Tests für beide Gruppen, ist der Effekt für die Vergleiche $/e:/-/\ə/$, $/e:/-/\epsilon/$ und $/\ə/-/\epsilon/$ signifikant für die Frauen ($p < .001$), bei den Männern haben sich die Effekte für die Vergleiche $/e:/-/\ə/$ und $/e:/-/\epsilon/$ als signifikant (jeweils $p = 0,002$ und $p < .001$) erwiesen. Bei letzterer Gruppe ist der Effekt für den Vergleich $/\ə/-/\epsilon/$ nicht signifikant ($p = 0,288$).

Nachfolgend wird die Realisierung von $/\ə/$ im Vergleich zu der von $/\epsilon/$ und $/e:/$ bei den Lernerinnen des Deutschen und der deutschen Kontrollgruppe in Kontrast gesetzt (Abbildung 20):

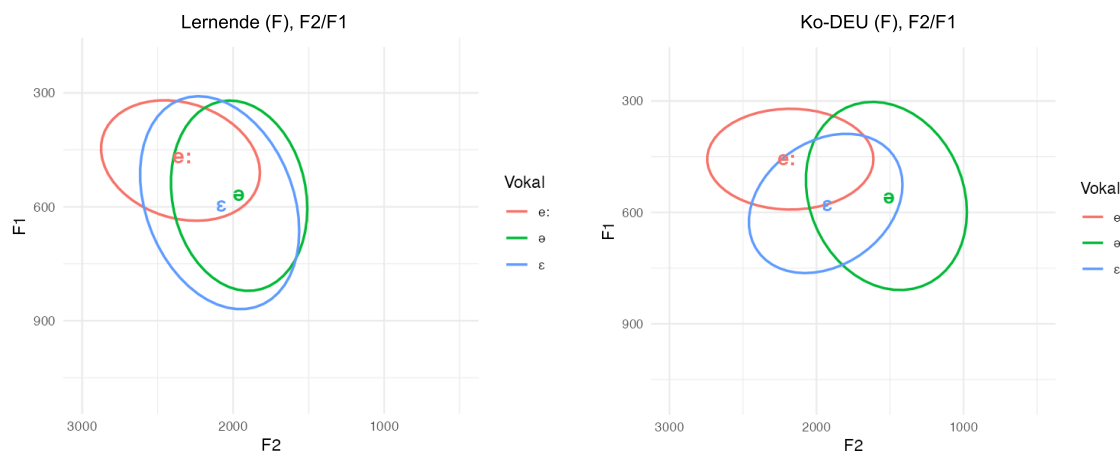


Abbildung 20: Dispersionellipsen (95%-Konfidenzniveau) und Medianwerte für $/\ə/$, $/e:/$ und $/\epsilon/$ für die Stichprobe der Frauen ("F"). Links: Lernerinnen. Rechts: Sprecherinnen mit L1 Deutsch

Auf deskriptiver Ebene fällt das unterschiedliche Verhältnis zwischen $/\ə/$ und $/\epsilon/$ bei beiden Gruppen auf: Einerseits überlappen sich die Streuungen in der Lernerinnengruppe, andererseits nur teilweise in der deutschen Kontrollgruppe. Im Allgemeinen ist bei den Lernerinnen außerdem zu beobachten, dass alle drei Kategorien höhere F2-Werte aufweisen als bei der Kontrollgruppe. Dies weist auf eher periphere Realisierungen im Vokalraum hin. Die Streuungen verlaufen somit in eine ähnliche Richtung, nämlich entlang der vorderen Bereiche des Vokalraums. Bei der deutschen Kontrollgruppe hingegen nehmen die Streuungen, insbesondere bei $/e:/$ und $/\epsilon/$, auch zentralere Positionen ein.

Um die Zielsprachlichkeit der Realisierungen bei der Lernerinnen zu überprüfen, wurden die statistische Modelle

$$\begin{aligned} \text{Imer F2} &= 1 + \text{Gruppe} + \text{Vokal} + \text{Gruppe:Vokal} + (1 \mid \text{Versuchsperson}), \\ \text{Imer F1} &= 1 + \text{Gruppe} + \text{Vokal} + \text{Gruppe:Vokal} + (1 \mid \text{Versuchsperson}) \end{aligned}$$

auf sowohl die F2- als auch die F1-Werte durchgeführt. Als abhängige Variablen wurden jeweils F1 und F2 angegeben, als Faktoren die Variable *Vokal* und *Gruppe*, während als Zufallsfaktor die Variable *Versuchsperson* angegeben wurde.

Die Analyse der F2-Werte hat signifikante ($p < .001$) Effekte für die Faktoren *Vokal*, *Gruppe* und ihre Interaktion geliefert. Der Post-Hoc Test für die Interaktion der Faktoren ergab signifikante ($p < .001$) Unterschiede zwischen den Gruppen für alle drei Kategorien. Diese Ergebnisse deuten auf eine signifikant unterschiedliche Verortung der Kategorien hinsichtlich der Anteriorität/Posteriorität der Artikulation in beiden Gruppen hin. Der Test hat außerdem bezüglich der Interaktion *Gruppe*Vokal* keinen Effekt für den Vergleich Kontrollgruppe /ɛ/ – Lernerinnen /ə/ ergeben. Dieses Ergebnis deutet darauf hin, dass sich die Kategorie /ə/ der Lernerinnen mit der Kategorie /ɛ/ der deutschen Kontrollgruppe in Bezug auf die horizontale Positionierung überschneidet.

Die Analyse der F1 Werte hat einen signifikanten ($p < .001$) Effekt für den Faktor *Vokal* und die Interaktion *Vokal*Gruppe* ergeben, aber keinen für den Faktor *Gruppe* ($p = 0,120$). Genauer betrachtet, hat der Post-Hoc Test keine signifikanten ($p = 1,000$) Unterschiede für alle drei Kategorien zwischen den Gruppen ergeben.

5.3.2.2 Realisierung von /ə/ in verschiedenen Positionen

In diesem Abschnitt werden die Ergebnisse zur Analyse der Variation der Werte des F1 und F2 in Bezug auf die Position von /ə/ im Wort präsentiert. Für die Analyse wurden die Positionen (Variable *Position*) durch folgende Stufen kodiert: *präfixal*, *Wortstamm*, *präfinal* und *final*. In der folgenden Abbildung (Abbildung 21) werden die zentrale Tendenz und die zugehörigen Streuungen für die F1- und F2-Werte in den vier Positionen bei Frauen und Männern angeführt.

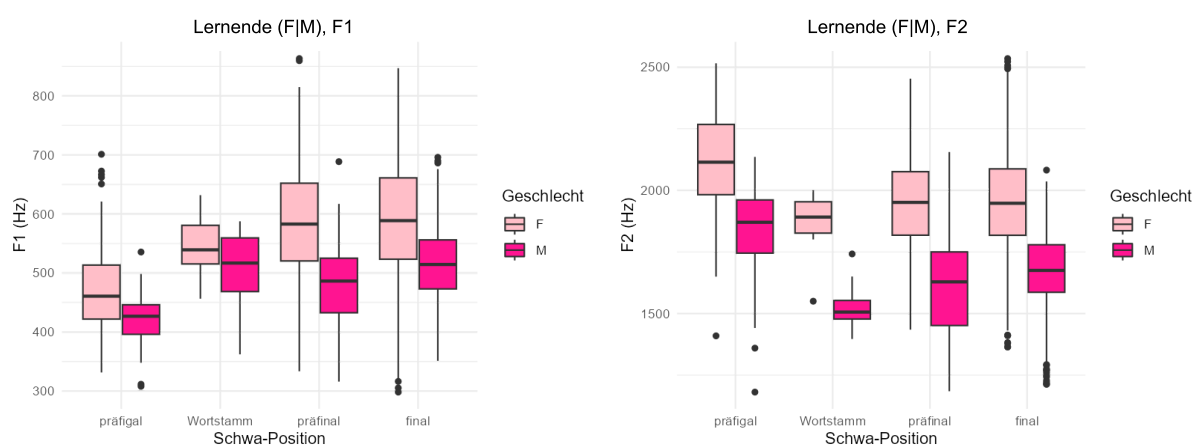


Abbildung 21: Vergleich der Datenstreuungen und Medianwerte für alle Positionen innerhalb der Frauen- ("F") und Männergruppe ("M") aus der Stichprobe der Lerner:innen. Links: Streuung der F1-Werte. Rechts: Streuung der F2-Werte.

Auf deskriptiver Ebene zeigen die Boxplots zwei Tendenzen für beide Geschlechtsgruppen bei den F1- und F2-Werten. Die Boxplots für die F1-Werte zeigen eine tendenzielle Steigerung der Werte – besonders regelmäßig bei den Frauen – je mehr sich die Position von /ə/ der finalen Position nähert, was auf einen höheren Öffnungsgrad bei der präfixalen als bei der finalen Position hindeutet. In Wortstamm-Position sind die F1-Werte bei den Männern höher als die in den anderen Positionen. Bei den Männern merkt man außerdem eine geringere Varianz bei der präfixalen Position als bei den anderen. Bei den Frauen sind sowohl die Varianz als auch die Medianwerte bei der präfinalen und finalen Position fast gleich, während die Varianz in der Wortstamm-Position am geringsten ausfällt. Bezüglich der F2-Werte zeigen die Balkendiagramme sowohl bei Frauen als auch bei Männern die höchsten Werte in präfixaler Position, die niedrigsten in Wortstamm-Position. In präfinaler und finaler Position sind sie insbesondere bei den Frauen vergleichbar und zwar einerseits höher als in Wortstamm-Position und niedriger als in präfixaler Position. Diese Ergebnisse deuten auf eine deutlich vordere Artikulation des /ə/ in präfixaler Position als in den anderen Positionen. Sowohl bei den Frauen als auch bei den Männern ist die Varianz in der Wortstammposition im Vergleich zu den anderen Positionen am geringsten.

Folgende statistische Modelle

$$lmer F1 \sim 1 + schwa_Position + (1 | Versuchsperson);$$

$$lmer F2 \sim 1 + schwa_Position + (1 | Versuchsperson)$$

wurden sowohl bei den Frauen als auch bei den Männern auf die Werte der F1 und F2 durchgeführt. Als abhängige Variable wurden *F1* und *F2* jeweils angegeben, während als Faktor die Variable *schwa_Position* und als Zufallsfaktor die Variable *Versuchsperson* angegeben.

Beginnend mit der Analyse der F2-Werte, hat sich sowohl bei den Frauen als auch bei den Männern der Effekt für den Faktor *schwa_Position* als signifikant ($p < .001$) erwiesen. Der Post-Hoc Test hat im Allgemeinen vergleichbare Ergebnisse für beide Geschlechtsgruppen geliefert. Bei beiden Gruppen haben die Vergleiche zwischen den Positionen *Wortstamm-präfixal* und *präfixal-präfinal* signifikante Effekte geliefert ($p < .001$), wobei die F2-Werte für die Wortstamm-Position signifikant niedriger sind als die für die präfixale Position und die für die präfixale Position signifikant höher sind als die für die präfinale. Diese Ergebnisse deuten auf eine zentralere Artikulation in Wortstamm-Position und auf eine eher vordere Artikulation in präfixaler Position hin. Bei den Frauen hat sich der Vergleich *Wortstamm-präfinal* als signifikant ($p = 0,047$) ergeben, in ähnlicher Weise hat sich der Vergleich *Wortstamm-final* bei den Männern als signifikant ($p = 0,010$) ergeben. In beiden Fällen resultieren die F2-Werte in Wortstamm-Position signifikant niedriger als die in präfinaler bzw. finaler Position.

Fortsetzend mit der Analyse der F1-Werte, hat sich der Effekt für den Faktor *schwa_Position* in beiden Geschlechtsgruppen als signifikant erwiesen ($p < .001$). Der Post-Hoc Test hat im Allgemeinen vergleichbare Ergebnisse für Frauen und Männer geliefert. Bei beiden Gruppen haben

die Vergleiche zwischen den Positionen *final-präfixal*, *final-präfinal* und *Wortstamm-präfixal* signifikante Effekte geliefert (jeweils $p < .001$, $p < .001$, $p = 0,015$ bei Frauen und $p < .001$ bei Männern). Die F1-Werte für die präfixale Position ergeben sich als signifikant niedriger als in präfinaler Position und die Werte für die finale und Wortstamm-Position ergeben sich als signifikant höher als die für die präfixale Position. Diese Ergebnisse deuten auf einen niedrigeren Öffnungsgrad bei der präfixalen Position als bei den obengenannten Positionen. Bei den Männern und zwar bei dem Vergleich *final-präfinal* ($p < .001$) sind die Werte für die finale Position signifikant höher als für die präfinale. Diese Ergebnisse deuten auf einen niedrigeren Öffnungsgrad in finaler als in präfinaler Position.

Nachfolgend wird die zentrale Tendenz mit zugehörigen Streuungen für die Realisierungen durch die Lernerinnen mit den von den L1-Sprecherinnen der deutschen Kontrollgruppe in Kontrast gesetzt (Abbildung 22).

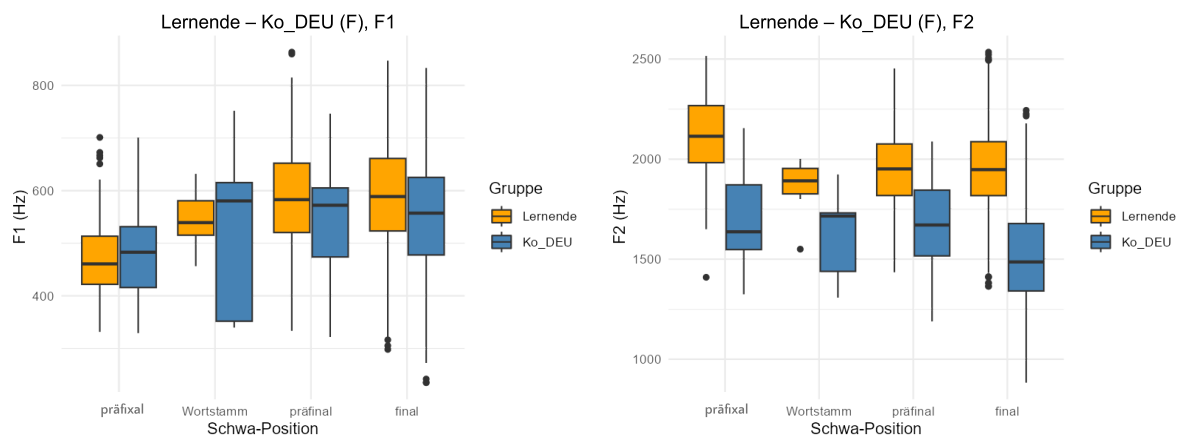


Abbildung 22: Vergleich der Datenstreuungen und Medianwerte für alle Positionen innerhalb der Frauengruppe ("F") unterteilt in Lernerinnen und Sprecherinnen mit L1 Deutsch. Links: Streuung der F1-Werte. Rechts: Streuung der F2-Werte.

Auf deskriptiver Ebene zeigt sich, dass die F1-Werte beider Gruppen über alle Positionen hinweg tendenziell ansteigen. Dabei lassen sich deutliche Unterschiede der Medianwerte in der präfixalen und finalen Position sowie in der Wortstamm-Position zwischen beiden Gruppen erkennen, während sich in der präfinalen Position nur leichte Abweichungen der Medianwerte zeigen. Bezüglich der F2-Werte sind sie bei der Gruppe der Lernerinnen generell deutlich höher als bei der Kontrollgruppe. Bei den Lernerinnen sind die Werte für die präfinale und finale Position nur leicht voneinander abweichend, während sie sich bei der Kontrollgruppe deutlich unterscheiden.

Um die Zielsprachlichkeit der Realisierungen von /ə/ in den verschiedenen Positionen zu überprüfen, wurden folgende statistische Modelle

lmer (F1 - 1 + schwa_Position + Gruppe + schwa_Position:Gruppe + (1 | VP und lmer F1 - 1 + schwa_Position + Gruppe + schwa_Position:Gruppe + (1 | Versuchsperson);
lmer (F2 - 1 + schwa_Position + Gruppe + schwa_Position:Gruppe + (1 | VP und lmer F2 - 1 + schwa_Position + Gruppe + schwa_Position:Gruppe + (1 | Versuchsperson)

jeweils auf die F1- und F2-Werte durchgeführt. Als abhängige Variablen wurden jeweils *F1* und *F2* angegeben, während als Faktoren die Variablen *schwa_Position* und *Gruppe* angegeben wurden und als Zufallsfaktor die Variable *Versuchsperson*.

Bezüglich der F2-Werte hat das Modell signifikante Effekte ($p < .001$) für beide Faktoren und ihre Interaktion geliefert. Bezugnehmend insbesondere auf die Ergebnisse des Post-Hoc Tests für die Interaktion *schwa_Position*Gruppe*, hat sich ergeben, dass die Realisierungen von /ə/ in präfixaler und finaler Position durch die Lernerinnen signifikant von denen der Kontrollgruppe abweichen ($p < .001$). Konkret ergeben sich die Realisierungen in diesen Positionen als signifikant niedriger in der Kontrollgruppe, was auf eine eher zentralere Artikulation hinweist. Für die Wortstamm-Position ist der Unterschied zwischen beiden Gruppen nicht signifikant ($p = 1,000$).

Bezüglich der F1-Werte hat das Modell einen signifikanten Effekt ($p < .001$) für den Faktor *schwa_Position* aber nicht für den Faktor *Gruppe* und ihre Interaktion geliefert (jeweils $p = 0,528$, $p = 0,098$).

5.3.2.3 CLI-Phänomene mit dem L1-phonologischen System (Italienisch) bei der Aussprache von /ə/

In diesem Abschnitt werden die Beziehungen zwischen der Aussprache der deutschen Phoneme durch die Lernerinnen und den Realisierungen der italienischen Phoneme durch die italienische Kontrollgruppe (Ko_ITA) untersucht. Im Folgenden werden die Realisierungen der deutschen Phoneme /ə/, /e:/ und /ɛ/ mit denen der italienischen Phoneme /e/ und /ɛ/ verglichen (Abbildung 23). Die Grafiken zu F2 werden zuerst aufgeführt, um einen ersten Überblick über die Positionierung im Vokalraum in Bezug auf Anteriorität/Posteriorität zu geben.

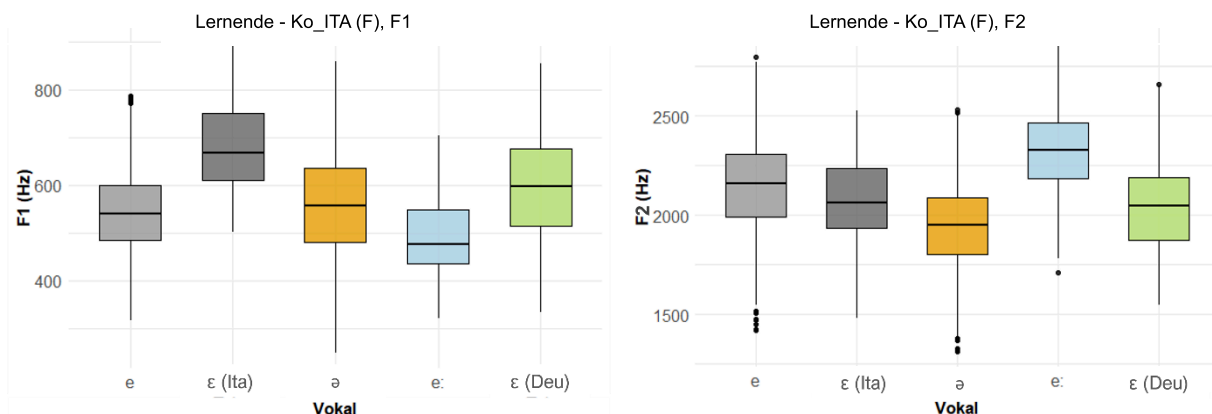


Abbildung 23: Vergleich der Datenstreuungen und Medianwerte für die italienischen (hell- und dunkelgrau) und die deutschen (orange, hellblau, grün) Vokale. Links: Streuung der F1-Werte. Rechts: Streuung der F2-Werte. Jede der zwei Grafiken folgt einer Links-nach-rechts-Logik: links die Ausgangssprache, rechts die Zielsprache. Die Daten beziehen sich auf die Frauengruppen ("F").

Auf deskriptiver Ebene überlappen sich die Realisierungen von /ə/ überwiegend mit denen von dem italienischen Vokal /ε/, was auf eine nahe Positionierung im Vokalraum hindeutet. Man kann außerdem beobachten, dass die Realisierungen von dem italienischen Phonem /ε/ mit denen des deutschen Phonems /ε/ fast völlig überlappen, was auf eine fast identische Positionierung im Vokalraum hindeutet. Trotzdem ist aus der zweiten Graphik zu entnehmen, dass sich beide Phoneme hinsichtlich der Höhe unterscheiden, indem das italienische Phonem tiefer (höhere F1-Werte) ausgesprochen wird als das deutsche. Was /ə/ angeht, wird es mit einem sehr ähnlichen Öffnungsgrad wie das italienische Phonem /e/ ausgesprochen.

Folgende statistische Modelle

$$lmer F2 \sim 1 + Vokal + (1 | Versuchsperson);$$

$$lmer F1 \sim 1 + Vokal + (1 | Versuchsperson)$$

wurde sowohl auf F2 als auch auf F1 durchgeführt. Als abhängige Variablen wurden jeweils *F2* und *F1* angegeben, als Faktor die Variable *Vokal* und als Zufallsfaktor die Variable *Versuchsperson*

Bezüglich F2, hat das Modell einen signifikanten Effekt ($p < .001$) für den Faktor *Vokal* ergeben. Der Post-Hoc Test hat tatsächlich die Abwesenheit eines signifikanten Unterschieds zwischen /ə/ und dem italienischen /ε/ bestätigt ($p = 0.352$) sowie auch zwischen dem deutschen und dem italienischen /ε/ ($p = 1,000$). Der Vergleich zwischen dem italienischen /e/ und dem deutschen /ε/ weist ebenso keinen signifikanten Unterschied ($p = 1,000$) auf.

Was die Analyse von F1 angeht, hat das Modell einen signifikanten Effekt ($p < .001$) für den Faktor *Vokal* ergeben. Der Post-Hoc Test hat tatsächlich die Abwesenheit einen signifikanten Unterschied

($p=1.000$) zwischen /ə/ und dem italienischen /e/ bestätigt. Außerdem hat das Modell keinen signifikanten Unterschied für den Vergleich /e/ und dem deutschen /ɛ/ geliefert ($p=0.081$).

5.3.2.4 Der Einfluss der Sprachkompetenz und -dominanz sowie des Sprachrepertoires auf die Realisierung von /ə/

In diesem Abschnitt wird die Realisierung von /ə/ in Abhängigkeit vom Grad der Sprachkompetenz im Deutschen (GER-Niveau) sowie von der sprachlichen Dominanz im Deutschen und Englischen innerhalb der Lerner:innengruppe untersucht. Zu diesem Zweck wird die Variation der F2-Werte in Bezug auf diese Variablen analysiert, um zu überprüfen, ob diese aussagekräftigen Hinweise auf eine kompetenzdynamische Variation der Produktion liefern. Für die vorliegende Analyse wurde der Fokus verstärkt auf den zweiten Formanten als Index für die Anteriorität/Posteriorität der Realisierung des /ə/ gelegt. Da die Zentralität des Phonems Schwa für die Lerner:innen die größte Schwierigkeit darstellt, wurde davon ausgegangen, dass F2 mehr noch als F1 einen unmittelbaren Beweis für einen möglichen positiven Einfluss der Breite des Repertoires auf seine Realisierung liefern könnte.

Folgendes statistisches Modell wurde auf die Frauen und auf die Männergruppe der Stichprobe der Lerner:innen getrennt angewandt.

$$\text{Imer } F2 - 1 + \text{GER} + \text{Dom_DEU} + \text{Dom_Eng} + \text{Sprachrepertoire} (1 | \text{Versuchsperson})$$

Die Variable *F2* wurde als abhängige Variable festgelegt, während die Variablen *GER*, *Dom_DEU*, *Dom_Eng* und *Sprachrepertoire* als Faktoren und die Variable *Versuchsperson* als Zufallsfaktor festgelegt wurden.

Sowohl bei den Frauen als auch bei den Männern resultierte kein signifikanter Effekt aller Variablen (Frauengruppe: *GER* $p=0,197$, *Dom_DEU* $p=0,875$, *Dom_Eng* $p=0,081$, *Sprachrepertoire* $p=0,334$; Männergruppe: *GER* $p=0,308$, *Dom_DEU* $p=0,643$, *Dom_Eng* $p=0,538$, *Sprachrepertoire* $p=0,378$).

Darüber hinaus werden die Ergebnisse der explorativen Untersuchung zu den beiden Probandinnen (DR_PROD, OP_PROD) berichtet, die über dialektale Kompetenzen in regionalen Varietäten des Italienischen aus dem südlichen Dialektgebietes verfügen. Die Medianwerte und Standardabweichungen wurden mit denen der übrigen Teilnehmerinnen verglichen. Die Ergebnisse dieser Analyse sind in Tabelle 8 aufgeführt.

Tabelle 8: Anzahl an Beobachtungen, Medianwerte (F2) und IQR für beide Gruppen.
DR_PROD (Frau) und OP_PROD (Frau) beziehen sich auf die zwei untersuchten Probandinnen

Formant	Anzahl an Beobachtungen	Median	IQR
<i>DR_PROD + OP_PROD</i>			
F2	116	1956	262
<i>L1 Italienisch ohne rel. Dialektkenntnisse</i>			
F2	1430	1964	278

Auf der deskriptiven Analyse resultiert, dass beide Teilnehmerinnen niedrigere F2-Werte aufweisen als die Kontrollgruppe. Dies spricht für eine zentralere Realisierung von /ə/ in erster Gruppe als bei der italienischen Kontrollgruppe ohne relevante Dialektkenntnisse.

5.3.2.5 Der Taskbedingte Einfluss auf die Realisierung von /ə/

In diesem Abschnitt wird auf den Einfluss der Art der Experimentaufgabe auf die Realisierung von /ə/ hinsichtlich der Variation der F1- und F2-Werte eingegangen. Die Analyse bezieht sich auf die experimentelle Gruppe der Lerner:innen. In der folgenden Abbildung (Abbildung 24) wird die Variation der F1- und F2-Werte für Frauen und Männer dargestellt.

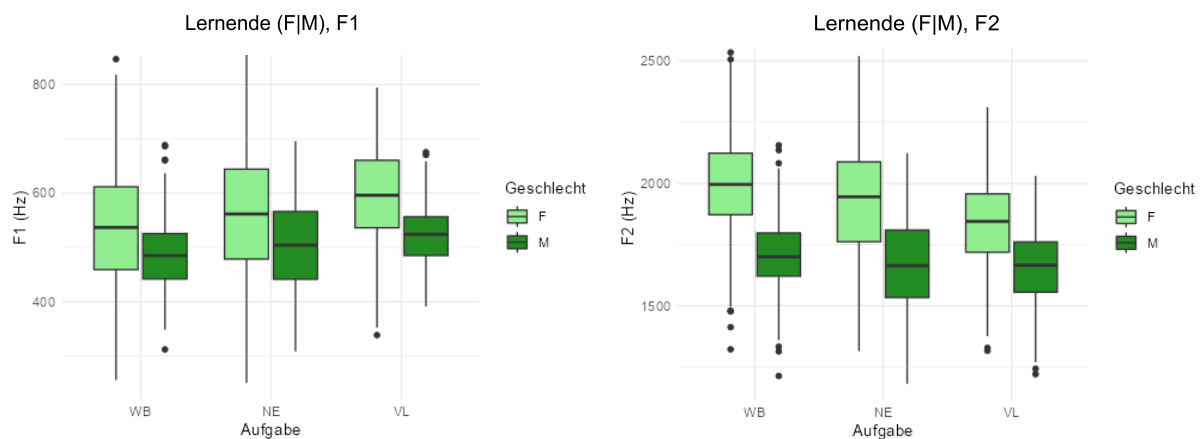


Abbildung 24: Links: Datenstreuung und Medianwerte der F1-Werte für Frauen ("F") und Männer ("M"). Rechts: Datenstreuung und Medianwerte der F2-Werte für Frauen ("F") und Männer ("M"). Die Aufgaben werden durch Abkürzungen auf Deutsch angegeben: "WB" für die Aufgabe "Wegbeschreibung", "NE" für die Aufgabe "Nacherzählung" und "VL" für die Aufgabe "Vorlesen".

Auf deskriptiver Ebene zeigt die erste Graphik einen Anstieg der Werte sowohl bei den Frauen als auch bei den Männern, je strenger die Aussprachebedingungen werden. Im Gegensatz dazu zeigt

die zweite Grafik eine umgekehrte Tendenz der F2-Werte, nämlich eine Abnahme. Diese Tendenz ist vor allem bei den Frauen zu beobachten.

Folgende statistische Modelle

$$lmer F2 \sim 1 + Aufgabe + (1 | Versuchsperson);$$

$$lmer F1 \sim 1 + Aufgabe + (1 | Versuchsperson)$$

wurden sowohl auf die Frauen als auch auf die Männer angewendet. Als abhängige Variablen wurden *F2* und *F1* angegeben, als Faktor die Variable *Aufgabe* (Stufen: Wegbeschreibung (WB), *Nacherzählung* (NE), *Vorlesen* (VL)) und als Zufallsfaktor die Variable *Versuchsperson*.

Was die F2-Werte angeht, hat das Modell sowohl bei den Frauen als auch bei den Männern einen signifikanten Effekt für den Faktor *Aufgabe* geliefert (jeweils $p < .001$, $p = 0.012$). Der Post-Hoc Test hat jedoch einen signifikanten Effekt für alle Aufgabenvergleiche (NE-VL, NE-WB, VL-WB) nur für die Gruppe der Frauen geliefert ($p < .001$), indem die F2-Werte bei NE höher sind als bei VL und beide niedriger als bei WB (WB > NE > VL). Diese Ergebnisse deuten darauf hin, dass die Artikulation im Vokalraum immer zentraler wird, je mehr die Freiheit des Sprechens durch die experimentellen Bedingungen eingeschränkt wird. Bei den Männern erweist sich nur der letzte Vergleich als signifikant auf ($p = 0.010$), indem die F2-Werte bei VL signifikant niedriger als die bei WB sind.

Bezüglich der Analyse der F1, hat sich der Effekt bei beiden Gruppen als signifikant ($p < .001$) ergeben. Bei den Frauen hat das Modell signifikante Effekte ($p < .001$) für alle Aufgabenvergleiche geliefert (NE-VL, NE-WB, VL-WB), indem die F1 Werte bei VL höher sind als NE und bei letzterer Aufgabe höher als bei WB (VL > NE > WB). Diese Ergebnisse deuten darauf hin, dass die Artikulation im Vokalraum bei VL und NE tiefer ist als bei WB. Bei den Männern ist die Signifikanz für die ersten zwei Vergleiche leicht niedriger als bei den Frauen (NE-VL: $p = 0.005$, NE-WB: $p = 0.037$, VL-WB: $p < .001$), trotzdem sehen die Verhältnisse gleich aus

5.3.3 Ergebnisse zur Realisierung von /y:/

In diesem Abschnitt werden die Ergebnisse in Bezug auf die Realisierung von /y:/ präsentiert. Es wird die potenzielle Unschärfe der Trennung zwischen den deutschen Vokalkategorien /y:/-/ʏ/-/i:/-/ɪ/-/u:/-/ʊ/ und der Einfluss des rechten und linken phonologischen Kontextes auf die Realisierung des Phonems /y:/ berücksichtigt. In diesem Zusammenhang werden die Ergebnisse für die experimentelle Gruppe mit denen für die Kontrollgruppe mit deutscher L1 verglichen, um die Zielsprachlichkeit der Realisierungen zu überprüfen. Sie werden außerdem mit den Realisierungen der italienischen Phoneme /i/ und /u/ verglichen, um eventuelle Kategorienüberlappungen und somit CLI-Phänomene zu identifizieren. Was die Angaben zu Sprachkompetenz und -dominanz betrifft, wird auf die mögliche Aussagekraft der Variablen Kompetenz im Deutschen gemessen an GER-Niveau und der Sprachdominanz eingegangen.

5.3.3.1 Realisierung von /y:/in Kontrast zu /ʏ/, /i:/, /ɪ/, /u:/ und /ʊ/

In diesem Abschnitt werden die Ergebnisse zur Realisierung von /y:/ durch die Lerner:innen im Vergleich zu den Realisierungen der benachbarten Phoneme /ʏ/, /i:/, /ɪ/, /u:/ und /ʊ/ präsentiert. Für die Analyse der Kategorientrennung sowohl innerhalb der experimentellen Gruppe als auch zwischen den Lernerinnen und den zwei Kontrollgruppen wurden die F2- und F3-Werte in Betracht gezogen. Die Werte des dritten Formanten werden in der Regel für die Lippenrundung als aussagekräftig gehalten. Für die Analyse des Einflusses des rechten und linken phonologischen Kontextes auf die Realisierung von /y:/ wurden die F3-Werte in Betracht gezogen. Der Fokus der Untersuchung liegt auf der Analyse der F3-Werte, da die Hypothese darauf aufbaut, den Einfluss des umgebenden konsonantischen Kontextes auf die Rundung des Phonems /y:/ zu bestimmen und damit potenzielle positive Effekte vorderer konsonantischer Artikulationspositionen zu verifizieren. Für die Analyse des Einflusses der Faktoren der Sprachkompetenz, -dominanz und des Sprachrepertoires auf die Realisierung von /y:/ wurden die F2- sowie F3-Werte analysiert. Im zuletzt genannten Fall konzentrierte sich die Analyse auf die Fähigkeit der Lernenden, /y:/ in vorderen Positionen zu artikulieren und das Merkmal der Lippenrundung adäquat umzusetzen. Dadurch sollte überprüft werden, ob eine höhere sprachliche Kompetenz mit einer vorderen bzw. gerundeten Produktion einhergeht, die näher an der standardsprachlichen Zielvorgabe liegt.

In der folgenden Abbildung (Abbildung 25 und 26) werden auf F3/F2 Graphiken die Dispersionsellipsen (95% Konfidenzniveau) für alle in diesem Abschnitt untersuchten Vokale für die Lerner:innen aufgeführt:

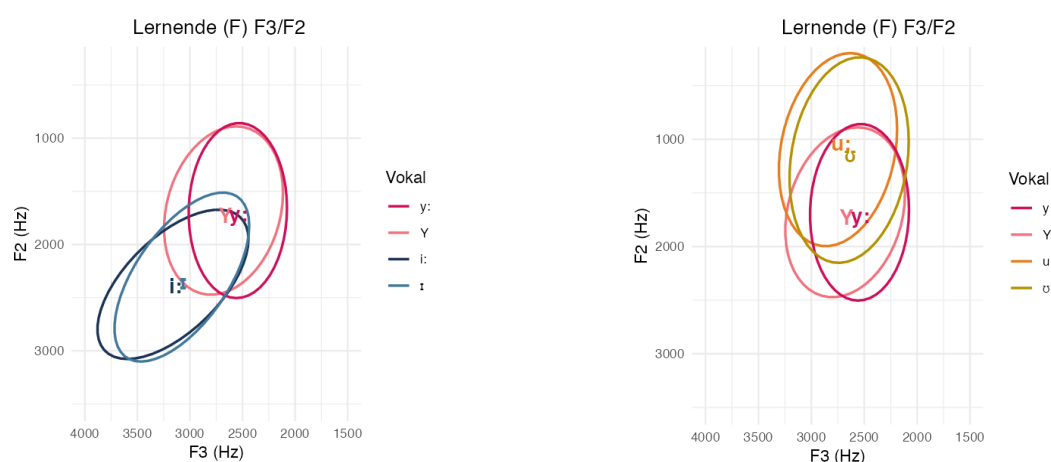


Abbildung 25: Dispersionsellipsen (95% Konfidenzniveau) mit Medianwerten zu F3/F2 für /y:/, /ʏ/, /i:/, /ɪ/ und /y:/, /ʏ/, /u:/, /ʊ/ für die Gruppe der Lernerinnen ("F"). Links: /y:/, /ʏ/, /i:/, /ɪ/. Rechts: /y:/, /ʏ/, /u:/, /ʊ/.

Was die Kontraste /y:/- /ʏ/ und /i:/- /ɪ/ in der Frauengruppe angeht, lässt sich im Streudiagramm beobachten, dass die Ellipsen der ungerundeten Vokale /i:/ und /ɪ/ im

Vergleich zu den gerundeten Vokalen /y:/ und /ʏ/ tendenziell zu höheren F3-Werten hin verschoben sind. Die Ellipsen der gerundeten Vokale zeigen eine ausgeprägte vertikale Ausdehnung entlang der F2-Achse sowie eine deutliche gegenseitige Überlappung. Diese Überlagerung deutet auf eine hohe Variabilität in der akustischen Positionierung dieser Phoneme sowie auf eine geringe Unterscheidung hinsichtlich des Merkmals der Gespanntheit innerhalb dieser Gruppe hin.

Bezüglich der Kontraste /y:/- /ʏ/ und /u:/- /ʊ/ lässt sich im Streudiagramm eine ausgeprägte vertikale Anordnung der Ellipsen erkennen, wobei die hinteren Vokale /u:/ und /ʊ/ im Vergleich zu den vorderen Gegenstücken /y:/ und /ʏ/ tendenziell niedrigere F2-Werte aufweisen. Es ist jedoch eine erhebliche Überlappung aller vier Kategorien festzustellen, insbesondere bei den Phonemen, die sich durch das Merkmal der Gespanntheit differenzieren lassen (/y:/- /ʏ/ ; /u:/- /ʊ/).

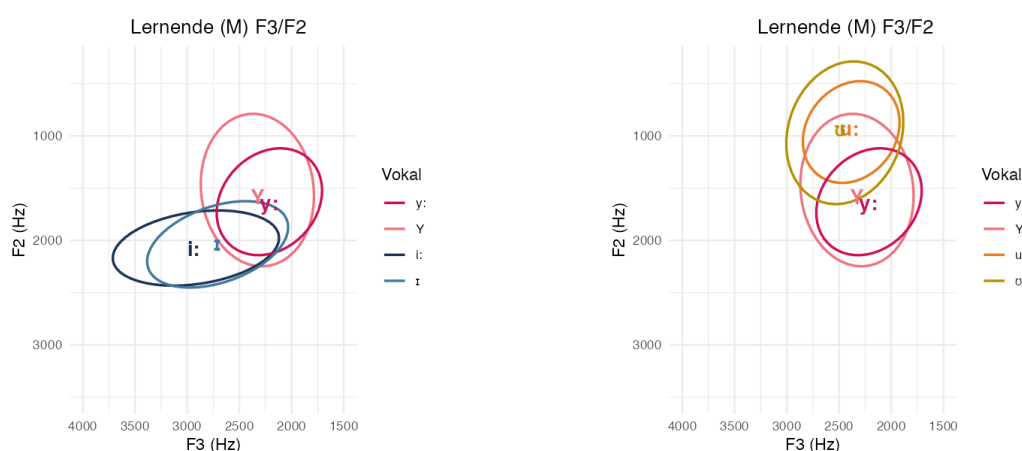


Abbildung 26: Dispersionsellipsen (95% Konfidenzniveau) mit Medianwerten zu F3/F2 für /y:/, /ʏ/, /i:/, /ɪ/ und /y:/, /ʏ/, /u:/, /ʊ/ für die Gruppe der Lerner ("M"). Links: /y:/, /ʏ/, /i:/, /ɪ/. Rechts: /y:/, /ʏ/, /u:/, /ʊ/.

Im Vergleich zu den zuvor analysierten Daten der Lernerinnen sind die Ellipsen der männlichen Probanden (Abbildung 26) in einem insgesamt niedrigeren Frequenzbereich verortet, was den physiologischen geschlechtsspezifischen Differenzen in der Stimmtraktlänge entspricht. Innerhalb dieses akustischen Raums besetzen die ungerundeten Vokale /i:/, /ɪ/ und die gerundeten Vokale /y:/, /ʏ/ differenzierbare Bereiche entlang der F3-Dimension, wobei die gerundeten Kategorien erwartungsgemäß zu niedrigeren F3-Werten tendieren. Auffällig ist die beträchtliche vertikale Ausdehnung der Ellipse für /ʏ/ auf der F2-Achse, welche die Fläche von /y:/ nahezu vollständig umschließt; dies deutet auf eine unzureichende phonetische Distinktion hinsichtlich des Merkmals [gespannt] hin. Eine analoge Tendenz zur Überlappung zeigt sich bei dem Vokalpaar /i:/- /ɪ/. Wie bereits bei der Gruppe der Lernerinnen beobachtet, ist somit eine ausgeprägte interkategoriale Interferenz festzustellen.

Bezüglich der Kontraste /y:/-/ʏ/ und /u:/-/ʊ/ wird in den Dispersionsellipsen (F3/F2) deutlich, dass die männlichen Probanden eine geringere Variabilität aufweisen als die Lernerinnen: Während die Ellipsen der Frauen eine starke vertikale Ausdehnung entlang der F2-Achse zeigen, sind die der Männer deutlich konzentrierter. Physiologisch bedingt liegt der gesamte Vokalraum der Männer in einem tieferen Frequenzbereich, zeigt jedoch eine präzisere kategoriale Trennung des vorderen gerundeten Vokala /y:/ von den anderen Kategorien. Auch in diesem Fall ist eine mangelhafte Trennung der Kategorien hinsichtlich des Merkmals [gespannt] festzustellen, was bei den Phonemen /u:/ und /ʊ/ besonders auffällig ist. ppe gruppiert sich sehr eng um den Median, was auf eine eher homogene Realisierung hindeutet.

Folgende statistische Modelle

$$\begin{aligned} \text{Imer } F2 &= 1 + \text{Vokal} + (1 | \text{Versuchsperson}), \\ \text{Imer } F3 &= 1 + \text{Vokal} + (1 | \text{Versuchsperson}) \end{aligned}$$

wurde jeweils sowohl auf die Frauen als auch auf die Männer angewendet. Als abhängige Variable wurden jeweils *F2* und *F3* angegeben, als Faktor die Variable *Vokal* und als Zufallsfaktor die Variable *Versuchsperson*.

Bezüglich der Analyse der *F2* angeht, hat das Modell sowohl für die Frauen als auch für die Männer einen signifikanten Effekt ($p < .001$) für den Faktor *Vokal* geliefert. Der Post-Hoc Text hat für die Vergleiche /ʏ/-/y:/ und /u:/-/ʊ/ keinen signifikanten Effekt für die Frauengruppe geliefert (jeweils $p = 1,000$, $p = 0,594$), während dies bei der Männergruppe nur für den Vergleich /u:/-/ʊ/ gilt ($p = 1,000$). Bei allen anderen Vergleichen sind die Effekte für beide Gruppen signifikant (Frauen: $p < .0001$ bei allen Vergleichen, Männer: $p < .001$ bei allen außer /ʏ/-/y:/, hier $p = 0,006$).

Die Analyse von *F3* ergab für die Variable *Vokal* einen signifikanten Effekt ($p < .001$) sowohl für die männliche als auch für die weibliche Gruppe. Der Post-Hoc-Test zeigt, dass beim Vergleich /ʏ/-/y:/ der Vokal /ʏ/ signifikant ($p < .001$) höhere Werte als /y:/ aufweist. Dies deutet auf einen geringeren Grad an Lippenrundung bei /ʏ/ im Vergleich zu /y:/ hin. Des Weiteren zeigt sich beim Vergleich /i:/-/y:/ eine signifikante Diskrepanz ($p < .001$) zugunsten von /i:/. Eine ähnliche Tendenz lässt sich auch bei /u:/-/y:/ beobachten, wobei /u:/ signifikant ($p < .001$) höhere Werte aufweist. Die Ergebnisse weisen darauf hin, dass der Grad der Lippenrundung bei /i:/ und /u:/ geringer ist als bei /y:/. Zudem weist /y:/ beim Vergleich /y:/-/ʊ/ signifikant ($p < .001$) niedrigere Werte als /ʊ/, was auf einen höheren Grad an Lippenrundung hindeutet. Es lässt sich feststellen, dass der Vokal /y:/ im Allgemeinen gerundeter ist, sowohl als vordere (/i:/, /ʏ/) als auch als hintere Vokale (/u:/, /ʊ/).

Nachfolgend werden die Realisierungen der Lernerinnen denen der Frauen mit L1-Deutsch gegenübergestellt (Abbildung 27 und 28).

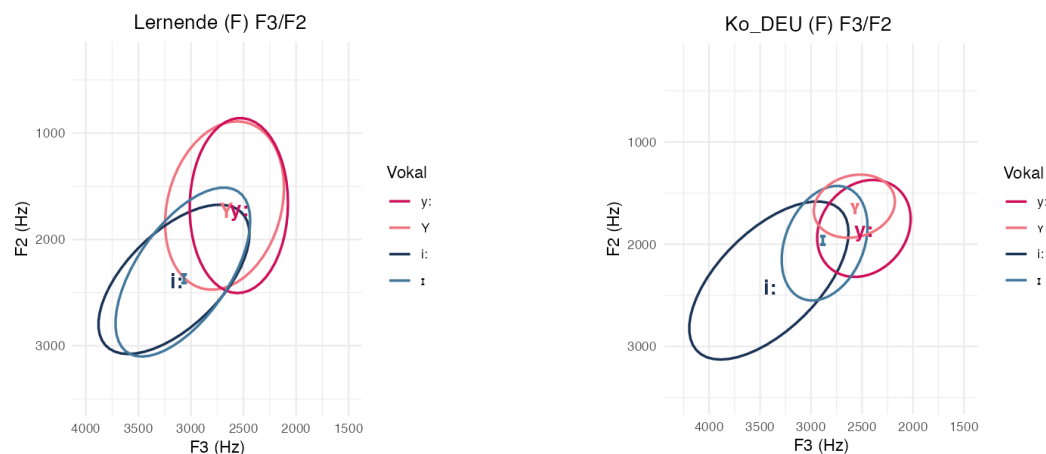


Abbildung 27: Dispersionsellipsen (95% Konfidenzniveau) mit Medianwerten zu F3/F2 für /y:/, /Y/, /i:/, /ɪ/ für die Frauengruppen ("F"). Links: Lernerinnen. Rechts: Sprecherinnen mit L1 Deutsch.

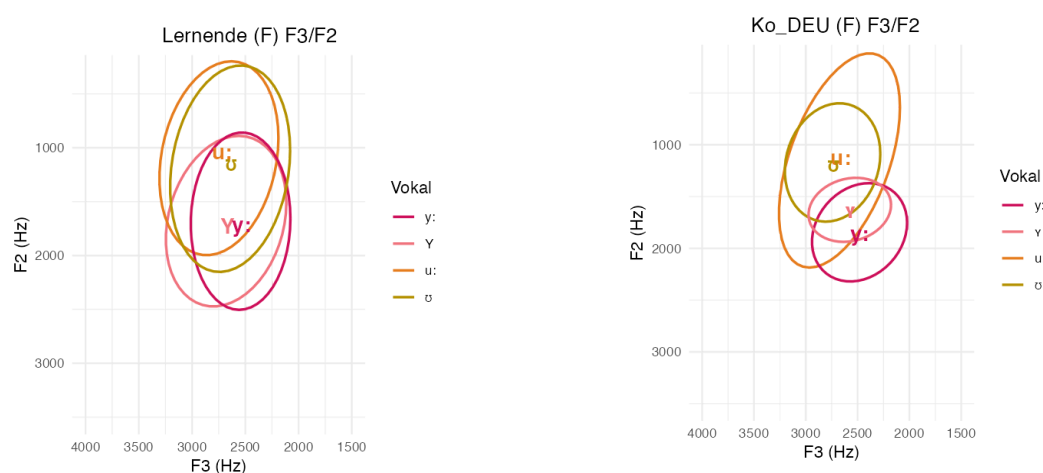


Abbildung 28: Dispersionsellipsen (95% Konfidenzniveau) mit Medianwerten zu F3/F2 für /y:/, /Y/, /u:/, /ʊ/ für die Frauengruppen ("F"). Links: Lernerinnen. Rechts: Sprecherinnen mit L1 Deutsch.

Der Vergleich der graphischen Grafiken des F3/F2-Vokalraums (Abbildung 27) offenbart deutliche Unterschiede in der phonetischen Realisierung und kategorialen Trennung zwischen den Lernerinnen und der weiblichen Kontrollgruppe mit L1-Deutsch. Während in beiden Gruppen die ungerundeten Vokale /i:/ und /ɪ/ eine diagonale Ausrichtung mit tendenziell höheren F3-Werten aufweisen, zeigt sich bei den gerundeten Vokalen /y:/ und /Y/ die erwartete akustische Absenkung der dritten Formante. Ein wesentlicher Kontrast liegt in der Variabilität und Konsistenz der Vokalproduktion. Die Ellipsen der Kontrollgruppe weisen eine deutlich geringere vertikale

Ausdehnung auf der F2-Achse auf als die der Lernerinnen, was auf eine stabilere und präzisere Artikulation der Zielvokale durch die L1-Sprecherinnen hindeutet. Im Gegensatz dazu lassen die Lernerdaten eine ausgeprägte interindividuelle Variabilität erkennen, die sich besonders in der weiten Streuung der gerundeten Qualitäten niederschlägt. Obgleich auch bei der Kontrollgruppe eine partielle Überlappung der Streubereiche – insbesondere der Phonemen /i:/, /ɪ/ und /i:/, /y:/, /ʏ/ – zu beobachten ist, tritt die Überlappung zwischen gespannten und ungespannten Vokalen bei den Lernerinnen sehr markanter vor.

Im Streuungsdiagramm zur Realisierung der Kontraste /y:/- /ʏ/ und /u:/- /ʊ/ bei der Gruppe der Lernerinnen (Abbildung 28) zeigt sich eine ausgeprägte vertikale Anordnung der Ellipsen entlang der F2-Achse. Während die hinteren Vokale /u:/ und /ʊ/ erwartungsgemäß niedrigere F2-Werte aufweisen als die vorderen Gegenstücke /y:/ und /ʏ/, ist eine erhebliche Überlappung zwischen allen vier Kategorien festzustellen. Dies deutet sowohl auf eine mangelhafte Differenzierung hinsichtlich der Gespanntheit als auch auf eine hohe Variabilität der Realisierung vorderer Vokale hin, die sich teilweise bis in den artikulatorischen Bereich der hinteren Vokale erstreckt. Hingegen ist solche interkategoriale Überlagerung (vorne-hinten) bei der Kontrollgruppe nur partiell.

Um die Zielsprachlichkeit der Realisierungen zu prüfen, wurden folgende statistische Modelle

$$\begin{aligned} \text{Imer F2} &= 1 + \text{Gruppe} + \text{Vokal} + \text{Gruppe:Vokal} + (1 | \text{Versuchsperson}), \\ \text{Imer F3} &= 1 + \text{Gruppe} + \text{Vokal} + \text{Gruppe:Vokal} + (1 | \text{Versuchsperson}) \end{aligned}$$

auf beide Gruppen angewendet. Als abhängige Variable wurden jeweils F2 und F3 angegeben, während als Faktor die Variablen *Vokal* und *Gruppe* und als Zufallsfaktor die Variable *Versuchsperson*.

Was die Analyse der F2 angeht, hat das Modell signifikante Effekte ($p < .001$) für die Variable *Vokal* und die Interaktion *Vokal*Gruppe* geliefert, während für die Variable *Gruppe* keinen signifikanten Effekt ($p = 0,450$). Der Post-Hoc Test für die Interaktion hat ergeben, dass es zwischen den Gruppen keinen signifikanten Unterschied ($p = 1,000$) bei den Phonemen /ʏ/, /i:/, /u:/ und /ʊ/ gibt, während dies bei /y:/ und /ɪ/ der Fall ist (jeweils $p = 0,007$, $p < .001$). Die F2 Werte des Phonems /y:/ erweisen sich als höher bei der Kontrollgruppe mit L1-Deutsch. Dieses Ergebnis deutet auf eine vordere Artikulation dieses Phonems bei der Kontrollgruppe als bei den Lernerinnen.

Bezüglich der Analyse der F3, hat das Modell signifikante Effekte ($p < .001$) für die Variable *Vokal* und die Interaktion *Vokale*Gruppe* ergeben, während keinen signifikanten Effekt ($p = 0,382$) für die Variable *Gruppe*. Der Post-Hoc Test für die Interaktion hat ergeben, dass die Phoneme /ʏ/, /u:/, /y:/ und /ʊ/ nicht signifikant unterschiedlich zwischen den beiden Gruppen sind, während es der Fall bei den Phonemen /ɪ/ ($p = 0,002$) und /i:/ ($p = 0,022$) ist, wobei die F3-Werte für /ɪ/ bei der weiblichen Kontrollgruppe signifikant niedriger sind als bei den Lernerinnen und die F3-Werte für /i:/ bei der Kontrollgruppe signifikant höher sind als bei den Lernerinnen. Diese Ergebnisse deuten

darauf hin, dass die Phoneme /ɪ/ und /i:/ der weiblichen Kontrollgruppe mit jeweils einem höheren und geringeren Grad an Lippenrundung ausgesprochen werden.

5.3.3.2 Realisierung von /y:/ in Abhängigkeit vom linken und rechten phonologischen Kontext

Um die Variation der Realisierung von /y:/ hinsichtlich des linken und rechten Erscheinungskontext wird die Analyse auf die F2- und F3-Werte durchgeführt. Für die Analyse wurden die (linken und rechten) phonologischen Kontexte wie folgt kodiert: *bilabial*, *labiodental*, *alveolar*, *uvular*, *velar*, *glottal*. Für die Analyse des linken Kontextes werden folgende Kontexte in Betracht gezogen, die von Frauen und Männern der experimentellen Gruppe geteilt sind: *labiodental*, *alveolar*, *velar*, *uvular*, *glottal*. Für die Analyse des rechten Kontextes werden aus denselben Gründen folgende Kontexte in Betracht gezogen: *bilabial*, *alveolar*, *velar*, *uvular*.

Es wird auf die Variation von den F2- und F3-Werten bei der Realisierung von /y:/ hinsichtlich des linken Kontextes eingegangen. Nachfolgend wird die zentrale Tendenz mit zugehörigen Streuungen für die Realisierungen von /y:/ abhängig von den verschiedenen linken Kontexten für die Lerner:innen präsentiert (Abbildung 29).

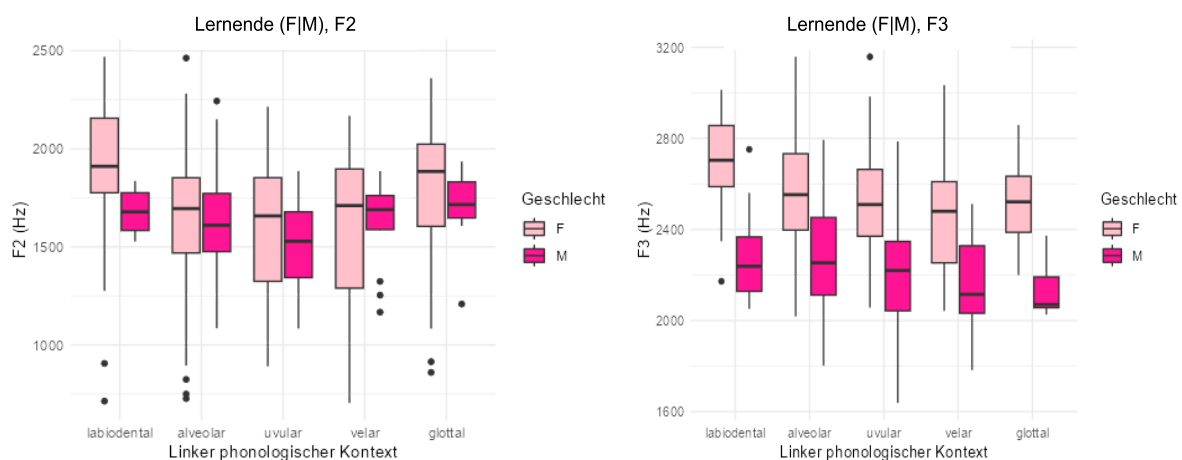


Abbildung 29: Datenstreuung und Medianwerte der F2- und F3-Werte über alle linken phonologischen Kontexte innerhalb der Stichprobe der Lerner:innen. Links: Streuung der F2-Werte für Frauen ("F") und Männer ("M"). Rechts: Streuung der F3-Werte für Frauen ("F") und Männer ("M").

Auf deskriptiver Ebene fällt auf, dass die geschlechtsspezifischen Unterschiede im Median des dritten Formanten ausgeprägter sind als die Unterschiede im zweiten Formanten.

Bezüglich der F2-Werte lassen sich für beide Geschlechter höhere Werte im labiodentalen und glottalen Kontext beobachten. Tendenziell sinken die F2-Werte, je näher der Artikulationsort an die Uvula rückt – mit einem anschließenden Anstieg, insbesondere bei den Frauen. Im velaren Kontext überschneiden sich die Medianwerte von Frauen und Männern. Bei den Frauen zeigen sich im

alveolaren, uvularen und velaren Kontext nur geringe Unterschiede zwischen den Medianwerten, während bei den Männern deutlichere Abweichungen erkennbar sind – insbesondere zwischen dem alveolaren und uvularen sowie dem uvularen und velaren Kontext. Diese Tendenzen deuten darauf hin, dass bei beiden Geschlechtsgruppen im labiodentalen und glottalen Kontext eher vordere Artikulationsorte realisiert werden, während im uvularen Kontext besonders hintere Artikulationen des /y:/ zu beobachten sind.

Bezüglich der F3-Werte kann beobachtet werden, dass sie bei den Frauen auffällig niedriger werden, je hinterer der linke phonologische Kontext im Vokalraum positioniert ist. Das gilt insbesondere für die ersten vier Kontexte. Diese Tendenz ist bei den Männern weniger auffällig. Insbesondere bei den Frauen deutet diese auf einen geringeren Grad an Lippenrundung bei bei vorderen als bei hinteren linken Artikulationsorten hin.

Insgesamt merkt man bei beiden Geschlechtsgruppen einen höheren F2- und F3-Medianwert beim labiodentalen linken Kontext. Dies deutet auf eine vordere und weniger gerundete Artikulation von /y:/ beim labiodentalen linken Kontext hin. Beim glottalen linken Kontext merkt man bei beiden Geschlechtsgruppen einen höheren F2-Medianwert und einen niedrigeren F3-Medianwert – Letzteres ist bei der Männergruppe besonders auffällig. Dies deutet auf eine vordere aber gleichzeitig mehr gerundete Artikulation von /y:/ beim glottalen linken Kontext.

Um zu überprüfen, inwieweit der linke Kontext signifikant bei der Artikulation des Phonems /y:/ ist, wurden die folgenden statistische Modelle

$$lmer F2 \sim 1 + y_links_Kontext + (1 | Versuchsperson)$$

$$lmer F3 \sim 1 + y_links_Kontext + (1 | Versuchsperson)$$

jeweils auf die Frauen- und die Männergruppe hinsichtlich der F2- und F3-Variation durchgeführt. Als abhängige Variable wurde jeweils F2 oder F3, als Faktor die Variable *y_links_Kontext* und als Zufallsfaktor die Variable *Versuchsperson* festgelegt.

Das Modell für die F2-Werte ergab einen signifikanten Effekt sowohl für die Frauen ($p = 0.008$) als auch für die Männer ($p = 0.049$), wobei der Effekt in der Männergruppe nur grenzwertig signifikant war und bei keinem der Paarvergleiche signifikante Unterschiede festgestellt werden konnten. Bei den Frauen hat der Vergleich *labiodental-alveolar* einen signifikanten Effekt erwiesen ($p=0.004$), der auf eine weiter vorne gelegene Artikulation im labiodentalen Kontext hinweist.

Das Modell für die F3-Werte hat einen signifikanten Effekt für die Frauen aber nicht für die Männer ergeben (jeweils $p=0,001$, $p=0,134$), wobei der Vergleich *labiodental-alveolar* einen signifikanten Effekt ($p=0,001$) bei den Frauen aufweist. Konkret zeigen sich höhere F3-Werte im labiodentalen Kontext im Vergleich zum alveolaren Kontext, was als einen geringeren Grad an Lippenrundung im labiodentalen Kontext zu interpretieren ist.

Die Verhältnisse zwischen den Realisierungen von /y:/ in Abhängigkeit vom linken Kontext und daher ihre Positionierung im Vokalraum werden im Vergleich zu denen der Sprecherinnen mit L1 Deutsch analysiert. Nachfolgend wird die zentrale Tendenz mit zugehörigen Streuungen für die Realisierungen von /y:/ abhängig vom linken Kontext für beide Gruppen aufgeführt (Abbildung 30):

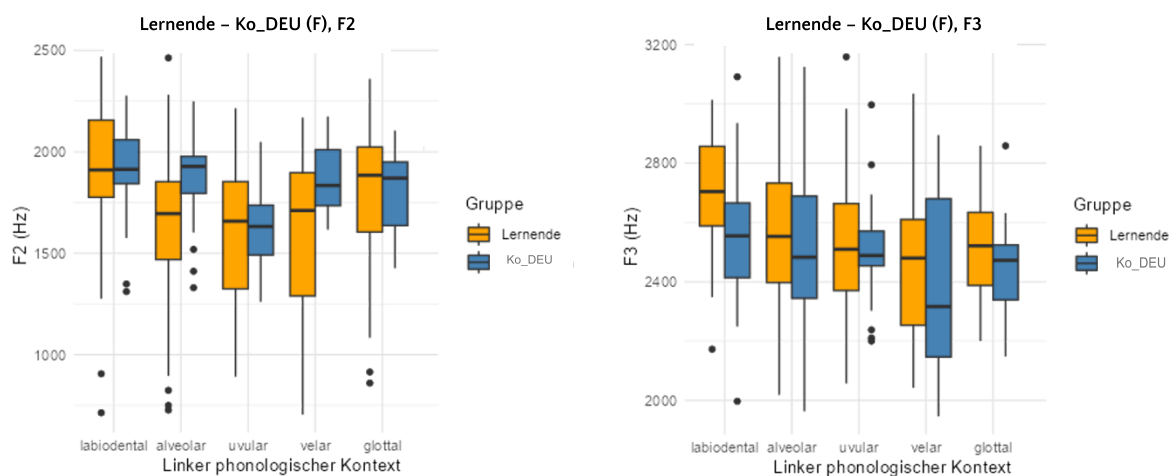


Abbildung 30: Vergleich der Datenstreuung und der Medianwerte (F2, F3) bei der Realisierung von /y:/ in den geteilten linken Kontexten. Die Daten beziehen sich auf die Lernerinnen ("F") und die Sprecherinnen ("F") mit L1 Deutsch. Links: Datenstreuung und Medianwerte der F2-Werte. Rechts: Datenstreuung und Medianwerte der F3-Werte

Bezüglich der F2-Werte lässt sich feststellen, dass beide Gruppen einer ähnlichen Tendenz folgen, wobei die Unterschiede der Medianwerte über alle Positionen hinweg bei der Lernergruppe stärker ausgeprägt sind als bei der Kontrollgruppe mit L1-Deutsch. Letztere zeigt zum Beispiel keine auffälligen Unterschiede der Medianwerte zwischen dem labiodentalen und dem alveolaren sowie zwischen dem velaren und dem glottalen linken Kontext. Insgesamt lassen sich niedrigere F2-Werte bei der Gruppe der Lernerinnen feststellen, was auf eine generell hintere Artikulation von /y:/ hinweist.

Bezüglich der F3-Werte lassen sich unterschiedliche Verhältnisse zwischen den Realisierungen von /y:/ in beiden Gruppen identifizieren. In der Kontrollgruppe konnte keine systematische Abnahme der F3-Werte beobachtet werden, wie sie bei den Lernerinnen zu beobachten ist. Bei dem uvularen und glottalen Kontext sind lediglich geringfügige Unterschiede der Medianwerte festzustellen, während dies bei dem labiodentalen, alveolaren und insbesondere bei dem velaren nicht der Fall ist. Beim letzteren ist der Medianwert bei der Gruppe der Lernerinnen deutlich höher, was auf einen geringeren Grad an Lippenrundung hindeutet. Insgesamt lassen sich bei der Gruppe der Lernerinnen höhere F3-Werte feststellen, was als ein generell geringerer Grad an Lippenrundung zu interpretieren ist.

Um die Zielsprachlichkeit der Verhältnisse zwischen den Realisierungen und dem Faktor *y_links_Kontext* wurden die folgenden statistischen Modelle

$$\text{Imer } F2 - 1 + \text{Gruppe} + y_links_Kontext + \text{Gruppe}:y_links_Kontext + (1 | \text{Versuchsperson})$$

$$\text{Imer } F3 - 1 + \text{Gruppe} + y_links_Kontext + \text{Gruppe}:y_links_Kontext + (1 | \text{Versuchsperson})$$

auf die Lernerinnen und die Sprecherinnen mit L1 Deutsch angewendet. Als abhängige Variablen wurden jeweils *F2* und *F3* untersucht, als Faktoren wurden die Variablen *Gruppe* und *y_links_Kontext* und als Zufallsfaktor die Variable *Versuchsperson* angegeben.

Das Modell für die *F2*-Werte ergab einen signifikanten Effekt sowohl für den Faktor *y_links_Kontext* ($p < .001$) als auch für die Interaktion *Gruppe*y_links_Kontext* ($p = 0,006$). Aufgrund des unzureichenden Signifikanzgrads der Interaktion wurde jedoch kein Post-Hoc-Test für die Interaktion durchgeführt. Bezüglich des Faktors *y_links_Kontext* hat der Post-Hoc Test einen signifikanten Unterschied bei den Vergleichen *alveolar-uvular* ($p < .001$), *glottal-uvular* ($p = 0,004$) und *labiodental-uvular* ($p < .001$) ergeben, wobei die *F2*-Werte für den alveolaren, glottalen und labiodentalen Artikulationsort höher sind. Diese Ergebnisse entsprechen den auffälligen Abweichungen an den Medianwerten, die in der Abbildung 18 zu sehen sind. Die Ergebnisse deuten auf eine weiter vorne Artikulation von /y:/ bei diesen drei linken Kontexten hin, wobei eine vordere Artikulation nach glottalem linkem Kontext ungewöhnlich ist.

Das Modell für die *F3*-Werte hat einen signifikanten Effekt sowohl für den Faktor *Gruppe* als auch für den Faktor *y_links_Kontext* (jeweils $p = 0,046$, $p < .001$) geliefert. Die *F3*-Werte bei der Gruppe der Lernerinnen ergeben sich als höher als die für die Kontrollgruppe, was als einen generell niedrigen Grad an Lippenrundung zu interpretieren ist. Dies entspricht dem Bild, das in Abbildung 18 zu sehen ist. Der Post-Hoc Test für den Faktor *y_links_Kontext* hat signifikante Effekte für die Vergleiche *glottal-labiodental* ($p = 0,002$), *labiodental-uvular* ($p = 0,041$), *labiodental-velar* ($p < .001$). Bei dem Vergleich *glottal-labiodental* ergeben sich die *F3*-Werte als signifikant niedriger beim glottalen linken Kontext als beim labiodentalen, was als einen höheren Grad an Lippenrundung beim glottalen linken Kontext zu interpretieren ist. Beim Vergleich *labiodental-uvular* ergeben sich die *F3*-Werte als höher beim labiodentalen linken Kontext als bei uvularen, was als einen niedrigeren Grad an Lippenrundung beim Ersten als beim Zweiten zu interpretieren ist. Beim Vergleich *labiodental-velar* sind die *F3*-Werte signifikant höher beim labiodentalen Kontext als beim velaren, was als einen niedrigeren Grad an Lippenrundung beim labiodentalen linken Kontext zu interpretieren ist.

Nachfolgend wird auf die Variation von den *F2* und *F3*-Werten bei der Realisierung von /y:/ hinsichtlich des rechten Kontextes eingegangen.

Folgende statistische Modelle

$$\text{Imer } F2 - 1 + y_rechts_Kontext + (1 | Versuchsperson)$$

$$\text{Imer } F3 - 1 + y_rechts_Kontext + (1 | Versuchsperson)$$

wurden auf die Frauen und auf die Männer der experimentellen Gruppe getrennt angewendet. Als abhängige Variablen wurde F2 bzw. F3 verwendet, als Faktor die Variable *y_rechts_Kontext* und als Zufallsfaktor die Variable *Versuchsperson*.

Sowohl bei den Frauen als auch bei den Männern lieferte das Modell keine signifikanten Effekte für die abhängige Variable F2 (jeweils $p = 0,427$ und $p = 0,976$). Analoge Ergebnisse zeigten sich für die F3-Werte (jeweils $p = 0,238$ und $p = 0,127$).

Um die Zielsprachlichkeit der Verhältnisse zwischen den Realisierungen und dem Faktor *y_rechts_Kontext* wurden folgende statistische Modelle

$$\text{Imer } F2 - 1 + Gruppe + y_rechts_Kontext + Gruppe:y_rechts_Kontext + (1 | Versuchsperson),$$

$$\text{Imer } F3 - 1 + Gruppe + y_rechts_Kontext + Gruppe:y_rechts_Kontext + (1 | Versuchsperson)$$

auf die Lernerinnen und die Sprecherinnen mit L1 Deutsch angewendet. Als abhängige Variablen wurden *F2* und *F3* getrennt untersucht, als Faktoren wurden die Variablen *Gruppe* und *y_rechts_Kontext* und als Zufallsfaktor die Variable *Versuchsperson* angegeben.

Hinsichtlich der F2 zeigt das Modell keine signifikanten Effekte für den Faktor *y_rechts_Kontext* und die Interaktion *y_rechts_Kontext*Gruppe* (jeweils $p=0,702$, $p=0,491$), jedoch einen signifikanten Effekt ($p=0,003$) für den Faktor *Gruppe*. In diesem Fall zeigen sich die F2-Werte bei der Gruppe der Lernerinnen signifikant niedriger als bei der Kontrollgruppe, was als Zeichen einer hinteren Artikulation von /y:/ bei den Lernerinnen zu interpretieren ist.

Hinsichtlich der F3 hat das Modell einen leicht signifikanten Effekt für den Faktor *y_rechts_Kontext* ergeben ($p=0,011$), jedoch nicht für den Faktor *Gruppe* ($p=0,141$) oder ihrer Interaktion ($p=0,329$).

5.3.3.3 CLI-Phänomene mit dem L1-phonologischen System (Italienisch) bei der Aussprache von /y:/

In diesem Abschnitt werden die Beziehungen zwischen der Aussprache der deutschen Phoneme durch die Lernerinnen und den Realisierungen der italienischen Phoneme durch die italienische Kontrollgruppe (Ko_ITA) untersucht. Im Folgenden werden die Realisierungen der deutschen Phoneme /y:/, /ʏ/, /i:/, /ɪ/, /u:/ und /ʊ/ mit denen der italienischen Phoneme /u/ und /i/ verglichen (Abbildung 31).

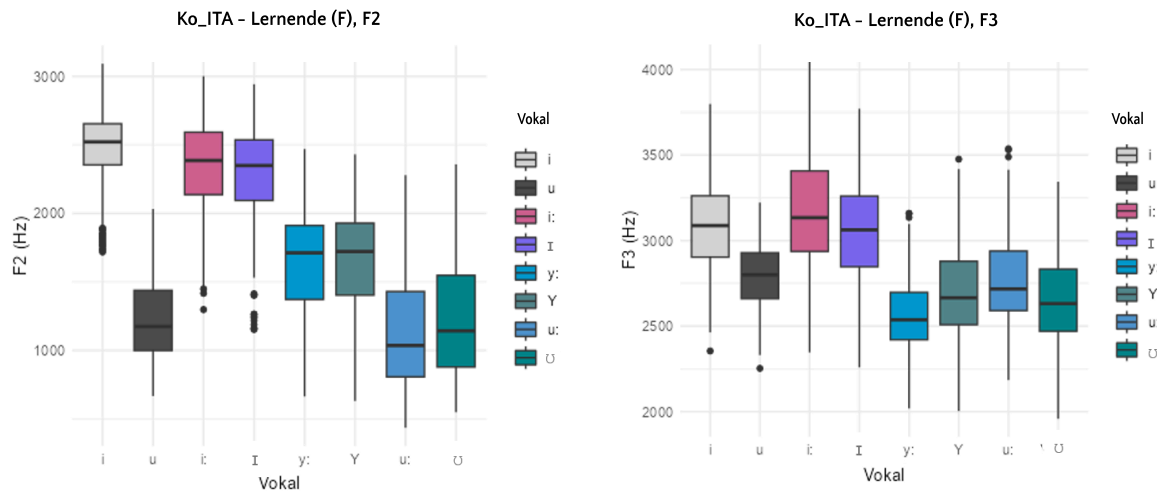


Abbildung 31: Datenstreuung der F2- und F3-Werte für die italienischen (hell- und dunkelgrau) und deutschen Vokale (unterschiedliche Farben). Links: Verteilung der F2-Werte. Rechts: Verteilung der F3- Werte. Die Daten beziehen sich auf die weibliche Stichprobe ("F").

Auf deskriptiver Ebene fällt bei der Streuung der F2-Werte auf, dass sich die Kategorie /y:/ mit keiner Kategorie des italienischen Sprachsystems überlappt, sich aber fast vollständig mit der deutschen Kategorie /Y/ überlappt, was auf eine fast identische Positionierung im Vokalraum hindeutet. Die Kategorie /i:/ des Deutschen überlappt sich mit der deutschen Kategorie /I/ und mit der italienischen /i/ und die Kategorie /u/ des Italienischen fällt in den Existenzbereich der deutschen Kategorien /u:/ und /ʊ/, die sich teilweise überlappen.

Bei den F3-Werten fällt auf, dass die deutschen Kategorien /y:/ und /Y/ relativ niedrige F3-Werte aufweisen, die auch die Realisierungen der deutschen Kategorie /u:/ kennzeichnen. Dies deutet auf einen relativ hohen Grad an Lippenrundung bei den deutschen Kategorien /y:/, /Y/ und /u:/ hin, wobei die Medianwerte bei /y:/ am niedrigsten sind und somit die Lippenrundung am höchsten ist. Parallel kann beobachtet werden, dass die Lippenrundung bei /i:/ und /I/ ähnlich ist (relativ niedrig) wie bei der italienischen Kategorie /i/.

Folgende statistische Modelle

$$\begin{aligned} \text{Imer } F2 &= 1 + \text{Vokal} + (1 | \text{Versuchsperson}), \\ \text{Imer } F3 &= 1 + \text{Vokal} + (1 | \text{Versuchsperson}) \end{aligned}$$

wurden sowohl auf F2 als auch auf F3 durchgeführt. Als abhängige Variablen wurden jeweils F2 und F3 angegeben, als Faktor die Variable *Vokal* und als Zufallsfaktor die Variable *Versuchsperson*.

Bezüglich F2 hat das Modell signifikante Effekte ($p < .001$) ergeben. Der Post-Hoc Test ergibt, dass bei den Vergleichen /i/-/y:/, /u/-/y:/, /Y/-/i/, /Y/-/u/, und /i/-/I/ ist der Effekt signifikant ($p < .001$), während das bei den Vergleichen /u/-/u:/ und /u/-/ʊ/ nicht der Fall ist (bei beiden $p = 1,000$).

Bezüglich F3 hat das Modell signifikante Effekte ($p < .001$) ergeben. Aus dem Post-Hoc Test ist es besonders interessant zu beobachten, dass während bei den Kontrasten /u/-/u:/ und /u/-/ʊ/ keinen signifikanten Unterschied an dem Grad der Lippenrundung festzustellen ist, die Lippenrundung bei dem Vergleich /u/-/y:/ signifikant ($p < .001$) ist. Außerdem ist der Effekt beim Vergleich /ʏ/-/u/ nicht signifikant ($p = 1,000$). Diese zwei letzten Ergebnisse deuten auf einen höhere Lippenrundung bei /y:/ als beim italienischen Phonem /u/ und auf eine vergleichbare Lippenrundung bei /ʏ/ und /u/.

5.3.3.4 Der Einfluss der Sprachkompetenz und -dominanz sowie des Sprachrepertoires auf die Realisierung von /y:/

In diesem Abschnitt wird die Realisierung von /y:/ in Abhängigkeit vom Grad der Sprachkompetenz (GER-Niveau) und Sprachdominanz im Deutschen sowie des Umfangs des Sprachrepertoires innerhalb der Lerner:innengruppe untersucht. Zu diesem Zweck wird die Variation der F2- und F3-Werte in Bezug auf diese Variablen analysiert, um zu überprüfen, ob diese aussagekräftigen Hinweise auf eine kompetenzdynamische Variation der Produktion liefern. Alle statistischen Modelle wurden sowohl auf die Frauen- als auch auf die Männergruppe angewandt.

Folgende statistische Modelle wurden auf die Frauengruppe der Stichprobe der Lerner:innen angewandt¹⁸:

$$\text{Imer F2} = 1 + \text{GER} + \text{Dom_DEU} + \text{Sprachrepertoire} \quad (1 \mid \text{Versuchsperson}),$$

$$\text{Imer F3} = 1 + \text{GER} + \text{Dom_DEU} + \text{Sprachrepertoire} \quad (1 \mid \text{Versuchsperson})$$

Die Variablen F2 und F3 wurden als abhängige Variablen angegeben, während die Variablen GER, Dom_DEU und Sprachrepertoire als Faktoren und die Variable Versuchsperson als Zufallsfaktor festgelegt wurden.

Was F2 angeht, resultierte bei dieser Gruppe kein signifikanter Effekt der angegebenen Variablen (GER $p = 0,077$; Dom_DEU $p = 0,757$; Sprachrepertoire $p = 0,300$).

Was F3 angeht, gelten ähnliche Ergebnisse für die Variable GER ($p = 0,794$) und Sprachrepertoire ($p = 0,332$), während die Variable Dom_DEU einen leicht signifikanten Effekt aufweist ($p = 0,031$). Die Parameterschätzungen verdeutlichen, dass bei einer geringen Dominanz des Deutschen (Stufe 5) die F3-Frequenz im Vergleich zu einer hohen Dominanz (Stufe 2) signifikant ansteigt. Akustisch

¹⁸ Aufgrund der ungleichen Stichprobengröße und der geringeren inhärenten Varianz im Teilkollektiv der Männer unterschied sich die mathematische Belastbarkeit der Modellstrukturen zwischen den Geschlechtern. Während bei den Frauen alle linguistischen und soziodemografischen Prädiktoren (GER, Dom_DEU, Sprachrepertoire) simultan in einem einzigen maximalen Modell für F2 und F3 konvergierten, führte dieser Ansatz bei den Männern zu einer perfekten Multikollinearität (bzw. zu einem singulären Fit). Um algorithmische Artefakte und verzerrte Parameterschätzungen auszuschließen, wurde für die männliche Stichprobe eine datengeleitete methodische Trennung vorgenommen: Die Effekte von Sprachkompetenz und Dominanz (GER, Dom_DEU) einerseits und der Effekt des sprachlichen Repertoires (Sprachrepertoire) andererseits wurden in zwei separate Modellsätze für F2 und F3 extrahiert. Dieses Vorgehen sichert die mathematische Validität und die Vergleichbarkeit der p-Werte für beide Geschlechter.

deutet dieser F3-Anstieg bei abnehmender Dominanz des Deutschen auf einen geringeren Lippenrundungsgrad bei der Vokalproduktion hin.

Folgende statistische Modelle wurden auf die Männergruppe der Stichprobe der Lerner:innen angewandt (siehe Fußnote 18):

$$\begin{aligned} & \text{Imer F2} - 1 + \text{GER} + \text{Dom_DEU} (1 | \text{Versuchsperson}), \\ & \text{Imer F3} - 1 + \text{GER} + \text{Dom_DEU} (1 | \text{Versuchsperson}), \end{aligned}$$

Als abhängige Variable wurden jeweils F2 und F3, als Faktoren die Variablen GER und Dom_DEU und als Zufallsfaktor die Variable Versuchsperson angegeben.

Was F2 angeht, zeigen die Ergebnisse keinen signifikanten Effekt weder für die Variable GER ($p=0,188$) noch für die Variable Dom_DEU ($p=0,266$)¹⁹.

Die Analyse von F3 lieferte ebenfalls keine signifikanten Ergebnisse (GER $p = 0,988$; Dom_DEU $p = 0,773$).

Für die Analyse des Effekts der Variable Sprachrepertoire auf F2 und F3 wurden folgende Modelle angewendet:

$$\begin{aligned} & \text{Imer F2} - 1 + \text{Sprachrepertoire} (1 | \text{Versuchsperson}), \\ & \text{Imer F3} - 1 + \text{Sprachrepertoire} (1 | \text{Versuchsperson}), \end{aligned}$$

Weder für F2 noch für F3 ergaben sich signifikante Effekte (F2 $p= 0,933$; F3 $p=0,943$).

5.3.3.5 Der taskbedingte Einfluss auf die Realisierung von /y:/

In diesem Abschnitt wird auf den Einfluss der Art der Experimentaufgabe auf die Realisierung von /y:/ hinsichtlich der Variation der F2- und F3-Werte eingegangen. Alle statistischen Modelle wurden sowohl auf die Frauen- als auch auf die Männergruppe angewandt. Nachfolgend werden die Datenstreuungen und die Medianwerte für beide Formanten und Geschlechtsgruppen dargestellt (Abbildung 32).

¹⁹ Das Modell weist einen singulären Fit auf. Die extrahierten Haupteffekte für GER und Dom_DEU werden aufgrund der mathematischen Zulässigkeit der p-Werte berichtet, sollten jedoch unter Berücksichtigung dieser geringen Stichprobenvarianz interpretiert werden.

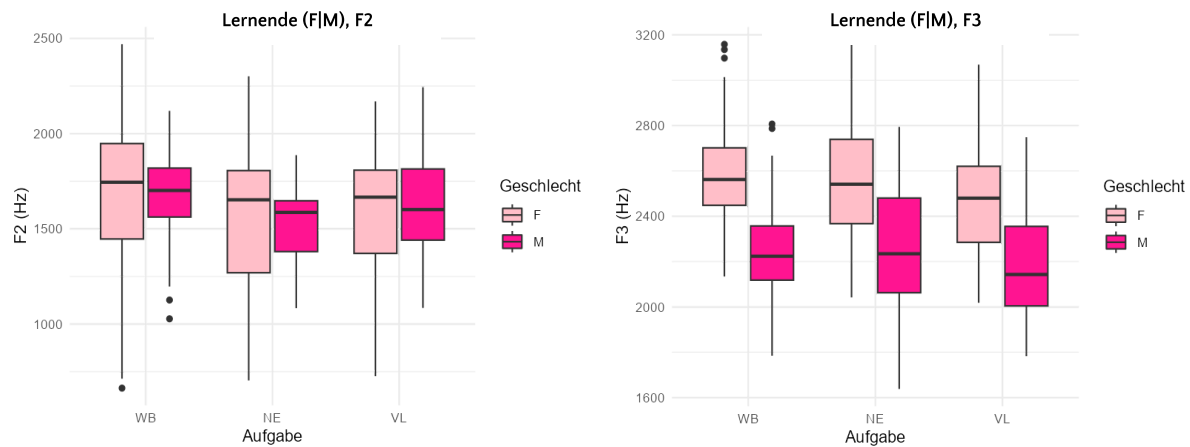


Abbildung 32: Datenstreuung und Medianwerte der F2- und F3-Werte für Frauen ("F") und Männer ("M") innerhalb der Stichprobe der Lerner:innen über alle Aufgaben. Links: Verteilung der F2-Werte. Rechts: Verteilung der F3-Werte.

Auf deskriptiver Ebene hinsichtlich der F2-Werte kann man beobachten, dass sowohl bei den Frauen als auch bei den Männern eine Abnahme an den Medianwerten über die ersten beiden Aufgaben hinweg zu verzeichnen ist. Im Vergleich zu NE zeigen die VL-Werte hingegen sehr ähnliche Medianwerte. Hinsichtlich der F3-Werte zeigen sich bei den ersten beiden Aufgaben ähnliche Medianwerte, während sie bei NL in beiden Geschlechtsgruppen niedriger sind als bei den ersten beiden Aufgaben.

Um der Effekt der Variable *Aufgabe* zu überprüfen, wurden für die folgenden statistischen Modelle

$$\begin{aligned} \text{Imer F2} &= 1 + \text{Aufgabe} + (1 | \text{Versuchsperson}), \\ \text{Imer F3} &= 1 + \text{Aufgabe} + (1 | \text{Versuchsperson}) \end{aligned}$$

jeweils als abhängige Variablen *F2* und *F3*, als Faktor die Variable *Aufgabe* und als Zufallsfaktor die Variable *Versuchsperson* angegeben.

Hinsichtlich der F2, hat das Modell einen signifikanten Effekt sowohl für die Frauen als auch für die Männer ergeben (jeweils $p=0,014$, $p=0,008$). Konkret ist der Effekt signifikant beim Vergleich WB-NE sowohl bei den Frauen als auch bei den Männern (jeweils $p=0.010$, $p=0,002$), wobei die F2-Werte bei WB signifikant höher sind als bei NE, was auf eine vordere Realisierung bei WB als bei NE hindeutet.

Hinsichtlich der F3 hat das Modell signifikante Effekte für die Frauen ($p=0,003$) jedoch nicht für die Männer ($p=0,063$) ergeben. Konkret ergibt sich bei den Frauen ein signifikanter Effekt ($p=0,031$) für den Vergleich VL-NE, wobei die F3-Werte bei VL signifikant niedriger sind als bei NE, was auf einen höheren Grad an Lippenrundung bei der VL-Aufgabe als bei der NE-Aufgabe hindeutet.

5.3.4. Synopse der wesentlichen Ergebnisse

Die zentralen statistischen Ergebnisse werden nachfolgend in einer Tabelle (Tabelle 9) zusammengefasst, um die detailliert beschriebenen Einzeleffekte der untersuchten Faktoren in einen Gesamtzusammenhang zu stellen.

Tabelle 9: Überblick der zentralen statistischen Ergebnisse der Produktionsstudie. Von links nach rechts: die im Mittelpunkt der Untersuchung stehenden Phoneme (/ə/, /y:/), die festen Faktoren, die abhängigen Variablen, die angewandten Modelle sowie die Ergebnisse (n.s. = nicht signifikant). In den Ergebnissen werden drei Analysearten unterschieden: (i) Intragruppenvergleich zwischen Frauen und Männern der Stichprobe der Lerner:innen (F|M), (ii) Intergruppenvergleich zwischen den weiblichen Lernerinnen (F) und der Kontrollgruppe mit Deutsch als L1 sowie (iii) Intergruppenvergleich zwischen den weiblichen Lernerinnen (F) und der Kontrollgruppe mit Italienisch als L1 (CLIPS-Korpus).

				Intragruppenvergleich (F M)	Vergleich mit Ko_DEU (F)	Vergleich mit Ko_ITA (F)
Vokal	Faktor	Abhängige Variable	Analyseebene	Signifikanz (p-Wert bzw. n.s.)		
/ə/	Vokal (/ə/-/e:/-/ε/)	F1, F2	Gesamtmodell	F2, F1 (F&M): signifikant	F2 (Vokal, Gruppe, Gruppe*Vokal): signifikant	F2 (Vokal): signifikant
			Post-Hoc	F2 (F&M): signifikant; F1 (F): signifikant (alle Vergleiche), F1(M): signifikant (/e:/-/ə/), signifikant (/e:/-/ε/)	F2 (Vokal*Gruppe, bei gleichen Phonemen): signifikant; F1 (Vokal, Vokal*Gruppe): signifikant	F2 (/ə/-/ε/ITA, /ε/DEU-/ε/ITA, /e/ITA-/ε/DEU): n.s.
	Position im Wort (präfixal, Wortstamm, präfinal, final)	F1, F2	Gesamtmodell	F2, F1 (F&M): signifikant	F2 (schwa_Position, Gruppe, schwa_Position*Gruppe): signifikant F1 (schwa_Position): signifikant	-
			Post-Hoc	F2 (F&M): signifikant (Wortstamm-präfixal, präfixal-präfinal), F2 (F): signifikant (Wortstamm-präfinal), F2 (M): signifikant (Wortstamm-final). F1 (F&M): signifikant (final-präfixal, final-präfinal), F1(F): signifikant (Wortstamm-präfixal), F1(M): signifikant (Wortstamm-präfixal),	F2 (schwa_Position*Gruppe): signifikant (präfixal, präfinal, final)	-
	Kompetenz im Deutschen (GER-Niveau; B1, B2, C1)	F2			-	-

	Dominanz im Deutschen (2,3,4,5)		Gesamtmodell	F2(F&M): n.s		
			Post-Hoc	-		
	Dominanz im Englischen (2,3,4)					
	Umfang Sprachrepertoire (3,4,5,6)					
	Art der Aufgabe (WB, NE, VL)	F1, F2	Gesamtmodell	F2(F): signifikant F2(M): signifikant F1(F&M): signifikant	-	-
			Post-Hoc	F2(F): signifikant (WB>NE>VL) F1(F&M): signifikant (VL>NE>WB)	-	-
/y:/	Vokal (/y:/-/ʏ/-/i:/-/ɪ/-/u:/-/ʊ/)	F2, F3	Gesamtmodell	F2(F&M): signifikant F3 (F&M): signifikant	F2(Vokal, Vokal*Gruppe): signifikant F3(Vokal, Vokal*Gruppe): signifikant	F2(Vokal): signifikant F3(Vokal): signifikant
			Post-Hoc	F2(F): signifikant (alle Vergleiche; /ʏ/-/y:/, /u:/-/ʊ/ n.s.) F2(M): signifikant (alle Vergleiche; /u:/-/ʊ/ n.s.; /ʏ/-/y:/ signifikant) F3 (F&M): /i:/./y:/ signifikant /y:/-/ʊ/: signifikant; /u:/-/y:/ signifikant	F2 (Vokal*Gruppe; /y:/ und /ɪ/): signifikant F3 (Vokal*Gruppe; /ɪ/, /i:/): signifikant	F2(u/ITA-/u:/DEU und /u/ITA-/ʊ/DEU): n.s.; F3(/u/ITA-/y:/DEU): signifikant, aber F3(/ʏ/DEU-/u/ITA): n.s.
	linker phonologischer Kontext (labiodental, alveolar, uvular, velar, glottal)	F2,F3	Gesamtmodell	F2(F,M): signifikant F3(F): signifikant	F2(links_Kontext): signifikant F2(Gruppe*links_Kontext): signifikant F3(links_Kontext): signifikant , F3(Gruppe): signifikant	-
			Post-Hoc	-	F2(alveolar-uvular): signifikant, F2(glottal-uvular): signifikant	-

					F2(labiodental-uvular): signifikant F3(glottal-labiodental): signifikant F3(labiodental-uvular): signifikant F3(labiodental-velar): signifikant	
	rechter phonologischer Kontext (labiodental, alveolar, uvular, velar)	F2, F3	Gesamtmodell	F2(F&M): n.s.; F3(F&M): n.s.	F2(Gruppe): signifikant; F3(rechts_y_Kontext): signifikant	-
			Post-Hoc	-	-	-
	Kompetenz im Deutschen (GER-Niveau; B1, B2, C1)	F2, F3			-	-
					-	-
	Dominanz im Deutschen (2,3,4,5)		Gesamtmodell	F2(F): n.s.; F3(F): n.s. mit Ausnahme von Dominanz im Deutschen F2 (M; Kompetenz + Dominanz): n.s. F3 (M; Sprachrepertoire): n.s.	-	-
	Umfang Sprachrepertoire (3,4,5,6)		Post-Hoc	-	-	-
	Art der Aufgabe (WB, NE, VL)	F2, F3	Gesamtmodell	F2(F, M): signifikant; F3(F): signifikant	-	-
			Post-Hoc	-	-	-

5.4 Diskussion

In diesem Abschnitt werden die Ergebnisse der vorliegenden Produktionsstudie anhand der Hypothesen diskutiert. Die Diskussion fängt mit einer intralinguistischen Perspektive an, die auf die Verhältnisse zwischen den Repräsentationen der deutschen Phoneme innerhalb der experimentellen Gruppe der Lerner:innen eingeht. Mit einer interlinguistischen Perspektive, die auf die CLI-Phänomene zwischen dem *Interlanguage*-System der Lernerinnen und dem italienischen und deutschen System vertreten von den zwei Kontrollgruppen wird dann die Diskussion abgeschlossen.

Bezüglich der Realisierung von beiden untersuchten Phonemen /ə/ und /y:/ kann man tendenziell von einer Peripheralisierung bzw. Dezentralisierung (vgl. Missaglia & Sendlmeier, 1999) der Realisierungen sprechen: die Lerner:innen tendieren, diese Phoneme in der Peripherie des Vokalraums zu realisieren und zwar fallen diese Kategorien neben anderen Kategorien des Deutschen in Positionen, die schon durch entsprechend ähnliche L1-Kategorien besetzt werden. Das zentrale Areal des Vokalraums bleibt daher grundsätzlich „leer“ und es werden periphere Vokalclusters gebildet. Außerdem sind die Lernerinnen generell nicht sensitiv gegenüber dem Merkmal der Gespanntheit. Das beobachtete Phänomen der Dezentralisierung der Vokale, insbesondere bei /i:/, /ɪ/, /e:/ und /ɛ/ ist im Einklang mit den Ergebnissen von Missaglia & Sendlmeier (1999), die dasselbe Phänomen bei der Realisierung /i:/, /ɪ/ und /ɛ/ festgestellt haben. In der vorliegenden Studie wurde diese Tendenz auch bei den hinteren Vokalen /u:/ und /ʊ/ beobachtet, während hingegen Missaglia & Sendlmeier (1999) eine Zentralisierung der Laute /u:/, /ʊ/ und /o:/ beobachtet haben. Die allgemeine Dezentralisierung und Clusterisierung der Lautkategorien in den Arealen, die schon in der L1 besetzt sind wird als plausibel erachtet: obwohl die Lerner:innen ein durchschnittlich theoretisch hohes Niveau im Deutschen aufweisen, untergeht noch die Lokalisierung der gelernten Kategorien der phonematischen Prägung der L1, was als ein Effekt der *phonetic equation* (Escudero, 2005) bzw. des Magnet-Effekts (Kuhl, 1991; Kuhl & Iverson, 1995) oder – nach Flege (1995) – der *equivalence classification* interpretiert werden kann. In diesem Sinne kann die H1 bestätigt werden. Die weiteren Hypothesen werden nachfolgend im näheren Detail behandelt.

Beginnt man mit den Verhältnissen zwischen den Lautkategorien innerhalb der experimentellen Gruppe der Lerner:innen, muss die Hypothese 1a), die eine Unschärfe der Trennung der deutschen Kategorien /ə/-/e/-/ɛ/ postuliert, grundsätzlich widerlegt werden. Diese Kategorien sind nämlich hinsichtlich ihrer horizontalen Positionierung im Vokalraum (F2) sowohl bei Frauen als auch bei Männern getrennt. Sie werden bei der Gruppe der Frauen hinsichtlich ihrer Höhe (F1) ebenfalls signifikant getrennt, bei den Männern ist das auch der Fall abgesehen vom Kontrast /ə/-/ɛ/. Das /ə/ scheint daher grundsätzlich seine eigene Position im Vokalraum eingenommen zu haben, trotzdem wird es bei den Männern noch als /ɛ/ ausgesprochen.

Bezüglich der Hypothese H1b), die ein mögliches Zusammenspiel der Merkmale [+vorne] und [+gerundet] bei der Aussprache von /y:/ und daher eine ähnliche Aussprache zu /i:/ oder /ɪ/ oder zu /ʏ/, /u:/ oder /ʊ/ und somit eine Unschärfe der Trennung dieser Lautkategorien postuliert, kann zum Teil bestätigt werden. Das Phonem /y:/ scheint einerseits sowohl bei den Frauen als auch bei den Männern seine eigene Position im Vokalraum eingenommen zu haben. Andererseits wird es mit der Kategorie /ʏ/ bei den Frauen assimiliert, was ein deutliches Zeichen der Insensibilität gegenüber dem Merkmal der Gespanntheit bzw. Länge ist. Aus diesem Grund kann festgestellt werden, dass bei den Frauen die Trennung dieser Kategorien nicht zustande gekommen ist. Die Insensibilität gegenüber dem Merkmal der Gespanntheit bzw. Länge ist auch bei den Kontrasten /u:-/ʊ/ und /i:-/ɪ/ zu beobachten, die sowohl bei den Frauen als auch bei den Männern eine sehr ähnliche Position im Vokalraum (F2) einnehmen und daher gleich ausgesprochen werden. Diese Ergebnisse sind im Einklang mit anderen Studien im Rahmen des Erwerbs des Deutschen als Fremdsprache durch erwachsene Lerner:innen (vgl. u.a. Zimmerer & Trovain, 2015 für die Konstellation Französisch-Deutsch; Lipinska, 2015 und Nimz & Khattab, 2020 für die Konstellation Polnisch-Deutsch). Des Weiteren ist die Hypothese H1b) in Bezug auf die Rundung (F3) des Phonems /y:/ zu widerlegen: Das Phonem /y:/ ist sowohl bei den Frauen als auch bei den Männern gerundeter als /i:/ und /u:/, /ʊ/. Dies lässt auf ein funktionelles Zusammenspiel zwischen den Merkmalen [+vorn] und [+gerundet] sowie auf eine Hyperkorrektur hindeuten.

Anhand der Vergleiche zwischen den Realisierungen der deutschen Phoneme durch die Lernerinnen und den entsprechenden ähnlichen Phonemen des Deutschen und Italienischen ausgesprochen von den weiblichen Kontrollgruppen sind neue Erkenntnisse zum Effekt der CLI entstanden, die mehr Licht auf die Verhältnisse zwischen diesen Systemen im Rahmen des Fremdspracherwerbs werfen.

Im Vergleich mit den Realisierungen durch die L1-Sprecherinnen des Deutschen zeigt sich, dass die Lernerinnen mehr Schwierigkeiten beim Aufbau zielsprachlich getrennter Lautkategorien für den Fall von /ə/ als für den Fall von /y:/ haben. Die Lautkategorien /ə/, /e:/ und /ɛ/ werden hinsichtlich der horizontalen Positionierung im Vokalraum (F2) zwar als drei getrennte Lautkategorien realisiert, die aber nicht zielsprachlich sind: die Unterschiede zu denen der Kontrollgruppe haben sich als signifikant erwiesen, indem sie tendenziell dezentralisierter artikuliert werden. Hinsichtlich des Grads der Öffnung werden sie aber anscheinend zielsprachlich realisiert. Unter den Lautkategorien /y:/, /ʏ/, /i:/, /ɪ/, /u:/ und /ʊ/ werden fast alle hingegen zielsprachlich ausgesprochen: /ʏ/, /i:/, /u:/ und /ʊ/ werden nämlich hinsichtlich der horizontalen Positionierung im Vokalraum zielsprachlich realisiert, während die Realisierung von /y:/ und /ɪ/ signifikant von der zielsprachlichen abweicht. Zudem wird /y:/ generell weiter vorne artikuliert im Vergleich zu L1-Realisierung. Das Merkmal [+gerundet] hat sich bei /y:/, /ʏ/, /u:/ und /ʊ/ als zielsprachlich gelernt ergeben, während die Phoneme /i:/ und /ɪ/ scheinen noch gewisse Schwierigkeiten zu bereiten: im Vergleich zu der Kontrollgruppe wird einerseits /i:/ zu viel und andererseits /ɪ/ zu wenig gerundet. Die Rundung von

vorderen Vokalen, die im Standarddeutschen theoretisch als ungerundet gelten, stellt kein systematisches Phänomen dar. Es existieren jedoch Nachweise (vgl. Salmons, 2012; Calvo Fernández, 2018), die eine Realisierung von /ɪ/ als gerundeten Vorderzungenvokal [ʏ] belegen (z. B. *nicht* > [nʏçt] bzw. [nʏft] statt [nɪçt]; vgl. Salmons, 2012), insbesondere unter jüngeren Sprecher:innen – wie auch in der in dieser Studie untersuchten Population. Soweit ersichtlich, liegen bislang keine Belege für eine systematische Beziehung zwischen /i:/ und /ɪ/ in Bezug auf F3-Werte vor, noch weniger im Hinblick auf die Aussprache des Deutschen durch italophone Lerner:innen. Dieser Befund bleibt daher interpretationsbedürftig.

Durch den Vergleich von den deutschen Phonemen /ə/, /e:/ und /ɛ/ mit den entsprechend ähnlichen Lautkategorien des Italienischen /e/ und /ɛ/ wird klar, dass die Lernerinnen das /ə/ tendenziell geschlossen und in einer ähnlichen horizontalen Position wie das italienische /ɛ/ (d.h. etwas zentralisierter als das italienische /e/ und ähnlicher zur Position von /ɛ/ im Vokalraum) realisiert wird. Außerdem überlappt sich die italienische Lautkategorie /e/ mit der deutschen Lautkategorie /ɛ/ hinsichtlich sowohl der F2 als auch der F1. Diese Ergebnisse können wie folgt interpretiert werden. Die Tatsache, dass die Lernerinnen das /ə/ aussprechen, indem sie artikulatorische Merkmale der entsprechenden italienischen Phoneme /e/ und /ɛ/ nutzen, deutet auf die Anwendung einer Lernstrategie hin: Die Tatsache, dass das italienische /ɛ/ artikulatorisch näher an /ə/ liegt als das italienische /e/, könnte das Fehlen signifikanter Unterschiede an den F2-Werten im Italienischen /ɛ/ und im von den Lerner:innen realisierten /ə/ erklären. Andererseits könnte die Tatsache, dass das /ə/ geschlossener als erwartet ausgesprochen wird, auf einen immer noch vorhandenen L1-Effekt hindeuten, der möglicherweise in der Orthografie verankert ist und demzufolge im Italienischen das Graphem <e> dem /e/ oder dem /ɛ/ entspricht (vgl. Bassetti, 2008; Bassetti & Atkinson, 2015). Der niedrige Grad an Kontrastivität der beiden letztgenannten Phoneme im Italienischen könnte den Lernerinnen eine gewisse strategische Flexibilität beim Lernen der Fremdsprache bieten. Diese Ergebnisse stimmen auch mit der früheren Beobachtung überein, dass das /ə/ und die deutsche /ɛ/ Kategorie hinsichtlich der Öffnung der Artikulation getrennt realisiert werden, was jedoch nicht ausschließt, dass dies im Bereich des italienischen /e/-Phonems verbleibt. Diese Ergebnisse sind mit der Assimilationshypothese von Maturi (2014) vereinbar.

Durch den Vergleich von den deutschen Phonemen /y:/, /ʏ/, /i:/, /ɪ/, /u:/ und /ʊ/ mit den entsprechend ähnlichen Lautkategorien des Italienischen /u/ und /i/ ist Folgendes klar geworden. Hinsichtlich der Anteriorität/Posteriorität der Artikulation überlappen sich die Kategorien /u/-/u:/ und /u/-/ʊ/: letztere entspricht dem Assimilationsmuster von Maturi (2014). Trotzdem werden die Kategorien /y:/-/i:/, /y:/-/u:/, /ʏ/-/i:/, /ʏ/-/y:/ und /i/-/ɪ/ unterschiedlich realisiert, was auf die Abwesenheit von CLI-Phänomenen bezüglich dieser Kategorien hindeutet. Hinsichtlich der Lippenrundung hat es sich ergeben, dass die hinteren Lautkategorien /u/-/u:/ und /u/-/ʊ/ gleich gerundet werden. Gleichzeitig ist es besonders interessant zu beobachten, dass während die

Lippenrundung von /ʏ/ zu der vom italienischen /u/ vergleichbar ist, ist die Rundung von /y:/ signifikant höher als die von /u/: Dies deutet darauf hin, dass im Vergleich zum italienischen phonologischen System /y:/ relativ stärker gerundet wird als /ʏ/. Verbindet man dieses Ergebnis mit den zuvor dargestellten Ergebnissen zum Vergleich mit dem deutschen phonologischen System, so kann man feststellen, dass Lernerinnen dieses Phonem zwar im Allgemeinen deutlich weniger runden als deutsche Sprecherinnen, dass sie aber das Merkmal [+gerundet] so weit wie möglich übertreiben, was als Phänomen der Hyperkorrektur angesehen werden kann.

Die Hypothese H2a), die postuliert, dass die Realisierung von /ə/ abhängig von der Position im Wort variiert, kann bestätigt werden. Die Variable der Position zeigt nämlich signifikante Effekte auf die Realisierung von /ə/ sowohl bei den Frauen als auch bei den Männern. Diese Ergebnisse unterscheiden sich von denen in der Studie von Hove (2020), in der kein signifikanter Effekt der Positionsvariable festgestellt wurde. Sie sind jedoch im Einklang mit Flemming & Johnson (2007) und Lass (2009), die einen signifikanten Effekt feststellen konnten: In Übereinstimmung mit den Ergebnissen von Flemming & Johnson (2007) und Lass (2009) zum Englischen, die signifikant höhere F1-Werte in der finalen Wortposition gefunden hatten, zeigen die Ergebnisse der vorliegenden Studie tatsächlich signifikant höhere Werte in finaler Position als präfixaler Position sowohl für Männer als auch für Frauen. Bei Männern – was aber bei Frauen nicht der Fall ist – sind außerdem die F1-Werte in finaler Position signifikant höher als in präfinaler Position. Es wurde festgestellt, dass sich die Verhältnisse hinsichtlich der F1 bei den Frauen aus der experimentellen Gruppe und der Kontrollgruppe mit L1 Deutsch nicht signifikant unterscheiden. Dies deutet auf eine zielsprachliche Realisierung im Hinblick auf die untersuchte Kontrollgruppe und auf vergleichbare Ergebnisse mit den Vorstudien hin.

In der vorliegenden Studie runden die Ergebnisse zur horizontalen Positionierung (F2) des /ə/ in den verschiedenen Positionen das Bild ab: Sowohl bei Frauen als auch bei Männern zeigt sich ein ähnlicher Zentralisierungstrend:

- In finaler Position wird das /ə/ zentraler realisiert als in präfinaler.
- Auch in präfixaler Position ist das /ə/ zentraler als in präfinaler.

Unterschiede zeigen sich jedoch in der Wortstamm-Position:

- Bei Frauen ist das /ə/ im Wortstamm zentraler als in präfinaler Position.
- Bei Männern ist es zentraler als in präfixaler Position.

Bezugnehmend auf die Ergebnisse zu den Frauen aus der experimentellen Gruppe und aus der Kontrollgruppe mit L1-Deutsch lässt sich konstatieren, dass die beobachteten Verhältnisse in den Realisierungen der Lernerinnen lediglich partiell in der Kontrollgruppe repliziert werden. Während bei den untersuchten L1-Sprecherinnen das /ə/ in der finalen Position tendenziell stärker zentral artikuliert wird als in präfixaler Position, werden keine signifikanten Unterschiede in den

Realisierungen in der Wortstamm-, präfinaler und präfixaler Position festgestellt. Diese Ergebnisse unterscheiden sich folglich von denen von Christen et al. (2010), die eine weniger zentrale Realisierung des /ə/ in der Finalposition als in der Präfinalposition festgestellt hatten. Es sei jedoch angemerkt, dass der Vergleich dieser früheren Studien mit Vorsicht zu genießen ist. Obwohl die Ergebnisse der Studie von Christen et al. (2010) in der Literatur mit denen von Flemming & Johnson (2007) und Lass (2009) verglichen wurden (vgl. Hove, 2020), muss betont werden, dass sie sich auf unterschiedliche Sprachsysteme beziehen (Schweizerdeutsch bzw. amerikanisches Englisch). Obwohl einige Variablen das /ə/ in verschiedenen Sprachen in ähnlicher Weise beeinflussen können, sind Vergleiche zwischen verschiedenen Sprachen (und verschiedenen Varietäten derselben Sprache) mit Vorsicht vorzunehmen.

Die Hypothese H2b), die postuliert, dass vordere Artikulationsorte sowohl im linken als auch im rechten phonologischen Kontext die Anteriorität und Rundung von /y:/ fördern sollten, kann nur teilweise bestätigt werden.

Hinsichtlich der Anteriorität der Artikulation von /y:/ d.h. der Variation der F2-Werte hat sich ergeben, dass der linke Kontext signifikant höhere Effekte – sowohl für Frauen als auch für Männer – auf die Artikulation hat als der rechte Kontext. Das gilt für die Gruppe der Lernerinnen sowie für die Gruppe der Sprecherinnen mit L1 Deutsch. Das ist in genereller Übereinstimmung mit Prieto (1993, vgl. auch Orkin, 1970): Hinsichtlich der Rundung von /y:/ d.h. der Variation der F3-Werte, hat sich nur der linke Kontext und zwar nur bei den Frauen als signifikant ergeben. Im Allgemeinen neigen die Lernerinnen dazu, das /y:/ weniger abzurunden als die Kontrollgruppe mit L1-Deutsch: Dies bestätigt die auch in Hypothese H1b) geäußerten Erwartungen, wonach das Merkmal [+gerundet] aufgrund seiner Inkompatibilität mit den fortgeschritteneren Artikulationspositionen in der L1 (Italienisch) möglicherweise nicht korrekt in das phonologische System der *Interlanguage* integriert wird.

Bezüglich der Variation der F2-Werte zeigte sich, dass an vorderen Artikulationsstellen (labiodental oder alveolar) des linken Konsonanten eine signifikant stärkere Vorverlagerung der Artikulation von /y:/ im Vergleich zu weiter hinteren Kontexten (glottal und uvular) entspricht. Dies bestätigt die Hypothese H2b) hinsichtlich des Einflusses des linken Kontextes auf die Vorverlagerung der Artikulation von /y:/. Eine Ausnahme stellt die signifikante Anteriorität dar, die im Zusammenhang mit einem glottalen Konsonanten (/h/) im Vergleich zu einem uvularen Konsonanten (/ʁ/) beobachtet wurde: Die Interpretation dieser Ergebnisse gestaltet sich als Herausforderung, die wahrscheinlich auf einem multifaktoriellen Hintergrund beruht. Als mögliche Erklärungsansätze sind beispielsweise die Tatsache zu nennen, dass beide linken Konsonanten neue Lautkategorien für die Lerner:innen darstellen (auch wenn /ʁ/ individuell im Rahmen von Rotazismus vorkommen kann), sowie die Notwendigkeit weiterer Untersuchungen in nachfolgenden Studien.

Die Ergebnisse der Analyse von F2 und F3 lassen sich besser interpretieren, wenn sie gemeinsam für die linken Kontexte betrachtet werden, bei denen sowohl für F2 als auch für F3 ein signifikanter Effekt festgestellt wurde. Die vorliegenden Untersuchungen zeigen, dass die Kontexte, die einen signifikanten Einfluss auf die Aussprache von /y:/ bezüglich sowohl F2 als auch F3 zeigen, der labiodentale und der glottale Kontext sind. Im Folgenden werden diese einzeln analysiert, um die zugrunde liegenden Ursachen zu ermitteln.

Hinsichtlich des labiodentalen Kontextes wurde beobachtet, dass einerseits eine weiter vorn liegende Artikulation von /y:/ begünstigt wird, andererseits jedoch die Rundung nicht gefördert wird. Infolgedessen wird die Hypothese H2b) für den labiodentalen Kontext nur partiell bestätigt. Eine mögliche Erklärung könnte in einem *feature weighting* liegen, bei dem in diesem Fall das Merkmal [+vorne] gegenüber [+gerundet] überwiegt. Zudem wird die Interpretation der Ergebnisse durch die Berücksichtigung der Studienergebnisse von Ferrero et al. (1978) klarer. In dieser Studie wurde der Einfluss des Vokals /u/ auf die vorhergehenden Konsonanten /p/ und /b/ in Bezug auf die Lippenprotrusion festgestellt – ein Merkmal, das sowohl diese Konsonanten als auch den Vokal /u/ im Italienischen kennzeichnet. Obwohl die Richtung des koartikulatorischen Einflusses in jener Studie entgegengesetzt zu der im vorliegenden Forschungskontext untersuchten ist, zeigt sich dennoch, dass die gemeinsame Präsenz des Merkmals [+bilabial] sowohl bei den Konsonanten als auch beim Vokal zu einer wechselseitigen Beeinflussung hinsichtlich der Lippenprotrusion führt. In der vorliegenden Studie ist der labiodentale Artikulationsort zwar anterior wie bei /p/ und /b/, aber es fehlt das Merkmal [+bilabial]. Infolgedessen ist es plausibel, dass das Fehlen dieses Merkmals die Rundung von /y:/ nicht unterstützt hat. Diese Beobachtung scheint, basierend auf den erhobenen Daten, unabhängig von der Gruppenzugehörigkeit (Lerner:innen / L1-Deutsch-Sprecherinnen) zu sein und deutet somit auf einen übergeordneten Mechanismus hin, der einer artikulatorischen Ökonomie folgt. Übertragen auf ein hypothetisches Vokaltrapez, so zeigt /y:/, wenn es von einem labiodentalen Konsonanten vorangegangen wird, eine Tendenz zur Annäherung an /i:/ oder /ɪ/.

Hinsichtlich des glottalen Kontexts (/h/) konnte festgestellt werden, dass dieser sowohl eine vordere Artikulation von /y:/ als auch eine verstärkte Rundung begünstigt. In diesem Fall kann die Hypothese H2b) vollständig bestätigt werden. Es ist jedoch erforderlich, dieses Resultat stärker zu kontextualisieren, um mögliche Ursachen und Einschränkungen besser zu verstehen. Im Datensatz, auf dem die Analyse basiert, tritt der glottale Kontext ausschließlich im Wort "Hügel" auf, dessen Vorkommen keineswegs vernachlässigbar ist und das im Rahmen der Wegbeschreibung erhoben wurde. In dieser Aufgabe war das Wort "Hügel" als fester Bestandteil schriftlich auf der Karte vorgegeben und musste während der Durchführung der Aufgabe zwingend verwendet werden. Die Ergebnisse zeigen, dass die Realisierung von /y:/ in "Hügel" die Merkmalskombination [+vorn] und [+gerundet] aufweist. Dies deutet darauf hin, dass diese Kombination erlernt wurde. Wie bereits angemerkt, manifestiert sich die angestrebte Rundung jedoch noch nicht zielsprachlich, da

sie noch nicht ausreichend ausgeprägt ist. In diesem Sinne lässt sich das Ergebnis im Rahmen der *Interlanguage* verorten: Es handelt sich um eine konfigurationsähnliche, aber noch nicht vollständig entwickelte Zielstruktur, was auf einen noch nicht abgeschlossenen oder unterbrochenen Lernprozess hindeuten könnte. Nichtsdestotrotz ist bei der Interpretation der Ergebnisse zum glottalen Kontext Vorsicht geboten, da diese sich auf lediglich ein einziges Wort beziehen. Darüber hinaus kann ein unterstützender Effekt der Orthografie nicht ausgeschlossen werden. Da diese jedoch nicht Gegenstand der vorliegenden Untersuchung ist, kann in dieser Hinsicht keine weiterführende Aussage getroffen werden – sie stellt jedoch einen vielversprechenden Ansatzpunkt für weiterführende Studien dar.

Hinsichtlich der Hypothese H3a), die einen Effekt der Dominanz im Englischen postuliert, muss diese widerlegt werden, weil sich dieser Faktor als nicht signifikant hinsichtlich der F2-Werte ergeben hat. Die fehlende Signifikanz dieses Ergebnisses lässt sich im Wesentlichen auf zwei verschiedene Weisen erklären. Zum einen ist es möglich, dass der statistische Test keine wesentlichen Unterschiede erfassen konnte, da die Variabilität im Dominanzniveau der Teilnehmer:innen überwiegend zwischen den Werten 2 und 3 lag, die eine mittel-hohe Dominanz des Englischen im sprachlichen Repertoire anzeigen. Der geringe "Abstand" zwischen diesen Niveaus könnte somit dazu geführt haben, dass das Modell keine substanziellen Effekte nachweisen konnte. Zum anderen lässt sich diese primär statistische Erklärung durch eine eher soziolinguistische Betrachtung ergänzen. Wie bei Leonidou (2020) und Gauci (2008) berichtet wird, ist der Englischunterricht im italienischen Kontext durch eine geringe Beachtung der Aussprache gekennzeichnet, was sich in ihrer untergeordneten Rolle im didaktischen Alltag widerspiegelt – insbesondere dann, wenn ein ausschließlich kommunikativer oder auf Aufgaben basierender Ansatz verfolgt wird (vgl. Gauci, 2008). Die durch intersystemische Unterschiede bedingten Lernschwierigkeiten werden somit durch einen kulturell konservativen Ansatz noch verstärkt, der eine zielsprachliche Aussprache im italienischsprachigen Kontext eher entmutigt (vgl. Leonidou 2020, S.158-159). Es handelt sich hierbei um weitreichende Beobachtungen, die das Gesamtbild der gewonnenen Ergebnisse durch einige zusätzliche Überlegungen bereichern können, ohne jedoch den Anspruch zu erheben, erschöpfende Erklärungen zu liefern. Zu diesem Zweck wären weitere gezielte Untersuchungen erforderlich, die diese Ergebnisse durch einen direkten Vergleich der Realisierung des englischen und deutschen /ə/ durch Lerner:innen unter Berücksichtigung der jeweiligen soziolinguistischen Kompetenzaspekte ergänzen.

In diese Hypothese fließt, wenn auch eher am Rande, die Annahme eines (positiven) Einflusses von Vorkenntnissen in Dialekten auf die Realisierung des Schwa ein. Untersucht wurden insbesondere zwei Proband:innen aus der Stichprobe, die angaben, Kompetenzen in Dialekten des südlichen Dialektgebietes zu besitzen (konkret angegeben als "abruzzesisch" bzw. "Dialekt von [...]“, Anm.) Neapel"). Diese explorative Analyse, die aufgrund des geringen Stichprobenumfangs auf einer rein deskriptiven Ebene blieb, zeigte, dass die Medianwerte von F2 bei den beiden Teilnehmerinnen mit

relevanten Dialektkenntnissen niedriger waren als bei der Kontrollgruppe, die aus den übrigen Teilnehmerinnen (ohne relevante oder ohne Dialektkenntnisse) bestand. Diese Ergebnisse deuten auf eine zentralere Realisierung von /ə/ bei den zwei Proband: innen und stellen nur ein erster Anstoß für weitere Untersuchungen in kontrollierteren Kontexten dar. Im Rahmen der vorliegenden Studie, in der die Variable der Kompetenz in einer L1-Varietät erst im Nachhinein berücksichtigt und nach einem ähnlichen Analyseverfahren wie in der Perzeptionsstudie behandelt wurde, stellen sie einen ersten Versuch dar, einen bislang unerforschten Ansatz zu verfolgen, der dialektale L1-Varietäten noch nicht systematisch als potenzielle Quellen von CLI im Fremdspracherwerb einbezieht.

Hinsichtlich der Hypothesen H3b) und H3c), die jeweils einen Effekt della Sprachkompetenz im Deutschen (GER-Niveau), der Dominanz sowie des Umfangs des Sprachrepertoires postulieren, müssen die Annahmen zum größeren Teil widerlegt werden, da sich diese Faktoren im Gesamtmodell als nicht signifikant erwiesen haben. Die einzige Ausnahme stellt die schwache Signifikanz der deutschen Sprachdominanz (*Dom_DEU*) für die Variation der F3-Werte bei der Gruppe der Lernerinnen dar. Aus akustisch-phonetischer Sicht lässt sich dieser F3-Anstieg bei abnehmender Sprachdominanz direkt als Modifikation des Artikulationsverhaltens interpretieren. Dieser Befund deutet auf einen abnehmenden Lippenrundungsgrad hin: Sprecherinnen mit einer schwächer ausgeprägten Dominanz der Zielsprache (Stufe 5) neigen demnach dazu, das inhärente artikulatorische Merkmal der Lippenrundung phonetisch zu reduzieren, was zu einer Erhöhung der dritten Formantenfrequenz führt. Die Ergebnisse zur nicht-Signifikanz der Sprachkompetenz im Deutschen stehen im Einklang mit den Ergebnissen von Gut (2010) und Gallardo del Puerto (2007). Ein möglicher Grund für die fehlende Signifikanz liegt in der begrenzten Differenzierungskraft der GER-Skala, wenn – wie hier – nur mittlere Stufen einbezogen werden. Zudem dürfte der nicht-immersive Erwerb und die eventuelle Unterbrechung des Lernens die *Interlanguage* stabilisiert und somit positionsbezogene Unterschiede entlang der Niveaus B1 bis C1 auf Produktionsebene nivelliert haben. Die fehlende Signifikanz der Rolle des Umfangs des Repertoires, anders als in der Ergebnisdiskussion der Studie von O'Brien & Smith (2010) nahegelegt, könnte auf den Unterschied im Lernkontext zurückzuführen sein: nicht immersiv in der vorliegenden Untersuchung, während er bei O'Brien & Smith (2010) immersiv war. Da jedoch diese Variable bislang nur wenig erforscht ist und auch die Studie von O'Brien & Smith (2010) lediglich eine mögliche Erklärung für ihre Befunde andeutet, erscheint es sinnvoll, diese Variable in zukünftig noch stärker kontrollierten Kontexten weiter zu untersuchen.

Hinsichtlich der letzten Hypothese H4), die einen Effekt der Aufgabebedingung postuliert, kann diese bestätigt werden. Sowohl bei /ə/ als auch bei /y:/ hat sich dieser Faktor sowohl bei Frauen als auch bei Männern für beide ersten Formanten als signifikant ergeben. Die Variation in Abhängigkeit von der Aufgabebedingung ist im Einklang mit den Aussagen in Siebenhaar & Hahn (2019), die solche Variabilität der experimentellen Daten hervorheben. Im Falle der Realisierung

von /ə/ weist die Frauengruppe die Tendenz Wegbeschreibung>Nacherzählung>Vorlesen hinsichtlich der F2-Werte: diese Ergebnisse deuten auf einer immer zentraleren Realisierung von /ə/ je mehr die Aufgabenbedingungen konditioniert sind. Dies könnte auch auf einen Effekt der Koartikulation bei freieren Sprechkontexten, wie bei der Wegbeschreibung (WB), hindeuten. Für die F1-Werte hat sich bei den Frauen die Tendenz Vorlesen>Nacherzählung>Wegbeschreibung ergeben, was darauf hindeutet, dass /ə/ in konditionierteren Kontexten eher offen ausgesprochen wird als in freieren Sprechkontexten. Bezüglich der Realisierung von /y:/ hat sich keine genaue Tendenz wie im Falle von /ə/ ergeben, trotzdem deuten die Ergebnisse auf eine vordere Realisierung bei der Wegbeschreibung als bei der Nacherzählung (NE) sowohl bei den Frauen als auch bei den Männern und auf eine höhere Rundung bei der Aufgabe "Vorlesen" als bei der Nacherzählung hin. Diese Ergebnisse weisen auf eine tendenziell Hyperartikulation von /y:/ in konditioniertem Aufgabenkontext hin.

Unter Berücksichtigung der Gesamtergebnisse zeigt sich, dass die phonologischen Repräsentationen von /ə/ und /y:/ bei den Lerner:innen stark durch die intersystemischen Unterschiede zwischen den konkurrierenden Lautsystemen geprägt sind. Diese Unterschiede üben einen signifikanten Einfluss aus – einerseits in Richtung des Ausgangssystems, andererseits in Richtung des Zielsprache-Systems, im Bestreben, eine zielsprachliche Lautung zu erreichen. Es sind somit primär sprachsysteminterne Faktoren, die die Ausgestaltung dieser Repräsentationen bestimmen, die hier als Momentaufnahme innerhalb eines andauernden oder unterbrochenen Prozesses visualisiert werden. Aus einer näheren Perspektive lassen sich mikroskopische Variationen der Lautrealisierungen beobachten, die auf spezifischere linguistische Einflussfaktoren zurückzuführen sind – etwa auf die Position innerhalb des Wortes oder auf den benachbarten konsonantischen Kontext. Diese Variationen deuten auf Strukturen hin, die nicht zielsprachlich im engeren Sinne sind, jedoch auch nicht als regelunbewusst im Hinblick auf das Zielsprachsystem gelten können. Vielmehr handelt es sich offenbar um ein temporäres System, das sich als charakteristisch für eine *Interlanguage* interpretieren lässt.

6. Übergreifende Diskussion

6.1 Einleitung zu den durchgeführten Studien

Das vorliegende Kapitel enthält eine umfassende Diskussion der Resultate der psycholinguistischen und phonetischen Studie, die in den Kapiteln 4 und 5 dargelegt wurden. Das Hauptziel dieses Kapitels besteht in der Zusammenfassung der Ergebnisse beider Studien in einem Gesamtüberblick.

Die vorliegenden Studien hatten die Untersuchung der Wahrnehmungs- und Produktionsfähigkeiten italienischer Deutschlerner:innen hinsichtlich der Vokale /ə/ und /y:/ zum Ziel. Die Studien fokussieren diese zwei Phoneme, da sie für italienische Deutschlerner:innen (vgl. Maturi, 2014) zu den schwierigsten Phonemen des Deutschen gezählt werden, die häufig von Assimilationsphänomenen betroffen sind und damit die Aussprache im Deutschen beeinträchtigen können (vgl. Missaglia & Sendlmeier, 1999; Missaglia, 2000; Missaglia, 2007). Durch die Studien sollte der Annahme nachgegangen werden, ob es zu einem negativen Transfer aus der L1 kommt. Dies bedeutet, dass Vokale des Deutschen wie ähnliche Vokale des Italienischen wahrgenommen bzw. realisiert werden und daher nicht zielsprachlich erkannt und produziert werden. Mechanismen, die dieses Phänomen erklären sollen sind beispielsweise der *magnet effect* (Kuhl, 1991; Kuhl & Iverson, 1995), die *equivalence classification* (Flege, 1995) oder die *phonemic equation* (Escudero, 2005). Bisherige Forschungen untersuchten insbesondere Sprachkonstellationen, in denen das Englische die Ausgangssprache war. Die Sprachkonstellation Italienisch–Deutsch ist bislang nur in begrenztem Umfang erforscht (vgl. Missaglia & Sendlmeier, 1999; Missaglia, 2000; Missaglia, 2007), sodass bisher weder akustische Daten zur Produktion deutscher Vokale durch italienische Deutschlerner:innen noch Erkenntnisse zur Wahrnehmung des /ə/-Lauts vorliegen. Bezüglich der Wahrnehmung von /y:/ durch Hörer:innen mit L1 Italienisch existieren bislang lediglich begrenzte Evidenzen (vgl. Missaglia, 2000). Das Hauptziel dieser Studien bestand daher in der Erweiterung der Kenntnisse über die Wahrnehmung und Produktion des Deutschen durch erwachsene Deutschlerner:innen mit L1 Italienisch.

6.2 Übergreifende Diskussion: Gemeinsamkeiten und Unterschiede bei der Wahrnehmung und Produktion von /ə/ und /y:/

Auf Grundlage der Ergebnisse beider Studien lässt sich feststellen, dass jede von ihnen einen jeweils unterschiedlichen, jedoch komplementären Beitrag leistet. Die Wahrnehmungsstudie ermöglicht eine Einschätzung der phonologischen Diskriminationsfähigkeit der Lerner:innen, wenn diese mit Zielstimuli (in diesem Fall deutsche Vokale) konfrontiert werden. Die Akkuratheit der

Antworten liefert Hinweise darauf, inwieweit die wahrgenommenen Lautkategorien den Zielkategorien entsprechen. Es lassen sich somit auch indirekte Rückschlüsse auf den Einfluss der L1 (Italienisch) auf die Wahrnehmung ziehen.

Die Produktionsstudie hingegen erlaubt eine Analyse sowohl der Annäherung der produzierten Vokale an die Zielkategorien des Deutschen als auch möglicher Ähnlichkeiten mit den Lautkategorien der L1. Darüber hinaus stellt sie durch die Erhebung konkreter akustischer Daten (F1-/F2-Koordinaten) eine Grundlage für die visuelle Verortung der deutschen Lautkategorien innerhalb der *Interlanguage* der Lerner:innen dar.

In den Resultaten beider Teilstudien lassen sich Parallelen erkennen, die nachfolgend detaillierter dargestellt werden.

Die vorliegenden Untersuchungen zeigen signifikante Unterschiede zwischen der Wahrnehmung und der Produktion der Kategorien /ə/ und /y:/ durch die Lerner:innen im Vergleich zu den jeweiligen Kontrollgruppen mit deutscher L1. Diese Ergebnisse bestätigen die Hypothese, dass die Ziel-Lautkategorien in beiden Modalitäten nicht erreicht wurden. In der perzeptiven Modalität zeigen sich signifikant niedrigere Akkuratheitswerte bei den Lerner:innen, während in der Sprachproduktion signifikant abweichende F1- und F2-Werte beobachtet werden, die sich nicht im zielsprachlichen Bereich befinden. Insbesondere für das /ə/ lässt sich eine ausgeprägte Resistenz gegenüber einer Annäherung an die Zielkategorie feststellen.

Im Perzeptionskontext zeigte sich, dass die Identifikation von /ə/ – im Kontrast zu ähnlichen Lauten – durch die Richtung des präsentierten Stimulus und damit durch salienzbedingte Relationen zwischen den Phonemen beeinflusst wird. Es konnte festgestellt werden, dass die Wahrnehmungsgenauigkeit signifikant sinkt, wenn die Stimuli vom Rand des Vokaltrapezes in Richtung seiner zentralen Zone verlaufen. Dieser Effekt tritt insbesondere bei der Bewegung von stärker salienten Positionen zu weniger salienten (wie z.B. /e:/ → /ə/) auf. Diese Beobachtung stützt die Prognosen der *Natural Referent Vowel*- Hypothese nach Polka & Bohn (2011). Im Rahmen des Produktionsexperiments wurde eine Tendenz der Lerner:innen beobachtet, das /ə/ artikulatorisch in Richtung periphererer Vokalräume zu verschieben. Dieser Befund steht in Einklang mit der von Missaglia und Sendlmeier (1999) beschriebenen Tendenz zur Clusterisierung in der Peripherie des Vokalraums (Dezentralisierung).

In beiden Experimenten manifestierten sich konsistente Assimilationsmuster. Auf perzeptiver Ebene wurde das deutsche /ə/ von den Lerner:innen ähnlich wie die deutschen Laute /ɛ/ und /e:/ wahrgenommen. Auf produktiver Ebene fiel ihre Realisierung des /ə/ mit derjenigen des Phonems /ɛ/, wie es von der deutschsprachigen Kontrollgruppe artikuliert wurde, zusammen. Eine Schwierigkeit zeigte sich bei den Lerner:innen auch im Umgang mit den dem /ə/ ähnlichen Kategorien /e:/ und /ɛ/, sowohl auf der Ebene der Wahrnehmung als auch auf der Ebene der Produktion. Im *Oddity-Discrimination-Task* manifestiert sich diese Diskrepanz in geringeren

Akkuratheitswerten im Kontrast /e:-ε/ im Vergleich zum Kontrast /e:-i:/. Im Produktionsexperiment sind die Realisierungen beider Kategorien – insbesondere in Bezug auf den F1-Wert – nicht zielsprachlich. Dies könnte auf einen Parallelismus hindeuten: Eine nicht-zielsprachliche Wahrnehmung geht möglicherweise mit einer nicht-zielsprachlichen Produktion einher.

Die Schwierigkeiten mit dem Laut /ə/ erwachsen aus der *Cross-Linguistic Influence* (CLI), wie sie aus dem Produktionsexperiment hervorgehen – insbesondere im Vergleich zwischen den Realisierungen der Lernerinnen und den akustisch-phonetisch ähnlichen Vokalen der italienischen Kontrollgruppe. Im Gegensatz zu /y:/, das keiner der beiden ähnlichen italienischen Kategorien (/i/ und /u/) zugeordnet wird, zeigt /ə/ Überschneidungen mit zwei L1-Kategorien: mit /ε/ in Bezug auf F2 sowie mit dem italienischen /e/ in Bezug auf F1.

In beiden Studien manifestierten sich diese phonologischen Phänomene weitgehend unabhängig von Faktoren wie der Deutschkompetenz, der Dominanz des Deutschen oder Englischen sowie der Größe des Sprachrepertoires – mit Ausnahme des leichten Einflusses der Dominanz des Deutschen auf die Realisierung der Rundung (F3). Diese Ergebnisse weisen darauf hin, dass strukturelle Diskrepanzen zwischen den jeweiligen Sprachsystemen eine zentrale Rolle in der Wahrnehmung und Produktion spielen.

Neben den bereits aufgeführten Parallelen lassen sich jedoch auch einige Diskrepanzen feststellen.

Insbesondere hinsichtlich der Wahrnehmung und Produktion von /y:/ zeigen sich divergierende Assimilationsmuster in den beiden Bereichen. Im *Oddity-Discrimination*-Task wird /y:/ tendenziell eher als ein hinterer (/u:/) als vorderer (/i:/) Vokal wahrgenommen. Die Akkuratheitswerte im Kontrast /y:-u:/ fallen demnach deutlich niedriger aus als im Kontrast /y:-i:/. Im Produktionsexperiment zeigt sich hingegen eine klare Unterscheidung (gemessen am F2-Wert) zwischen der Realisierung von /y:/ durch die Lerner:innen und den entsprechenden ähnlichen Kategorien des Deutschen (/i:/, /ɪ/, /u:/, /ʊ/), wie sie von der Kontrollgruppe produziert werden. Eine Ausnahme bildet jedoch der Laut /ʏ/, bei dem sich die Realisierung der Lerner:innen mit der der Kontrollgruppe überschneidet. Dies lässt auf eine fehlende Sensitivität gegenüber dem Merkmal der Gespanntheit schließen, welche im *Oddity-Discrimination*-Task aus methodischen Gründen nicht erkennbar ist. Diese reduzierte Sensitivität manifestiert sich zudem in anderen Kontrasten, die im Rahmen des Produktionsexperiments untersucht wurden (/i:-ɪ/, /u:-ʊ/). Die Ergebnisse stehen in Einklang mit den Beobachtungen von Missaglia und Sendlmeier (1999) und bestätigen damit Befunde aus anderen Sprachkonstellationen.

Das Produktionsexperiment offenbarte differenzierte Realisierungsmuster der beiden Laute /ə/ und /y:/. jeweils in Bezug auf die Position im Wort und auf den linken und rechten konsonantischen Kontext. Einerseits hat sich der Laut /ə/ als positionssensitiv erwiesen. Andererseits zeigt sich, dass der linke konsonantische Kontext einen besonders starken Einfluss auf die Realisierung von /y:/ hat.

Neben diesen kontextbedingten Differenzierungen bietet das Produktionsexperiment einen Mehrwert, indem es nicht nur die Analyse der Annäherung an die Zielkategorien des Deutschen, sondern auch einen direkten Vergleich zwischen den produktiven Kategorien der *Interlanguage* und jenen des Italienischen ermöglicht. Die Einbeziehung einer vergleichbaren Kontrollgruppe von italienischen L1-Sprecherinnen ermöglichte demzufolge nicht nur die Ermittlung der Annäherung der Lernerinnen an die deutsche Sprache, sondern auch die Differenzierung von ihrer L1.

Die Analyse des Produktionsexperiments ergab zudem eine signifikante Variation in der Art der Realisation von /ə/ und /y:/ in Abhängigkeit von der spezifischen Aufgabe, die im Rahmen des Experiments gestellt wurde. In stärker kontrollierten Aufgaben wird /ə/ zentralisierter und /y:/ mit ausgeprägter Rundung realisiert. Bei der Aufgabe "Wegbeschreibung", die eine Annäherung an spontane Sprache beinhaltet, war die Artikulation hingegen weniger zentral bzw. weiter vorn. Diese Ergebnisse bestätigen frühere Beobachtungen (vgl. Hahn & Siebenhaar, 2019).

Das differenzierte Bild, das sich aus den Ergebnissen des Produktionsexperiments ergibt, ergänzt die Erkenntnisse aus dem Wahrnehmungsexperiment und eröffnet vielversprechende Perspektiven für zukünftige Forschungen, insbesondere an der Schnittstelle zwischen experimenteller Linguistik und Fremdsprachendidaktik. Anlass zu dieser Überlegung gibt der im Rahmen der Analyse der Produktionsdaten identifizierte mögliche Lernmechanismus.

Zusammenfassend lässt sich festhalten, dass die beiden Studien – jeweils durch ihre spezifischen methodischen Ansätze – die zentrale Rolle sprachstruktureller Faktoren, insbesondere der intersystemischen Unterschiede zwischen den konkurrierenden Vokalsystemen, aufgezeigt haben. Diese Einflüsse konnten jedoch nicht in gleichem Maße auf die untersuchten Faktoren zu Sprachkompetenz, -dominanz und Sprachrepertoire übertragen werden. Es ist jedoch anzumerken, dass dieses Ergebnis auf die analysierten Stichproben beschränkt ist. Daher wird zu weiterführenden Untersuchungen in diesem Bereich angeregt.

6.3 Einordnung der Ergebnisse innerhalb des theoretischen Frameworks

Die im Rahmen der vorliegenden Studien gewonnenen Erkenntnisse lassen sich unmittelbar auf die im theoretischen Rahmen (Kap. 2.2.4.1 und 2.2.4.2) dargelegten Modelle beziehen, wobei die Ergebnisse sowohl theoretische Bestätigungen als auch systemimmanente Divergenzen aufzeigen.

6.3.1 Einordnung in Modelle der L2-Wahrnehmung und -Produktion

Hinsichtlich des Vokals /y:/ stützen die Ergebnisse die Full Copying Hypothesis des L2LP (Escudero, 2005). Die Tendenz der Lerner:innen, /y:/ im Wahrnehmungsexperiment primär als Hinterzungenvokal (/u:/) zu perzipieren, bestätigt die Annahme, dass das initiale Wahrnehmungssystem der L1 vollständig auf die Zielsprache übertragen wird. Die Lernenden nutzen hierbei die im italienischen System vorhandenen akustischen Grenzwerte, um den

unbekannten deutschen Stimulus zu kategorisieren. Dennoch offenbaren die Daten eine signifikante Diskrepanz zu der von Escudero postulierten spiegelbildlichen Verknüpfung von Wahrnehmung und Produktion: Während die perzeptive Diskriminierung weiterhin stark durch den L1-Filter dominiert wird, zeigt die Produktion bereits eine messbare Abweichung in Richtung der Zielkategorie. Diese empirische Beobachtung legt nahe, dass die phonologischen Repräsentationen in der *Interlanguage* für beide Modalitäten nicht zwangsläufig zeitgleich aktualisiert werden und die Produktion möglicherweise früher durch artikulatorische Adaptation optimiert wird als die perzeptiven Kategorien.

Im Falle des Schwa-Lauts verdeutlichen die Daten die Grenzen einer einfachen *Single Category Assimilation* im Sinne des PAM-L2 (Best & Tyler, 2007). Da das italienische System keine zentrale Vokalkategorie aufweist, erfolgt die Assimilation nicht an ein einzelnes Phonem-Label, sondern verteilt sich differenziert nach akustischen Dimensionen (F1 vs. F2) auf die Kategorien /e/ und /ɛ/. Im Sinne des SLM-r (Flege & Bohn, 2021) lässt sich dies darauf zurückführen, dass die *phonetic distance* zwischen dem deutschen Input und den nächstgelegenen L1-Kategorien nicht als ausreichend wahrgenommen wird, um eine neue Kategorie zu bilden. Ein abweichendes *feature weighting* führt stattdessen dazu, dass die Lerner:innen das /ə/ innerhalb des bereits existierenden L1-Vokalraums verorten, anstatt ein neues Zielareal zu erschließen. Diese "geteilte" Wahrnehmung des /ə/ erschwert zudem den im L2LP beschriebenen Lernprozess durch Fehlerkorrektur. Da kein eindeutiger *Mismatch* zu einer einzelnen L1-Kategorie vorliegt, sondern eine Ambiguität zwischen mehreren Kategorien besteht, verbleibt die Kategorisierung bei den Lerner:innen instabil. Dies erklärt die in dieser Arbeit konstatierte Resistenz des /ə/ gegenüber einer zielsprachlichen Annäherung im Vergleich zu /y:/.

Zusammenfassend lässt sich festhalten, dass die Ergebnisse die Grundannahmen der Modelle zum *Initial State* stützen, jedoch die Notwendigkeit unterstreichen, diese um Mechanismen zu erweitern, die sowohl die Modalitäten-Entkoppelung als auch die multi-kategoriale Überlappung bei mangelnden Äquivalenzstrukturen abbilden.

6.3.2 Einordnung in L3-Modelle und Ausblick auf eine integrierte LX-Phonologieforschung

Die Einordnung der Ergebnisse in die aktuellen Modelle der Drittsprachenforschung (*Third Language Studies*) erfordert eine differenzierte Betrachtung, die über eine rein komparative Analyse hinausgeht. Es ist festzustellen, dass viele dieser Ansätze, wie das *Typological Primacy Model* (TPM) oder die *L2 Status Factor Hypothesis*, in einem Forschungsdiskurs entstanden sind, der primär morphosyntaktische oder lexikalische Transferprozesse fokussierte. Vor diesem Hintergrund stellen die vorliegenden phonetischen Daten eine Herausforderung für die universelle Anwendbarkeit dieser Modelle dar, bieten jedoch gleichzeitig einen wertvollen Ausgangspunkt für

deren Erweiterung auf die phonologische Ebene. Ein wesentlicher Vorzug dieser relativ jungen Forschungsrichtung ist die Integration der äußerlinguistischen Variablen in die linguistische Analyse, was zu einer höheren ökologischen Validität der theoretischen Annahmen führt, da die Mehrsprachigkeit hier als dynamisches und kontextabhängiges System begriffen wird.

Dennoch zeigt die vorliegende Untersuchung auch die Grenzen einer direkten modelltheoretischen Übertragbarkeit auf, insbesondere da ohne eine experimentelle Erhebung der phonetischen Realisierung der L2 (Englisch) Aussagen über einen selektiven Transfer im Sinne des *Linguistic Proximity Model* (LPM) nur unter Vorbehalt getroffen werden können. Anstatt eine erzwungene Übereinstimmung zwischen den Daten und den Modellen herzustellen, verdeutlicht die Analyse vielmehr die bestehenden Schnittstellen und Diskontinuitäten zwischen *Second Language*- und *Third Language*-Ansätzen. Die beobachtete Dominanz der L1-Strukturen bei der Vokalproduktion legt nahe, dass phonetisch-artikulatorische Prozesse möglicherweise anderen Gesetzmäßigkeiten unterliegen als syntaktische Strukturen, was eine stärkere Verzahnung beider Forschungszweige erforderlich macht. Ziel dieser Einordnung ist es daher nicht nur, die Passfähigkeit der Ergebnisse zu prüfen, sondern vielmehr Impulse für eine theoretische Integration zu geben, die sowohl die kognitive Architektur der Phonologie als auch die soziolinguistische Komplexität des Drittsprachenerwerbs gleichermaßen berücksichtigt.

7. Fazit, Limitationen und Ausblick

7.1 Fazit

Die vorliegende Arbeit hatte die Intention, die Wahrnehmung und Artikulation deutscher Vokale durch Lerner:innen mit L1 Italienisch zu erforschen. Der Fokus der Untersuchung lag dabei insbesondere auf den Vokalen /ə/ und /y:/. Die Kombination aus einem Perzeptionsexperiment und einer akustischen Analyse von Produktionsdaten ermöglichte es, differenzierte Erkenntnisse über *Interlanguage*-Phänomene, *Cross-Linguistic Influence* (CLI) sowie kategoriale Resistenz im laufenden bzw. unterbrochenen Lernprozess zu gewinnen.

Die Ergebnisse zeigen, dass trotz einiger Parallelen zwischen Perzeption und Produktion deutlich unterschiedliche Assimilationsphänomene auftreten können. Während /y:/ in der Perzeption häufig mit /u:/ verwechselt wurde, wies es in der Produktion bereits ziemlich eigenständige Merkmale – obwohl nicht zielsprachlich – auf. Demgegenüber zeigte sich das /ə/ in beiden Modalitäten resistent gegenüber einer vollständigen Integration ins Zielsprachensystem. In der Produktion wurde deutlich, dass diese Resistenz nicht als bloßes Defizit zu verstehen ist, sondern vielmehr als Ausdruck einer stabilisierten, temporären *Interlanguage*-Kategorie, welche sich flexibel an die Anforderungen der LX annähert, jedoch stark durch L1-Einflüsse geprägt bleibt.

Ein weiterer relevanter Aspekt dieser Arbeit liegt in der Berücksichtigung anderer linguistischer Variablen, die sich als mitverantwortlich für die beobachteten phonetischen Muster herausstellen. Sowohl die Position des /ə/ im Wort als auch der konsonantische Kontext von /y:/ wirkten sich signifikant auf die lautliche Realisierung aus. Diese Beobachtungen unterstreichen die Notwendigkeit, phonetische Lernprozesse nicht losgelöst von sprachstrukturellen zu betrachten.

Die Faktoren Sprachkompetenz, Dominanz des Deutschen, Dominanz des Englischen sowie Sprachrepertoire zeigten im Allgemeinen keine signifikanten Effekte. Eine Ausnahme bildeten ein schwacher Interaktionseffekt zwischen der Dominanz des Englischen und dem Stimulusblock bei der Wahrnehmung von /ə/ sowie ein geringfügig signifikanter Effekt der Dominanz des Deutschen auf die Realisierung der Rundung (F3) bei /y:/. Das Ausbleiben signifikanter Effekte bei den meisten Faktoren sowie die beobachteten schwachen Effekte konnten sowohl auf methodischer als auch auf soziolinguistischer bzw. fremdsprachendidaktischer Ebene eingeordnet werden. Dies stellt einen weiteren Beitrag der vorliegenden Arbeit dar und liefert zugleich einen Anstoß, sprachwissenschaftliche Analysen künftig verstärkt durch soziolinguistische Untersuchungen zu ergänzen, um ein umfassenderes Bild des sprachlichen Verhaltens von Sprecher:innen bzw. Lerner:innen zu gewinnen. Dabei sollten nicht nur konkurrierende Sprachsysteme, sondern auch die sozialen und didaktischen Kontexte berücksichtigt werden, in denen dieses sprachliche Verhalten stattfindet.

7.2 Limitationen und Ausblick

Obwohl die durchgeführten Studien unter Berücksichtigung der verfügbaren Mittel und des zeitlichen Rahmens mit größtmöglicher Sorgfalt geplant und umgesetzt wurden, besteht Raum für die Reflexion über bestimmte Limitationen. Diese sollten insbesondere mit Blick auf mögliche zukünftige Entwicklungen berücksichtigt werden.

Ein methodischer Schwachpunkt beider Studien – sowohl der Perzeptions- als auch der Produktionsstudie – liegt in der fehlenden Erhebung und Analyse der Perzeption und Produktion in der jeweiligen L1 der Lerner:innen. Eine solche Ergänzung hätte einen noch direkteren Vergleich mit den individuellen (perzeptiven und produktiven) Kategorien der Versuchspersonen ermöglicht.

Ein weiterer kritischer Aspekt betrifft das Design des Perzeptionsexperiments, wobei eine alternative Konzeption die Validität und Interpretierbarkeit der Reaktionsdaten signifikant steigern könnte. Die Implementierung eines fixierten Schemas, wie etwa eines AAX-Paradigmas (oder AXB), wäre als methodische Alternative in Betracht zu ziehen. Ein wesentlicher Vorteil dieses Paradigmas liegt in der Reduktion der kognitiven Last: Durch die strukturelle Vorhersehbarkeit der Stimuli können Probanden ihre Aufmerksamkeitsressourcen gezielter auf die akustische Diskriminationsleistung fokussieren, was die statistische Sensitivität zur Ermittlung von Wahrnehmungsschwellen erhöhen kann. Dennoch wurde im Rahmen dieser Arbeit bewusst ein variables Stimulus-Design gewählt, um gravierenden methodischen Verzerrungen entgegenzuwirken. Fixierte Schemata bergen das Risiko eines ausgeprägten Lerneffekts sowie der Entwicklung von Antwortstrategien. Bei hoher Redundanz neigen Probanden dazu, rein akustische Mustererkennungsstrategien zu entwickeln, anstatt eine genuine linguistische Verarbeitung zu leisten. Die in dieser Studie eingesetzte Variabilität zielt darauf ab, die ökologische Validität zu wahren: Sie zwingt das auditive System zu einer kontinuierlichen Rekalibrierung, wodurch sichergestellt wird, dass die erhobenen Reaktionsdaten die tatsächliche perzeptive Kategorisierung widerspiegeln und nicht als Artefakt einer prozeduralen Automatisierung zu deuten sind.

Um potenzielle konfundierende Variablen zu minimieren, wäre die Verwendung von Stimmen desselben Geschlechts oder einer akustisch weniger markanten männlichen Stimme zielführend. Darüber hinaus bietet der Einsatz synthetischer Stimmen eine methodisch vorteilhafte Alternative. Moderne *Text-to-Speech-Systeme* (TTS), basierend auf neuronalen Architekturen, erreichen mittlerweile einen Grad an Natürlichkeit, der dem menschlichen Original nahezu gleichkommt. Der entscheidende wissenschaftliche Mehrwert liegt hierbei in der parametrischen Manipulierbarkeit: Synthetische Stimmen erlauben eine präzise Kontrolle prosodischer und akustischer Merkmale, wodurch perzeptive Verzerrungen durch idiosynkratische Sprechermerkmale systematisch ausgeschlossen werden können.

Angesichts der hohen Diversität innerhalb der untersuchten experimentellen Gruppe der Deutschlerner:innen erscheint es besonders wünschenswert, in künftigen Studien detailliertere Analysen zu den möglichen Transferdynamiken – sowohl positiver als auch negativer Art – zwischen den Repertoiresprachen und der Zielsprache durchzuführen. Eine stärker auf das Individuum fokussierte Perspektive könnte in diesem Zusammenhang einen besonders fruchtbaren Ansatz darstellen. Dafür wäre die systematische und detaillierte Erhebung der sprachlichen Profile der Teilnehmer:innen, wie sie im Rahmen dieses Projekts bereits angelegt wurde, eine zentrale Grundlage. Eine weiterführende typologische Klassifizierung der jeweiligen Sprachrepertoires könnte dabei den nächsten Schritt bilden. Auf diese Weise ließen sich potenzielle Quellen positiven oder negativen Transfers aus zuvor erworbenen Sprachen noch präziser identifizieren. Eine derart differenzierte Analyse würde wesentlich zu einem vertieften, individualisierten Verständnis phonologisch-phonetischer Lernprozesse in mehrsprachigen Kontexten beitragen.

Über die theoretische Ebene hinaus bieten die Ergebnisse dieser Arbeit konkrete Anknüpfungspunkte für die Sprachlehrforschung, insbesondere angesichts der in der aktuellen kommunikativen Methodik häufig konstatierten Marginalisierung der angewandten Phonetik (vgl. z.B. Pasunova, 2007; Breitzkreuz et al., 2001; Dieling, 2000). Da der Fokus moderner Fremdsprachendidaktik primär auf der funktionalen kommunikativen Kompetenz liegt, kommt einer systematischen und evidenzbasierten Ausspracheschulung oft eine untergeordnete Rolle zu. Hier könnten die vorliegenden Erkenntnisse zur Implementierung eines gezielten Phonetic Training beitragen.

Ein besonders vielversprechender Ansatz ist das *High Variability Phonetic Training* (HVPT, vgl. Uchihara et al., 2025). Im Gegensatz zu herkömmlichen, oft monotonen Wiederholungsübungen konfrontiert HVPT die Lerner:innen mit einer Vielzahl von Sprecher:innen und Kontexten. Dies unterstützt die kategoriale Perzeption, indem es den Lerner:innen hilft, die phonematisch relevanten Merkmale von rein idiosynkratischen akustischen Variationen zu abstrahieren. Die Integration solcher computergestützter Trainingsmodule in den DaF/DaZ-Unterricht würde es ermöglichen, die diagnostizierten *Interlanguage*-Muster proaktiv zu adressieren.

Für die Praxis bedeutet dies eine Abkehr von der bloßen imitativen Korrektur hin zu einer perzeptionsbasierten Sensibilisierung. Durch den Einsatz experimentell validierter Stimuli in digitalen Lernplattformen könnten Lehrkräfte den Lerner:innen helfen, jene ‚perzeptiven Filter‘ der L1 zu überwinden, die den Erwerb schwieriger Kontraste wie /y:/ vs. /u:/ systematisch behindern. Damit leistet die vorliegende Arbeit einen Beitrag zur Entwicklung evidenzbasierter Curricula, die phonetische Korrektheit nicht als peripheres Element, sondern als essenzielle Grundlage für soziale Integration und professionelle Teilhabe begreifen.

Ein weiterer wesentlicher Aspekt für die pädagogische Umsetzung liegt in der gezielten Förderung der metasprachlichen Kompetenz der Lerner:innen. Da moderne Lerngruppen im Bereich DaF/DaZ

zunehmend durch multilinguale Profile und heterogene Lernbiografien geprägt sind, sollte die Vermittlung phonetischer Inhalte über rein imitative Ansätze hinausgehen. Ein vielversprechender Weg ist die Implementierung eines kontrastiven phonetischen Ansatzes, der die Lerner:innen dazu befähigt, bewusste Vergleiche zwischen den Lautsystemen ihrer Erst- oder weiteren Fremdsprachen und der Zielsprache Deutsch anzustellen. Indem Lerner:innen eine kognitive Sensibilität für artikulatorische Differenzen entwickeln, können sie vorhandene sprachliche Ressourcen strategisch nutzen, um komplexe deutsche Phoneme effizienter zu erschließen. Dieser Fokus auf das Sprachbewusstsein (*language awareness*) transformiert den Lernprozess von einer unbewussten Nachahmung hin zu einer reflektierten, evidenzbasierten Aneignung. Die Integration experimenteller Daten in den Unterricht dient dabei nicht nur der Diagnose, sondern auch der Transparenz: Sie verdeutlicht den Lerner:innen die zugrunde liegenden Mechanismen ihrer eigenen *Interlanguage*. Ein solcher Ansatz fördert nicht nur die phonetische Präzision, sondern stärkt auch die Lernerautonomie in einem wissenschaftlich fundierten, modernen Fremdsprachenunterricht.

Danksagung

Ich möchte allen Menschen meinen aufrichtigen Dank aussprechen, die mich bei der Ausarbeitung dieser Dissertation unterstützt haben.

Mein besonderer Dank gilt Frau Prof. Dr. Christina Kauschke und Frau Prof. Dr. Ulrike Domahs von der Philipps-Universität Marburg sowie Frau Prof. Dr. Birgit Alber von der Freien Universität Bozen. Ein herzlicher Dank geht auch an Herrn Prof. Dr. Alessandro Vietti, der – obwohl er keine offizielle Betreuungsfunktion innehatte – meine Arbeit im Bereich der experimentellen Phonetik mit seiner wertvollen Expertise bereichert hat.

Diese Dissertation ist das Ergebnis einer fruchtbaren internationalen Zusammenarbeit, ermöglicht durch meine Einbindung in den *International Doctoral Track in Linguistics* (IDT).

Mein besonderer Dank gilt auch Herrn Prof. Dr. Stefan Rabanus, meinem Ansprechpartner an der Universität Verona im Rahmen des IDT, für seine beständige und verlässliche Unterstützung – sowohl während meines Aufenthalts in Verona als auch im gesamten Verlauf meiner Promotion.

Ich danke außerdem allen Dozierenden der Universitäten Verona und Bozen sowie der internationalen Partnerinstitutionen, die ich im Rahmen meines Forschungsaufenthalts sowie der Summer und Winter Schools kennenlernen durfte.

Mein herzlicher Dank gilt auch Herrn Prof. Dr. Marc van Oostendorp und Herrn Prof. Dr. Luis Miguel Rojas-Berscia sowie dem gesamten Kollegium der Radboud University, die mich im Rahmen meines ERASMUS+ – Praktikums mit großer Herzlichkeit und Offenheit in ihre Forschungsaktivitäten eingebunden haben.

In diesem Zusammenhang danke ich ebenfalls Herrn Prof. Dr. Marijn van't Veer und Frau Prof. Dr. Silke Hamann von der Universität Amsterdam, die meine unermüdliche Neugier für die experimentelle Phonetik zusätzlich inspiriert haben.

Mein Dank gilt darüber hinaus allen Wissenschaftlerinnen und Wissenschaftlern, die am Linguistischen Kolloquium und am Internationalen Workshop teilgenommen haben, für den bereichernden intellektuellen Austausch. Besonders in Erinnerung geblieben sind mir die Gespräche mit Frau Prof. Dr. Adele Goldberg und Herrn Prof. Em. Dr. James E. Flege.

Mein herzlicher Dank gilt auch meinen studentischen Hilfskräften, insbesondere Frau Christina Duda, für ihre Leidenschaft, Ausdauer und stetige Lernbereitschaft – und dafür, dass sie mich im letzten Jahr dieses Weges mit unermüdlicher Positivität begleitet hat.

Nicht zuletzt danke ich allen Menschen, die in verschiedenster, aber stets bedeutender Weise an dieses Projekt geglaubt und seine Umsetzung unterstützt haben.

Ein besonderer Dank gilt den Teilnehmerinnen und Teilnehmern der Experimente in Marburg und an der Universität Verona: Ich danke Ihnen nicht nur für Ihre Teilnahme, sondern auch für Ihre überraschende Begeisterung – Sie haben mir mehr zurückgegeben, als man sich von einer ersten Feldstudie je erhoffen kann. Ohne Ihren Beitrag wäre dieses Herkulesunternehmen bloß eine gute Idee geblieben.

Literatur

- Albano Leoni, F. (2006). *CLIPS. Corpora e lessici di italiano parlato e scritto (AILab UrbanECO Research Center, University of Naples Federico II)* [Dataset]. <http://www.clips.unina.it/home>
- Alber, B. (2022). *Linguistik des Deutschen-kompakt und kontrastiv*. QuiEdit.
- Atkinson, D. (1987). The mother tongue in the classroom: a neglected resource?. *ELT journal*, 41(4), 241–247.
- Bachman, L., & Palmer, A. (2010). *Language Assessment in Practice*. Oxford University Press.
- Baker, W., & Trofimovich, P. (2006). Perceptual paths to accurate production of L2 vowels: The role of individual differences. *International Review of Applied Linguistics*, 44, 231–250.
- Baker, W., Trofimovich, P., Flege, J. E., Mack, M., & Halter, R. (2008). Child–Adult Differences in Second–Language Phonological Learning: The Role of Cross–Language Similarity. *Language and Speech*, 51(4), 317–342. <https://doi.org/10.1177/0023830908099068>
- Balboni, P. E. (2014). *Didattica dell'italiano come lingua seconda e straniera*. Loescher.
- Barbour, S., & Stevenson. (1998). *Variation im Deutschen: Soziolinguistische Perspektiven*. De Gruyter Mouton.
- Bardel, C., & Falk, Y. (2007). The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research*, 23(4), 459–484.
- Bardel, C., & Falk, Y. (2012). The L2 status factor and the declarative/ procedural distinction. In A. Cabrelli, S. Flynn, & J. Rothman (Hrsg.), *Third language acquisition in adulthood* (S. 61–78). John Benjamins Publishing Company.
- Ballmer, T. T. (1981). Die Dynamik von Erwartungen und Bewertungen. In C. F. Graumann & H. G. Herrmann (Hrsg.), *Karl Bühlers Axiomatik: Verfahren, Ergebnisse und Ausblicke* (S. 171–193). Klostermann.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.
- Bassetti, B. (2008). Orthographic input and second language phonology. In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (S. 191–206). Multilingual Matters.
- Bassetti, B., & Atkinson, N. (2015). Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics*, 36, 67–91.
- Berríos, J., Swain, A., & Fricke, M. (2023). Implementing the map task in applied linguistics research: What, how, and why. *Research Methods in Applied Linguistics*, 2(3), 100081. <https://doi.org/10.1016/j.rmal.2023.100081>
- Berruto, G. (1987). *Sociolinguistica dell'italiano contemporaneo*. La Nuova Italia Scientifica.
- Berruto, G. (2008). Sul parlante nativo (di italiano). In H.–I. Radatz & R. Schlösser (Hrsg.), *Donum Grammaticorum: Festschrift für Harro Stamerjohann* (S. 1–14). Max Niemeyer Verlag.
- Berruto, G. (2012). *Sociolinguistica dell'italiano contemporaneo* (Nuova ed. aggiornata). Carocci.

- Bertinetto, P. M., & Loporcaro, M. (2005). The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. *Journal of the International Phonetic Association*, 35(2), 131–151. <https://doi.org/10.1017/S0025100305002148>
- Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Hrsg.), *Speech perception and linguistic experience: Issues in cross-language research* (S. 171–204). York Press.
- Best, C. T., & Faber, A. (2000, Juli). Developmental increase in infants' discrimination of nonnative vowels that adults assimilate to a single native vowel. In *International Conference on Infant Studies* (S. 16–19).
- Best, C. T., Singh, L., Bouchard, J., Connelly, G., Cook, A., & Faber, A. (1997, April 3). Developmental changes in infants' discrimination of non-native vowels that assimilate to two native categories. *Meeting of the Society of Research in Child Development*, Providence, RI.
- Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In *Language Experience in Second Language Speech Learning. In Honor of James Emil Flege* (S. 13–34).
- Birdsong, D. (2018). Plasticity, Variability and Age in Second Language Acquisition and Bilingualism. *Frontiers in Psychology*, 9, 81.
- Bloch, C., Kaiser, A., Kuenzli, E., Zappatore, D., Haller, S., Franceschini, R., ..., & Nitsch, C. (2009). The age of second language acquisition determines the variability in activation elicited by narration in three languages in Broca's and Wernicke's area. *Neuropsychologia*, 47(3), 625–633.
- Blommaert, J., & Backus, A. (2013). Superdiverse Repertoires and the Individual. In I. D. Saint-Georges & J.-J. Weber (Hrsg.), *Multilingualism and Multimodality* (S. 11–32). Sense Publishers. https://doi.org/10.1007/978-94-6209-266-2_2
- Boudelaa, S. (2015). The differential time course for consonant and vowel processing in Arabic: Implications for language learning and rehabilitation. *Frontiers in Psychology*, 5, 1557.
- Bozena, P., & Levy, R. (2014). The role of abstraction in non-native speech perception. *Journal of phonetics*, 46, 147–160.
- Bradlow, A. R., & Pisoni, D. B. (1999). Recognition of spoken words by native and nonnative listeners: Talker-, listener- and item-related factors. *The Journal of the Acoustical Society of America*, 106(4), 2074–2085.
- Breitkreutz, J. A., Derwing, T. M., & Rossiter, M. J. (2001). Pronunciation teaching practices in Canada. *TESL Canada Journal*, 19(1), 51–61.
- Brière, E. (1966). An investigation of phonological interference. *Language*, 42(4), 768–796.
- Browman, C., & Goldstein, L. (1992). „Targetless“ schwa: An articulatory analysis. In: Docherty G.J., Ladd D.R. (Hrsg.). *Gesture, Segment, Prosody. Papers in Laboratory Phonology* (S. 26–67). Cambridge University Press.
- Brown, G., Anderson, A., Shillock, R., & Yule, G. (1983). *Teaching talk: Strategies for production and assessment*. Cambridge University Press.
- Bußmann, H. (1990). *Lexikon der Sprachwissenschaft*. (2. völlig neu bearbeitete Auflage). Kröner (Kröners Taschenausgabe 452).
- Calabrese, A. (1987). The interaction of phonological rules and filters in Salentino. *North East Linguistics Society*, 17(1), 79–98.
- Calabrese, A. (1988). Towards a theory of phonological alphabets. MIT.

- Calvo Fernández, M. (2018). Vowel change in English and German: A comparative analysis [Thesis for the Degree in English Studies].
- Canepari, L. (1999). *MaPI. Manuale di pronuncia italiana*. Zanichelli.
- Canepari, L. (2007). *A handbook of pronunciation: English, Italian, French, German, Spanish, Portuguese, Russian, Arabic, Hindi, Chinese, Japanese, Esperanto*. Lincom Europa.
- Carbotti, S. (2015). App per l'infanzia: Linee guida per una progettazione efficace. *Form@re. Open Journal per la formazione in rete*, 5(1), 159–169.
- Cardoso, W. (2011). The development of coda perception in second language phonology: A variationist perspective. *Second Language Research*, 27(4), 433–465.
- Cavirani, E. (2015). Modelling phonologization: Vowel reduction and epenthesis in Lunigiana dialects [Dissertation]. LOT.
https://pure.knaw.nl/ws/portalfiles/portal/1582629/Full_text_dissertationB.pdf
- Cenoz, J. (2005). Learning a third language: Cross-linguistic influence and its relationship to typology and age. In B. Hufeisen & R. Fouser (Hrsg.), *Introductory L3 readings* (S. 1–9). Stauffenberg Verlag.
- Ceracchi, M. (2007). Lingua madre e lingua straniera: Le implicazioni psicoaffettive dell'apprendimento. *Studi di Glottodidattica*, 1, 19–34.
- Chan, A. Y. W. (2014). The perception and production of English speech sounds by Cantonese ESL learners in Hong Kong. *Linguistics*, 52, 35–72.
- Chilla, S., & Fox-Boyer, A. (2016). *Zweisprachigkeit, Bilingualität: Ein Ratgeber für Eltern* (2., überarbeitete Auflage). Schulz-Kirchner Verlag.
- Cho, M.-H., & Lee, S. (2023). The perception–production link varies with stages of L2 development and vowel properties. *Poznan Studies in Contemporary Linguistics*, 59(2), 315–340.
- Christen, H., Guntern, M., Hove, I., & Petkova, M. (2010). Hochdeutsch in aller Munde. Eine empirische Untersuchung zur gesprochenen Standardsprache in der Deutschschweiz. *Zeitschrift für Dialektologie und Linguistik. Beihefte*.140.
- Ciccolone, S. (2016). Italiano e tedesco in contatto: Alcune osservazioni macro- e microsociolinguistiche in Alto Adige. *Quaderns d'Italia*, 21, 27–43.
- Clements, G. N. (2003). Feature economy in sound systems. *Phonology*, 20(3), 287–333.
- Cook, V. (2002). 1. Background to the L2 User. In V. Cook (Hrsg.), *Portraits of the L2 User* (S. 1–28). Multilingual Matters.
- Crocco, C. (2017). Everyone has an accent. Standard Italian and regional pronunciation. In M. Cerruti, C. Crocco, & S. Marzo (Hrsg.), *Towards a New Standard* (S. 89–117). De Gruyter Mouton.
- D'Achille, P. (2003). *L'italiano contemporaneo*. Il Mulino.
- Dal Negro, S. (2011). Notes on the mother tongue ideology: Evidence from multilingual environments. In W. Wiater & G. Videsott (Hrsg.), *New theoretical perspectives in multilingualism research* (S. 185–202). Peter Lang.
- Darcy, I., & Krüger, F. (2012). Vowel perception and production in Turkish children acquiring L2 German. *Journal of Phonetics*, 40(4), 568–581.

- de Angelis, G. (2007a). Factors Affecting Non-native Language Influence. In G. de Angelis (Hrsg.), *Third or Additional Language Acquisition* (S. 19–40). Blue Ridge Summit: Multilingual Matters.
- de Angelis, G. (2007b). The Multilingual Learner and Speaker. In G. de Angelis (Hrsg.), *Third or Additional Language Acquisition* (S. 1–18). Blue Ridge Summit: Multilingual Matters.
- de Bot, K. (2012). Rethinking multilingual processing: From a static to a dynamic approach. In J. Cabrelli, S. Flynn, & J. Rothman (Hrsg.), *Third language acquisition in adulthood* (Bd. 46). John Benjamins Publishing.
- de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic view as a complementary perspective. *Bilingualism: Language and Cognition*, 10(1), 51–55.
- de Houwer, A. (2009). Introducing Bilingual First Language Acquisition. In C. Baker (Hrsg.), *Bilingual First Language Acquisition*. Multilingual Matters & Channel View Publications.
- Dell'Aquila, V., & Iannaccaro, G. (2006). *Survey Ladins: Usi linguistici nelle valli ladine* (Progetto SPELL). Centro di documentazione e ricerca per il plurilinguismo.
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins Publishing.
- Dewaele, J.-M. (1998). Lexical inventions: French interlanguage as L2 versus L3. *Applied Linguistics*, 19(4), 471–490.
- Dewaele, J.-M., Bak, T. H., & Ortega, L. (2021). Why the mythical “native speaker” has mud on its face. In N. Slavkov, S. Melo-Pfeifer, & N. Kerschhofer-Puhalo (Hrsg.), *The Changing Face of the “Native Speaker”* (S. 25–46). De Gruyter Mouton.
- Dieling, H., & Hirschfeld, U. (2000). *Phonetik lehren und lernen* (Fernstudieneinheit 21). Langenscheidt.
- Dietz, G. (1999). Ausspracheschulung für fortgeschrittene ungarische Deutschlerner—Ein Unterrichtskonzept. In *Jahrbuch der ungarischen Germanistik 1998* (S. 111–132). Gondolat Kiadó Kör.
- Dietz, G. (2015). Didaktik der Ausspracheschulung für französischsprachige DaF-LernerInnen. Verfügbar über <https://opus.bibliothek.uniaugsburg.de/opus4/frontdoor/index/index/docId/70290>
- Dietz, G. & Tronka, C. (1988). *Ausspracheschulung für fortgeschrittene Deutschlerner*. Enzyklopädie.
- Dizionari Garzanti Linguistica. (2025). Verfügbar über <https://www.garzantilinguistica.it/>
- Ecke, P. (2001). Lexical retrieval in a third language: Evidence from errors and tip-of-the-tongue states. In J. Cenoz, B. Hufeisen, & U. Jessner (Hrsg.), *Cross-linguistic aspects of L3 acquisition* (S. 90–114). Multilingual Matters.
- Ecke, P., & Hall, C. J. (2000). Lexikalischer Fehler in Deutsch als Drittsprache: Translexikalischer Einfluss auf 3 Ebenen der mentalen Repräsentation. *Deutsch als Fremdsprache*, 37(1) 30–36.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech Perception in Infants. *Science*, 171(3968), 303–306.
- Elsner, D. (2020). Fremdsprachen. In I. Gogolin, A. Hansen, S. McMonagle, & D. Rauch (Hrsg.), *Handbuch Mehrsprachigkeit und Bildung* (S. 39–45). Springer EV.

- Erdener, V. D., & Burnham D.K. (2005). The role of audiovisual speech and orthographic information in nonnative speech production. *Language Learning*, 55(2), 191–228.
- Escudero, P. (2009). The linguistic perception of SIMILAR L2 sounds. In P. Boersma & S. Hamann (Hrsg.), *Phonology in Perception* (S. 151–190). De Gruyter Mouton.
- Escudero, P. (2005). Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization. [Dissertation.]. Utrecht University. LOT Dissertation Series 113
- Escudero, P., & Yazawa, K. (2024). The second language linguistic perception model. In M. Amengual (Hrsg.), *The Cambridge handbook of bilingual phonetics and phonology* (S. 173–195). Cambridge University Press.
- Europarat (Hrsg.). (2010). *Common European framework of reference for languages: Learning, teaching, assessment* (10th print). Cambridge University Press.
- Fabbro, F. (2001). The Bilingual Brain: Cerebral Representation of Languages. *Brain and Language*, *The Bilingual Brain*, 79(2), 211–222.
- Favata, G. (2020). Monolinguismo versus multilinguismo e plurilinguismo: l'Italia e la Francia a confronto. *Italiano LinguaDue*, 12(1), 277–287.
- Ferrero, F. E., Magno-Caldognetto, E., Vaggies, K., & Lavagnoli, C. (1978). Some acoustic characteristics of the Italian vowels. *Journal of Italian Linguistics Amsterdam*, 3(1), 87–96.
- Flege, J. E. (1995). Second Language Speech Learning. Theory, Findings, and Problems. In W. Strange (Hrsg.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. York Press.
- Flege, J. E. (2022). A new method for eliciting speech samples for production research. Verfügbar über <https://doi.org/10.13140/RG.2.2.36660.32640>
- Flege, J. E., Aoyama, K., & Bohn, O.-S. (2021). The Revised Speech Learning Model (SLM-r) Applied. In R. Wayland (Hrsg.), *Second Language Speech Learning* (1. Aufl., S. 84–118). Cambridge University Press.
- Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America*, 106(5), 2973–2987. <https://doi.org/10.1121/1.428116>
- Flege, J. E. & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *The Journal of the Acoustical Society of America*, 70(5), 1257–1271.
- Flege, J. E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, 40(4), 467–491.
- Flemming, E., & Johnson, S. (2007). Rosa's roses: Reduced vowels in American English. *Journal of the International Phonetic Association*, 37(1), 83–96.
- Flinz, C. (2014). Mehrsprachigkeit: Ein Überblick. Konsequenzen für den DaF-Unterricht. Plurilingualism: An overview. Consequences for the teaching of German as a foreign language. In L. Ciepielewska-Kaczmarek & G. Gorąca-Sawczyk (Hrsg.), *Glottodidaktik früher, heute und morgen* (S. 15–27). Wydawnictwo Naukowe UAM.
- Flynn, S., Foley, C., & Vinnitskaya, I. (2004). The cumulative-enhancement model of language acquisition: Comparing adults' and children's patterns of development in first, second and third language acquisition of relative clauses. *International Journal of Multilingualism*, 1(1), 3–16.

- Fouser, R. (2001). Too close for comfort? Sociolinguistic transfer from Japanese into Korean as an L3. In J. Cenoz, B. Hufeisen, & U. Jessner (Hrsg.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (S. 149–169). Multilingual Matters.
- Francis, W. (1999). Cognitive integration of language and memory in bilinguals: Semantic representation. *Psychological Bulletin*, 125(2), 193–222.
- Gallardo del Puerto, F. (2007). Is L3 phonological competence affected by the learner's level of bilingualism?. *International Journal of Multilingualism*, 4(1), 1–16.
- Gallina, F. (2021). La Giornata internazionale della lingua madre e il parlante madrelingua. Verfügbar über <https://www.linguisticamente.org/la-giornata-internazionale-della-lingua-madre-e-il-parlante-madrelingua/>
- Gauci, P. (2008). Interferenza fonetica. Un'analisi degli errori più comuni commessi da parlanti di madrelingua italiana nella pronuncia dell'inglese. *Studi di glottodidattica*, 2(1), 89–111.
- Ghyselen, A.-S., Speelman, D., & Plevoets, K. (2020). Mapping the Structure of Language Repertoires: On the Use of Sociolinguistic Methods. *Zeitschrift Für Dialektologie Und Linguistik*, 87(2), 202–249.
- Glück, H. (2000). Metzler Lexikon Sprache (2. Aufl.). Springer.
- Grande Dizionario Hoepli. (2022). Verfügbar über <https://grandidizionari.hoepli.it/> , https://www.grandidizionari.it/Dizionario_Italiano/parola/M/madrelingua.aspx?query=madrelingua
- Grassi, C., Sobrero, A. A. & Telmon, T. (2003). *Fondamenti di dialettologia italiana*. Laterza.
- Gut, U. (2010). Cross-linguistic influence in L3 phonological acquisition. *International Journal of Multilingualism*, 7(1), 19–38.
- Gut, U., Kopečková, R., & Nelson, C. (2023). *Phonetics and Phonology in Multilingual Language Development* (1. Aufl.). Cambridge University Press.
- Hahn, M., & Siebenhaar, B. (2019). Schwa unbreakable – Reduktion von Schwa im Gebrauchsstandard und die Sonderposition des ostdeutschen Sprachraums. *Germanistische Linguistik*, 241, 215–236.
- Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical Evidence: A Test of the Critical-Period Hypothesis for Second-Language Acquisition. *Psychological Science*, 14(1), 31–38.
- Hammarberg, B. (2001). Roles of L1 and L2 in L3 production and acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Hrsg.), *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives* (S. 21–41). Multilingual Matters.
- Hammarberg, B., & Hammarberg, B. (2005). Re-setting the basis of articulation in the acquisition of new languages: A third-language case study. In B. Hufeisen & R. Fouser (Hrsg.), *Introductory readings in L3* (S. 11–18). Stauffenberg Verlag.
- Hammarberg, B., Williams, S., & Hammarberg, B. (1993). A study of third language acquisition. In *Problem, process, product in language learning* (S. 60–70). Stockholm University.
- Haugen, E. (1991). The „mother tongue“. In R. L. Cooper & B. J. Spolsky (Hrsg.), *The Influence of Language on Culture and Thought: Essays in Honor of Joshua A. Fishman's Sixty-Fifth Birthday* (S. 75–84). De Gruyter Mouton
- Hayes-Harb, R., & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching*, 54(3), 297–326.

- Heinz, M., & Schmid, S. (2021). *Phonetik und Phonologie des Italienischen: Eine Einführung für Studierende der Romanistik*. De Gruyter Mouton.
- Helgason, P., Kohler, K., & Kohler, K. (1996). Vowel deletion in the Kiel corpus of spontaneous speech. *Sound Patterns in Spontaneous Speech. Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel*, 30, 115–157.
- Hentschel, G. (1986). *Vokalperzeption und natürliche Phonologie: Eine kontrastive Untersuchung zum Deutschen und Polnischen*. Otto Sagner
- Hochmann, J.-R., Benavides-Varela, S., Nespor, M., & Mehler, J. (2011). Consonants and vowels: Different roles in early language acquisition. *Developmental Science*, 14(6), 1445–1458.
- Hoffmann, C. (2001). Towards a description of trilingual competence. *The International Journal of Bilingualism*, 5(1), 1–17.
- Hollmach, U. (1998). Zur Realisierung der Fortisplosive in standardsprachlichen Nachrichten und Gesprächen. In A. Biege & I. Bose (Hrsg.), *Theorie & Empirie in der Sprechwissenschaft* (S. 85–94). Dausien.
- Hove, I. (2020). The quality of schwa in Swiss German. *Zeitschrift Für Dialektologie Und Linguistik*, 87(1), 3–24.
- Hulstijn, J. H. (2011). Language proficiency in native and nonnative speakers: An agenda for research and suggestions for second-language assessment. *Language Assessment Quarterly*, 8(3), 229–249
- Il Sabatini Coletti—Dizionario della lingua italiana. (2018). Verfügbar über https://dizionari.corriere.it/dizionario_italiano/
- Isbell, D. R. (2016). The perception–production link in L2 phonology. *MSU Working Papers in Second Language Studies*, 7(1), 57–67.
- ISTAT. (2017). Anno 2015. L'uso della Lingua Italiana, dei Dialetti e delle Lingue Straniere (S. 1–22) [Report–Statistische]. Istat. <http://www.istat.it>
- Jia, G., Strange, W., Wu, Y., Collado, J., & Guan, Q. (2006). Perception and production of English vowels by Mandarin speakers: Age-related difference vary with amount of L2 exposure. *Journal of the Acoustical Society of America*, 119(2), 1118–1130.
- Kecskes, I., & Papp, T. (2000). *Foreign Language and Mother Tongue* (1. Aufl.). Psychology Press.
- Kellerman, E. (1983). Now You See It, Now You Don't. In S. Glass & L. Selinker (Hrsg.), *Language Transfer in Language Learning* (S. 112–134). Newbury House.
- Kim, C. W., & Park, S. G. (1995). Pronunciation problems of Australian students learning Korean: Intervocalic liquid consonants. *Australian Review of Applied Linguistics*, 12(1), 183–202.
- Kim, J. (2013). A longitudinal study of Korean vowel production by Chinese learners of Korean. *Phonetics and Speech Sciences*, 5(2), 71–79.
- Kisilevsky, B. S., Hains, S. M. J., Lee, K., Xie, X., Huang, H., Ye, H. H., Zhang, K., & Wang, Z. (2003). Effects of Experience on Fetal Voice Recognition. *Psychological Science*, 14(3), 220–224.
- Kisler, T., Reichel, U. D., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, 45, 326–347.

- Kleiner, S., & Knöbl, R. (2023). *Duden–Das Aussprachewörterbuch: 144.000 Aussprachen und Betonungen von Wörtern und Namen* (Bd. 6). Duden.
- Kohler, K. J., & Rodgers, J. (2001). Schwa deletion in German read and spontaneous speech. *Spontaneous German Speech: Symbolic Structures and Gestural Dynamics*, 35, 97–123. Institut für Phonetik und digitale Sprachverarbeitung, Universität Kiel.
- König, K. (2016). Erstsprache – Herkunftssprache – Muttersprache. Sprachbiographische Zugriffe von Deutsch-TürkInnen auf den Ausdruck Muttersprache. In M. Raml (Hrsg.), *Wanderer zwischen den Welten. Deutsch-türkische Sprachbiographien* (S. 269–294). Königshausen & Neumann.
- Kopečková, R., Marecka, M., Wrembel, M., & Gut, U. (2016). Interactions between three phonological subsystems of young multilinguals: The influence of language status. *International Journal of Multilingualism*, 13(4), 426–443.
- Krämer, M. (2009). *The phonology of Italian*. Oxford University Press.
- Krech, E.-M. (1982). *Großes Wörterbuch der deutschen Aussprache*. Bibliographisches Institut.
- Krech, E.-M., Stock, E., Hirschfeld, U. & Anders, L.-C. (2009). *Großes Wörterbuch der deutschen Aussprache* (1. Aufl.). De Gruyter.
- Kuhl, P. K. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, 50(2), 93–107.
- Kuhl, P. K., & Iverson, P. (1995). Linguistic experience and the “perceptual magnet effect”. In W. Strange (Hrsg.) *Speech perception and linguistic experience: Issues in cross-language research* (S. 121–154). York Press.
- Kuhl, P. K., & Miller, J. D. (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *The Journal of the Acoustical Society of America*, 63(3), 905–917.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception by infants by 6 months of age. *Science*, 255(5044), 606–308.
- Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press. <https://books.google.com/books?id=ZzYGAQAIAAJ>
- Lameli, A. (2004). *Standard und Substandard: Regionalismen im diachronen Längsschnitt* (Bd. 128). Franz Steiner Verlag.
- Lange, R., Sell, B., Terada, M., Belz, M., Mooshammer, C., & Lüdeling, A. (2024). Schwa realisation in verbal inflection in two dialogue registers of German spontaneous speech. *Zeitschrift Für Sprachwissenschaft*, 43(2), 237–266.
- Lass, R. (2009). On schwa: Synchronic prelude and historical fugue. In D. Minkova (A c. Di), *Phonological weakness in English: From old to present-day English* (S. 47–77). Houndmills: Palgrave Macmillan.
- Leather, J. (2003). Phonological Acquisition in Multilingualism. In M. del Pilar García Mayo & M. L. G. Lecumberri (Hrsg.), *Age and the acquisition of English as a foreign language* (S. 23–58). Multilingual Matters.
- Leadgeway, A. (2016). The dialects of southern Italy. In *The Oxford Guide to the Romance Languages* (S. 246–269). Oxford University Press.

- Lee, S., & Cho, M.-H. (2020). The impact of L2-learning experience and target dialect on predicting English vowel identification using Korean vowel categories. *Journal of Phonetics*, 82, 100983. <https://doi.org/10.1016/j.wocn.2020.100983>
- Lee, S., & Rhee, S.-C. (2019). A comparison of normalized formant trajectories of English vowels produced by American men and women. *Phonetics and Speech Sciences*, 11(2), 1–13.
- Lenneberg, E. H. (1967). The Biological Foundations of Language. *Hospital Practice*, 2(12), 59–67.
- Leonidou, A. D. (2020). Incertezze ortografiche e ipercorrettismo. Esempi tratti dall'italiano contemporaneo. In C. Ossenkop, G. Veldre-Gerner (Hrsg.), *Fehler-Abweichung-Variation. Sprachnormen aus linguistischer und didaktischer Sicht* (S. 149–165). Peter Lang.
- Lipinska, D. (2015). Production of L3 Vowels: Is it Possible to Separate Them From L1 and L2 Sounds? *Research in Language*, 13(1), 77–89.
- Lüttenberg, V. D. (2010). Mehrsprachigkeit, Familiensprache, Herkunftssprache. *Begriffsvielfalt und Perspektiven für die Sprachdidaktik*, 2, 299–315.
- Maddieson, I. (2013). Vowel Quality Inventories. In D. Matthew & M. Haspelmath (Hrsg.), *The World Atlas of Language Structures Online* (WALS Online (v2020.4)). Max Planck Institute for Evolutionary Anthropology. Verfügbar über <http://wals.info/chapter/2>.
- Marello, C. (2009). *L'italiano disciplinare attraverso la lingua madre dello straniero* (Bd. 1). Guerra.
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). Language Experience and Proficiency Questionnaire (LEAP-Q). *Journal of Speech, Language, and Hearing Research*, 50, 940–967.
- Maturi, P. (2014). *I suoni delle lingue, i suoni dell'italiano: Nuova introduzione alla fonetica* (3. ed). Il Mulino.
- Mayr, R., & Escudero, P. (2010). Explaining individual variation in L2 perception: Rounded vowels in English learners of German. *Bilingualism: Language and Cognition*, 13(3), 279–297.
- Medin, D. L., & Barsalou, L. W. (1987). Categorical processes and categorical perception. In S. Harnard (Hrsg.), *Categorical Perception*. Cambridge University Press.
- Ménard, L., Schwartz, J.-L. & Boë, L.-J. (2007). Articulatory-acoustic relationships during vocal tract growth for French vowels: Analysis of real data and simulations with an articulatory model. *Journal of Phonetics*, 35(1), 1–19.
- Milazzo, R. (2015). Madrelingua e italiano L2: Un'indagine su bilinguismo e personalità. *Italiano LinguaDue*, 7(2), 36–50.
- Mioni, A. M. (1976). *Fonematica delle lingue germaniche. Inglese, tedesco e neerlandese in contrasto con l'italiano*. Patron.
- Mioni, A. M. (1993). Fonetica e fonologia. In A. A. Sobrero (Hrsg.), *Introduzione all'italiano contemporaneo. Le strutture* (S. 101–139). Laterza.
- Missaglia, F. (2000). Distorted perception and production of L2-German vowels by adult Italian learners. In J. Drevillon, J. Vivier, A. Salinas (Hrsg.), *Proceedings of the VIth International Congress of the International Society of Applied Psycholinguistics (ISAPL)*, 305–314.
- Missaglia, F. (2007). Prosodic Training for Adult Italian Learners of German. The Contrastive Prosody Method. In J. Trouvain, U. Gut (Hrsg.), *Non-Native Prosody. Phonetic Description and Teaching Practice*. (S. 236–258). De Gruyter Mouton.

- Missaglia, F., & Sendlmeier, W. (1999). Die Realisierung deutscher Vokale durch italienische Muttersprachler: Eine experimentalphonetische Untersuchung. *Zeitschrift für Fremdsprachenforschung*, 10(1), 73–95.
- Möhle, D. (1989). Multilingual interaction in foreign language production. In H. Dechert & M. Raupach (Hrsg.), *Interlingual processes* (S. 179–194). Narr.
- Mora, J. C. (2021). Assessing L2 vowel production gains after high-variability phonetic training: Acoustic measurements vs. perceptual judgements. 3rd International Symposium on Applied Phonetics (ISAPh 2021), 9–18. <https://doi.org/10.21437/ISAPh.2021-2>
- Müller, N. (2006). Emerging complementizers. German in contact with French / Italian. In C. Lefevre & C. Jourdan (Hrsg.), *L2 Acquisition and Creole Genesis* (S. 145–165.). Benjamins.
- Müller, N., Kupisch, T., Schmitz, K., Cantone, K. F., & Arnaus Gil, L. (2023). *Einführung in die Mehrsprachigkeitsforschung: Deutsch – Französisch – Italienisch – Spanisch* (4. Aufl.). Gunter Narr Verlag.
- Myers-Scotton, C. (2006). *Multiple Voices: An Introduction to Bilingualism* (1. Aufl.). Wiley-Blackwell.
- Nespor, M., Peña, M., & Mehler, J. (2003). On the Different Roles of Vowels and Consonants in Speech Processing and Language Acquisition. *Lingue e linguaggio, Rivista semestrale*, 2, 203–230.
- Neufeld, G. (1988). Phonological asymmetry in second-language learning and performance. *Language Learning*, 38, 531–60.
- Neufeld, G. G. (1997). Phonological asymmetry in L2 perception and production. In J. Leather & A. James (Hrsg.), *New Sounds 97: Proceedings of the Third International Symposium on the Acquisition of Second-Language Speech* (S. 257–261). University of Klagenfurt.
- Neuhold, P. (2022). Heteroglossisch denken und experimentieren Notizen zu den Begrenzungen der Begriffe Muttersprache, *Herkunftssprache und Erstsprache*, 92, 35–39.
- Nimz, K. (2011, August 17). Vowel Perception and Production of Late Turkish Learners of L2 German. ICPhS XVII, Hong Kong.
- Nimz, K. (2015). The Perception of Vowel Quality and Quantity by Turkish Learners of German as a Foreign Language. In E. Raimy & C. E. Cairns (Hrsg.), *The Segment in Phonetics and Phonology*. John Wiley & Sons.
- Nimz, K. (2022). *Sound perception and production in a foreign language: Does orthography matter?* IDEALAB.
- Nimz, K., & Khattab, G. (2015). L2 sound perception: Does orthography matter? In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*, Glasgow, UK.
- Nimz, K., & Khattab, G. (2020). On the role of orthography in L2 vowel production: The case of Polish learners of German. *Second Language Research*, 36(4), 623–652.
- Nitsch, C. (2007). Mehrsprachigkeit: Eine neurowissenschaftliche Perspektive. In T. Anstatt (Hrsg.), *Mehrsprachigkeit bei Kindern und Erwachsenen: Erwerb, Formen, Förderung* (S. 47–68). Attempto.
- O'Brien, M. G., & Smith, L. C. (2010). Role of first language dialect in the production of second language German vowels. *International Review of Applied Linguistics in Language Teaching*, 48, 297–330.

- Odlin, T. (1989). *Language Transfer: Cross-Linguistic Influence in Language Learning*. Cambridge University Press.
- Oh, E. (2019). The relationship between “native-like” L2 vowel production and perceptual judgments enhancement by native listeners. *Linguistic Research, 언어연구*, 36(2), 241–261.
- Oh, E., & Shin, E. (2023). Effects of age of L2 acquisition and L2 experience on the production of English vowels by Korean speakers. *Phonetics and Speech Sciences*, 15(3), 9–16.
- Oksaar, E. (2003). *Zweitspracherwerb. Wege zur Mehrsprachigkeit und zur interkulturellen Verständigung*. Kohlhammer.
- Panušová, M. (2007). Phonetik in DaF-Lehrwerken. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 12(2), 1–22.
- Paradis, M. (2008). Language and communication disorders in multilinguals. In B. Stemmer & H. A. Whitaker (Hrsg.), *Handbook of the Neuroscience of Language*, (S. 341–349). Elsevier.
- Paradis, M., Goldblum, M. C., & Abidi, R. (1982). Alternate antagonism with paradoxical translation behavior in two bilingual aphasic patients. *Brain and Language*, 15(1), 55–69.
- Paschke, P. (2009). Untersuchungen zur Kodifizierung der Standardaussprache in Deutschland. *Informationen Deutsch als Fremdsprache*, 36(2–3), 207–212.
<https://doi.org/10.1515/infodaf-2009-2-338>
- Pätzold, M., & Simpson, A. P. (1997). Acoustic analysis of German vowels in the Kiel Corpus of Read Speech. In *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung Universität Kiel* (Bd. 32, S. 215–247).
- Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. f., Dupoux, E., ..., & Mehler, J. (1996). Brain processing of native and foreign languages. *NeuroReport*, 7(15), 2439–2444.
- Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, 24(2), 175–184.
- Polivanov, E. (1969). La perception des sons d’une langue étrangère. *Le Cercle de Prague* (Change 3), 111–114.
- Polka, L. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *The Journal of the Acoustical Society of America*, 97(2), 1286–1296.
- Polka, L., & Bohn, O.-S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *The Journal of the Acoustical Society of America*, 100(1), 577–592.
- Polka, L., & Bohn, O.-S. (2003). Asymmetries in vowel perception. *Speech Communication*, 41(1), 221–231.
- Polka, L., & Bohn, O.-S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, 39(4), 467–478.
- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 20(2), 421–435.
- Polka, L., Bohn, O.-S., & Molnar, M. (2005). Natural referent vowels guide the development of vowel perception. *The Journal of the Acoustical Society of America*, 117(4_Supplement), 2398.
- Polka, L., Molnar, M., Zhao, T. C., & Masapollo, M. (2021). Neurophysiological Correlates of Asymmetries in Vowel Perception: An English–French Cross-Linguistic Event-Related Potential Study. *Frontiers in Human Neuroscience*, 15, 607148.

- Predeck, K. (2022). *Vowel Length in German: Use of Quality and Quantity in the Perception of Long and Short Vowels in German*. California Digital Library.
- Propp, V. (1976). Study of the Folktale: Structure and History. *Dispositio*, 1(3), 277–292.
- Puig-Mayenko, E., Cunnings, I., Bayram, F., Miller, D., Tubau, S., & Rothman, J. (2018). Language Dominance Affects Bilingual Performance and Processing Outcomes in Adulthood. *Frontiers in Psychology*, 9, 1–16.
- Ricci Garotti, F. (2017). Deutsch als dritte Sprache für bilinguale Lernende. In M. C. Moroni & F. Ricci Garotti (Hrsg.), *Brücken schlagen zwischen Sprachwissenschaft und DaF-Didaktik* (S. 95–130). Peter Lang.
- Richter, R. (2011). Ausspracheabweichungen chinesischer Deutschlerner und ihre kommunikative Relevanz. *Zeitschrift für interkulturellen Fremdsprachenunterricht*, 16(2), 176–184
- Riehl, C. M. (2014). *Sprachkontaktforschung* (3., überarbeitete Auflage). Attempto.
- Ringbom, H. (1987). *The role of the mother tongue in foreign language learning*. Multilingual Matters Limited.
- Ringbom, H. (2002). Levels of transfer from L1 and L2 in L3 acquisition. In J. Ytsma & M. Hooghiemstra (Hrsg.), *Proceedings of the Second International Conference on Trilingualism*. Fryske Academie.
- Ringbom, H. (2005). L2-transfer in third language acquisition. In B. Hufeisen & R. Fouser (Hrsg.), *Introductory L3 readings* (S. 71–82). Stauffenberg Verlag.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605.
- Rothman, J. (2011). L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second language research*, 27(1), 107–127.
- Salmons, J. (2012). *History of German: What the past reveals about today's language*. Oxford University Press.
- Savoia, L., & Maiden, M. (1997). Vowel Systems. In M. Maiden & M. Parry (Hrsg.), *The dialects of Italy*. Routledge.
- Schiel, F. (1999). Automatic phonetic transcription of non-prompted speech. In *Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS 1999)* (S. 607–610). San Francisco, CA.
- Schiel, F. (2015). A statistical model for predicting pronunciation. In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*. Glasgow, UK.
- Schirmunski, V. M. (mit Wiesinger, P.). (2010). *Deutsche Mundartkunde: Vergleichende Laut- und Formenlehre der deutschen Mundarten* (L. Naiditsch, Hrsg.). Peter Lang.
- Schmidt, J. E., & Herrgen, J. (2011). *Sprachdynamik: Eine Einführung in die moderne Regionalsprachenforschung*. Erich Schmidt Verlag.
- Schneider, S. (2015). *Bilingualer Erstspracherwerb*. Ernst Reinhardt Verlag.
- Schultheiss, L. K. G. (2008). *Cross-Language Perception of German Vowels by Speakers of American English*. BYU ScholarsArchive.
- Schwartz, J.-L., Abry, C., Boë, L.-J., Ménard, L., & Vallée, N. (2005). Asymmetries in vowel perception, in the context of the Dispersion–Focalisation Theory. *Speech Communication*, 45(4), 425–434.

- Seid, Y. (2019). The impact of learning first in mother tongue: Evidence from a natural experiment in Ethiopia. *Applied Economics*, 51(6), 577–593.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10, 209–241.
- Selinker, L., & Gass, S. M. (Hrsg.). (1983). *Language Transfer in Language Learning*. John Benjamins Pub. Co.
- Sendlmeier, W. F. (1981). Der einfluß von qualität und quantität auf die perzeption betonter vokale des deutschen. *Phonetica*, 38(5–6), 291–308.
- Sendlmeier, W. F., & Seebode, J. (2006). Formantkarten des deutschen Vokalsystems. Verfügbar über https://www.kw.tuberlin.de/fileadmin/a01311100/Formantkarten_des_deutschen_Vokalsystems_01.pdf
- Sharwood-Smith, M. (1994). *Second language learning: Theoretical foundations*. Longman
- Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, 3(3), 243–261.
- Siebenhaar, B., & Hahn, M. (2019). Vowel Space, Speech Rate and Language Space. In *Proceedings of the 19th International Congress of Phonetic Sciences*. Melbourne, Australia.
- Smith, L. C., & Baker, W. (2011). L2 Dialect Acquisition of German Vowels: The Case of Northern German and Austrian Dialects. *Poznań Studies in Contemporary Linguistics*, 47(1), 120–132.
- Sortieren und erzählen: Bildergeschichten für 3- 8-Jährige—Jahreszeiten*. (2016). Verlag an der Ruhr.
- Strange, W., Levy, E., & Law II, F. F. (2009). Cross-language categorization of French and German vowels by naïve American listeners. *Journal of the Acoustical Society of America*, 126(3), 1461–1476.
- Stratton, J. M. (2025). The effects of production training on speech perception in L2 learners of German. *Journal of Phonetics*, 108, 1–15.
- Thon, W., & Dommelen, W. v. (1992). PHONDAT90: Rechnerverarbeitbare Sprachaufnahmen eines umfangreichen Korpus des Deutschen. In K. J. Kohler (Hrsg.), *Phonetisch-akustische Datenbasis des Hochdeutschen: Kieler Arbeiten zu den PHONDAT-Projekten 1989–1992*, (Bd. 26, S. 41–79). Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel.
- Treffers-Daller, J. (2019). What Defines Language Dominance in Bilinguals? *Annual Review of Linguistics*, 5(1), 375–393. <https://doi.org/10.1146/annurev-linguistics-011817-045554>
- Triulzi, M., Winter, C., & Maahs, I. M. (2023). Erstsprache, Muttersprache, Herkunftssprache oder Familiensprache?. *Universitäts- und Landesbibliothek Darmstadt*, 57–87. Verfügbar über <https://doi.org/10.48694/ZIF.3644>
- Trofimovich, P., Baker, W., & Mack, M. (2001). Context- and experience-based effects on the learning of vowels in a second language. *Studies in the Linguistic Sciences*, 31(2), 167–186.
- Trumper, J. (1997). Calabria and southern Basilicata. In M. Maiden & M. Parry (Hrsg.), *The dialects of Italy*. Routledge.

- Tsukada, K., Birdsong, D., Mack, M., Sung, H., Bialystok, E., & Flege, J. (2005). Release Bursts in English Word-Final Voiceless Stops Produced by Native English and Korean Adults and Children. *Phonetica*, 61(2–3), 67–83.
- Tversky, A., & Gati, I. (1978). Studies of similarity. In E. Rosch & B. Bloom Lloyd, *Cognition and Categorization*. Lawrence Elbaum Associate. Verfügbar über <https://pages.ucsd.edu/~scoulson/203/tvgati.pdf>
- Uchihara, T., Karas, M., & Thomson, R. I. (2025). High variability phonetic training (HVPT): A meta-analysis of L2 perceptual training studies. *Studies in Second Language Acquisition*, 47(3), 794–827.
- van den Noort, M., Struys, E., Kayoung, K., Bosch, P., Mondt, K., van Kralingen, R., Lee, M., & van de Craen, P. (2013). Multilingual processing in the brain. *International Journal of Multilingualism*, 11(2), 182–201.
- van Dijk, M., & van Geert, P. (2007). Wobbles, humps and sudden jumps: A case study of continuity, discontinuity and variability in early language development. *Infant and Child Development*, 16(1), 7–33.
- Vanhove, J. (2013). The Critical Period Hypothesis in Second Language Acquisition: A Statistical Critique and a Reanalysis. *PLoS ONE*, 8(7), e69172. Verfügbar über <https://doi.org/10.1371/journal.pone.0069172>
- Verspoor, M., Lowie, W., & Bot, K. D. (2008). 3. Input and Second Language Development from a Dynamic Perspective. In T. Piske & M. Young-Scholten (Hrsg.), *Input Matters in SLA* (S. 62–80). Multilingual Matters.
- Videsott, G., Herrnberger, B., Hoenig, K., Schilly, E., Grothe, J., Wiater, W., Spitzer, M. & Kiefer, M. (2010). Speaking in multiple languages: Neural correlates of language proficiency in multilingual word production. *Brain and Language*, 113(3), 103–112.
- Vignuzzi, U. (2010). Aree linguistiche. In Enciclopedia Treccani. Zuletzt abgerufen am 18.05.2025 über [https://www.treccani.it/enciclopedia/aree-linguistiche_\(Enciclopedia-dell%27Italiano\)/](https://www.treccani.it/enciclopedia/aree-linguistiche_(Enciclopedia-dell%27Italiano)/)
- Vildomec, V. (1963). *Multilingualism: General Linguistics and Psychology of Speech*. Sijthoff.
- Vogel, T. (1992). Englisch und Deutsch gibt es immer Krieg, Sprachverarbeitungsprozesse beim Erwerb des Deutschen als Drittsprache. *Zielsprache Deutsch*, 23(2), 95–99.
- Weinberger, S. (2015). *Speech Accent Archive* [Dataset]. Verfügbar über <http://accent.gmu.edu>
- Westergaard, M., Mitrofanova, N., Mykhalyk, R., & Rodina, Y. (2017). Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism*, 21(6), 666–682.
- Wiese, R. (2000). *The Phonology of German*. Oxford University Press.
- Williams, S., & Hammarberg, B. (1998). Language switches in L3 production: Implications for a polyglot speaking model. *Applied Linguistics*, 19(3), 295–333.
- Wrembel, M. & Cabrelli Amaro, J. (Hrsg.). (2018). *Advances in the Investigation of L3 Phonological Acquisition*. Routledge.
- Wurzel, W. U. (1981). Phonologie: Segmentale Struktur. In K. E. Heidolph, W. Flämig & W. Motsch (Hrsg.), *Grundzüge einer deutschen Grammatik* (S. 898–990). Akademie Verlag.
- Yadav, D. M. K. (2014). Role of Mother Tongue in Second Language Learning. *International Journal of Research*, 1(11), 572–582.

Zimmerer, F., & Reetz, H. (2011). Same or different? Schwas in naturally spoken German. In *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII)* (S. 2324–2327), Hong Kong, China, 17–21 August 2011.

Zimmerer, F., & Trouvain, J. (2015). Perception of French speakers' German vowels. In *Proceedings of the 16th Annual Conference of the International Speech Communication Association (Interspeech 2015)*, Dresden, Germany, September 6–10, 2015. Verfügbar über <https://doi.org/10.21437/Interspeech.2015-395>

Software

Audacity(R) (Version 3.4.2). (2023). [Software]. <http://audacity.sourceforge.net/>

Boersma, P., & Weenink, D. (2024). *PRAAT, doing phonetics by computer* (Version 6.4.22) [Software]. <https://www.fon.hum.uva.nl/praat/>

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51, 195–203.

The jamovi project. (2025). *Jamovi (Version 2.6)* [Software]. <https://www.jamovi.org>

Plagiarismuserklärung (Versicherung gemäß § 10 Abs. 1 d)

Hiermit versichere ich, dass ich die vorgelegte Dissertation mit dem Titel

*Dynamik und Stabilität der Repräsentationen im Sprachkontakt Italienisch-Deutsch:
eine psycholinguistische und phonetische Analyse*

selbstständig, ohne unerlaubte Hilfe Dritter angefertigt und andere als die in der Dissertation angegebenen Hilfsmittel nicht benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen sind, habe ich als solche kenntlich gemacht. Dritte waren an der inhaltlich-materiellen Erstellung der Dissertation nicht beteiligt; insbesondere habe ich hierfür nicht die Hilfe eines Promotionsberaters in Anspruch genommen. Kein Teil dieser Arbeit ist in einem anderen Promotions- oder Habilitationsverfahren verwendet worden. Mit dem Einsatz von Software zur Erkennung von Plagiaten bin ich einverstanden.

Marburg, den 14.07.2026
(Ort, Datum)


Alexia Despina Leonidou
(Unterschrift mit Vor- und Zuname)

Appendix

Stimuli für das Perzeptionsexperiment

Die zwei Versionen der Stimulidantesätze
(Block /ə/ + Block /y:/; Block /y:/ + Block /ə/)
für das Perzeptionsexperiment

Trial	Abfolge Stimuli (Block 1, Block 2)
Testphase	Hase, Hose, Hase
Testphase	Bone, Biene Biene
Testphase	eee
Testphase	uaa
Testphase	iuu
Testphase	aea
1	iia
2	əəə
3	iəə
4	əeə
5	ɛɛə
6	eee
7	eee
8	ɛiɛ
9	əəə
10	eee
11	eee
12	əeə
13	əəə
14	ɛii
15	iii
16	eeə
17	əəə
18	eii
19	iii
20	ɛɛɛ
21	ɛɛɛ
22	eɛe
23	əəə
24	əəe
25	əəɛ
26	əəi
27	eee
28	ɛɛɛ
29	ɛɛɛ

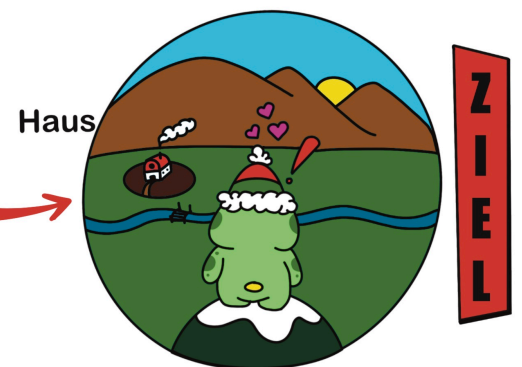
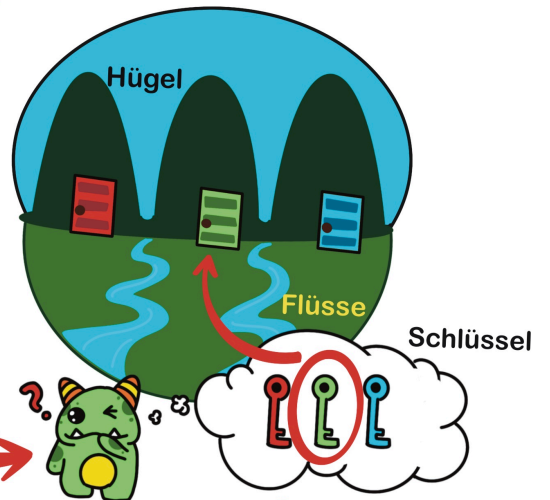
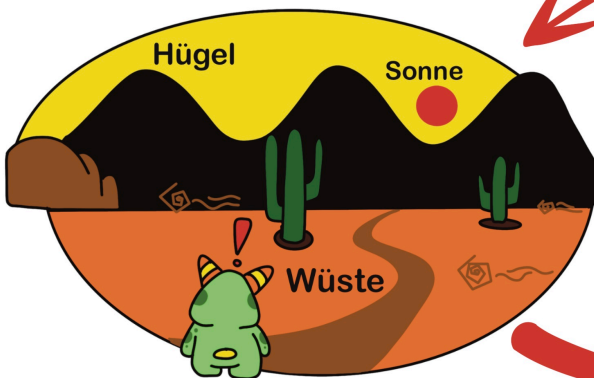
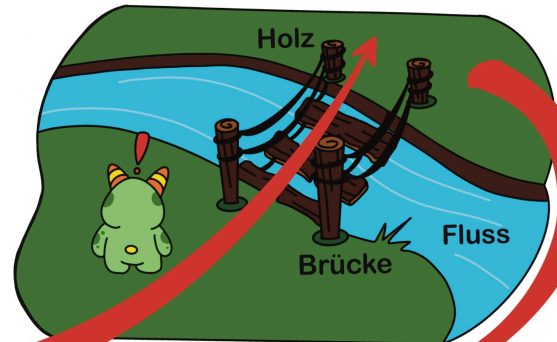
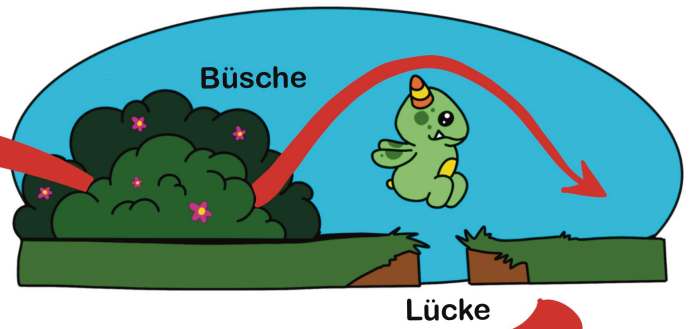
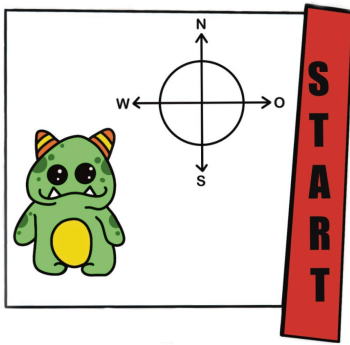
30	ɛeɛ
31	iii
32	iee
33	ɛee
34	ɛɛɛ
35	iɛɛ
36	eəe
37	iɛi
38	eee
39	iie
40	iii
41	eɛɛ
42	eei
43	əii
44	ɛɛɛ
45	əəə
46	ɛəɛ
47	iii
48	iii
49	iie
50	uuu
51	uiu
52	uyy
53	eie
54	eye
55	uuu
56	iei
57	eei
58	yey
59	yyy
60	yee
61	eee
62	uuu
63	ygi
64	iii
65	yyy
66	uee
67	uuu
68	iyy
69	yyy
70	iii
71	yyu
72	iii

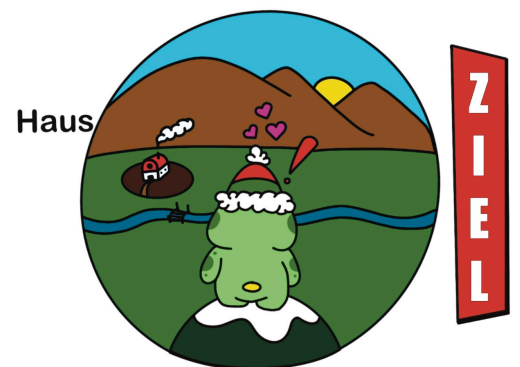
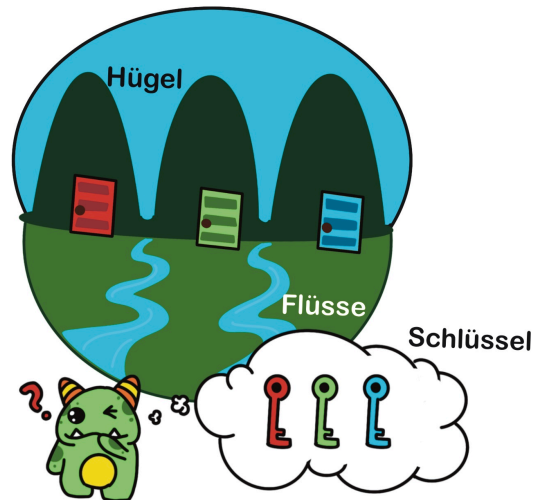
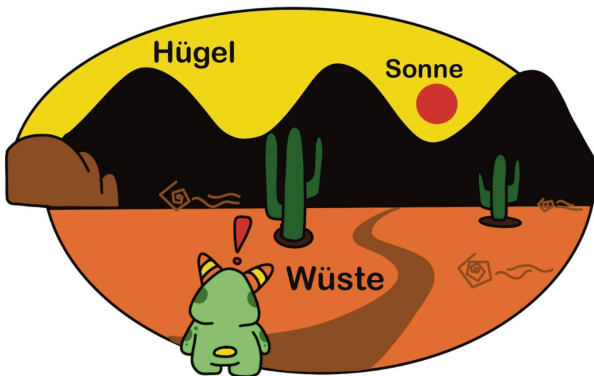
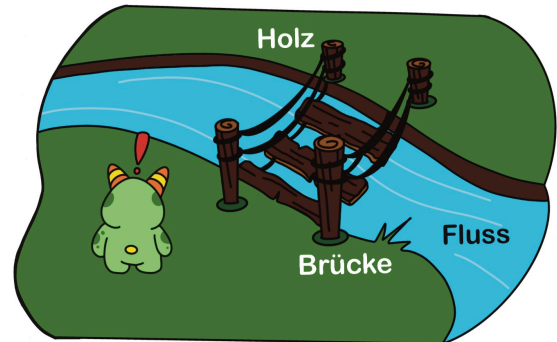
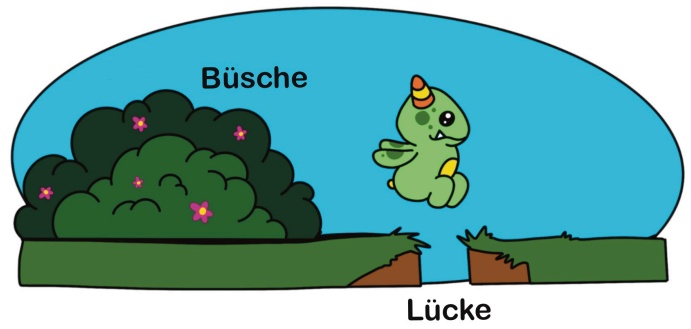
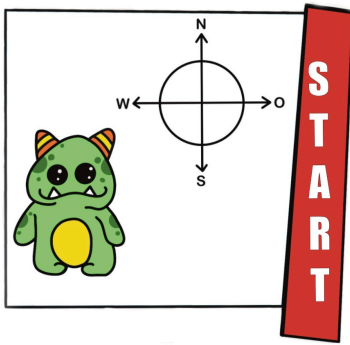
73	iii
74	uii
75	iyi
76	uuu
77	eee
78	euu
79	iui
80	eee
81	yiy
82	uuy
83	uue
84	uuu
85	yyy
86	eeu
87	yyy
88	eyy
89	eee
90	yuu
91	iii
92	iii
93	eee
94	iuu
95	yyy
96	eee

Trial	Abfolge Stimuli (Block /y:/, Block /ə/)
Testphase	Hase, Hose, Hase
Testphase	Bone, Biene Biene
Testphase	eee
Testphase	uaa
Testphase	iuu
Testphase	aea
1	iie
2	uuu
3	uiu
4	uyy
5	eie
6	eye
7	uuu
8	iei
9	eei
10	yey
11	yyy
12	yee
13	eee

14	uuu
15	y yi
16	iii
17	yyy
18	uee
19	uuu
20	i iy
21	yyy
22	iii
23	y yu
24	iii
25	iii
26	uii
27	i yi
28	uuu
29	eee
30	euu
31	i ui
32	eee
33	y iy
34	u uy
35	u ue
36	uuu
37	yyy
38	eeu
39	yyy
40	eyy
41	eee
42	yuu
43	iii
44	iii
45	eee
46	iuu
47	yyy
48	eee
49	i iə
50	ə əə
51	i iə
52	ə əə
53	ɛ ɛə
54	eee
55	eee
56	ɛ iɛ
57	ə əə
58	eee
59	eee

60	əɛə
61	əəə
62	ɛii
63	iii
64	eeə
65	əəə
66	ɛii
67	iii
68	ɛɛɛ
69	ɛɛɛ
70	eeɛ
71	əəə
72	əəɛ
73	əəɛ
74	əəi
75	eee
76	ɛɛɛ
77	ɛɛɛ
78	ɛeɛ
79	iii
80	iee
81	ɛee
82	ɛɛɛ
83	iɛɛ
84	eəɛ
85	iɛi
86	eee
87	iie
88	iii
89	eɛɛ
90	eei
91	əii
92	ɛɛɛ
93	əəə
94	ɛəɛ
95	iii
96	iii





Geschichten für die Aufgabe "Nacherzählung"

Geschichte 1: Stürmischer Spaziergang

Neele geht wie üblich mit ihrer Hündin spazieren. Es ist Herbst. Ein stürmischer Wind weht und bläst die bunten Blätter durch die Luft. Ein älterer Herr mit schickem neuem Hut und grüner Jacke läuft ihnen über den Weg. Neele und ihre Hündin beobachten, wie ein plötzlicher Windstoß seinen Hut erfasst und ihn hoch über seinen Kopf hinwegpustet. Kurzerhand rennt Neeles Hündin los, um den fliegenden Hut wieder einzufangen. Sie hüpfte hoch in die Luft und schafft es, sich den Hut zu schnappen. Stolz über den Erfolg geben Neele und die Hündin dem Mann mit der grünen Jacke seinen Hut zurück. Ein wenig betrübt darüber, dass der Hut durch den Wind und den Hundebiss nun überhaupt nicht mehr neu aussieht, nimmt er ihn entgegen. Trotzdem ist der Mann mit der grünen Jacke glücklich, seinen Hut wiederzuhaben. „Danke für euren Einsatz“, sagt er zu Neele und ihrer Hündin und setzt seinen Spaziergang im trüben, stürmischen Herbstwetter fort. „Bitte, gerne!“, ruft Neele noch hinterher. Müde vom Kampf gegen den Wind machen sich auch Neele und ihre Hündin auf den Weg zurück nach Hause.

Geschichte 2: Turbulente Schlittenfahrt

Syleiman und Pauline wollen Schlitten fahren. Mit einem Holzschlitten im Schlepptau stapfen sie durch den verschneiten Wald. Sie sind warm eingepackt; ihre Mützen und Handschuhe schützen sie vor dem kühlen Wind. Syleiman zieht den Schlitten den Berg hinauf. Sobald sie oben angekommen sind, setzen sie sich auf den Schlitten – Syleiman vorne, Pauline hinten. Los geht die Fahrt. Die beiden düsen über den Schnee. „Pass auf, da kommt ein Baum!“, ruft Pauline plötzlich. Doch Syleiman schafft es nicht rechtzeitig, den Schlitten umzulenken. Mit voller Wucht krachen sie gegen den Baum. Zum Glück hat sich niemand verletzt, aber der Baum ist zerbrochen. Syleiman rückt seine verrutschte Mütze zurecht und prüft, ob der Schlitten noch zu gebrauchen ist. „Keine Chance, der ist hinüber“, stellt er betrübt fest. „Wie sollen wir jetzt den Berg hinunterkommen?“, überlegen die beiden. Da hat Pauline eine Idee. Anstelle des Schlittens kann man sich doch auch auf ein abgebrochenes Stück des Baumstamms setzen! Gesagt, getan. Und tatsächlich, die Idee funktioniert überraschend gut! Auf dem Baumstamm sitzend düsen die beiden weiter den Berg hinunter. Die Schlittenfahrt ist gerettet.

Akustische Daten zu den Realisierungen der deutschen Vokale durch die Lerner:innen („Lernende“) und die Kontrollgruppe der Sprecher:innen mit L1 Deutsch („Ko_DEU“):

Lerner:innen des Deutschen ("Lernende")

Formanten (Hz)	Vokal	Anzahl	Median	IQR	Geschlecht
F1	a	272	795	164.7	W
	a	126	678	136.7	M
	a:	104	791	159.8	W
	a:	53	686	91.8	M
	e:	109	471	113.1	W
	e:	45	469	91.6	M
	ɛ	143	593	161.4	W
	ɛ	63	500	99.6	M
	ɪ	192	410	82.3	W
	ɪ	102	352	62.3	M
	i:	204	389	82.3	W
	i:	76	334	78.3	M
	o:	72	459	62.9	W
	o:	48	450	78.2	M
	ʊ	253	442	132.4	W
	ʊ	83	377	93.8	M
	u:	154	413	80	W
	u:	56	346	45.8	M
	ʏ	275	398	86.9	W
	ʏ	143	345	58.2	M
	y:	346	399	77	W
	y:	152	340	62.2	M
	ə	1532	557	157.1	W
	ə	752	504	96.3	M
F2	a	272	1507	221.5	W
	a	126	1313	152.5	M
	a:	104	1431	225.4	W
	a:	53	1290	146.4	M
	e:	109	2333	291	W
	e:	45	1938	250.9	M
	ɛ	143	2048	315.3	W
	ɛ	63	1816	233.6	M
	ɪ	192	2349	444.4	W
	ɪ	102	2026	258.7	M
	i:	204	2383	456.2	W
	i:	76	2077	235.4	M
	o:	72	949	305.1	W
	o:	48	850	187.9	M
	ʊ	253	1143	666.6	W
	ʊ	83	933	417.5	M
	u:	154	1032	631.8	W
	u:	56	924	279.5	M
	ʏ	275	1726	512	W

	Y	143	1580	498.1	M
	y:	346	1711	537	W
	y:	152	1637	308.6	M
	ə	1532	1949	283.2	W
	ə	752	1679	226.4	M
F3	a	272	2719	399.5	W
	a	126	2365	591.6	M
	a:	104	2584	309.2	W
	a:	53	2340	375.6	M
	e:	109	3041	287.8	W
	e:	45	2629	339.1	M
	ɛ	143	2929	319.8	W
	ɛ	63	2551	489.2	M
	ɪ	192	3060	404.9	W
	ɪ	102	2709	404.5	M
	i:	204	3132	466.5	W
	i:	76	2939	576	M
	o:	72	2685	316.7	W
	o:	48	2424	502.1	M
	ʊ	253	2643	361.9	W
	ʊ	83	2489	379.6	M
	u:	154	2719	355.2	W
	u:	56	2402	332.2	M
	Y	275	2666	370.8	W
	Y	143	2327	357.8	M
	y:	346	2536	280.7	W
	y:	152	2218	321.8	M
	ə	1532	2878	279.1	W
	ə	752	2511	444.1	M

Kontrollgruppe Sprecher:innen mit L1 Deutsch ("Ko_DEU")

Formanten (Hz)	Vokal	Anzahl	Median	IQR	Geschlecht
F1	a	125	792	92.4	W
	a	12	702	50.3	M
	a:	23	788	115.7	W
	a:	5	585	103.5	M
	e:	245	454	79.6	W
	e:	46	373	45.2	M
	ɛ	120	576	98.3	W
	ɛ	49	485	83.5	M
	ɪ	73	457	117.4	W
	ɪ	30	366	103.4	M
	i:	48	426	176.5	W
	i:	11	276	27.9	M
	o:	49	424	48.4	W
	o:	6	380	19.7	M
	ʊ	231	498	102.3	W

	ʊ	59	385	73.1	M
	u:	70	403	78.3	W
	u:	32	298	72.8	M
	Y	149	457	81.5	W
	Y	25	344	57.6	M
	y:	226	352	101.5	W
	y:	46	303	91.1	M
	ə	518	557	146.8	W
	ə	126	439	89	M
F2	a	125	1395	228.8	W
	a	12	1259	144	M
	a:	23	1441	185.3	W
	a:	5	1260	28.6	M
	e:	245	2204	354.2	W
	e:	46	1935	240.9	M
	ɛ	120	1923	304.5	W
	ɛ	49	1768	204.7	M
	ɪ	73	1941	376.9	W
	ɪ	30	1861	254.2	M
	i:	48	2417	521.6	W
	i:	11	2057	172.7	M
	o:	49	717	202.8	W
	o:	6	626	222.5	M
	ʊ	231	1184	347.4	W
	ʊ	59	990	477.3	M
	u:	70	1121	684.2	W
	u:	32	996	465.8	M
	Y	149	1635	182.4	W
	Y	25	1617	136.7	M
	y:	226	1852	305	W
	y:	46	1617	277.1	M
	ə	518	1505	349.2	W
	ə	126	1488	320.4	M
F3	a	125	2718	239.9	W
	a	12	2331	197.2	M
	a:	23	2832	330.7	W
	a:	5	2392	235.4	M
	e:	245	2901	388.3	W
	e:	46	2593	317.1	M
	ɛ	120	2806	303.7	W
	ɛ	49	2538	208.3	M
	ɪ	73	2884	213.8	W
	ɪ	30	2585	167.3	M
	i:	48	3402	397.3	W
	i:	11	2820	568.3	M
	o:	49	2861	277.9	W
	o:	6	2412	59.8	M
	u:	70	2662	330.8	W

u:	32	2331	330.9	M
ʊ	231	2737	269.7	W
ɪ	59	2378	207.1	M
ʏ	149	2571	246.4	W
ʎ	25	2308	261.7	M
y:	226	2480	235	W
y:	46	2287	485.1	M
ə	518	2848	331.8	W
ə	126	2429	195.7	M

Akustische Daten zu den Realisierungen der italienischen Vokale durch die Kontrollgruppe der Sprecher:innen mit L1 Italienisch:

Kontrollgruppe Sprecher:innen mit L1 Italienisch ("Ko_ITA")

Kontingenzgruppe 1: Deutschsprachige mit L1-Katholisch (N=171)						
Formanten (Hz)	Vokal	Anzahl	Median	IQR	Geschlecht	
F1	a	1115	800	161.6	W	
	a	521	654	119.6	M	
	e	793	541	114.4	W	
	e	403	465	93.6	M	
	ɛ	98	668	139.8	W	
	ɛ	72	550	68.7	M	
	i	579	423	72.5	W	
	i	238	328	60.1	M	
	o	881	563	144	W	
	o	395	477	88	M	
	ɔ	80	715	108.6	W	
	ɔ	36	577	72.5	M	
	u	177	444	93.2	W	
	u	102	391	85.2	M	
	F2	a	1115	1589	257.1	W
		a	521	1374	156.9	M
e		793	2163	313	W	
e		403	1796	290	M	
ɛ		98	2064	303.2	W	
ɛ		72	1730	229.3	M	
i		579	2522	299.8	W	
i		238	2062	224.4	M	
o		881	1257	317.1	W	
o		395	1137	226.2	M	
ɔ		80	1220	177.6	W	
ɔ		36	1075	168.2	M	
u		177	1175	438.9	W	
u		102	994	362	M	
F3		a	1115	2707	381.3	W
		a	521	2481	241.2	M
	e	793	2927	286.4	W	
	e	403	2577	227.8	M	

ε	98	2841	226.9	W
ε	72	2603	182.3	M
i	579	3088	358.1	W
i	238	2620	265.8	M
o	881	2814	307.3	W
o	395	2462	253.4	M
ɔ	80	2691	453.9	W
ɔ	36	2364	200.8	M
u	177	2800	268	W
u	102	2484	207.3	M

Northwestern Bilingualism & Psycholinguistics Research Laboratory

Marian, Blumenfeld, & Kaushanskaya (2007).

Übersetzung von Larissa Weigel & Monica Gonzalez-Marquez (2008); Adaptation by Marilyn Hall

Fragebogen zu Ermittlung von Sprachkenntnissen (LEAP-Q)

Versuchspersonennummer				Datum	
Alter		Geburtsdatum		Männlich ☐	Weiblich ☐

(1) Listen Sie alle Sprachen, die Sie beherrschen, **in der Reihenfolge Ihres Kenntnisstands**, auf:

1	2	3	4	5
---	---	---	---	---

(2) Listen Sie alle Sprachen, die Sie beherrschen, **in der Reihenfolge nach Zeitpunkt des Erlernens**, auf (beginnend mit Ihrer Muttersprache):

1	2	3	4	5
---	---	---	---	---

(3) Geben Sie an, in welchem Maß Sie *derzeit durchschnittlich* (in %) mit jeder der Sprachen konfrontiert sind.

(Insgesamt sollten Ihre Angaben 100% ergeben):

Sprache					
Prozentualer Anteil:					

(4) Sie möchten einen Text lesen, der in allen von Ihnen beherrschten Sprachen verfügbar ist. Wie häufig (in %) würden Sie sich beim Lesen für welche Sprache entscheiden? Gehen Sie davon aus, dass das Original in einer Ihnen unbekannten Sprache geschrieben wurde. *(Insgesamt sollten Ihre Angaben 100% ergeben):*

Sprache					
Prozentualer Anteil:					

(5) Sie unterhalten sich mit jemandem, der alle von Ihnen gesprochenen Sprachen gleichermaßen beherrscht. Wie häufig (in % und anteilig an der Gesamtzeit) würden Sie welche Sprache im Gespräch wählen?

(Insgesamt sollten Ihre Angaben 100% ergeben):

Sprache					
Prozentualer Anteil:					

(6) Welchen Kulturkreisen fühlen Sie sich zugehörig? Geben Sie auf einer Skala von 0 bis 10 an, wie sehr Sie sich dem jeweiligen Kulturkreis zugehörig fühlen (mögliche Kulturkreise sind der deutsche, türkische, arabische, russische, katholische, etc.):

Kulturkreis					
	(für Skala hier klicken)	(für Skala hier klicken)	(für Skala hier klicken)	(für Skala hier klicken)	(für Skala hier klicken)

(7) Dauer Ihrer formalen Schul- und Ausbildung (in Jahren) _____

Kreuzen Sie das Feld an, das Ihrem höchsten Bildungsstand entspricht:

- | | | |
|--|---|-------------------------------------|
| ☐ ohne Schulabschluss | ☐ Hochschulreife | ☐ Promotion |
| ☐ Hauptschulabschluss | ☐ Berufsausbildung | ☐ Sonstiges: |
| ☐ Realschulabschluss | ☐ Hochschulabschluss | ☐ |

(8) Zeitpunkt des Zuzugs nach Deutschland, falls zutreffend _____

Wenn Sie in einem anderen Land gelebt haben, wann und wie lange haben Sie dort gelebt?

(9) Leiden/litten Sie an einer Beeinträchtigung des Sehvermögens ☐, Hörvermögens, ☐, Sprachvermögens ☐, oder Lernvermögens ☐ ? (Bitte alle zutreffenden ankreuzen.) Erläutern Sie bitte die zutreffenden (einschließlich Maßnahmen zur Abhilfe). _____

Sprache

Dies ist meine (bitte aus dem Drop-Down-Menü wählen) Sprache.

Alle folgenden Fragen beziehen sich auf Ihre Kenntnisse.

(1) Alter, in dem Sie ...:

erlernten:	anfangen	fließend zu sprechen:	anfangen, in	zu lesen:	anfangen, fließend in	zu lesen:

(2) Wie viele Jahre und Monate hielten Sie sich in den jeweiligen Sprachumgebungen auf:

	Jahre	Monate
In einem Land, in dem gesprochen wird		
In einer Familie, in der gesprochen wird		
In einer Schule/einem Arbeitsumfeld, in dem gesprochen wird		

(3) Bitte bewerten Sie auf einer Skala von 0 bis 10 Ihre **Fähigkeiten** in für die Bereiche Sprechen, Verstehen und Lesen. Verwenden Sie hierfür jeweils das Drop-Down-Menü:

Sprechen	(zur Bewertung hier klicken)	Verstehen gesprochener Sprache	(zur Bewertung hier klicken)	Lesen	(zur Bewertung hier klicken)

(4) Bitte bewerten Sie auf einer Skala von 0 bis 10, wie hoch der Beitrag der folgenden Faktoren zum Spracherwerb von war:

Umgang mit Freunden	(für Drop-Down-Menü hier klicken)	Sprachlabor/Selbstunterricht	(für Drop-Down-Menü hier klicken)
Umgang mit der Familie	(für Drop-Down-Menü hier klicken)	Fernsehen	(für Drop-Down-Menü hier klicken)
Lesen	(für Drop-Down-Menü hier klicken)	Radio hören	(für Drop-Down-Menü hier klicken)

(5) Bitte bewerten Sie, in welchem Ausmaß Sie in derzeit den folgenden Faktoren ausgesetzt sind:

Umgang mit Freunden	(für Drop-Down-Menü hier klicken)	Radio/Musik hören	(für Drop-Down-Menü hier klicken)
Umgang mit der Familie	(für Drop-Down-Menü hier klicken)	Lesen	(für Drop-Down-Menü hier klicken)
Fernsehen	(für Drop-Down-Menü hier klicken)	Sprachlabor/Selbstunterricht	(für Drop-Down-Menü hier klicken)

(6) Wie stark ist Ihrer Meinung nach Ihr Akzent, wenn Sie sprechen?

(für Drop-Down-Skala hier klicken)

(7) Wie häufig werden Sie beim Sprechen von anderen auf Grund Ihres Akzents als Nicht-Muttersprachler erkannt?

(für Drop-Down-Menü hier klicken)

Sprache:

Dies ist meine (bitte aus dem Drop-Down-Menü wählen) Sprache.

Alle folgenden Fragen beziehen sich auf Ihre Kenntnisse.

(1) Alter, in dem Sie...:

erlernten:	anfangen fließend zu sprechen:	anfangen, in zu lesen:	anfangen, fließend in zu lesen:

(2) Wie viele Jahre und Monate hielten Sie sich in den jeweiligen Sprachumgebungen auf:

	Jahre	Monate
In einem Land, in dem gesprochen wird		
In einer Familie, in der gesprochen wird		
In einer Schule/einem Arbeitsumfeld, in dem gesprochen wird		

(3) Bitte bewerten Sie auf einer Skala von 0 bis 10 Ihre **Fähigkeiten** in für die Bereiche Sprechen, Verstehen und Lesen. Verwenden Sie hierfür jeweils das Drop-Down-Menü:

Sprechen	(zur Bewertung hier klicken)	Verstehen gesprochener Sprache	(zur Bewertung hier klicken)	Lesen	(zur Bewertung hier klicken)

(4) Bitte bewerten Sie auf einer Skala von 0 bis 10, wie hoch der Beitrag der folgenden Faktoren zum Spracherwerb von war:

Umgang mit Freunden	(für Drop-Down-Menü hier klicke	Sprachlabor/Selbstunterricht	(für Drop-Down-Menü hier klicke
Umgang mit der Familie	(für Drop-Down-Menü hier klicke	Fernsehen	(für Drop-Down-Menü hier klicke
Lesen	(für Drop-Down-Menü hier klicke	Radio hören	(für Drop-Down-Menü hier klicke

(5) Bitte bewerten Sie, in welchem Ausmaß Sie in derzeit den folgenden Faktoren ausgesetzt sind:

Umgang mit Freunden	(für Drop-Down-Menü hier klicke	Radio/Musik hören	(für Drop-Down-Menü hier klicke
Umgang mit der Familie	(für Drop-Down-Menü hier klicke	Lesen	(für Drop-Down-Menü hier klicke
Fernsehen	(für Drop-Down-Menü hier klicke	Sprachlabor/Selbstunterricht	(für Drop-Down-Menü hier klicke

(6) Wie stark ist Ihrer Meinung nach Ihr Akzent, wenn Sie sprechen?

(für Drop-Down-Skala hier klicken)

(7) Wie häufig werden Sie beim Sprechen von anderen auf Grund Ihres Akzents als Nicht-Muttersprachler erkannt?

(für Drop-Down-Menü hier klicken)

Sprache:

Dies ist meine (bitte aus dem Drop-Down-Menü wählen) Sprache.

Alle folgenden Fragen beziehen sich auf Ihre Kenntnisse.

(1) Alter, in dem Sie...:

erlernten:	anfangen	fließend zu sprechen:	anfangen, in	zu lesen:	anfangen, fließend in	zu lesen:

(2) Wie viele Jahre und Monate hielten Sie sich in den jeweiligen Sprachumgebungen auf:

	Jahre	Monate
In einem Land, in dem gesprochen wird		
In einer Familie, in der gesprochen wird		
In einer Schule/einem Arbeitsumfeld, in dem gesprochen wird		

(3) Bitte bewerten Sie auf einer Skala von 0 bis 10 Ihre Fähigkeiten in für die Bereiche Sprechen, Verstehen und Lesen. Verwenden Sie hierfür jeweils das Drop-Down-Menü:

Sprechen	(zur Bewertung hier klicken)	Verstehen gesprochener Sprache	(zur Bewertung hier klicken)	Lesen	(zur Bewertung hier klicken)

(4) Bitte bewerten Sie auf einer Skala von 0 bis 10, wie hoch der Beitrag der folgenden Faktoren zum Spracherwerb von war:

Umgang mit Freunden	(für Drop-Down-Menü hier klicke	Sprachlabor/Selbstunterricht	(für Drop-Down-Menü hier klicke
Umgang mit der Familie	(für Drop-Down-Menü hier klicke	Fernsehen	(für Drop-Down-Menü hier klicke
Lesen	(für Drop-Down-Menü hier klicke	Radio hören	(für Drop-Down-Menü hier klicke

(5) Bitte bewerten Sie, in welchem Ausmaß Sie in derzeit den folgenden Faktoren ausgesetzt sind:

Umgang mit Freunden	(für Drop-Down-Menü hier klicke	Radio/Musik hören	(für Drop-Down-Menü hier klicke
Umgang mit der Familie	(für Drop-Down-Menü hier klicke	Lesen	(für Drop-Down-Menü hier klicke
Fernsehen	(für Drop-Down-Menü hier klicke	Sprachlabor/Selbstunterricht	(für Drop-Down-Menü hier klicke

(6) Wie stark ist Ihrer Meinung nach Ihr Akzent, wenn Sie sprechen?

(für Drop-Down-Skala hier klicken)

(7) Wie häufig werden Sie beim Sprechen von anderen auf Grund Ihres Akzents als Nicht-Muttersprachler erkannt?

(für Drop-Down-Menü hier klicken)

Sprache:

Dies ist meine (bitte aus dem Drop-Down-Menü wählen) Sprache.

Alle folgenden Fragen beziehen sich auf Ihre Kenntnisse.

(1) Alter, in dem Sie...:

erlernten:	anfangen	fließend zu sprechen:	anfangen, in	zu lesen:	anfangen, fließend in	zu lesen:

(2) Wie viele Jahre und Monate hielten Sie sich in den jeweiligen Sprachumgebungen auf:

	Jahre	Monate
In einem Land, in dem gesprochen wird		
In einer Familie, in der gesprochen wird		
In einer Schule/einem Arbeitsumfeld, in dem gesprochen wird		

(3) Bitte bewerten Sie auf einer Skala von 0 bis 10 Ihre **Fähigkeiten** in für die Bereiche Sprechen, Verstehen und Lesen. Verwenden Sie hierfür jeweils das Drop-Down-Menü:

Sprechen	(zur Bewertung hier klicken)	Verstehen gesprochener Sprache	(zur Bewertung hier klicken)	Lesen	(zur Bewertung hier klicken)

(4) Bitte bewerten Sie auf einer Skala von 0 bis 10, wie hoch der Beitrag der folgenden Faktoren zum Spracherwerb von war:

Umgang mit Freunden	(für Drop-Down-Menü hier klicke	Sprachlabor/Selbstunterricht	(für Drop-Down-Menü hier klicke
Umgang mit der Familie	(für Drop-Down-Menü hier klicke	Fernsehen	(für Drop-Down-Menü hier klicke
Lesen	(für Drop-Down-Menü hier klicke	Radio hören	(für Drop-Down-Menü hier klicke

(5) Bitte bewerten Sie, in welchem Ausmaß Sie in derzeit den folgenden Faktoren ausgesetzt sind:

Umgang mit Freunden	(für Drop-Down-Menü hier klicke	Radio/Musik hören	(für Drop-Down-Menü hier klicke
Umgang mit der Familie	(für Drop-Down-Menü hier klicke	Lesen	(für Drop-Down-Menü hier klicke
Fernsehen	(für Drop-Down-Menü hier klicke	Sprachlabor/Selbstunterricht	(für Drop-Down-Menü hier klicke

(6) Wie stark ist Ihrer Meinung nach Ihr Akzent, wenn Sie sprechen?

(für Drop-Down-Skala hier klicken)

(7) Wie häufig werden Sie beim Sprechen von anderen auf Grund Ihres Akzents als Nicht-Muttersprachler erkannt?

(für Drop-Down-Menü hier klicken)

Sprache:

Dies ist meine (bitte aus dem Drop-Down-Menü wählen) Sprache.

Alle folgenden Fragen beziehen sich auf Ihre Kenntnisse.

(1) Alter, in dem Sie....:

erlernten:	anfangen fließend zu sprechen:	anfangen, in zu lesen:	anfangen, fließend in zu lesen:

(2) Wie viele Jahre und Monate hielten Sie sich in den jeweiligen Sprachumgebungen auf:

	Jahre	Monate
In einem Land, in dem gesprochen wird		
In einer Familie, in der gesprochen wird		
In einer Schule/einem Arbeitsumfeld, in dem gesprochen wird		

(3) Bitte bewerten Sie auf einer Skala von 0 bis 10 Ihre **Fähigkeiten** in für die Bereiche Sprechen, Verstehen und Lesen. Verwenden Sie hierfür jeweils das Drop-Down-Menü:

Sprechen	(zur Bewertung hier klicken)	Verstehen gesprochener Sprache	(zur Bewertung hier klicken)	Lesen	(zur Bewertung hier klicken)

(4) Bitte bewerten Sie auf einer Skala von 0 bis 10, wie hoch der Beitrag der folgenden Faktoren zum Spracherwerb von war:

Umgang mit Freunden	(für Drop-Down-Menü hier klicken)	Sprachlabor/Selbstunterricht	(für Drop-Down-Menü hier klicken)
Umgang mit der Familie	(für Drop-Down-Menü hier klicken)	Fernsehen	(für Drop-Down-Menü hier klicken)
Lesen	(für Drop-Down-Menü hier klicken)	Radio hören	(für Drop-Down-Menü hier klicken)

(5) Bitte bewerten Sie, in welchem Ausmaß Sie in derzeit den folgenden Faktoren ausgesetzt sind:

Umgang mit Freunden	(für Drop-Down-Menü hier klicken)	Radio/Musik hören	(für Drop-Down-Menü hier klicken)
Umgang mit der Familie	(für Drop-Down-Menü hier klicken)	Lesen	(für Drop-Down-Menü hier klicken)
Fernsehen	(für Drop-Down-Menü hier klicken)	Sprachlabor/Selbstunterricht	(für Drop-Down-Menü hier klicken)

(6) Wie stark ist Ihrer Meinung nach Ihr Akzent, wenn Sie sprechen?

(für Drop-Down-Skala hier klicken)

(7) Wie häufig werden Sie beim Sprechen von anderen auf Grund Ihres Akzents als Nicht-Muttersprachler erkannt?

(für Drop-Down-Menü hier klicken)

Northwestern Bilingualism & Psycholinguistics Research Laboratory

Per citare: Marian, Blumenfeld, & Kaushanskaya (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech Language and Hearing Research*, 50 (4), 940-967.
Traduzione di Luca Bevacqua e Roberta Spelgorzi (2017), University of Edinburgh.

Versione di Sarah von Grebmer zu Wolfsthurn, Niels O. Schiller e Leticia Pablos Robles (2019), Leiden University Center for Linguistics

Questionario su esperienza e competenza linguistica (LEAP-Q)

Cognome		Nome		Data di oggi	
Età		Data di nascita		Uomo ☐	Donna ☐

(1) Elenca tutte le lingue che conosci **in ordine di dominanza**:

1	2	3	4	5
---	---	---	---	---

(2) Elenca tutte le lingue che conosci **in ordine di acquisizione** (per prima, la tua lingua nativa):

1	2	3	4	5
---	---	---	---	---

(3) Indica la percentuale di tempo in cui sei *attualmente* e *in media* esposto a ogni lingua.

(La somma delle percentuali deve essere 100%):

Indica qui le lingue:					
Indica qui le percentuali:					

(4) Potendo scegliere di leggere un testo disponibile in tutte le lingue che conosci, in che percentuale di casi sceglieresti di leggerlo in ciascuna lingua? Considera che il testo originale sia scritto in una lingua a te sconosciuta.

(La somma delle percentuali deve essere 100%):

Indica qui le lingue:					
Indica qui le percentuali:					

(5) Potendo scegliere di parlare con una persona che è ugualmente fluente in tutte le tue lingue, in che percentuale di casi sceglieresti ciascuna lingua? Indica la percentuale di tempo totale.

(La somma delle percentuali deve essere 100%):

Indica qui le lingue:					
Indica qui le percentuali:					

(6) Indica le culture con cui ti identifichi. In una scala da 0 a 10, valuta la misura in cui ti identifichi con ciascuna cultura. (Per esempio, culture regionali, cultura italiana, europea, o mediterranea):

Indica qui le culture:					
	(clicca qui per l'elen	(clicca qui per l'elen	(clicca qui per l'elen	(clicca qui per l'elen	(clicca qui per l'elen

(7) Per quanti anni hai studiato in un contesto formale? _____

Seleziona il livello di formazione più alto che hai ottenuto:

- | | | |
|--|--|---|
| ☐ Solo educazione obbligatoria | ☐ Laurea di primo livello | ☐ Specializzazione post-laurea |
| ☐ Scuola superiore di secondo grado | ☐ Laurea magistrale | ☐ Dottorato di ricerca |
| ☐ Formazione professionale | ☐ Master | ☐ Altro: |

(8) Hai mai avuto problemi di vista ☐, problemi di udito ☐, disturbi del linguaggio ☐ o disturbi dell'apprendimento ☐ ?
(Seleziona tutte le opzioni pertinenti). Se sì, specifica qui i dettagli (incluse eventuali correzioni):

Lingua:

Questa è la mia (seleziona dall'elenco) lingua.

Le seguenti domande si riferiscono alla tua conoscenza di: .

(1) Età in cui:

<i>hai iniziato ad apprendere la lingua:</i>	<i>sei diventato fluente:</i>	<i>hai iniziato a leggere nella lingua:</i>	<i>sei diventato fluente nella lettura:</i>

(2) Indica il numero di anni e mesi che hai passato in ognuno dei seguenti ambienti:

	Anni	Mesi
Un paese dove la lingua è parlata		
Una famiglia dove la lingua è parlata		
Un ambiente scolastico o lavorativo dove la lingua è parlata		

(3) In una scala da 0 a 10, seleziona dai menù il tuo *livello di competenza* nel parlare, comprendere e leggere la lingua:

Parlare	(clicca qui per l'elenco)	Comprendere la lingua parlata	(clicca qui per l'elenco)	Leggere	(clicca qui per l'elenco)
---------	---------------------------	-------------------------------	---------------------------	---------	---------------------------

(4) In una scala da 0 a 10, seleziona quanto i fattori seguenti hanno contribuito all'apprendimento della lingua:

Interazione con amici	(clicca qui per l'elenco)	Studio autodidatta	(clicca qui per l'elenco)
Interazione in famiglia	(clicca qui per l'elenco)	Televisione	(clicca qui per l'elenco)
Lettura	(clicca qui per l'elenco)	Radio	(clicca qui per l'elenco)

(5) Indica in che misura sei attualmente esposto alla lingua nei seguenti contesti:

Interazione con amici	(clicca qui per l'elenco)	Radio e musica	(clicca qui per l'elenco)
Interazione in famiglia	(clicca qui per l'elenco)	Lettura	(clicca qui per l'elenco)
Televisione	(clicca qui per l'elenco)	Laboratorio di lingua / studio autodidatta	(clicca qui per l'elenco)

(6) Nella tua percezione, il tuo accento straniero nella lingua è:

(clicca qui per l'elenco)

(7) Indica con quale frequenza gli altri ti identificano come parlante non nativo in base al tuo accento:

(clicca qui per l'elenco)

Lingua:

Questa è la mia (seleziona dall'elenco) lingua.

Le seguenti domande si riferiscono alla tua conoscenza di: .

(1) Età in cui:

<i>hai iniziato ad apprendere la lingua:</i>	<i>sei diventato fluente:</i>	<i>hai iniziato a leggere nella lingua:</i>	<i>sei diventato fluente nella lettura:</i>

(2) Indica il numero di anni e mesi che hai passato in ognuno dei seguenti ambienti:

	Anni	Mesi
Un paese dove la lingua è parlata		
Una famiglia dove la lingua è parlata		
Un ambiente scolastico o lavorativo dove la lingua è parlata		

(3) In una scala da 0 a 10, seleziona dai menù il tuo *livello di competenza* nel parlare, comprendere e leggere la lingua:

Parlare	(clicca qui per l'elenco)	Comprendere la lingua parlata	(clicca qui per l'elenco)	Leggere	(clicca qui per l'elenco)
---------	---------------------------	-------------------------------	---------------------------	---------	---------------------------

(4) In una scala da 0 a 10, seleziona quanto i fattori seguenti hanno contribuito all'apprendimento della lingua:

Interazione con amici	(clicca qui per l'elenco)	Studio autodidatta	(clicca qui per l'elenco)
Interazione in famiglia	(clicca qui per l'elenco)	Televisione	(clicca qui per l'elenco)
Lettura	(clicca qui per l'elenco)	Radio	(clicca qui per l'elenco)

(5) Indica in che misura sei attualmente esposto alla lingua nei seguenti contesti:

Interazione con amici	(clicca qui per l'elenco)	Radio e musica	(clicca qui per l'elenco)
Interazione in famiglia	(clicca qui per l'elenco)	Lettura	(clicca qui per l'elenco)
Televisione	(clicca qui per l'elenco)	Laboratorio di lingua / studio autodidatta	(clicca qui per l'elenco)

(6) Nella tua percezione, il tuo accento straniero nella lingua è:

(clicca qui per l'elenco)

(7) Indica con quale frequenza gli altri ti identificano come parlante non nativo in base al tuo accento:

(clicca qui per l'elenco)

Lingua:

Questa è la mia (seleziona dall'elenco) lingua.

Le seguenti domande si riferiscono alla tua conoscenza di: .

(1) Età in cui:

<i>hai iniziato ad apprendere la lingua:</i>	<i>sei diventato fluente:</i>	<i>hai iniziato a leggere nella lingua:</i>	<i>sei diventato fluente nella lettura:</i>

(2) Indica il numero di anni e mesi che hai passato in ognuno dei seguenti ambienti:

	Anni	Mesi
Un paese dove la lingua è parlata		
Una famiglia dove la lingua è parlata		
Un ambiente scolastico o lavorativo dove la lingua è parlata		

(3) In una scala da 0 a 10, seleziona dai menù il tuo *livello di competenza* nel parlare, comprendere e leggere la lingua:

Parlare	(clicca qui per l'elenco)	Comprendere la lingua parlata	(clicca qui per l'elenco)	Leggere	(clicca qui per l'elenco)
---------	---------------------------	-------------------------------	---------------------------	---------	---------------------------

(4) In una scala da 0 a 10, seleziona quanto i fattori seguenti hanno contribuito all'apprendimento della lingua:

Interazione con amici	(clicca qui per l'elenco)	Studio autodidatta	(clicca qui per l'elenco)
Interazione in famiglia	(clicca qui per l'elenco)	Televisione	(clicca qui per l'elenco)
Lettura	(clicca qui per l'elenco)	Radio	(clicca qui per l'elenco)

(5) Indica in che misura sei attualmente esposto alla lingua nei seguenti contesti:

Interazione con amici	(clicca qui per l'elenco)	Radio e musica	(clicca qui per l'elenco)
Interazione in famiglia	(clicca qui per l'elenco)	Lettura	(clicca qui per l'elenco)
Televisione	(clicca qui per l'elenco)	Laboratorio di lingua / studio autodidatta	(clicca qui per l'elenco)

(6) Nella tua percezione, il tuo accento straniero nella lingua è:

(clicca qui per l'elenco)

(7) Indica con quale frequenza gli altri ti identificano come parlante non nativo in base al tuo accento:

(clicca qui per l'elenco)

Lingua:

Questa è la mia (seleziona dall'elenco) lingua.

Le seguenti domande si riferiscono alla tua conoscenza di: .

(1) Età in cui:

<i>hai iniziato ad apprendere la lingua:</i>	<i>sei diventato fluente:</i>	<i>hai iniziato a leggere nella lingua:</i>	<i>sei diventato fluente nella lettura:</i>

(2) Indica il numero di anni e mesi che hai passato in ognuno dei seguenti ambienti:

	Anni	Mesi
Un paese dove la lingua è parlata		
Una famiglia dove la lingua è parlata		
Un ambiente scolastico o lavorativo dove la lingua è parlata		

(3) In una scala da 0 a 10, seleziona dai menù il tuo *livello di competenza* nel parlare, comprendere e leggere la lingua:

Parlare	(clicca qui per l'elenco)	Comprendere la lingua parlata	(clicca qui per l'elenco)	Leggere	(clicca qui per l'elenco)
---------	---------------------------	-------------------------------	---------------------------	---------	---------------------------

(4) In una scala da 0 a 10, seleziona quanto i fattori seguenti hanno contribuito all'apprendimento della lingua:

Interazione con amici	(clicca qui per l'elenco)	Studio autodidatta	(clicca qui per l'elenco)
Interazione in famiglia	(clicca qui per l'elenco)	Televisione	(clicca qui per l'elenco)
Lettura	(clicca qui per l'elenco)	Radio	(clicca qui per l'elenco)

(5) Indica in che misura sei attualmente esposto alla lingua nei seguenti contesti:

Interazione con amici	(clicca qui per l'elenco)	Radio e musica	(clicca qui per l'elenco)
Interazione in famiglia	(clicca qui per l'elenco)	Lettura	(clicca qui per l'elenco)
Televisione	(clicca qui per l'elenco)	Laboratorio di lingua / studio autodidatta	(clicca qui per l'elenco)

(6) Nella tua percezione, il tuo accento straniero nella lingua è:

(clicca qui per l'elenco)

(7) Indica con quale frequenza gli altri ti identificano come parlante non nativo in base al tuo accento:

(clicca qui per l'elenco)

Lingua:

Questa è la mia (seleziona dall'elenco) lingua.

Le seguenti domande si riferiscono alla tua conoscenza di: .

(1) Età in cui:

<i>hai iniziato ad apprendere la lingua:</i>	<i>sei diventato fluente:</i>	<i>hai iniziato a leggere nella lingua:</i>	<i>sei diventato fluente nella lettura:</i>

(2) Indica il numero di anni e mesi che hai passato in ognuno dei seguenti ambienti:

	Anni	Mesi
Un paese dove la lingua è parlata		
Una famiglia dove la lingua è parlata		
Un ambiente scolastico o lavorativo dove la lingua è parlata		

(3) In una scala da 0 a 10, seleziona dai menù il tuo *livello di competenza* nel parlare, comprendere e leggere la lingua:

Parlare	(clicca qui per l'elenco)	Comprendere la lingua parlata	(clicca qui per l'elenco)	Leggere	(clicca qui per l'elenco)
---------	---------------------------	-------------------------------	---------------------------	---------	---------------------------

(4) In una scala da 0 a 10, seleziona quanto i fattori seguenti hanno contribuito all'apprendimento della lingua:

Interazione con amici	(clicca qui per l'elenco)	Studio autodidatta	(clicca qui per l'elenco)
Interazione in famiglia	(clicca qui per l'elenco)	Televisione	(clicca qui per l'elenco)
Lettura	(clicca qui per l'elenco)	Radio	(clicca qui per l'elenco)

(5) Indica in che misura sei attualmente esposto alla lingua nei seguenti contesti:

Interazione con amici	(clicca qui per l'elenco)	Radio e musica	(clicca qui per l'elenco)
Interazione in famiglia	(clicca qui per l'elenco)	Lettura	(clicca qui per l'elenco)
Televisione	(clicca qui per l'elenco)	Laboratorio di lingua / studio autodidatta	(clicca qui per l'elenco)

(6) Nella tua percezione, il tuo accento straniero nella lingua è:

(clicca qui per l'elenco)

(7) Indica con quale frequenza gli altri ti identificano come parlante non nativo in base al tuo accento:

(clicca qui per l'elenco)