

Simone Heine

Fremd- und Zweitsprachenlernerfolg und seine Erklärung durch Erwerbsalter, kognitive, affektiv-motivationale und sozio-kulturelle Variablen

Eine empirische Studie

kassel
university
press

Es heißt für Kommunikation und Sprache, Deutsch als Fremdsprache.
13142024, Mail: simone.heine@kassel-univ.de

Das Englische ist insgesamt 74 Varianten des Englischen
wie auch Dialekte, Kriol und Pidgin-Varianten neben den
Sprachen aufgeführt werden.¹¹ Interessant ist der Blick auf die
Lernen in high context (L1 varieties)¹² (ca. von 21 aufgabebasierten
in ihrer: Australian English, Malay English, Colloquial Singapore
(wie South Asian English) und indigenized L2 varieties¹³ (zu den 17
Sprachen gehören bspw. Indian English, Hong Kong English, Black
English). Aus den Definitionen für diese Eingruppierung geht hervor,
um eine nicht-native, indigenisierte Varietät handelt es über keine
Anzahl von muttersprachlichen Sprechern verfügt. Die
Neutralität dieser Definition, wie im nächsten Abschnitt bestritten, um
in, was hetero-diskursive qualitativen Abgrenzungen stark
like (wie Bra) Kachru would make fantastic claims [...] aus Indian
South Asian English' are varieties of English, their speakers are not
ers of English' (U. Singh 1998: 16).

Interaktion und Dekonstruktion
interdisziplinärer Grund der
legte darin, dass die
sprachliche Trans-
g, „Mother
es von

Im Rahmen des Forschungsprojekts werden untersucht:

- Sprachkompetenzprofile (Aguado, Grotjahn, Schliak 2005, 2007, 2010)
- motivationale Gründe für den L2-Erwerb: L2 Motivational Self System; Ideal L2 Self/Dörnyei 2009)
- Sprachlerneignung (Meara 2005)
- Selbstwertschätzung der Sprachlerneignung (Dörnyei 2010), die für das L2-Lernen bedeutungsvollen biografischen Hintergründe, wie bspw. berufliche Anforderungen, persönliche/familiäre Beziehungen, soziale Kontexte etc. (Tuop et al. 1994, Moyer 2004)
- Lernwege, Lernstrategien (Bimmel 2010)
- Lerntyp (Aguado/Riemer 2010)
- Unterrichtserfahrungen, Unterrichtsqualität (Funk 2010)
- Lehrt- und Lerntraditionen (Hering 2010)
- subjektive Theorien der Lernenden über das L2-Lernen, insbesondere über den Einfluss des Faktors Alter auf den Lernerfolg

Muttersprachliche Kompetenz" als Messlatte für den L2-Lernerfolg?

„[...] when nonnative speakers have been shown to perform as well as a native speaker on a test, the cry goes up for yet another test“ (Davies 2003: 213). Davies (2011) hinterfragt die muttersprachliche Kompetenz unter anderem im Hinblick auf Sprachprüfungen und Testformate. Ihm zufolge handelt es sich bei der native speaker genau so um eine Idealisierung wie bei der standard language (ebd.: 306). Davies schlägt vor, bei der Leistungsbewertung zukünftig vom Kompetenzprofil idealer L2-Lernender auszugehen und nicht mehr von dem/der lange Zeit unhinterfragt gebliebenen Muttersprachler_in (ebd.: 296).

Auch Hulstijn (2011; 2012) stellt die idealisierte muttersprachliche Sprachkompetenz generell in Frage und nimmt stattdessen eine Unterscheidung in Basic Language Cognition (BLC) und Higher Language Cognition (HLC) vor, um die jeweilige Sprachkompetenz exakter zu beschreiben. Während die BLC v.a. Standardwortschatz und -Phrasen umfasst und einem hohen Grad der Automatisierung unterliegt, beinhaltet die HLC Wortschatz und Phrasen niedriger Frequenz und bezieht sich auf gesprochene und geschriebene Äußerungen (vgl. Hulstijn 2011: 230-231)

„(Mutter-) Sprache und Identität
„Who is a native speaker? What is it one is a native speaker of? What is not far from the „who am I?“ question“ (Davies 2011: 292)

„These terms [mother tongue, native] are no innocuous intuitions: They are the divisive implements of ethnolinguistic nationalism.“ (Bonfiglio 2010: 7)

Simone Heine

**Fremd- und Zweitsprachenlernerfolg
und seine Erklärung durch Erwerbssalter,
kognitive, affektiv-motivationale und
sozio-kulturelle Variablen**

Eine empirische Studie

Die vorliegende Arbeit wurde vom Fachbereich Geistes- und Kulturwissenschaften der Universität Kassel als Dissertation zur Erlangung des akademischen Grades eines Doktors der Philosophie (Dr. phil.) angenommen.

Gutachter: Prof. Dr. Karin Aguado
Prof. Dr. Ulrich Steinmüller

Tag der mündlichen Prüfung: 8. Februar 2016

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

Zugl.: Kassel, Univ., Diss. 2016
ISBN 978-3-7376-0272-3 (print)
ISBN 978-3-7376-0273-0 (e-book)
DOI: <http://dx.medra.org/10.19211/KUP9783737602730>
URN: <http://nbn-resolving.de/urn:nbn:de:0002-402735>

© 2017, kassel university press GmbH, Kassel
www.upress.uni-kassel.de

Printed in Germany

Danksagung

Diese Dissertation wäre nicht möglich gewesen ohne die Unterstützung zahlreicher Personen, denen ich an dieser Stelle danken möchte. Zuerst möchte ich meine Testpersonen erwähnen, deren Namen ich natürlich nicht nennen kann. Ohne ihre Bereitschaft, sich selbst und ihre Zeit einzubringen, wäre diese Arbeit nicht durchführbar gewesen. Bedanken möchte ich mich bei den Chinesischen Samstagsschulen, die mich bei der Suche nach Testpersonen unterstützen und an denen ich einen Teil der Datenerhebungen durchführen konnte. Namentlich sind das die „HuaDe China Schule e. V.“; die „Yi-Zhi-Chinesische-Kulturschule“ und die „Chinesische Schule Berlin e. V.“ Des weiteren möchte ich mich an dieser Stelle bei den folgenden Personen bedanken, die mitgewirkt haben, in dem sie bspw. die Fragebögen pilotierten, das Phonetik-*Rating* übernahmen oder fachliche Ansprechpartner_innen waren und mir durch ihre Rückmeldung wesentliche Impulse gaben: Luna Amer, Dr. Olga Averina, Anke Börsel, Tanja Butschek, Eva Diamantstein, Dr. Andreas Guder, Felix Golcher, Dr. Daniela Hartmann, Nataliya Iakovleva, Natalya Johnson, Svenja Neupert, Grazia Kampmann, Michail-Antonis Kioupliotis, Insa Kortenkamp, Dr. Constanze Saunders, Charles Saunders, Julia Schaaf, Dr. Almut Schön, Prof. Dr. Julia Settineri, Hengyan Tian, Johanna Varuzza und Kerstin Zimmermann.

Großer Dank gilt meiner Familie und insbesondere meinem Mann Jörg Nitschke, dessen Ermutigung, Verständnis und Unterstützung ich auf keinen Fall missen wollte.

Mit großem Dank an Prof. Dr. Karin Aguado und Prof. Dr. Rüdiger Grotjahn, die mir erlaubten, die von ihnen entwickelten Instrumente des Forschungsprojektes „Erwerbsalter und Sprachlernerfolg“ für die Datenerhebungen der Phase I einzusetzen und die von ihnen erhobenen Daten im Rahmen dieser Arbeit auszuwerten und zu verwenden.

Mit großem Dank an Prof. Dr. Karin Aguado und Prof. Dr. Ulrich Steinmüller, die meine Doktorarbeit betreuten, mir entscheidende Hinweise gaben und mir stets die nötige Sicherheit vermittelten, auf dem richtigen Weg zu sein. Mein Dank gilt Jun.-Prof. Dr. Christine Czinglar und Prof. Dr. Rüdiger Grotjahn, die mich als Kommissionsmitglieder unterstützten und mir wertvolle Rückmeldungen auf theoretische und praktische Fragen gaben.

Großer Dank gilt in wertschätzender Erinnerung Prof. Dr. Torsten Schlak (†), der mich zu diesem Forschungsprojekt inspirierte und ermutigte.

1. Einleitung	7
2. Forschungsstand Faktor ‚Alter‘	12
2.1 Kritische Periode(n)	12
2.2 Studien zum Faktor Alter.....	18
2.3 Forschungslücke.....	35
2.4 Was ist Alter? Was unterscheidet ältere von jüngeren Lerner_innen?	37
2.4.1 Neurophysiologische Reifungsprozesse	39
2.4.2 Motivation und Erwerbsantrieb.....	46
2.4.3 Sprachlerneignung	53
2.4.4 Die Natur-Kultur-Debatte.....	59
2.5 ‚Muttersprachliche Kompetenz‘.....	63
2.5.1 Die ‚muttersprachliche Kompetenz‘ als Messlatte für den L2-Lernerfolg	64
2.5.2 Konstruktion und Dekonstruktion eines Begriffs	67
2.5.3 Der ‚monolinguale Habitus‘ und seine Folgen.....	71
2.5.4 Alternativen zum Begriff der ‚Muttersprache‘	73
2.6 Einfluss der Ausgangssprache	76
3. Die Studie	83
3.1 Forschungsansatz.....	83
3.2 Vorarbeiten	85
3.3 Ziele und Forschungsfragen	87
3.4 Forschungsdesign.....	88
3.4.1 Testpersonen	88
3.4.2. Instrumente	92
3.4.2.1 Grammatikalitätstesttests	92
3.4.2.2 C-Tests und S-C-Tests	96
3.4.2.3 Fragebogen zum Sprachenprofil.....	99
3.4.2.4 TESTATT-Skalen.....	100

3.4.2.5 Phonetik-Test.....	103
3.4.2.6 LLAMA-Tests	107
3.4.2.7 Fragebogen biografische Hintergrundinformationen und Motivation.....	109
3.5 Durchführung der Untersuchung.....	115
3.5.1 Rahmenbedingungen.....	115
3.5.2 Proband_innen-Suche	116
4. Ergebnisse.....	120
4.1 Auswertung Phase I.....	122
4.1.1 Grammatical Judgement Tests	122
4.1.2 C-Tests und S-C-Tests	128
4.1.3 Gruppenvergleiche.....	133
4.1.4 Korrelationen.....	135
4.1.5 Phonetik-Rating.....	139
4.2 Auswertung Phase II.....	144
4.2.1 Fragebogen biografische Hintergrundinformationen und Motivation	144
4.2.2 LLAMA-Tests	148
4.3 Beantwortung der Forschungsfragen.....	150
4.4 Diskussion der Ergebnisse.....	164
5. Zusammenfassung	171
6. Abkürzungsverzeichnis.....	175
7. Literatur	177
8. Anhang.....	200

1. Einleitung

„Je früher, desto besser“ weiß schon der Volksmund. Tatsächlich wird jedoch der Einfluss des Faktors Alter auf den Spracherwerb in der Forschung kontrovers diskutiert. Einen der Hauptstränge innerhalb dieser wissenschaftlichen Auseinandersetzung stellt die Diskussion um die sogenannte *Critical Period Hypothesis* (CPH) dar (vgl. Kap. 2.1 *Kritische Periode(n)*). Die Befürworter der CPH gehen von einer oder mehreren kritischen Perioden oder auch sensiblen Phasen aus: Zeitfenster, in denen verschiedene Aspekte der menschlichen Sprache wie Phonetik, Morphosyntax, Lexik etc. besonders gut erworben werden können. Nach dem Ende dieser Phasen – i. d. R. spätestens mit dem Abschluss der Pubertät – können demnach die genannten Phänomene nur noch unvollständig erlernt werden. Zurückgeführt werden diese Unterschiede auf biologische Reifungsprozesse, wie etwa den oft angenommenen Verlust der neuronalen Plastizität des Gehirns, die die neurobiologische Grundlage für Lernen und Erinnern resp. Gedächtnis darstelle. Schon Lenneberg, auf den die CPH bezogen auf den Spracherwerb zurückgeht, stellte fest:

„Altersgebundenes Auftreten von Verhalten kann auf einer solchen Vielzahl von Faktoren beruhen, dass dieses Phänomen als solches nur von begrenztem heuristischen Wert ist.“ (Lenneberg 1972: 217)

Aus diesem Grund werden neben den biologischen und physiologischen Reifungsprozessen individuelle Faktoren im Zusammenhang mit der Frage nach dem Einfluss des Alters auf das Fremdsprachenlernen diskutiert (vgl. Kap. 2.4 *Was ist Alter? Was unterscheidet ältere von jüngeren Lerner_innen?*): Motivation, Spracherlernung, Persönlichkeitsprofile, Lerntypen und -strategien, Bildungshintergrund etc. Neben der Diskussion um individuelle Differenzen werden auch die Rahmenbedingungen untersucht: Qualität und Quantität des Inputs, Häufigkeit des informellen Sprachkontakts, Lehr- und Lernkontexte, allgemeine gesellschaftliche Gegebenheiten und der Einfluss der typologischen Ähnlichkeiten zwischen Ausgangs- und Zielsprache.

Ein weiterer wesentlicher Strang der Diskussion ist die sogenannte *Nature-Nurture*-Debatte bzw. die Diskussion zwischen der nativistischen Position (bspw. Chomsky 2007; Pinker 1994) – Sprache sei dem Menschen angeboren, der Mensch habe jedoch nur beschränkt Zugriff auf die sogenannte Universalgrammatik – und der interaktionistischen Position (bspw. Tomasello 2008) – Sprache sei erworben.

Die Frage nach dem Einfluss des Faktors Alter auf den Erwerb von Fremd- bzw. Zweitsprachen (im Folgenden L2) ist zweifelsohne ein interdisziplinäres Thema an der Schnittstelle von Neuropsychologie, -psychologie und -linguistik sowie der Spracherwerbsforschung. In der internationalen Sprachlehrforschung gehört die Frage nach dem Einfluss des Alters auf den Sprachlernerfolg zu den zentralen Forschungsfragen mit entsprechend großer Anzahl an Publikationen zum Thema.

Angesichts der hohen Praxisrelevanz des Themas bezüglich Curricula und didaktisch-methodischer Herangehensweisen ist es bedauerlich, dass die komplexe wissenschaftliche Debatte meist nur verkürzt in das Bewusstsein der Öffentlichkeit und der politischen Entscheidungsträger_innen¹ vordringt. Schmelter (2010) zeigt dies anhand verschiedener Curricula und Rahmenlehrpläne zum frühen Fremdsprachenbeginn in Grundschulen. Pagonis spricht in diesem Zusammenhang sogar von Mythenbildung, wenn eine Hypothese wie etwa „je früher, desto besser“ aus dem wissenschaftlichen Diskussionszusammenhang gerissen und als wissenschaftliche Erkenntnis rezipiert werde (Pagonis 2009a: 113).

Der Einfluss des Faktors Alter ist in Bezug auf die zunehmende Relevanz des lebenslangen Lernens bedeutungsvoll. Gleichwohl behandelt diese Arbeit nicht die spezifischen Bedürfnisse und methodisch-didaktischen Erfordernisse bei Senior_innen, die Fremdsprachen lernen (vgl. Berndt 2001, 2003; Raasch 2013; Sobotta 2013; Wegmann & Pomino 2010). Ebenso wenig geht es um den frühen Fremdsprachenbeginn an Schulen, obwohl beides zweifelsohne sowohl forschungsmethodisch herausfordernde wie gesellschaftlich relevante Themen sind. Im Rahmen dieses Dissertationsprojekts werden die Bedingungen erforscht, die Lernprozesse bei Erwachsenen gelingen lassen, um die Faktoren zu identifizieren, die individuelle Leistungsunterschiede beim L2-Lernen von sog. Späterwerber_innen beeinflussen (Erwerbsbeginn [*Age of Onset*: AO] 16 Jahre oder älter, vgl. Kap. 3.4.1 *Testpersonen* zur Festlegung dieser Altersgrenze). Die Ziele dieser Arbeit sind:

1. Messung und Vergleich der Sprachkompetenz von Personen mit bilingualem Erstspracherwerb (AO = 0–3 Jahre), Späterwerber_innen (AO ≥ 16)

¹ Der Unterstrich/*Gender Gap* ist eine aus dem Bereich der *Queer*-Theorie stammende Variante des Binnen-I und dient als Mittel der sprachlichen Darstellung aller sozialen Geschlechter und Geschlechtsidentitäten (vgl. URL 1; 2014).

und monolingualen Erstsprachler_innen

2. Spezifikation der Profile erfolgreicher Späterwerber_innen unter Berücksichtigung von kognitiven, affektiv-motivationalen und sozio-kulturellen Einflussgrößen

Eine Zielstellung des Projekts ist, die Sprachkompetenz mit einem Schwerpunkt auf Morphosyntax, Phonetik und dem grammatikalischen Beurteilungsvermögen erwachsener L2-Lerner_innen mit den Kompetenzprofilen sogenannter monolingualer Erstsprachler_innen zu vergleichen und mit denen von Personen mit bilinguaem Erstspracherwerb. Die Erhebung einer mehrsprachigen Vergleichsgruppe stellt eine Innovation vor dem Hintergrund der Diskussion um den Begriff ‚muttersprachliche Kompetenz‘ dar (vgl. Kap. 2.5 ‚*Muttersprachliche Kompetenz*‘).² Aktuell hinterfragt Davies (2011: 296) die muttersprachliche Kompetenz v. a. im Hinblick auf Sprachprüfungen und Testformate. Er schlägt vor, zukünftig vom Kompetenzprofil idealer L2-Lerner_innen auszugehen und nicht von dem/der ‚Muttersprachler_in‘. Der Vergleich zwischen Späterwerber_innen und Personen mit bilinguaem Erstspracherwerb verspricht mehr Fairness bei der Leistungsbemessung von L2-Lerner_innen.

Das Dissertationsprojekt besteht aus zwei Phasen und zielt darauf ab, einen Beitrag zur Beantwortung der Frage nach den Erfolgsbedingungen für L2-Späterwerber_innen zu leisten und Profile erfolgreicher Lerner_innen zu erstellen. In einem Mehr-Methoden-Forschungsdesign werden die Instrumente zeitlimitierte (S-C-Tests) und nicht zeitlimitierte C-Tests, *Grammatical Judgement Tests* (GJT), Phonetik-Tests sowie TESTATT-Skalen (Phase I) durch Sprachlerneignungstests und Motivationsskalen mit ausgewählten Proband_innen (Phase II) ergänzt. In beiden Phasen werden biographische Hintergrundinformationen erhoben. Das Forschungsdesign dient der Exploration der individuellen Variablen, die die Leistungsunterschiede bzgl. des L2-Erwerbs beeinflussen. Nach Auswertung der Ergebnisse der ersten Phase findet in der zweiten Phase eine Vertiefung statt.

Um die Phonetik zu testen, sprechen die Proband_innen zwei Minuten frei. Diese phonetischen Proben werden von einem fünfköpfigen Team einem Rating unterzo-

² Die Begriffe ‚Muttersprache‘, ‚muttersprachlich‘ etc. werden im Rahmen dieser Arbeit durch modalisierende Anführungszeichen markiert, um die Problematik des Begriffs anzudeuten (vgl. Kap. 2.5. „Muttersprachliche Kompetenz“).

gen. Der GJT bezieht sich auf besondere Lernschwierigkeiten der deutschen Sprache (vgl. Sopata 2004). Die allgemeine Sprachkompetenz mit den Schwerpunkten Lexik und Morphologie wird mittels zeitlimitierter und nicht zeitlimitierter C-Tests erhoben. Dies bietet den Vorteil, dass neben der allgemeinen lexikalischen und morphosyntaktischen Kompetenz auch die Bestimmung der prozeduralen und deklarativen Wissensanteile möglich wird. Mit dem von Grotjahn, Aguado und Schlak (2010) seit 2005 entwickelten zeitlimitierten C-Test (S-C-Test) können sehr hohe sprachliche Kompetenzprofile differenziert gemessen und sog. Deckeneffekte³ vermieden werden (ebd.: 298).

Bezüglich des Lerner_innenprofils sind v. a. Motivation, Extraversion/Introversion und Fremdsprachenverwendungsangst von Interesse, die mit Hilfe der Motivations- und Extraversionsskalen des TESTATTs (*Tests and Attitudes Scales for the Year Abroad*) ermittelt werden (Grotjahn 2004). Ergänzt werden die Profile in der zweiten Phase durch die Ergebnisse der Sprachlerneignungstests und einer weiteren Fragebogenerhebung zur Motivation und zu biografischen Hintergründen. Bei den Sprachlerneignungstests wird auf die LLAMA-Tests (Meara 2005) zurückgegriffen, bei denen es sich um ein erprobtes und gut zu handhabendes Instrument handelt. Der Fragebogen besteht aus zwei Teilen: Im ersten Teil werden weitere biografische Fakten und Angaben zum Gebrauch der L1, der L2 sowie weiterer Sprachen ermittelt. Im zweiten Teil wird eine adaptierte Version der Skalen zur Erhebung des *L2 Motivational Self Systems* (Dörnyei 2009a) eingesetzt.

Getestet werden 20 Späterwerber_innen (AO ≥ 16) der deutschen Sprache mit der Ausgangssprache Mandarin, die seit mindestens fünf Jahren in Deutschland leben (Length of Residence [LOR] ≥ 5) und über einen akademischen Bildungshintergrund verfügen. Die Vergleichsgruppe besteht aus 22 Personen mit doppeltem Erstspracherwerb (AO = 0–3) mit den Sprachen Mandarin bzw. Kantonesisch und Deutsch. Die Datenerhebungen für die erste Phase fanden vom 23.01.2013 bis 23.11.2013 statt. Phase II wurde vom 01.03.2014 bis 25.03.2014 erhoben.

Das Projekt richtet den Fokus auf Mandarin, weil Deutsch und Mandarin typologisch sehr verschieden sind. Mandarin, eine sinotibetische Sprache, wird zu den analyti-

³ Wenn bspw. alle getesteten Personen aufgrund ihrer hohen Sprachkompetenz 99 % richtig machen, sprechen wir von einem Deckeneffekt.

schen Sprachen gezählt, während Deutsch als indoeuropäische Sprache zu den synthetischen Sprachen gehört (Zhao 2010: 627). Hinzu kommt, dass Mandarin eine Tonsprache ist. Insofern ist auch im bezüglich des Einflusses des Faktors Alter viel diskutierten Feld der Phonetik eine beträchtliche Differenz vorhanden.

Die hier vorliegende Arbeit schließt an das Forschungsprojekt „Erwerbsalter und Sprachlernerfolg“ von Aguado, Grotjahn & Schlak (2005, 2007) an und greift in der Phase der quantitativen Erhebungen auf das Standardinstrumentarium zurück, das in diesem Projekt verwendet und teilweise dafür entwickelt wurde. Bei der Auswertung der Daten können die von Aguado, Grotjahn und Schlak erhobenen Daten genutzt werden.

Erwartet wird, dass sich die drei untersuchten Gruppen – ‚Muttersprachler_innen‘, Personen mit bilinguaem Erstspracherwerb (2L1) und Späterwerber_innen – bezüglich ihrer Sprachkompetenz unterscheiden. Dabei sollen die ‚Muttersprachler_innen‘ die besten Ergebnisse erzielen, wobei die Unterschiede zwischen ihnen und den Personen mit 2L1 minimal sind, während sich beide Gruppen deutlich von den Späterwerber_innen abheben. Des Weiteren wird erwartet, dass die Daten bei den Späterwerber_innen mehr variieren als in den anderen Gruppen. Außerdem wird mit einem kleinen Anteil (5–10 %) der Späterwerber_innen gerechnet, die im Leistungsspektrum der ‚Muttersprachler_innen‘ und Personen mit 2L1 abschneiden. Schließlich wird davon ausgegangen, dass die typologische ‚Nähe‘ und ‚Ferne‘ von Ausgangs- und Zielsprache Einfluss auf den Lernerfolg ausübt.

Die vorliegende Arbeit gliedert sich in einen theoretischen und einen empirischen Teil. Nach einer Einführung in die Diskussion um die kritische Periode für den Fremdspracherwerb und der Präsentation einschlägiger Studien zum Faktor Alter wird die Forschungslücke herausgearbeitet. Dem schließt sich ein Kapitel an, das die neurophysiologischen Grundlagen von Reifungsprozessen zusammenfasst, einen Überblick über Motivation und Sprachlernerneigung gibt und die Natur-Kultur-Debatte umreißt. Danach wird der Begriff der ‚muttersprachlichen Kompetenz‘ als solcher und v. a. im Hinblick auf das Ziel des L2-Lernens diskutiert sowie die Debatte um den Einfluss der Ausgangssprache auf das Erlernen der Zielsprache. Im empirischen Teil werden der Forschungsansatz, die Forschungsfragen, das Forschungsdesign und die eingesetzten Instrumente sowie die Testpersonen dargestellt und be-

schrieben. Des Weiteren wird der Forschungsprozess dokumentiert. Schließlich werden die Ergebnisse präsentiert, interpretiert und diskutiert.

2. Forschungsstand Faktor ‚Alter‘

Im Rahmen dieser Forschungsarbeit wird in den meisten Fällen nicht zwischen Fremdsprachenlernen und Zweitsprachenerwerb differenziert (vgl. Henrici & Riemer 2007: 39), sondern verallgemeinernd von L2-Erwerb gesprochen. Dies hat einerseits seinen Grund darin, dass in der diskutierten Forschungsliteratur sowohl das Fremdsprachenlernen als auch der Zweitsprachenerwerb untersucht wird. Andererseits können die meisten der im Rahmen dieses Projekts untersuchten Späterwerber_innen sowohl als Fremdsprachenlerner_innen als auch als Zweitspracherwerbende gelten, weil sie formalen Deutschunterricht im Heimatland besuchten und seit mindestens fünf Jahren in Deutschland leben. Ahrenholz (2008) definiert Zweitsprachenerwerb in und durch Kommunikation in bedeutsamen sozialen Kontexten und Fremdsprachenlernen in institutionellen Kontexten in meist nichtdeutschsprachiger Umgebung sowohl im Inland als auch im Ausland. Allerdings merkt er an, dass sich eine klare Trennung im Zeitalter der Globalisierung nicht aufrechterhalten lässt:

„Im Zeitalter von Internet, Satellitenforschung, Billigflügen und umfangreichen Austauschprogrammen für Studierende ist diese Inland-Ausland-Unterscheidung allerdings ansatzweise aufgelöst und wir haben zum Teil komplexe, kombinierte Spracherwerbssituationen eigener Art [...]“ (Ahrenholz 2008: 10)

Wenn es erforderlich ist zu differenzieren, werden die Begriffe entsprechend ausgeschrieben.

Im folgenden Kapitel wird die Diskussion der kritischen Periode(n) des Fremdsprachenerwerbs dargestellt. Dabei wird sowohl die Genese dieser Diskussion herausgearbeitet als auch gezeigt, welche Perioden für den Fremdsprachenerwerb angenommen werden.

2.1 Kritische Periode(n)

2005 hat das *Science Magazine* die Klärung des Zusammenhangs von Alter bzw. Kritischer Periode und Spracherwerb zu den 125 großen Fragen der Wissenschaft für das nächste Vierteljahrhundert gezählt (vgl. URL 2; 2005: o. S.). Die Diskussion

um die sogenannte *Critical Period Hypothesis* (CPH) für den L2-Erwerb dominierte lange Zeit die Forschungsarbeiten. Der Begriff Kritische Periode oder *Critical Period* (CP) stammt ursprünglich aus der Entwicklungspsychologie resp. Entwicklungsbiologie und meint eine Lebensphase mit definiertem Anfangs- und Endpunkt (*cut-off point* oder *offset*), in der äußere Faktoren die Entwicklung stark beeinflussen können. In diesem Zusammenhang wohl am bekanntesten ist das von dem Verhaltensbiologen Konrad Lorenz entdeckte Lernen durch Prägung bzw. die Nachfolgeprägung bei Graugänsen. In einem kurzen Zeitfenster nach dem Schlüpfen ‚lernen‘ die Graugänse, wer ihre Mutter ist. Mutter kann in diesem Fall jede_r werden, der/die sich in der Nähe der frisch geschlüpften Gänse aufhält: ein Fußball, eine Attrappe, ein Mensch etc. Ebenfalls häufig zitiert ist die Prägephase von Katzen bezüglich des Sehvermögens: „Katzen bleiben blind, wenn sie in ihren ersten Lebensmonaten keinen Zugang zu visuellen Informationen erhalten.“ (Grotjahn & Schlak 2010: 868)

Mit *Sensitive Period* ist häufig ein Zeitfenster ohne *cut-off point*, sondern mit allmählichem Ende gemeint. Während dieser Periode ist eine erhöhte Sensitivität gegenüber einem bestimmten Stimulus vorhanden. Erfolgreiches und nachhaltiges Lernen findet statt, wenn dieser Stimulus auftritt. Eine CP wird in den meisten Fällen zusätzlich wie folgt definiert: Wenn dieser Stimulus zum rechten Zeitpunkt ausbleibt, kann sich das betroffene Organ nicht in der vorgesehenen Weise entwickeln (vgl. hierzu Dörnyei 2009b: 237). In den Diskussionen um die CPH innerhalb der Spracherwerbsforschung werden die beiden Begriffe differenziert und undifferenziert verwendet (vgl. Schlak 2003: 18–19). Die Befürworter_innen der CPH gehen von einer oder mehreren kritischen Perioden bzw. auch sensiblen Phasen aus. In diesen entwicklungsbiologisch bedingten Zeitfenstern können demnach die verschiedenen Aspekte der menschlichen Sprache (Phonetik, Morphosyntax, Lexik etc.) besonders gut erworben werden. Hingegen können nach dem Ende dieser Phasen die genannten Phänomene nur noch unvollständig erlernt werden (vgl. Aguado 2010: 39).

Die folgenreiche Annahme einer CP für den Spracherwerb geht auf Eric Lenneberg (1967) zurück und wird in einer Vorform Penfield (Penfield & Roberts 1959) zugeschrieben. Penfields Epilog zu *Speech and Brain Mechanisms* ist ein Plädoyer für die Förderung mehrsprachiger Gesellschaften: „Bilingualism is not a handicap to a country. It has been a great benefit to mankind. As multilingualism has, also“ (ebd.: 255). Das Gewicht dieser Äußerung wird verstehbar, wenn der zeitgeschichtliche

Kontext einer tendenziell defizitären Sichtweise von Mehrsprachigkeit selbst in einem klassischen Einwanderungsland wie Kanada bedacht wird. Die bestmögliche Weise, Mehrsprachigkeit erfolgreich zu fördern, ist Penfield zufolge ein Lernbeginn im Alter von vier bis zehn Jahren; er spricht hier von einer „ideal period for language beginning“ (ebd.: 253). Neben dem frühen Lernbeginn ist die Lernmethode von entscheidender Wichtigkeit, denn nur durch „Language learning by the direct method“ oder „Language learning by the mother’s method“⁴ (ebd.: 253–254) etwa an einer bilingualen Schule, können Kinder erfolgreich eine oder mehrere Fremdsprachen erwerben (vgl. ebd.: 255–256). Penfields Epilog trägt einen eher essayistischen Charakter mit Verweisen auf seine neurologischen Forschungen – diese jedoch nur andeutend, nicht genau ausführend – sowie auf die kursorischen Wahrnehmungen der Sprachentwicklung und des L2-Erwerbs seiner mehrsprachig aufwachsenden Kinder. Diese Feststellung mindert seinen Beitrag allerdings keineswegs, denn spätere Studien werden seine Vermutungen bekräftigen.

Lenneberg, ein Linguist, Psychologe und Neurologe, formulierte acht Jahre nach Penfield aufgrund seiner Beobachtungen bei Aphasie-Patient_innen die Annahme einer kritischen Periode für den Spracherwerb: Es zeigte sich, dass die von Aphasie betroffenen Patient_innen zwischen vier und zehn Jahren zwar ähnliche Symptome wie die Erwachsenen aufweisen, jedoch viel bessere Prognosen haben. So wird die überwältigende Mehrheit der Kinder wieder gesund, während bei den Erwachsenen nur 30 % genesen (vgl. Lenneberg 1972: 182).⁵ Lenneberg führte diese Krankheitsverläufe auf Reifungsprozesse des Gehirns und dabei v. a. auf die Lateralisierung zurück: Die laterale Funktionsgliederung beginne ab dem 20. Lebensmonat, wobei strukturiert werde, welche Teile des Gehirns welche Aufgaben übernehmen. Mit der Pubertät komme die Lateralisierung zum Abschluss. Da im Kindesalter beide Hemisphären noch nicht ausreichend spezialisiert seien, könne die rechte Hemisphäre bei aphasischer Schädigung der linken Hemisphäre die Sprachfunktion vollständig

⁴ Penfields Begriffe entsprechen dem impliziten Lernen und der Immersion, die im Bereich Deutsch als Fremdsprache (DaF)/Deutsch als Zweitsprache (DaZ) bzw. Sprachlernforschung gebräuchlich sind.

⁵ Die englische Originalausgabe von *Biological Foundations of Languages* erschien bereits 1967, fünf Jahre vor der deutschen Übersetzung.

übernehmen (vgl. ebd.: 190). Spracherwerbsprozesse seien reifungsabhängig, weil ein Zusammenhang zwischen Sprachentwicklung und Lateralisierung bestehe:

„Daher lässt sich ferner nachweisen, dass zu der Zeit, zu der der primäre Spracherwerb aufhört, das Gehirn seinen Reifezustand erreicht hat und die zerebrale Lateralisation unumkehrbar festgelegt ist. [...] Als Arbeitshypothese nehmen wir deshalb an, dass die allgemeinen nichtspezifischen Stufen der Hirnreifung Vorbedingungen und begrenzende Faktoren für die Sprachentwicklung darstellen.“ (ebd.: 208)

An dieser Stelle ist es wichtig darauf hinzuweisen, dass Lenneberg von *primärem Spracherwerb* spricht, für den eine CP bestehe, *nicht* vom Erwerb einer L2. Der Beginn der kritischen Periode zeichne sich durch hohe Plastizität des Gehirns und zerebrale Unreife bzw. Undifferenziertheit aus, während an ihrem Ende der Abschluss des Stadiums der neuronalen Plastizität und die Lateralisierung der Funktionen stehe (vgl. ebd.: 217; zur Diskussion aktueller neurophysiologischer Erkenntnisse vgl. Kap. 2.4.1 *Neurophysiologische Reifungsprozesse*). Diese durch hirneurophysiologische Reifungsprozesse bewirkte CP verunmögliche nicht den Erwerb von Fremdsprachen nach Abschluss der Pubertät, wirke sich aber auf die Bedingungen dieses Erwerbs aus:

„So wird das Bild schon durch unsere Fähigkeit, Fremdsprachen zu erlernen verwirrt: Die meisten Individuen mit durchschnittlicher Intelligenz sind nach Beginn des zweiten Jahrzehnts in der Lage, eine zweite Sprache zu lernen, wenngleich Blockierungen der Spracherlernung nach der Pubertät schnell in erhöhtem Maße auftreten. Auch der automatische Erwerb einer bestimmten Sprache allein dadurch, dass man in ihrem Bereich lebt, scheint nach diesem Alter verlorenzugehen. Fremdsprachen müssen nach der Pubertät mit bewusster und angestrebter Mühe gelehrt und gelernt werden, und fremde Akzente können nicht mehr leicht überwunden werden. Doch *kann* man auch mit vierzig Jahren lernen, sich in einer Fremdsprache zu unterhalten.“ (ebd.: 217)

Indirekt lässt sich daraus der automatische („ohne angestrebte Mühe“) Erwerb einer L2 in Immersionssituationen, der eine akzentfreie Aussprache inkludiert, als Kriterien für die CP ableiten (vgl. Herschensohn 2007: 140).

Verschiedene Studien zu sogenannten Wolfskindern (vgl. den Überblick in Patkowski 2013: 6f) sowie Erhebungen zum Erwerb einer Gebärdensprache bei Gehörlosen im Erwachsenenalter (vgl. bspw. Mayberry 1993) scheinen zu belegen, dass die CPH für den Erstspracherwerb zutrifft. Muñoz & Singleton (2011) weisen jedoch darauf hin, dass die Datengrundlage der Argumentation mit dem unvollständigen Erwerb der L1 bei Wolfskindern nicht eindeutig interpretierbar sei. Der spätere unvollständige L1-Erwerb sei nur eine von mehreren Folgen der Vernachlässigung in-

nerhalb der ersten fünf Lebensjahre, neben einem kleineren Kopfumfang, einer kleineren Größe des Gehirns und der eingeschränkten Fähigkeit, soziales Verhalten zu erlernen (vgl. ebd.: 3).⁶ Diesbezüglich aussagekräftiger hingegen sind die Daten der Gehörlosen, die bedingt durch die geografische Abgeschiedenheit ihres Wohnorts erst später beginnen konnten, eine Gebärdensprache zu lernen, bei denen sonst aber keine kognitiven Einschränkungen oder emotionale Vernachlässigung vorliegen:

„The results showed unambiguously that the group with no early language was unable to acquire English grammar well even after many years of daily usage.“ (Dörnyei 2009b: 239)

Tab. 1, die dem Überblicksartikel von Singleton (2005: 273) entnommen ist, vermittelt einen Eindruck des Variantenreichtums der Forschung zur CPH.

Tab. 1: Übersicht der vorgeschlagenen Offsets einer CP (Tab. aus Singleton 2005: 273).

Table 1. *Summary of a range of proposals for CP termini*

Penfield and Roberts (1959)	Offset: age 9
Lenneberg (1967)	Onset: age 2 Offset: puberty
Molfese (1977)	Offset for phonetics/phonology: age 1
Seliger (1978)	Offset for phonetics/phonology: puberty
Diller (1981)	Offset for phonetics/phonology: age 6-8
Scovel (1988)	Offset for phonetics/phonology: age 12
Johnson and Newport (1989)	Offset of phase 1: age 7 Offset of phase 2: puberty
Long (1990)	Offset of phase 1: age 7 Offset of phase 2 for phonetics/phonology: age 12
Ruben (1997)	Offset of phase 2 for morphosyntax: age 15 Onset for phonetics/phonology: 6th month of foetal life Offset for phonetics/phonology: age 1 Offset for syntax: 4th year of life Offset for semantics 15th/16th year of life
Hyltenstam and Abrahamsson (2003)	Offset: shortly after birth

Singleton zeigt die Vielgestaltigkeit der Forschung zur CPH bzgl. der verschiedenen untersuchten sprachlichen Ebenen und der Bandbreite der gefundenen *Offsets* der CP. Bedingt ist diese Vielfalt durch:

⁶ Dessen ungeachtet wertet Patkowski (2013) fünf der gut dokumentierten Fälle mit verspätetem L1-Erwerbsbeginn aus, um damit für qualitativ andere Prozesse des Parametersetzens zu argumentieren, die ihm zufolge durch die CP bedingt seien.

- die verschiedenen zum Einsatz kommenden Methoden und Testverfahren
- die unterschiedlichen sprachlichen Bereiche, die untersucht werden
- die Heterogenität der Proband_innengruppen und ihrer L2-Lernsituationen (gesteuert vs. ungesteuert, Zahl der Kontaktstunden etc.).
- forschungsmethodische Probleme, bspw. zu geringe Fallzahlen, fragwürdige statistische Verfahren etc. (vgl. Grotjahn & Schlak 2013: 34).

Angesichts eines knappen halben Jahrhunderts Diskussionen über die CPH bezogen auf den L2-Erwerb bemerkt Singleton:

„As it stands it is like the mythical hydra, whose multiplicity of heads and capacity to produce new heads rendered it impossible to deal with.“ (Singleton 2005: 280)

Uneinigkeit herrscht auch darüber, was eine CP belegen resp. widerlegen kann. Long zufolge wäre die CPH durch das Auffinden einer Person mit spätem Erwerbsbeginn und ‚muttersprachlichen Kompetenzen‘ falsifiziert (vgl. Long 1990: 253). Herschensohn bemerkt hierzu, dass eine Person eine zu schmale Datenbasis darstellt. Besser wäre es, mehrere außerordentliche Späterwerber_innen zu finden oder frühe Lerner_innen, die die ‚muttersprachliche Kompetenz‘ nicht erreichen (vgl. Herschensohn 2007: 142). DeKeyser & Larson-Hall zufolge ließen sich sehr viele Missverständnisse und Streitpunkte lösen, wenn die ursprünglich von Lenneberg (s. o.) intendierte Definition der Bedingungen für den L2-Erwerb Berücksichtigung finden: automatischer Spracherwerb durch bloße Immersion (vgl. DeKeyser & Larson-Hall 2005: 89). In der folgenden Zusammenfassung der Diskussion wird allerdings nicht dieser engen Definition gefolgt, da sonst die Diskussion in ihrer ganzen Bandbreite nicht wiedergegeben werden kann.

Trotz der Unterschiedlichkeit der Studien lassen sich zusammenfassend Alterseffekte bezüglich des Lerntempos (*rate of acquisition*) und des letztendlich erreichten Sprachstands (*ultimate attainment*, im Folgenden UA) des L2-Erwerbs empirisch feststellen (vgl. Grotjahn 2005; Davies 2011). Im *Barcelona Age Factor Project* wurden in verschiedenen Studien 2068 Schüler_innen untersucht, die mit 8, 11, 14 oder 18+ anfangen, Englisch zu lernen (vgl. Muñoz 2006b). Dabei wurde festgestellt, dass ältere Schüler_innen und junge Erwachsene eine L2 schneller lernen als jüngere, wenn es um das Lernen in einer Unterrichtssituation (gesteuertes L2-Lernen) geht. Einige Studien weisen das höhere Tempo erwachsener Lerner_innen oder pubertierender Lerner_innen auch für Kontexte ungesteuerten L2-Erwerbs nach, v. a. die Be-

reiche Lexik, Pragmatik und Semantik betreffend (vgl. DeKeyser & Larson-Hall 2005). Bezogen auf das UA wurde für ungesteuerte Erwerbskontexte jedoch festgestellt, dass Kinder bis zu einem Alter von ca. sechs Jahren den erwachsenen Lerner_innen in der Regel überlegen sind, was die erreichte phonetische Kompetenz betrifft (Grotjahn & Schlak 2010: 871).

Im Zusammenhang mit der CPH werden neurobiologische Reifungsprozesse diskutiert, etwa die Lateralisierung der Gehirnfunktionen (vgl. Lenneberg 1967) und der Verlust der neuronalen Plastizität des Gehirns, die die neurobiologische Grundlage für Lernen und Erinnern bzw. Gedächtnisleistung darstellt. Letzteres wurde mittlerweile jedoch durch neuere Forschungsergebnisse widerlegt (vgl. Nitsch 2007: 51). Andere Erklärungsansätze argumentieren entweder mit altersgemäßen unumkehrbaren Entwicklungen der kognitiven Strukturen oder mit entscheidenden Veränderungen der affektiv-motivationalen, psycho-sozialen und sozio-kulturellen Einflussgrößen, die ebenfalls durch das Lebensalter geprägt sind (vgl. die ausführliche Diskussion in Kap. 2.4 *Was ist Alter? Was unterscheidet ältere von jüngeren Lerner_innen?*).

Nachdem in diesem Unterkapitel der Diskussionszusammenhang der kritischen Periode(n) für den Fremdsprachenerwerb aufgezeigt wurde, werden im nächsten Unterkapitel die wichtigsten Studien zum Faktor Alter im Rahmen der Fremdspracherwerbsforschung vorgestellt.

2.2 Studien zum Faktor Alter

„Almost everybody has a view of the role of age in language learning, and, curiously enough, most of these views are wrong.“ (Dörnyei 2009b: 233)

Forschungsarbeiten zum Einfluss des Faktors Alter auf das L2-Lernen und zur Falsifikation der CPH gibt es in großer Zahl. Im folgenden Überblick wird der Schwerpunkt auf die klassischen, häufig zitierten und teilweise replizierten Studien gelegt. Darüber hinaus werden Arbeiten mit besonderer Relevanz für die vorliegende Studie dargestellt. Einen sehr guten Überblick über die Diskussionsstränge und den Forschungsstand verschaffen Birdsong 2005; Cook & Singleton 2014; Dörnyei 2009b; Grotjahn & Schlak 2010, 2013; Long 2005; Muñoz & Singleton 2011; Singleton 2005; Singleton & Ryan 2004.

Mark Patkowskis (1980) Studie untersucht den Unterschied zwischen dem prä- und postpubertierenden Beginn des L2-Erwerbs bzgl. des Erwerbs syntaktischer Kompetenz. Seine Hypothese lautet:

„[...] that full, native-like acquisition of syntax in a nonnative language can be achieved only if learning begins before the age of fifteen years.“ (ebd.: 451)

Patkowski lässt Transkriptionen von Gesprächen mit 67 Migrant_innen von zwei Ratern auf einer Skala von 0 = *no ability* bis 5 = *native like performance* bewerten. Die abhängige Variable ist die Beurteilung der syntaktischen Kompetenz der Proband_innen. Unabhängige Variablen sind AO, LOR, informeller Sprachkontakt und formeller Unterricht. Patkowski findet einen starken Haupteffekt für AO; die anderen unabhängigen Variablen zeigen keine Wirksamkeit (vgl. ebd.: 454).⁷

Mit den Ergebnissen kann sowohl für als auch gegen die CPH argumentiert werden. Einerseits scheinen sie einen systematischen Einfluss des Einreisealters auf das erreichte Leistungsniveau in der L2 zu zeigen. So schneiden in der präpubertierenden Gruppe (n = 33) alle bis auf eine_n mit 4+ oder 5 ab, während die Ergebnisse in der postpubertierenden Gruppe (n = 34) von 2+ bis 5 breiter streuen, wobei die meisten (n = 12) einen Wert von 3+ erreichen. Andererseits belegen sie gleichzeitig, dass der Einfluss des Faktors Alter nicht determinierend ist, da es bei den unter 15-jährigen eine Ausnahme gibt, die mit 3 abschneidet, und 14,7 % der über 15-jährigen Lerner_innen mit 4+ oder 5 bewertet werden.

Die Studie von Johnson & Newport (1989) gehört zu den Studien, auf die sich die meisten berufen, die für die Gültigkeit der Hypothese „je früher, desto besser“ für den L2-Erwerb argumentieren. Johnson und Newport differenzieren die CPH in zwei Versionen. *The exercise hypothesis* besagt: Wenn die ausgezeichnete Fähigkeit der Menschen, in frühem Alter Sprachen zu erwerben, nicht in dieser Zeit trainiert wird, wird sie mit der Reifung verschwinden oder sich verringern. Folglich sind Kinder den erwachsenen Lerner_innen überlegen, was den Erstspracherwerb anbelangt, beide Gruppen haben aber gleiche Chancen, was den Erwerb oder das Lernen von weiteren Sprachen betrifft. *The maturational state hypothesis* hingegen sagt aus, dass

⁷ Um Oyamas (1976) Studie zu replizieren und die phonetische Kompetenz der Proband_innen mit-einzubeziehen, bewerteten die Rater Auszüge der Bandaufnahmen nach Akzent. Hierbei wurden Oyamas Ergebnisse ebenfalls bestätigt und für AO ein Haupteffekt gefunden (vgl. ebd.: 459).

diese frühe Fähigkeit des Spracherwerbs (L1 oder L2) durch Reifung verschwindet bzw. sich verringert (vgl. ebd.: 64). Johnson und Newport untersuchten insgesamt 46 koreanische und chinesische Migrant_innen, die Studierende oder Lehrende an der *University of Illinois* waren, im Alter von 3–39 Jahren in die USA einwanderten und bereits 3–26 Jahre in den USA lebten, mittels eines *Grammatical Judgement Tests* (GJT). Chinesisch und Koreanisch als L1 der Proband_innen wurden ausgewählt „because of their typological dissimilarity to English“ (ebd.: 68). Da bei der Auswertung keine Unterschiede zwischen den Ergebnissen für die beiden Herkunftssprachen gefunden wurden, werden die Ergebnisse zusammengefasst präsentiert. Diejenigen mit einem AO < 7 schnitten bezüglich der Grammatikalitätsurteile genauso erfolgreich ab wie amerikanische Erstsprachler_innen. Bei denjenigen mit einem AO ≥ 17 ließ sich kein regelhafter Zusammenhang zwischen ihrem Alter und dem erzielten Testergebnis erkennen. In der Gruppe mit AO = 7–15 wurden die Testergebnisse in Korrelation mit dem Einreisealter schwächer. Diese Ergebnisse unterstützen die *maturational state hypothesis*: „Human beings appear to have a special capacity for acquiring language in childhood, regardless of whether the language is their first or second“ (ebd.: 95).

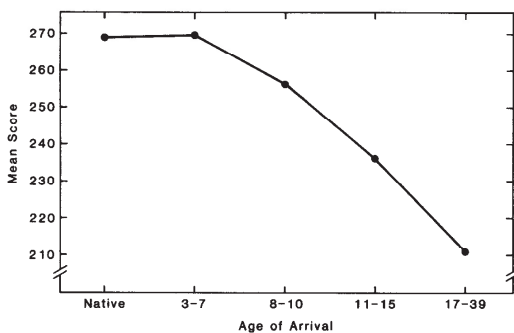


Abb. 1: Zusammenhang der GJT-Ergebnisse und des Einreisealters in die USA (Abb. aus Johnson & Newport 1989: 79).

Im Unterschied zu Lennebergs Beschreibung einer CP setzen Johnson und Newport den Beginn des *Offsets* bei einem Alter von 7–8 Jahren an, nicht erst in der Pubertät. Johnson und Newport teilen die Daten in zwei Gruppen auf (Gruppe I: AO = 3–15; Gruppe II: AO = 17–39). Ihre Auswertung der Daten ergibt eine starke inverse lineare Relation zwischen AO und Sprachkompetenz ($r = -.87$, $p < .01$) bis zum Alter

von 15 Jahren. Für Gruppe II kann keine Korrelation ($r = -.16$, $p < .05$) festgestellt werden (vgl. ebd.: 79–80). Abb. 1 zeigt die Beziehung der Ergebnisse des GJT (Mittelwerte) zum Einreisealter in die USA.

Bialystok & Hakuta (1994) nehmen in ihrer Reanalyse der von Johnson und Newport erhobenen Daten eine Auswertung der Gesamtzahl der Proband_innen vor und unterteilen sie nicht in Gruppen. Sie zeigen, dass sich der *cut-off* von 15 auf 20 Jahre verschiebt und die starke inverse Korrelation ($r = -.87$) erhalten bleibt. In der Gruppe der Personen mit einem AO ≥ 20 verringert sich die Korrelation ($r = -.49$) (vgl. ebd.: 69). Abb. 2 zeigt die Beziehung zwischen GJT-Ergebnissen und dem Einreisealter der Testpersonen nach der Re-Analyse, die einen deutlich anderen Verlauf aufweist.

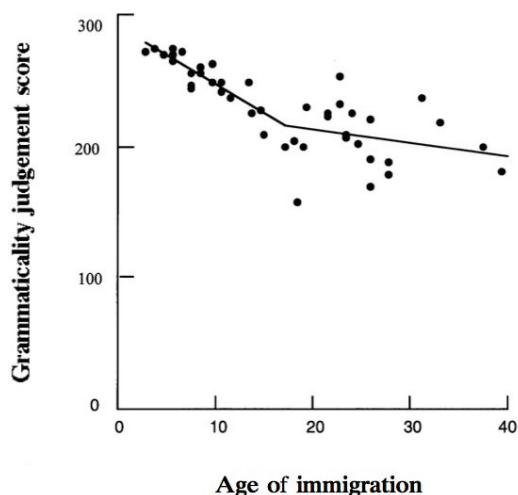


Abb. 2: Grafische Darstellung der Reanalyse der Johnson- und Newport-Daten von Bialystok & Hakuta 1994. (Abb. aus Bialystok 1997: 122. Aufgrund der besseren Darstellung in der Onlineausgabe der Zeitschrift wird auf diese spätere Darstellung zurückgegriffen).

Bialystok und Hakuta erklären die gefundenen Unterschiede zwischen denjenigen, die vor dem Alter von 20 Jahren einreisten, und denjenigen, die älter als 20 waren, mit dem Alter zum Zeitpunkt der Datenerhebung. Während die jüngere Gruppe zum Zeitpunkt des Tests etwa 19 Jahre alt war, war die ältere Gruppe im Durchschnitt 37. Angesichts des Testformats – 276 Items, die jeweils zwei Mal vorgelesen werden mit 1–2 Sekunden Zeit zwischen den Wiederholungen und 3–4 Sekunden Pause zwi-

schen den Items – führen Bialystok und Hakuta einen Teil des Ergebnisses zurück auf „the stage of the subjects in the aging process when they were tested“ (ebd.: 70).⁸ Des Weiteren führen sie die Art des von den Altersstufen bzw. Lebensabschnitten abhängigen Inputs an: während die Jüngeren noch formalen Englischunterricht und Grammatikunterricht an der Schule erhielten, ist davon auszugehen, dass die Älteren weniger Gelegenheiten hatten, Unterricht zu besuchen (vgl. ebd.: 70–71). Die Frage, warum die Testpersonen mit einem AO ≤ 7 genauso wie die getesteten ‚Muttersprachler_innen‘ abschneiden, beantworten Bialystok und Hakuta dahingehend, dass sie eventuell tatsächlich monolinguale ‚Muttersprachler_innen‘ seien: „[...] since only their English was tested in this study, we don’t even know how bilingual they were“ (ebd.: 71). Bialystoks und Hakutas Reanalyse der Daten ist frappierend, führt sie doch vor Augen, wie biegsam statistische Ergebnisse sind und wie stark Gruppenbildungen auf die Ergebnisse einwirken.

Auch Kellerman (1995) weist in seinem Aufsatz „Age before Beauty. Johnson and Newport revisited“ auf mehrere Probleme der Johnson- und Newport-Studie hin. Bezüglich der Methode ist fraglich, was in einem GJT eigentlich gemessen wird. Ohne zusätzliche Informationen, bspw. Gedankenprotokolle, Unterstreichungen oder Korrekturen, können wir nicht wissen, was eine Testperson als ungrammatisch an einem Beispielsatz empfand (vgl. ebd.: 220). Ein weiterer Kritikpunkt Kellermans beschäftigt sich mit der Kohärenz der ausgewählten Beispielsätze des GJT, die teilweise formale und teilweise funktionale Aspekte testeten (vgl. ebd.: 224). Am schwersten wiegt jedoch, was Kellerman bezüglich der Signifikanz der Ergebnisse der Johnson- und Newport-Studie anmerkt. Beispielsweise konnte Wuijtswinkel (1994, zit. nach Kellerman 1995: 225) zeigen, dass sich der gefundene Alterseffekt bei Englischlerner_innen mit einem AO ≥ 12 und der L1 Niederländisch nicht bestätigt. Vielmehr spielt dabei die Ähnlichkeit des Englischen und Niederländischen eine wichtige Rolle. Des Weiteren scheinen die Menge des Inputs und die Sprachlerneignung wesentliche Faktoren zu sein. Wuijtswinkels Studie weist bezüglich des von Johnson und Newport verwendeten GJTs auch einen Deckeneffekt nach.

Johnsons und Newports Studie ist auch insofern ein Klassiker, weil sie mehrmals re-

⁸ DeKeyser (2000) errechnet, dass die Korrelation zwischen AO und Testergebnis robust bleibt ($r = -.54$; $p < .001$), nachdem das Alter bei der Datenerhebung auspartialisiert worden ist (vgl. ebd.: 514).

pliziert wurde. Johnson (1992) selbst wiederholte seine Studie mit denselben Proband_innen ein Jahr später. Allerdings wurden die zu beurteilenden Sätze des GJT's dieses Mal in geschriebener Form präsentiert und nicht als Audio. Die gefundenen Korrelationen fallen weniger stark aus als in der ersten Studie: $r = -.73$ ($p < .001$) für die Gruppe der Testpersonen mit einem $AO \leq 17$ und $r = -.54$ ($p < .01$) für die Gesamtgruppe.

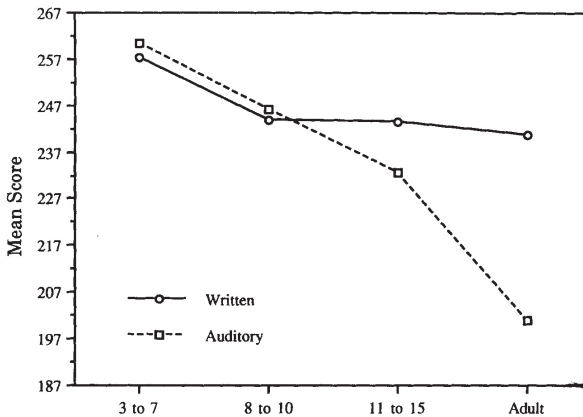


Abb. 3: Die Mittelwerte der Testergebnisse des schriftlichen und auditiven Tests, gebündelt nach Altersgruppen (Abb. aus Johnson 1992: 234).

Die Ergebnisse scheinen Bialystok und Hakuta rechtzugeben, die einen Alterseffekt bei der Datenerhebung der durchschnittlich 20-Jährigen gegenüber den 40-Jährigen vermuteten, und zwar aufgrund der hohen Anforderungen und des knappen Zeitlimits der Audiopräsentation des GJT's. Abb. 3 zeigt, dass der Mittelwert der Gruppe mit $AO = 17-39$ beim schriftlichen GJT deutlich höher ist als beim auditiven GJT. Problematisch an dieser Grafik wie auch den Darstellungen in Johnson & Newport (1989) ist der Umstand, dass das arithmetische Mittel allein nur wenig aussagekräftig ist, so lange wir nichts über die Streuung der Daten wissen. Gerade bezogen auf den Alterseffekt, der eine höhere individuelle Varianz bei den Ergebnissen mit zunehmendem Alter erwarten lässt, ist es fraglich, welche Aussagekraft ein Mittelwert hat (vgl. hierzu auch Birdsong 2005: 117).

Der Vergleich der Standardabweichungen zeigt, dass die Werte der älteren Testpersonen allgemein weiter streuen und bei der auditiven Testpräsentation erheblich weiter streuen als die der jüngeren Gruppe (vgl. die Tabellen in Johnson & Newport 1989: 78 und Johnson 1992: 228). Der Vergleich zeigt auch, dass die ‚muttersprachliche‘ Vergleichsgruppe im schriftlichen Test eine viel höhere Standardabweichung hat als im auditiven. Obwohl das schlechtere Abschneiden der ‚Muttersprachler_innen‘ im Vergleich zur Gruppe AO = 3–7 nicht signifikant wird (vgl. Johnson: 229), wären weitere Informationen zu den ‚muttersprachlichen‘ Proband_innen an dieser Stelle wünschenswert, um die größere Datenstreuung bei ihnen zu erklären. Vermuten ließe sich, dass die ‚muttersprachliche‘ Vergleichsgruppe bspw. über inhomogene Bildungshintergründe verfügt und sich insofern Unterschiede bemerkbar machen zwischen den Kompetenzen Hören und Lesen (vgl. die Unterscheidung von *Basic Language Cognition* und *Higher Language Cognition* (Hulstijn 2011) in Kap. 2.5.1 *Die ‚muttersprachliche Kompetenz‘ als Messlatte für den L2-Lernerfolg*).

DeKeyser (2000) benutzt Johnson und Newports Testbatterie in leicht variiertes Form, um Bley-Vromans *Fundamental Difference Hypothesis* (FDH) (1988) zu überprüfen (ausführlicher zur FDH in Kap. 2.4.4 *Die Natur-Kultur-Debatte*).⁹ DeKeyser's Proband_innen unterscheiden sich allerdings stark von denen von Johnson und Newport: Ihre L1 ist Ungarisch, ihr Bildungshintergrund und sozialer Status ist nicht homogen und reicht von der Beendigung der Schulpflicht bis zur Promotion bzw. von Arbeiter_innen bis zu hochqualifizierten Expert_innen. Auch DeKeyser stellt eine starke Korrelation zwischen AO und Lernerfolg im morphosyntaktischen Bereich fest. Die jüngere Gruppe (n = 15; AO < 16) erreicht homogen gute bis sehr gute Ergebnisse mit einer Korrelation von $r = -.26$ (ns) zwischen AO und GJT-Ergebnis. Die Lernerfolge der älteren Lerner_innen (n = 42; AO > 16) variieren individuell. Die Korrelation liegt bei $r = -.04$ (ns), für die Gesamtgruppe (n = 57) beträgt sie $r = -.63$ ($p < .001$). Trotz der nicht signifikanten Korrelationen in der jüngeren und älteren Gruppe gelten DeKeyser's Ergebnisse als übereinstimmend mit denen von Johnson & Newport (1989) und als Evidenz für die Robustheit der Effekte einer CP (vgl. Birdsong 2005: 118). Von den sechs Testpersonen mit einem AO > 16, die so gut wie

⁹ In DeKeyser's (2000: 500) Formulierung lautet die FDH wie folgt: „According to this hypothesis, adults can no longer rely on the innate mechanisms for implicit language acquisition and must, therefore, rely on alternative problem-solving mechanisms.“

die jüngere Gruppe abschneiden, erzielten fünf ebenfalls überdurchschnittliche Ergebnisse bei der mittels MLAT (*Modern Language Aptitude Test*, Carroll & Sapon 1959) gemessenen Sprachlerneignung (bzw. grammatischen Sensitivität, s. u.), und zwar unabhängig vom sonstigen Bildungsniveau. Insofern lässt sich ein Einfluss der Sprachlerneignung auf die erzielten Testergebnisse bei erwachsenen Lerner_innen annehmen, auch wenn die Zahl der Proband_innen zu klein ist, um dies wirklich konstatieren zu können. Im Gegenzug weisen die Ergebnisse der jüngeren Testgruppe keinen signifikanten Zusammenhang zwischen guten Testergebnissen im GJT und überdurchschnittlichen Ergebnissen im MLAT aus.¹⁰ Schließlich zeigt DeKeyser durch eine Fehleranalyse, dass die erwachsenen Lerner_innen sich von den jüngeren durch ihren expliziten Lernstil unterscheiden und morphosyntaktische Aspekte, die salient sind, erfolgreich erwerben, im Gegensatz zu nicht salienten Aspekten (vgl. ebd.: 515–517). Damit ist für ihn die FDH bewiesen: Erwachsene lernen eine L2 nur dann erfolgreich, wenn sie auf explizite, analytische Problemlösestrategien zurückgreifen (vgl. ebd.: 518). Für DeKeyser ist somit auch Evidenz für die CPH gegeben, allerdings will er sie in einem engen Sinn definiert wissen: „that is, applying only to implicit learning of abstract structures“ (ebd.: 518).

Birdsong & Molis (2001) replizierten Johnsons und Newports Studie. Sie untersuchten insgesamt 61 Proband_innen mit der L1 Spanisch ($r = -.77$; $p < .0001$). Die jüngere Gruppe ($n = 29$; $AO < 16$) erreicht sämtlich sehr gute Ergebnisse ($r = -.24$; $p = .22$); die ältere Gruppe ($n = 32$; $AO > 17$) schneidet schlechter ab und variiert individuell stärker ($r = -.69$, $p < .0001$). Der von Birdsong & Molis (2001) berechnete Kurvenverlauf unterscheidet sich deutlich von den Ergebnissen aus Johnson & Newport (1989), wie Abb. 4 zeigt. Zu hinterfragen bleibt an dieser Stelle, wie viel explikative Kraft die Regressionslinie für die älteren Lerner_innen angesichts der Streuweite der Daten hat. Festzuhalten bleibt, dass bei Johnson & Newport (1989), DeKeyser (2000) und auch Birdsong & Molis (2001) das AO die stärkste Voraussagekraft auf das Testergebnis hat.

¹⁰ Die Korrelationen zwischen den Ergebnissen des MLATS und des GJTs: jüngere Gruppe ($r = .07$; ns) und ältere Gruppe ($r = .33$; $p < .05$).

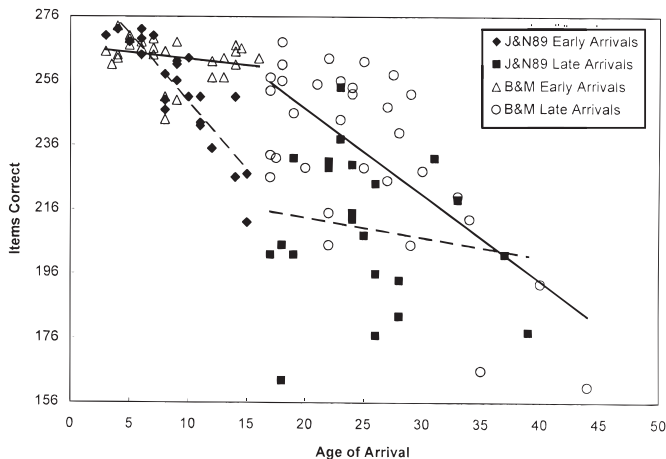


Abb. 4: Vergleich der Ergebnisse der Studien von Birdsong & Molis 2001 (durchgezogene Linie) und Johnson & Newport 1989 (gestrichelte Linie), gebündelt nach zwei Altersgruppen (AO < 16 und AO > 16) (Abb. aus Birdsong & Molis 2001: 240).

Im *Barcelona Age Factor Project* (vgl. Muñoz 2006a, 2006b, 2006c) werden Geschwindigkeit (*rate*) und Ergebnisse des Fremdsprachenlernens von Schüler_innen-Gruppen mit unterschiedlichem Erwerbsbeginn verglichen. Ein in Katalonien implementiertes neues Curriculum für die Primar- und Sekundarstufe reformierte u. a. den Beginn des Englischunterrichts an den Schulen, der von der 6. Klasse (AO = 11) in die 3. Klasse (AO = 8) vorverlegt wurde. Die sich daraus ergebende Chance, zwei Schüler_innen-Gruppen mit unterschiedlichem Erwerbsbeginn parallel untersuchen zu können, nutzen die Studien, die im Zusammenhang des *Barcelona Age Factor Projects* (1995–2004) (vgl. die Beiträge in Muñoz 2006b) entstanden.

Zusätzlich zu den beiden Hauptgruppen mit AO = 8 (Gruppe A, $n = 164^{11}$) und AO = 11 (Gruppe B, $n = 107$) wurden Lerner_innen mit AO = 14 (Gruppe C, $n = 21$) und AO ≤ 18 (Gruppe D, $n = 67$) untersucht, wobei die Gruppen C und D eine andere Fremdsprache (i. d. R. Französisch) zuerst lernten und entsprechend später mit Englisch anfangen (vgl. Muñoz 2006a: 14f.). Die Schüler_innen sind zumeist bilingual mit den Sprachen Spanisch und Katalanisch aufgewachsen. Die Daten wurden an

¹¹ Die Anzahl der Proband_innen variiert zu den drei Messzeitpunkten. Die hier wiedergegebenen Zahlen entsprechen der Anzahl bei der ersten Messung (vgl. Muñoz 2006a: 15).

22 Schulen erhoben, um Lehrer_inneneffekte und vergleichbare sozioökonomische Hintergründe der Schüler_innen zu kontrollieren (vgl. Muñoz 2006c: 83). Schüler_innen mit außerschulischem Kontakt zur englischen Sprache wurden in der Auswertung nicht berücksichtigt, um die Vergleichbarkeit der Fälle zu gewährleisten. Die Schüler_innen haben zwei bis drei Unterrichtsstunden Englisch pro Woche. Im alten Curriculum werden insgesamt 750 Stunden in sieben Jahren unterrichtet, das neue Curriculum sieht 800 Stunden verteilt auf zehn Jahre vor (vgl. Muñoz 2006a: 14).

In einer Kombination von longitudinaler Studie und Querschnittsstudie wurden die Gruppen nach 200, 416 und 726¹² Unterrichtsstunden mit einer umfangreichen Testbatterie untersucht. Folgende Formate kamen dabei zum Einsatz: Fragebogen (Sprachgebrauch und Extraunterricht), Grammatiktest (Multiple Choice), Cloze-Test (in Englisch, Katalanisch und Spanisch), Diktate (in Englisch, Katalanisch und Spanisch), Hörverstehen, Schreiben, mündliches Erzählen, mündliches Interview, phonetische Imitation, phonetische Diskrimination und Rollenspiel (vgl. Muñoz 2006a: 16 und Muñoz 2006c: 84).

Die Studie zeigt, dass v. a. die Schnelligkeit des Lernens (*rate*) mit dem Alter bei Lernbeginn steigt. Die besten Ergebnisse werden von den erwachsenen Lerner_innen erzielt, die schwächsten Ergebnisse von den frühen Englischlerner_innen. In einigen Tests erreichen die erwachsenen Lerner_innen nach 200 Stunden Englischunterricht mehr Punkte als die frühen nach 726 Stunden; nach 416 Stunden liegen die Erwachsenen in allen Bereichen deutlich vor Gruppe A nach 726 Stunden. Dies verdeutlicht klar das zentrale Ergebnis der Studie: die Vorteile erwachsener Lerner_innen bezogen auf die Lerngeschwindigkeit im gesteuerten Fremdsprachenunterricht. Auch ein Vergleich der Gruppen A und B (Erwerbsbeginn acht bzw. elf Jahre) weist die Überlegenheit der älteren Lerngruppe auf: Bereits nach 200 Unterrichtsstunden liegt Gruppe B vor Gruppe A; nach 416 Stunden wird dieser Abstand signifikant (vgl. auch die Zusammenfassung der Studie in Schlak 2006: 16f.). Muñoz (vgl. 2006a: 33) führt in ihrer Conclusio die Überlegenheit der älteren Lerner_innen auf deren kognitive Entwicklung zurück, eventuell auch auf die Vorteile der *Explicit Learning Mechanisms*, die sich mit zunehmendem Alter entwickeln. Sie vermutet Folgendes: Sobald die Jüngeren in ihrer kognitiven Entwicklung aufholen, gleichen

¹² Für Gruppe C und Gruppe D liegen keine Daten des dritten Messzeitpunktes vor.

sich die Ergebnisse an die ältere Vergleichsgruppe unter der Bedingung an, dass von derselben Zeit und Unterrichtsweise ausgegangen wird. Mit diesem Ergebnis kann auf empirischer Basis den Vorhersagen früherer Studien widersprochen werden. Diese gingen auf der Grundlage des Spracherwerbs unter Immersionsbedingungen davon aus, dass jüngere Lerner_innen auch im Kontext des Fremdsprachenunterrichts den älteren überlegen wären:

„In sum, these findings suggest that second language learning success in a foreign language context may be as much a function of exposure as of age. Exposure needs to be intense and to provide an adequate model. Initial age of learning seems to be more relevant for skills that can be acquired implicitly, whereas age at learning can be seen as a factor explaining the rate of learning of most skills.“ (Muñoz 2006a: 35)

Muñoz (2006c: 89) stellt in ihrer Zusammenfassung und Diskussion der Ergebnisse heraus, dass die älteren Lerner_innen schneller sind und die jüngeren dies bis zur 726. Stunde nicht aufholen können. Dies betrifft alle getesteten Bereiche außer Hörverstehen und phonetische Imitation. Dies könne, so Muñoz, darauf zurückgeführt werden, dass diese Fertigkeiten eventuell nicht gleichermaßen sensibel auf individuelle kognitive Unterschiede und die Herausbildung akademischer Kompetenzen reagieren. Im Anschluss an Bley-Vromans *Fundamental Difference Hypothesis* (vgl. Kap. 2.4.4 *Die Natur-Kultur-Debatte*) führt Muñoz die gefundenen Unterschiede darauf zurück, dass die Jüngeren nicht von ihrem *Implicit Learning Mechanism* profitieren könnten, weil sie nicht genug Input hätten (vgl. ebd.: 90).

Im Zusammenhang der CPH und der kontroversen Diskussion darüber, ob Spätler_innen ein muttersprachliches Niveau erreichen können (s. Kap. 2.5.1 *Die ‚muttersprachliche Kompetenz‘ als Messlatte für den L2-Lernerfolg*), untersuchen Abrahamsson & Hyltenstam (2009) in einer umfangreichen Studie 195 Bilinguale mit den Sprachen Schwedisch und Spanisch. Ihr Erwerbsbeginn lag zwischen 1 und 47 Jahren und sie schätzen sich selbst als muttersprachlich ein. In der ersten Phase des Projekts wird die *perceived nativeness* der Testpersonen durch eine Jury bewertet (vgl. die ausführliche Darstellung des Phonetik-Ratings in Kap. 3.4.2.5 *Phonetik-Test*). In einer zweiten Phase (*scrutinized nativeness*) werden 41 der im ersten Schritt als ‚muttersprachlich‘ bewerteten Testpersonen – deren AO möglichst gleichmäßig in der Spanne zwischen eins und 20+ liegt – mit einer aus 20 Tests bestehenden Batterie intensiv untersucht.

Eine zehnköpfige Jury aus schwedischen ‚Muttersprachler_innen‘ bewertet kurze, spontan geäußerte Sprechproben im Hinblick auf die *perceived nativeness*. Von den 107 frühen Sprachlerner_innen (AO ≤ 11 Jahre) wurden 62 % von neun der zehn Rater als ‚muttersprachlich‘ bewertet. Nur drei Personen dieser Gruppe wurden von allen Ratern als ‚nicht muttersprachlich‘ beurteilt. In der Gruppe der Spätlerner_innen (AO = 12–47) wurden 6 % von neun der zehn Juror_innen als ‚muttersprachlich‘ eingeschätzt (mit einem AO von 12–17 Jahren), während 41 % von allen zehn Ratern als nicht-muttersprachlich bewertet wurden (vgl. ebd.: 272–273). Die Auswertung zeigt, dass das AO bei weitem am besten die *perceived nativeness* vorhersagen kann (vgl. ebd.: 288).

In der zweiten Phase wurden die folgenden Instrumente eingesetzt, um Sprachproduktion, Sprachverstehen, Morphosyntax und den Umgang mit formelhaften Sprachstrukturen zu testen (vgl. ebd.: 276f.): *Production and perception of voice onset time (VOT)*; *speech perception in noise*; *grammaticality judgment*; *grammatical, lexical, and semantic inferencing*; *formulaic language (idioms and proverbs)*. Bei der Auswertung dieser zehn Tests stellt sich heraus, dass nur drei der 41 ausgewählten Testpersonen in allen Tests im Bereich der 15 ‚Muttersprachler_innen‘ abschnitten, die dieselben Tests absolvierten. Diese drei hatten einen AO von drei, sieben und acht Jahren:

„In other words, only a few of the early learners and *none* of the late learners exhibited actual, linguistic nativelikeness across a broad range of tasks when their performance was scrutinized in detail.“ (ebd.: 287)

Ein weiteres Ergebnis dieser Untersuchung besagt, dass die hauptsächlichsten Veränderungen bezüglich der erreichten Bewertungen in der Test-Batterie zur *scrutinized nativeness* mit einem AO ≤ 12 Jahren zusammenhängen (vgl. Abb. 5). Unter den jüngeren Lerner_innen (hier AO ≤ 11) sind die phonetischen Lernerfolge (VOT-Produktion und -Perzeption, Perzeption mit Störgeräuschen) verbreiteter als muttersprachliche grammatische Intuition. Bei den älteren Lerner_innen (AO > 11) ist es genau umgekehrt. Diese Unterschiede sind jedoch nicht signifikant (vgl. ebd.: 286). Abrahamsson und Hyltenstam betonen zudem, dass sie daraus nicht ableiten, dass bestimmte sprachliche Bereiche ab einem bestimmten Alter nicht mehr lernbar seien. Des Weiteren halten sie in ihrer Auswertung fest, dass es zwar fünf der 88 älteren Lerner_innen gibt, die in allen Tests muttersprachlich abschneiden. Sie betonen

aber gleichzeitig, dass diese fünf alle einen AO von 12–17 Jahre haben (vgl. ebd.: 288). Sie halten es für äußerst unwahrscheinlich, eine_n ältere_n Lerner_in (hier älter als 17) zu finden, die/der in allen (oder den meisten) Bereichen ‚muttersprachlich‘ abschneide (vgl. ebd.: 289).

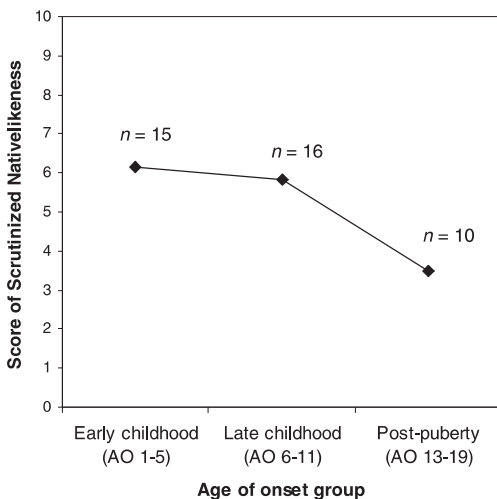


Abb. 5: Durchschnittliche ‚Scrutinized nativeness‘-Ergebnisse der Testpersonen, aufgeteilt in drei AO-Kategorien (Abb. aus Abrahamsson & Hyltenstam 2009: 282).

Ein zentrales Ergebnis der Studie ist, dass das frühe L2-Lernen nicht automatisch zur ‚Muttersprachlichkeit‘ führt. Abrahamsson und Hyltenstam zufolge sind die Instrumente der Studien, deren Ergebnisse zeigen, dass alle Lerner_innen im Kindesalter ‚muttersprachliches‘ Niveau erreichen (bspw. der von Johnson & Newport 1989 verwendete GJT), zu leicht und bringen deswegen Deckeneffekte hervor (vgl. ebd.: 290f.).

Abrahamsson (2012) merkt zu dieser Studie an, dass mit dieser umfangreichen Testbatterie „any advanced and apparently nativelike adult L2- learner could be revealed as only superficially nativelike (or near-native)“ (ebd.: 197, vgl. auch die Diskussion in Kap. 2.5 ‚Muttersprachliche Kompetenz‘). In seiner Studie von 2012 ist das Forschungsdesign entsprechend schlanker: 200 Testpersonen mit L1 Spanisch und L2 Schwedisch (n = 101 mit AO = 1–15; n = 99 mit AO = 16–30) und 20 schwe-

dische ‚Muttersprachler_innen‘ legen einen auditiven GJT ab und werden bezüglich ihrer Perzeption der VOT getestet. Der Einfluss des Erwerbsbeginns wird durch die kombinierte Auswertung der phonetischen und grammatischen Intuition errechnet. Darüber hinaus möchte die Studie Hinweise bezüglich der Unterschiedlichkeit der Spracherwerbsmechanismen von Kindern und Erwachsenen empirisch erheben. Der Alterseffekt ist sowohl in der einzelnen als auch der kombinierten Auswertung der Tests deutlich (vgl. ebd.: 200f.), wobei die Scheidelinie bei AO = 13 liegt (vgl. ebd.: 208). Für Abrahamsson ist damit die CPH in der Formulierung von Lenneberg (1967) und Johnson & Newport (1989) bestätigt. Er ist aber mit den Schlussfolgerungen bezüglich der zur Anwendung kommenden Lernmechanismen vorsichtig, da nur die junge Lerner_innen-Gruppe eine positive Korrelation zwischen den beiden Tests aufwies. Bei den ‚Muttersprachler_innen‘ und den späten Lerner_innen konnte keine Beziehung zwischen den Testergebnissen nachgewiesen werden (vgl. ebd.: 208). Es lasse sich nur mit äußerster Vorsicht vermuten, dass die bei den frühen Lerner_innen gefundene Korrelation auf die Nutzung eines impliziten Lernmechanismus‘ verweise, mit dem die Sprache als Ganzes „from mere exposure“ (Lenneberg 1967) erworben werde könne.

Dimroth (2007) untersucht die Erwerbsverläufe von Kindern und Jugendlichen, die Deutsch als Zweitsprache lernen, auf der Grundlage des DaZ-AF-Korpus.¹³ Die longitudinale Fallstudie DaZ-AF erfasst die sprachliche Entwicklung der beiden Schwestern Nastja (AO = 8;7) und Dascha (AO = 14;2) mit der L1 Russisch, die beide vor ihrer Ankunft in Deutschland keinen Kontakt zur deutschen Sprache hatten. Ihr ungesteuerter Erwerb des Deutschen wird durch wöchentlich stattfindende Aufzeichnungen von freien Gesprächen mit erwachsenen oder gleichaltrigen deutschen ‚Muttersprachler_innen‘ über 18 Monate hinweg dokumentiert (vgl. ebd.: 124). Die Transkriptionen dieser Spontansprachdaten wurden von einer Reihe von Forscher_innen hinsichtlich des Erwerbs von morphosyntaktischen Eigenschaften ausgewertet (Bast 2003, Dimroth 2008, Dimroth & Haberzettel 2008, Pagonis 2009b¹⁴), Dimroth (2007) beruht auf diesen Arbeiten.

¹³ Das Korpus Deutsch als Fremdsprache – Altersfaktor (DaZ-AF-Korpus) wurde von 2000 bis 2002 unter der Leitung von Ursula Stephany und Christine Dimroth erhoben.

¹⁴ Pagonis (2009b) wird im Kap. 2.4.2 *Motivation und Erwerbsantrieb* ausführlich dargestellt.

Bezüglich der Wortstellung hält Dimroth (ebd.: 124f.) fest, dass beide Lernerinnen zunächst eine Wortstellung verwendeten, die der „Basisvarietät“ von Erwachsenen ähnelte: Erwachsene Lerner_innen bilden in den frühen Stadien des ungesteuerten L2-Erwerbs typischerweise eine Basisvarietät aus, die von den Prinzipien der Informationsgliederung geprägt ist, über einfache, meist pragmatisch motivierte Wortstellungsregeln verfügt und keine produktive Flexionsmorphologie kennt. Nach eineinhalb Jahren¹⁵ besteht zwischen beiden weder hinsichtlich der Verwendung von Auxiliaren noch der Anhebung finiter Verben ein Unterschied. Trotz verschiedener Erwerbsreihenfolgen sind beide dabei nicht mehr von der zielsprachlichen Norm zu unterscheiden. Ein Unterschied besteht bei deklarativen Hauptsätzen, die Dascha nicht konsequent zielsprachlich umsetzt. Beide Lernerinnen erwerben die Anhebung finiter Verben, Negation und Inversion, wobei Dascha nur manche dieser Regeln konsequent anwendet. Dimroth leitet aus den gefundenen Ergebnissen ab, dass Altersunterschiede nicht nur den Endstand betreffen, sondern bereits die Entwicklung der Lerner_innensprachen beeinflussen (vgl. ebd.: 127).

Bei der Flexionsmorphologie (vgl. ebd.: 127f.) zeigt sich, dass die Subjekt-Verb-Kongruenz von beiden in derselben Reihenfolge erworben wird, Nastja allerdings schneller ist. Auch bei den Vergangenheitsformen ist Nastja schneller. Die Pluralformen werden von Nastja so gut erworben, dass sie sich nach 18 Monaten nicht mehr von der zielsprachlichen Verwendung unterscheiden, während Dascha Übergeneralisierungen vornimmt. Beim Erwerb der attributiven Adjektive zeigen sich schon früh gravierende Unterschiede. Dimroth merkt hierzu an, dass der kommunikative Nutzen dieses Merkmals beschränkt ist:

„Ob und wie sich die Altersfaktoren auf den Zweitspracherwerb auswirken, kann also nicht pauschal vorausgesagt werden, sondern muss immer im Hinblick auf den einzelnen Lerngegenstand – als Teil des zielsprachlichen *und* des lernersprachlichen Systems – spezifiziert werden. Überspitzt kann man festhalten, dass Eigenschaften der zielsprachlichen Grammatik, die häufig, auffällig und kommunikativ relevant, zugleich aber nicht zu komplex und nicht redundant sind, in allen Lebensaltern erworben werden.“ (ebd.: 134)

Czinger (2014) legt die jüngste Untersuchung dieses Korpus vor. Sie vergleicht den Erwerbsverlauf und die Erwerbsgeschwindigkeit der beiden Schwestern bei drei zentralen Problembereichen der Verbstellung des Deutschen: Distanzstellung des

¹⁵ Dimroth weist darauf hin, dass eineinhalb Jahre ein zufällig gewählter Endstand ist, der nicht mit dem Ende der Entwicklung gleichgesetzt werden kann (vgl. ebd.: 127).

finiten und infiniten Verbteils in der Satzklammer (XV-Stellung), Verb-Zweit-Stellung (V2-Stellung) mit Inversion, Verb-End-Stellung (VE-Stellung) im Nebensatz. Darüber hinaus erarbeitet sie ein multifaktorielles Modell, um den Einfluss des AO auf den ungesteuerten L2-Erwerb zu erfassen. Das Besondere an diesem Modell ist, dass zwischen jüngeren und älteren Kindern, Jugendlichen und Erwachsenen klar unterschieden wird und viele Einflussfaktoren integriert sind: Plastizität des Gehirns, kognitive Reife, Veränderung der Motivation, Bildungssprache, Sprachlerneignung, kommunikative Bedürfnisse (vgl. ebd.: 25f.).

Czinglar zeigt, dass die jüngere Lernerin Nastja die Verbstellungsregeln schneller erwirbt als die ältere Dascha. Am Ende des Beobachtungszeitraums ist Nastjas Verbstellung im Gegensatz zu Daschas zielsprachlich (vgl. ebd.: 247–248). Daschas Erwerbtempo ist zu keinem Zeitpunkt schneller als das von Nastja. Czinglar zeigt, dass sich drei der vier von ihr aufgestellten explanativen Hypothesen bestätigen lassen: (1) Der Erwerb der Finitheit unterstützt den V2-Erwerb; (2) die ältere Lernerin hält länger am negativen Transfer aus der L1 fest; (3) es besteht ein Zusammenhang zwischen dem Erwerb der Verbstellung und des Verblexikons. Nastjas Verblexikon weist sowohl bezogen auf die *Types* als auch auf die Anzahl der seltenen Verben eine größere Diversität auf (vgl. ebd.: 248f.).

Schließlich möchte ich an dieser Stelle noch zwei Studien zusammenfassen, die bei der Erstellung des Fragebogens für die zweite Phase eine besondere Rolle gespielt haben, nämlich Ioup, Boustagui, El Tigi, Manal & Moselle (1994) und Moyer (2004).

Ioup et al. (1994) untersuchten den außerordentlichen Lernerfolg einer Engländerin in Ägypten (AO = 21, L1 Englisch, L2 Arabisch) in einem ungesteuerten Spracherwerbskontext mit ausreichendem zielsprachlichem Input und Feedback auf ihre Äußerungen. Der Sprachstand wurde mit Tests in den drei Bereichen *speech production*, *accent identification* und *grammatical intuition* gemessen (vgl. ebd.: 79–91), die die Probandin alle mit überdurchschnittlichem Erfolg und bis auf wenige Ausnahmen im Bereich der ‚muttersprachlichen‘ Kontrollgruppe absolvierte. Ioup et al. führen diesen Erfolg in erster Linie auf die Sprachlerneignung der Lernerin zurück, nennen aber auch andere Faktoren wie ihre Extravertiertheit, ihren Sprachlerntypus sowie ihre persönlichen biografischen Daten.

loup et al. (ebd.: 92) referieren, dass die Lernerin verschiedene Sprachlernstrategien anwendete. Beispielsweise konzentrierte sie sich von Anfang an trotz des naturalistischen Settings, in dem sie lernte, auf die Flexionsformen, weil ihr bewusst war, dass sie eine wichtige Rolle spielen. Als Dozentin für Englisch als Zweitsprache war die Testperson sensibilisiert für die Rolle der Grammatik beim L2-Erwerb. Außerdem bemühte sie sich aktiv um Feedback ihres Umfelds zu ihren Äußerungen und Notizen und versuchte es dann auch umzusetzen (vgl. ebd.). Zu den besonderen biografischen Merkmalen in Bezug auf ihren Sprachlernerfolg gehört der Umstand, dass sie einen ägyptischen Lebenspartner mit ausschließlich arabischsprachiger Verwandtschaft hat. Sie gab zu Protokoll, dass es ihr wichtig sei, sich mit der Verwandtschaft unterhalten zu können. Darüber hinaus waren für ihre berufliche Tätigkeit Arabischkenntnisse notwendig (vgl. ebd.: 77). Bezüglich der Rolle, die Talent bzw. Sprachlerneignung (bei loup et al. werden beide Begriffe angeführt) bezüglich des Lernerfolgs spielt, muss erwähnt werden, dass sie nicht mit einem der gängigen Testverfahren gemessen wurden, sondern aufgrund von Selbstaussagen der Probandin: Beispielsweise sei ihr L1-Erwerb besonders zügig vonstatten gegangen, in ihrer Familie sei das Geschwind-Cluster anzutreffen (Linkshändigkeit, Allergien, Zwillinge), sie sei für Mathematik und Zahlen völlig unbegabt, sie habe ein gutes Gedächtnis, könne Codes gut entschlüsseln und habe schon immer ein großes Talent gehabt, Stimmen zu imitieren. Die Zusammenführung von Talent und Sprachlerneignung ist problematisch und im Rahmen der Fremdsprachenerwerbsforschung unüblich, wo in der Regel auf die Sprachlerneignung fokussiert wird (vgl. Kap. 2.4.3 *Sprachlerneignung*).

Auch Moyer (2004) stellt in ihrer Studie Faktoren dar, die den L2-Lernerfolg von Erwachsenen erklären können. Die von ihr untersuchten 25 in Berlin lebenden Testpersonen sind alle zwischen 25 und 35 Jahre alt, studieren und ihre LOR variiert zwischen sechs und 22 Jahren. Ihre Lernsituation wird als optimal bezeichnet, weil sie sowohl formalen Unterricht als auch informellen Sprachkontakt haben. Ihre Ausgangssprachen repräsentieren diverse kulturelle Hintergründe: Polnisch, Russisch, Ukrainisch, Slowakisch, Französisch, Türkisch,ritisches und amerikanisches Englisch (vgl. ebd.: 54–55). Moyer führte mit ihren Proband_innen ein umfangreiches Testset durch, das die Bereiche *biological-experiental*, *social-psychological*, *instructional-cognitive*, *experiental-social* und *linguistic performance* abdeckt. Wie zu erwar-

ten, ist das AO signifikant für die Bewertung der sprachlichen Performanz. Mit einer Interpretation des Einflusses der LOR hält sich Moyer zurück, weil LOR mit vielen anderen Faktoren korreliert (vgl. ebd.: 75).

Bei den sozial-psychischen Faktoren erwies sich v. a. die Motivation als starke Einflussgröße, wobei die persönliche Motivation stärker als die berufliche wirkte (vgl. ebd.: 76). Bezüglich der *instructional-cognitive* Faktoren wurde der Sprachunterricht als hochsignifikante Einflussgröße gefunden. Bezüglich der Unterrichtsmethoden erwies sich der Grammatik-Übersetzungs-Unterricht nur wenig förderlich für die Fertigkeit Sprechen (vgl. ebd.: 77). Bei den *experiential-social* Faktoren wurde die Möglichkeit, mit L1-Sprecher_innen zu kommunizieren, signifikant. Andere Faktoren wie bspw. Feedback erreichten nicht das Signifikanzniveau (vgl. ebd.). Allerdings ist eine *native-like performance* auch möglich mit niedrigerem Kontakt zu L1-Sprecher_innen: So erreichte eine Frau mit langem formalen Unterricht und breit angelegter Motivation ohne ausgeprägte Kontakte ein *native-like level* (vgl. ebd.: 78). Moyer gibt das Profil eines besonders erfolgreichen Falls wieder: eine hohe L2-Lernmotivation und ein Selbstbild als erfolgreiche L2-Lernerin bzw. -Lerner, eine integrative Motivation im Sinne einer Bleibeabsicht, eine sprachorientierte Identität sowie intensiver Unterricht und Feedback über Jahre, einhergehend mit informellem Sprachkontakt (vgl. ebd.: 120f.).

In diesem Unterkapitel wurden die einschlägigen wie auch die für diese Arbeit besonders relevanten Studien vorgestellt und diskutiert. Diesem Überblick folgt im nächsten Unterkapitel die Herausarbeitung der Forschungslücke, zu deren Schließung die hier vorliegende Arbeit einen Beitrag leistet.

2.3 Forschungslücke

Vergegenwärtigt man sich die geometrischen Darstellungen der CPH, lässt sich eine von Birdsong (2005: 113) als *stretched Z* bezeichnete Figur erkennen, die ihm zufolge eine prototypisch verlaufende CP abbildet (vgl. Abb. 6).

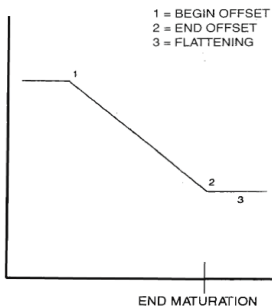


Abb. 6: Geometrische Eigenschaften eines prototypischen Verlaufs einer CP (Abb. aus Birdsong 2005: 113).

Demgegenüber stellen van Boxtel, Bongaerts & Coppens (2005: 356) den *onset* als ansteigende Kurve dar. Dennoch lässt sich in ihrer Darstellung ein *stretched Z* erkennen (vgl. Abb. 7).¹⁶

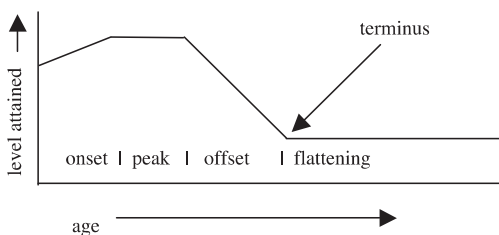


Abb. 7: Geometrische Eigenschaften einer CP (Abb. aus van Boxtel et al. (2005: 356).

Neben der problematischen Diskussion bezüglich der CPH im Allgemeinen (s. o.) ist an diesen Darstellungen v. a. das flache Auslaufen (*flattening*) des erreichten Sprachstandes zu hinterfragen. Angesichts zahlreicher Studien¹⁷, die für erwachsene L2-Lerner_innen individuell sehr unterschiedliche erreichte Sprachniveaus ausweisen, liegt folgende Formulierung zum Verlauf der Sprachkompetenz innerhalb der CPH näher:

¹⁶ Sowohl Birdsong als auch van Boxtel et al. sprechen in diesem Zusammenhang von der CP des Spracherwerbs, diskutieren diese dann jedoch im Kontext des L2-Erwerbs.

¹⁷ Vgl. etwa die ‚klassischen‘ Untersuchungen von Johnson & Newport (1989) oder die aktuelle Untersuchung von Abrahamsson & Hyltenstam (2009).

„Die erreichte Sprachkompetenz lässt während der kritischen Periode biologisch bedingt linear nach, es kommt zu einem „cut-off point“ in der statistischen Auswertung am Endpunkt der kritischen Periode, der Späterwerb nach Abschluss der kritischen Periode ist primär von sehr unterschiedlichen Umweltbedingungen abhängig, und die durchschnittlich erreichte Sprachkompetenz lässt dadurch diskontinuierlich nach.“ (Schlak 2003: 18)

Anstatt sich wiederholt an der Falsifizierung der CPH abzuarbeiten, die aufgrund der verschiedenen untersuchten sprachlichen Register und dabei eingesetzten Forschungsmethoden äußerst vielgestaltig ist, empfiehlt es sich, die Faktoren zu eruiieren, die den erfolgreichen Spracherwerb im Erwachsenenalter unterstützen:

„Rather than providing more evidence for the (non-)existence of a critical period, future research should be aimed at trying to find out what factor(s) contribute(s) most to language learning success and what are the best ways to influence learners' sensitivity to non-salient aspects of the input.“ (van Bostel et al. 2005: 378)

Dementsprechend wird im Folgenden das Augenmerk auf die verschiedenen Erklärungsansätze bezüglich der Altersunterschiede beim Erlernen einer Fremdsprache gelegt. Zweifellos beeinflusst Alter den Sprachlernerfolg, es wäre aber zu zeigen, dass dies nicht aufgrund einer CP geschieht, sondern wegen der Assoziierung des Alters mit sozialen, psychologischen, erzieherischen und anderen Faktoren (vgl. Marinova-Todd, Marshall & Snow 2000: 28).

Nachdem in diesem Kapitel die Forschungslücke herausgearbeitet wurde, wird im nächsten Kapitel dargestellt, was unter Alter verstanden wird und inwiefern es wichtig ist, neben den biologischen Aspekten auch sozial-konstruktivistische Begriffsdimensionen zu berücksichtigen.

2.4 Was ist Alter? Was unterscheidet ältere von jüngeren Lerner_innen?

Neben der chronologischen Dimension von Alter, die in Jahren, Monaten, Tagen, Stunden etc. gemessen wird, besitzt Alter auch eine soziale Dimension. Berndt (2003), die die spezifischen Erfordernisse des L2-Erwerbs von Senior_innen erforschte, weist darauf hin, dass in der Gerontologie unterschieden wird zwischen kalendarischem bzw. chronologischem Alter und dem biologisch-funktionellem Alter (Vitalität), die mit der psychischen Realität des individuell erlebten Alters verkoppelt sind.

Im Zusammenhang der Frage, ob *Chunks* lebenslang erlernbar sind, bezieht sich Aguado (2013) ausdrücklich auf einen um die soziale Dimension erweiterten Altersbegriff. Ähnlich wie ‚Geschlecht‘ sei ‚Alter‘ keine rein biologische Größe, sondern größtenteils sozial konstruiert, d. h. von zahlreichen Umgebungsvariablen abhängig. Bezogen auf den L2-Erwerb bedeutet das:

„Von besonderer Relevanz sind hier die altersbedingt unterschiedlich gelagerten kommunikativen Bedürfnisse der lernenden Individuen sowie die an sie von der Umgebung gestellten Erwartungen und Anforderungen im Hinblick auf Korrektheit, Flüssigkeit und Komplexität ihres Sprachgebrauchs.“ (ebd.: 135)

Zu den zahlreichen Unterschieden zwischen kindlichen und erwachsenen L2-Lerner_innen gehören bspw. die Quantität und Qualität des Inputs, Wiederholung und Anwendung der zu lernenden Sequenzen, der Einfluss des abgeschlossenen bzw. noch nicht abgeschlossenen L1-Erwerbs, der Automatisierungsgrad der L1, soziokulturelle Adaption an die Zielkultur und der inhaltliche Aufmerksamkeitsfokus der Erwachsenen (vgl. ebd.: 135–138).

Andrew (2012) betrachtet in ihrer Studie die soziale Konstruktion von Alter im Kontext des Englisch-als-Fremdsprache-Unterrichts in Mexiko. Durch die Auswertung von Interviews und den narrativen Erzählungen ihrer sieben Proband_innen sowie der Unterrichtsbeobachtungen (vgl. ebd.: 69) rekonstruiert sie, wie sich ihre Proband_innen in die drei allgemeinen Kategorien ältere, mittlere und jüngere Erwachsene eingruppiert und eingruppiert werden. Diese Kategorien haben jedoch nur einen schwachen Bezug zum chronologischen Alter. So ist die Grenze zwischen den Mittleren und Älteren eher durch das Beenden der Lohnarbeit markiert und nicht durch das Erreichen eines gewissen Lebensalters (vgl. ebd.: 153). Alter ist untrennbar verkettet mit anderen Subjektpositionen wie bspw. Geschlecht, beruflicher Identität, Ethnizität und sozialem Status (vgl. ebd.: 155). Der in den westlichen Gesellschaften dominante Altersdiskurs des Verfalls findet sich auch in Mexiko wieder. Der damit einhergehende *Ageismus* manifestiert sich auf vielen Ebenen (vgl. ebd.: 157). Alle Proband_innen teilen die Auffassung, dass sie es als erwachsene Lerner_innen einer Fremdsprache viel schwerer haben als Kinder. Gleichzeitig sind sie davon überzeugt, dass altersbedingte Nachteile durch Motivation, Hingabe und harte Arbeit kompensiert werden können (vgl. ebd.: 154).

Im Zusammenhang des Einflusses des Faktors Alter auf das Fremdsprachenlernen werden physiologische bzw. biologische sowie soziale Aspekte diskutiert. Die soziale Dimension untersucht das Selbstbild der Lerner_innen, ihre Motivation, ihre Imitations- und Identifikationsbereitschaft sowie die Quantität und Qualität des sprachlichen Inputs. In den folgenden Unterkapiteln der vorliegenden Studie wird zunächst ein Überblick über neurowissenschaftliche Studienergebnisse gegeben. Danach werden die Themen Motivation und Erwerbsantrieb sowie Sprachlerneignung diskutiert, weil sie für die vorliegende Studie entscheidend sind. Beschlossen wird das Kapitel mit einem Einblick in die Grundsatzdebatte zwischen Empirismus und Nativismus, denn erst vor diesem Hintergrund lassen sich einige der geschilderten Diskussionen einordnen.

2.4.1 Neurophysiologische Reifungsprozesse

„In a sense, the question “What do we see when we look at a brain scan?” epitomizes the problem we are investigating. Functional neuroimages seem to provide visual diagnoses, and tell us why we are the way we are. They have become pictures of the self at the expense of public awareness that they result from complex processing of computer data, and could look totally different.“ (Ortega & Vidal 2007: 257)

Nachdem lange Zeit eine Sichtweise vorherrschte, die gehirnphysiologische Alterungsprozesse v. a. defizitär beschrieb, wurden in den letzten Jahren neue Erkenntnisse gesammelt. Diese zeigen, dass die Beschreibung neurophysiologischer Funktionsweisen als kontinuierliche Abwärtsbewegung nach Überschreiten des 25. Lebensjahres in ihrer Absolutheit nicht zutreffend ist. Dass das menschliche Gehirn ein Leben lang anpassungsfähig bleibt und in der Lage ist, neue synaptische Verbindungen hervorzubringen (Neuroplastizität), gehört dabei zu den prominentesten Entdeckungen in diesem Zusammenhang.¹⁸

Gleichzeitig ist Vorsicht geboten, was den Erkenntnisgewinn neurowissenschaftlicher Ergebnisse anbelangt, denn gerade die bildgebenden Verfahren – Funktionelle Magnetresonanztomographie (fMRT), Positronen-Emissions-Tomographie (PET) etc. – sind in den letzten Jahren aus forschungsmethodologischen und erkenntnis-

¹⁸ Das große Interesse der Öffentlichkeit an diesen Ergebnissen der Hirnforschung zeigen populärwissenschaftliche Veröffentlichungen wie „Je älter desto besser. Überraschende Erkenntnisse aus der Hirnforschung“ (Pöppel & Wagner 2010) und „Was wir sind und was wir sein könnten. Ein neurobiologischer Mutmacher“ (Hüther 2011).

theoretischen Gründen kritisiert worden.¹⁹ Ein illustres Beispiel für die dringende Notwendigkeit statistischer Korrekturverfahren bei der Auswertung von fMRT-Scans sind die Ergebnisse von Bennett, Baird, Miller & Wolford (2009, 2010). Ihre eindrückliche Studie zeigt die Reaktionen ihres Probanden auf Stimuli, die verschiedene inklusive und exklusive soziale Situationen zeigen. Das Besondere an der Untersuchung: Der Proband im fMRT-Scanner war ein Lachs und zum Zeitpunkt der Untersuchung bereits tot. Bennet et al. wollten mit ihrer Untersuchung auf die dringliche Notwendigkeit statistischer Korrekturen bei den bildgebenden Verfahren hinweisen, um solche Artefakte aufgrund von falsch positiven Signalen zu vermeiden. Darüber hinaus überprüften Bennett et al. (2010: 2) die im Jahr 2008 veröffentlichten Studien von einschlägigen Zeitschriften wie bspw. *Human Brain Mapping* oder *Cerebral Cortex*. Sie zeigten, dass 26–40 % der Studien ihre Ergebnisse ohne statistische Korrektur für multiple Vergleiche auswerteten und veröffentlichten. Die Reliabilität von fMRT-Ergebnissen wird von Bennett & Miller (2010) diskutiert, indem sie Studien zitieren, nach denen die interindividuellen Unterschiede größer sind als die intraindividuellen beim Erledigen verschiedener kognitiver Aufgaben. Es zeigte sich, dass die Ergebnisse einer Person, die zwei verschiedene Aufgaben verrichtete, sich ähnlicher waren als die Daten von zwei verschiedenen Testpersonen, die dieselbe Aufgabe lösten (vgl. ebd.: 145).

Des Weiteren stellen Bennet und Miller in ihrer Zusammenschau von 63 fMRT-Test-Retest-Artikeln fest, dass die Überlappung von ‚aktiven Regionen‘ im Durchschnitt nur knapp 30 % beträgt, obwohl die Experimente mit derselben Person im selben Scanner durchgeführt werden (vgl. ebd.: 150). Neben den methodischen Schwierigkeiten diskutieren sie eine Reihe von *Within-subject*-Gründen, die die Wiederholbarkeit der Studien erschweren. Diese sollen hier ausführlich wiedergegeben werden, weil sie einen Einblick in die Herausforderungen der neurowissenschaftlichen Forschung geben: Bereits während einer Sitzung kann es zu Unterschieden in der Aufmerksamkeit und Erregbarkeit der Testperson kommen. Die Testperson kann auch einen Wechsel der kognitiven Strategie vornehmen und von einem Moment auf den

¹⁹ In seiner Streitschrift „Neuromythologie“ (2012) vertritt der Pharmakologe Felix Hasler zahlreiche Einwände gegen die Deutungshoheit der Neurowissenschaften. Dieser unterhaltsam geschriebenen Einführung in die Debatte verdanke ich zahlreiche Literaturimpulse sowie eine Einführung in die grundsätzlichen Begrifflichkeiten der Neurowissenschaften.

anderen ein anderes Entscheidungskriterium wählen, um die gestellte Aufgabe zu lösen. Schließlich sind Lerneffekte zu erwarten, die bei der Auswertung irreführend sein können, weil sie zu veränderten Aktivierungsmustern führen. Bei einem zeitlichen Abstand zwischen Test und Retest ist zu erwarten, dass sich die physiologischen und kognitiven Veränderungen der Testperson noch gravierender auswirken (vgl. ebd.: 137).

In die Kritik geraten sind die bildgebenden Verfahren auch bezüglich der verblüffend hohen Korrelationen zwischen den Messungen individueller Differenzen der Persönlichkeit, der Emotion und sozialen Kognition und Messungen der Gehirnaktivität durch fMRT-Scans. Vul, Harris, Winkelman & Pashler (2009) und Vul & Pashler (2012) kamen zu dem Ergebnis, dass in mehr als 50 % der untersuchten Studien statistische Berechnungen zur Anwendung kamen, die die Ergebnisse verzerren und zu hohe Korrelationen errechnen – Voodoo-Korrelationen, so der ursprüngliche Titel des Aufsatzes (vgl. Vul et al. 2009: 285).

Die mittels bildgebender Verfahren erzielten Ergebnisse der sozialen Neurowissenschaften müssen kritisch hinterfragt werden und sollten den Gütekriterien empirischer Forschung standhalten. Bezüglich der Objektivität lässt sich festhalten, dass es keine verbindlichen Richtlinien für die zur Anwendung kommenden Auswertungsverfahren gibt, die Ergebnisse folglich von subjektiven Entscheidungen der Forscher_innen beeinflusst werden. Die Reliabilität erwies sich ebenfalls als problematisch. Button, Ioannidis, Mokrysz, Nosek, Flint, Robinson & Munafò (2013) zeigen, dass eine zu kleine Stichprobe, die zu einer geringen statistischen *Power* führt, nicht nur zur Folge hat, dass vermutliche Effekte nicht das Signifikanzniveau erreichen und deswegen unentdeckt bleiben. Hinzu kommt, dass die gefundenen Effekte eventuell nicht die tatsächlichen Verhältnisse widerspiegeln und sich höchstwahrscheinlich nicht reproduzieren lassen. Schließlich regen sich auch bezüglich des dritten Gütekriteriums, der Validität, Zweifel, und es stellt sich die Frage, inwiefern die Lokalisierung des gestiegenen Sauerstoffverbrauchs hochkomplexe Denkprozesse zu erklären hilft (Hasler 2012: 43; 54–55). Trotz alledem ist die Auseinandersetzung mit den Ergebnissen neurowissenschaftlicher Forschung im Zusammenhang dieses Forschungsprojekts aufschlussreich und erforderlich.

Neuronale Plastizität, die Modulation der Anzahl und Eigenschaften der Synapsen und des Netzwerkes, ist die Grundlage von Lernen und Gedächtnis (Nitsch 2007:

50). Auf molekularer Ebene entspricht Lernen Veränderungen in der Stärke synaptischer Verbindungen (synaptische Plastizität) und der Bildung neuer Neuronen (Neurogenese) in bestimmten Gehirnarealen. Auf biochemischer Ebene wurde gezeigt, dass bestimmte Neurotransmitter wie bspw. Dopamin und Serotonin, Proteine und Hormone wie Östrogen u. a. modulierende Einflüsse auf das Gedächtnis haben (vgl. Stemmer 2010).

Während man zu Zeiten Lennebergs noch davon ausging, dass die neuronale Plastizität nur in Kindheit und Jugend gegeben sei und sich auf bestimmte Hirnareale, *Hippocampus* und *Amygdala*, beschränke, ist heute unbestritten:

„Neuronale Plastizität ist als natürliche Reaktion auf Änderungen in der Umwelt lebenslang vorhanden und betrifft alle Hirngebiete, wenngleich sicherlich sekundäre und multimodale Rindenareale eine höhere Plastizitätsrate aufweisen.“ (Nitsch 2007: 51)

Das Altern hat neurophysiologische Veränderungen zur Folge: Die Gehirnmasse schrumpft aufgrund der Abnahme der Dichte der *substantia alba* und eines Rückgangs der Neuronengröße und -dichte (vgl. Stemmer 2010: 8). Gleichzeitig lassen sich kognitive Veränderungen im Zusammenhang mit Alterungsprozessen feststellen, die v. a. die Verarbeitungsgeschwindigkeit, Funktionen des Arbeitsgedächtnisses, Aspekte des Langzeitgedächtnisses und ausführende Kontrollprozesse betreffen (vgl. ebd.). Im Zusammenhang mit den Fragestellungen dieses Forschungsprojekts ist aber von besonderer Relevanz, dass trotz der neurobiologischen Alterungsprozesse die kognitive Performanz individuell sehr verschieden ist (vgl. ebd.: 8, 15). Stemmer zufolge beeinflussen die folgenden Faktoren Lernen und Erinnerungsvermögen in positiver Weise und tragen dazu bei, biologische Alterung zu kompensieren: mittlerer oder hoher Bildungsgrad, soziale, mentale und physische Aktivität, lebenslanger Bilingualismus, soziales Engagement und Vermeidung sozialer Isolation, kognitiv stimulierende Tätigkeiten (Trainingseffekte), Vermeidung von akutem und/oder chronischem Stress sowie ausreichend Schlaf (vgl. ebd.: 19–20). Ebenfalls vorteilhaft könne sich ihr zufolge das *Scaffolding* beim L2-Lernen in erwachsenem Alter auswirken. Die Nervenbahnen-,Gerüste', die beim Lernen zunächst provisorisch angelegt werden, differenzieren sich mit der Zeit und mit fortschreitendem Lernen aus. Später dienen diese ‚Gerüste‘ als eine Art von *Backup*-System, das bei besonderen Herausforderungen aktiviert wird und eventuell biologische Alterung kompensieren kann (vgl. ebd.: 15–16).

Wiederholt beschäftigten sich neurowissenschaftliche Studien mit der Frage, ob sich die Verarbeitung von Sprache(n) bei Einsprachigen, Zweisprachigen oder Mehrsprachigen unterscheidet und ob dabei das Erwerbsalter eine Rolle spielt. Richtungweisend für die Diskussion war die Studie von Kim, Relkin, Lee & Hirsch (1997), der zufolge Personen mit 2L1 und/oder Frühbilinguale (n = 6, AO = „infancy“²⁰) beim freien Erzählen die gleichen Subareale im Broca-Areal²¹ aktivieren. Dagegen prozessiert die Vergleichsgruppe (AO > 10²²) die beiden Sprachen in jeweils anderen Arealen. Dieses Ergebnis konnte von der interdisziplinären Basler Projektgruppe *Multilingual-brain* reproduziert werden (vgl. Nitsch 2007: 56). Personen mit 2L1 (n = 20; AO = 0–3 Jahre) „aktivieren überlappende Subareale im Broca-Areal, und zwar nicht nur für ihre L1 und L2, sondern auch für die spät (nach dem 9. Lebensjahr) gelernte L3“ (ebd.: 57). Späte Mehrsprachige (nach dem 9. Lebensjahr) aktivieren kleine getrennte Subareale, woraus geschlossen wird, dass sich frühe Mehrsprachigkeit auf den Erwerb weiterer Sprachen auswirkt (vgl. ebd.). Nitsch (2007: 55) fasst mehrere Studien zusammen, bei denen bildgebende Verfahren (PET und fMRT) eingesetzt wurden. Die Studien zeigen keine Unterschiede in der regionalen Aktivierung bei Mehrsprachigen mit hoher Sprachkompetenz in der L2, weder bezüglich der Produktion noch der Perzeption, dem Sprachverständnis auf Satzebene. Unterschiede finden sich nur bei starken Kompetenzunterschieden zwischen den Sprachen und wenn die Sprachen auf ganz verschiedenen Altersstufen erworben wurden.

Andere Untersuchungen kommen zu abweichenden Ergebnissen.²³ Perani, Abutalebi, Paulesu, Brambati, Scifo, Cappa & Fazio (2003) weisen per fMRT nach, dass Personen mit 2L1 (n = 11; AO = 0–3) mit den Sprachen Spanisch und Katalanisch verschiedene Hirnareale aktivieren, je nachdem ob sie Spanisch oder Katalanisch verarbeiten. Allerdings ist darauf hinzuweisen, dass die Testpersonen ihre erste

²⁰ Die Informationen über den AO der Proband_innen sind wenig spezifisch: „Six subjects (‘early’ bilinguals) were exposed to two languages during infancy, and six subjects (‘late’ bilinguals) were exposed to a second language in early adulthood.“ (Kim et al. 1997: 174).

²¹ Das Broca-Areal ist eine Region der Großhirnrinde, meist in der linken Hemisphäre, die für die Sprachproduktion zuständig ist. Benannt wurde es nach seinem Entdecker, dem französischen Chirurgen Paul Broca. Die andere Hauptregion des Sprachzentrums im Gehirn ist das Wernicke-Areal (benannt nach dem deutschen Neurologen Claus Wernicke), das für das Sprachverstehen von Bedeutung ist (vgl. Glück 2010a, 2010b).

²² „The mean age of initial exposure to the second language was 11.2 years“ (Kim et al. 1997: 174).

²³ Einen Überblick über die Studien bieten Zippel (2009: 61 f.) und Zappatore (2006: 78–79).

Sprache ab Geburt und ihre zweite Sprache ab drei Jahren erwerben (vgl. ebd.: 172). Einerseits fallen sie unter die hier verwendete Definition von 2L1, andererseits unterscheidet sich der AO jeweils. Wie erheblich dieser Unterschied das Ergebnis beeinflusst, ist schwer zu beurteilen, weil die Forschungslage nicht eindeutig ist.

Die Ergebnisse der von Stowe & Sabourin (2005) ausgewerteten sieben Studien, die sich mit der Lokalisierung der durch L1 und L2 aktivierten Gehirnareale beschäftigen, unterscheiden sich stark: Drei Studien finden gemeinsame Sprachverarbeitungsareale, die anderen vier finden sowohl gemeinsame als auch sich unterscheidende Areale (vgl. ebd.: 332). Dabei ist zu bemerken, dass – ähnlich wie schon bei der Diskussion um die CPH gezeigt – jeweils verschiedene Altersgrenzen gesetzt werden (AO zwischen drei und 18+), unterschiedliche Stimuli verwendet werden und auch die Sprachkompetenz der Testpersonen erheblich differiert (vgl. ebd.: 332). Des Weiteren weisen Stowe und Sabourin darauf hin, dass neben der Frage der Lokalisation der sprachlichen Verarbeitung von L1, L2 etc. die Frage nach der Verwendungsart dieser neuronalen Ressourcen entscheidend ist (vgl. ebd.: 336). Basierend auf der Auswertung mehrerer EKP-Studien²⁴ schlussfolgern sie: „in general the same neural resources are used in processing L2 and L1, even when the L2 is learned relatively late“ (ebd.: 348). Gelegentlich verzögerte Reaktionen der Späterwerber_innen und ein festgestellter erhöhter Energiebedarf bzw. Aktivierungsgrad bei der Verarbeitung der L2 sowie eine teilweise übermäßige Inanspruchnahme, die bei weniger kompetenten L2-Sprecher_innen gefunden wurde, führen zu „questions about the role of proficiency and of L1 on the degree to which the linguistic system used in L2 processing is similar to that used in L1“ (ebd.: 348).

Der Bericht von Andrews, Frigau, Voyvodic-Casabo, Voyvodic & Wright (2013) über das erste Jahr einer longitudinalen Studie unterstützt den Ansatz, dass die Sprachkompetenz Aktivierungsareal und Aktivierungsgrad beeinflusst. Er bietet neben der Auswertung von fMRT-Scans auch Informationen zu Sprachkontakt und Dauer, Sprachkompetenz, zuvor gelernten Sprachen und dem Alter bei Erwerbsbeginn. Damit versuchen die Autor_innen der Studie teilweise die Forderungen De Bots (2008) einzulösen. Dieser diskutiert verschiedene Kontextualisierungen, deren adä-

²⁴ Ereigniskorrelierte Potentiale (EKP): durch Wahrnehmung oder kognitive Prozesse beeinflusste Wellenformen, die mit Elektroenzephalogrammen (EEG) gemessen werden können.

quate Operationalisierung nötig ist, um die Ergebnisse neurologischer Untersuchungen fruchtbar zu machen für das Verstehen multilingualer Prozesse: Alter des Lernbeginns, Sprachkompetenz, Intensität des Sprachkontakts, Sprachlernmotivation, Sprachlerneignung, Einstellungen der Lerner_innen gegenüber der L1, L2, L3 etc., andere Sprachen, die vor der Untersuchung gelernt wurden, sowie der Bildungsgrad der Testpersonen (De Bot 2008: 117–119). Andrews et al. (2013) untersuchen 5 Testpersonen mit der L1 Englisch in den USA, die intensiv Russisch lernen (AO = 19–20) und während dieser Zeit auch zehn Wochen Intensivunterricht und Immersion in St. Petersburg erfahren. Innerhalb eines Jahres, in dem die Testpersonen ohne Vorkenntnisse das Niveau B1 (GER) erreichen, werden drei fMRT Scans durchgeführt (Monat 1, Monat 8, Monat 12, weitere Scans sind geplant). Sowohl die Testergebnisse des russischen Sprachtests (Sprachkompetenz) als auch die Zeit beeinflussen signifikant die Aktivierungsgrade und Areale (vgl. Andrews 2013: 868). Die Studie ist insofern interessant, als sie die erste ist, die ein longitudinales Forschungsdesign aufweist und neben den fMRT-Messungen Daten zur Entwicklung der Sprachkompetenz in die Auswertung miteinbezieht. Trotzdem bleibt das Ergebnis erwartbar und in seiner unspezifischen Formulierung sehr allgemein (vgl. ebd.: 868–869).²⁵

Beenden möchte ich dieses Kapitel mit der Vorstellung der Zwei-Faktoren-Hypothese von Pulvermüller & Schumann (1994). Dieser Aufsatz ist zwar angesichts des Anwachsens der neurobiologischen Erkenntnisse in den letzten 20 Jahren teilweise veraltet,²⁶ die Zwei-Faktoren-Hypothese ist aber nach wie vor ein relevanter Erklärungsansatz für die gefundenen Sprachkompetenzunterschiede zwischen frühen und späten Lerner_innen. Pulvermüller und Schumann zufolge müssen zwei Bedingungen erfüllt sein, um in vollem Umfang eine Sprache zu erwerben:

²⁵ Als Beispiel sei dieser Auszug aus der Zusammenfassung der Ergebnisse zitiert. Meines Erachtens nach fehlen die Angaben der p-Werte, die konkrete Benennung der Areale etc.: „Secondary results of the analysis include the following results: 1. The time effect is significant for Russian-rest; average activation levels change across the different sets of measurements. 2. The time effect is not significant for English-rest; average activation levels do not change across the different sets. 3. There is a significant hemisphere effect. 4. The Middle Occipital Gyrus, used as an internal statistical standard, shows a lack of effect as expected. 5. Different regions show variation in activation patterns“ (vgl. Andrews et al. 2013: 869).

²⁶ Eine genaue Auseinandersetzung bzw. Aktualisierung des neurobiologischen Kenntnisstands bleibt an dieser Stelle aus, da sie den Expert_innen auf diesem Gebiet überlassen werden soll.

„The first condition is that the learner is motivated to learn the language. The second condition is that the learner is equipped with the ability to acquire grammatical knowledge.“ (Pulvermüller & Schumann 1994: 687–688)

Eine hohe Motivation und positive Einstellungen gegenüber dem Lernen einer L2 beeinflussen die Dopamin-Ausschüttung, die wiederum gebraucht wird, um neue Verbindungen im *Amygdala*-Mittelhirnareal zu schaffen (vgl. ebd.: 717). Die Abnahme der *ability to acquire grammatical knowledge* erklären die Autoren durch die eingeschränkte neuronale Plastizität, die durch die abgeschlossene Myelination verursacht wird, und daraus folgende Schwierigkeiten beim Erwerb von grammatikalischen Strukturen. Pulvermüller und Schumann machen deutlich, dass sie die CPH unterstützen (vgl. ebd.: 722). Sie erklären mithilfe der Zwei-Faktoren-Hypothese den Abfall des Grammatikererwerbs bei steigendem AO und die größere Varianz der grammatikalischen Kenntnisse bei Späterwerber_innen (vgl. ebd.: 721).

Nach der Darstellung der Erkenntnisse zur Gehirnreifung aus neurophysiologischer Perspektive folgt im nächsten Unterkapitel ein Überblick über die Motivation in Bezug auf die Fremdspracherwerbsforschung. Dieser wird durch die Beschreibung des Konzepts Erwerbsantrieb ergänzt, da damit altersspezifische Unterschiede erfasst werden können.

2.4.2 Motivation und Erwerbsantrieb

Seit den späten 50er Jahren wird die Frage, welchen Einfluss die Motivation auf den Erfolg beim Zweit- und Fremdsprachenlernen hat, empirisch untersucht (vgl. Rost-Roth 2010: 877). Dörnyei (2005) zufolge lassen sich dabei drei Perioden voneinander unterscheiden:

Die sozialpsychologische Periode (1959–1990), die v. a. durch die Theorie von Gardner & Lambert (1959) geprägt ist. Diese war für die Fremdsprachenforschung folgenreich, insbesondere durch ihre Unterscheidung in eine integrative und eine instrumentelle Motivation. Integrativ Motivierte lernen eine neue Sprache, um mehr über die Sprecher_innen und ihre kulturellen Hintergründe zu erfahren, um Menschen treffen und kennenlernen zu können. Instrumentell Motivierte hingegen erlernen eine Sprache eher aus utilitaristischen Motiven, etwa aus beruflichen Gründen (vgl. ebd.: 267). Gardner und Lambert wiesen nach, dass integrative Motivation den

Sprachlernerfolg stärker befördert als die instrumentelle Motivation (vgl. ebd.: 271; vgl. Rost-Roth 2010: 878).²⁷

In der *kognitiven Periode* (1990er Jahre) gehen die Motivationspsycholog_innen davon aus, dass das Denken einer Person über die eigenen Fähigkeiten, Möglichkeiten, Einschränkungen und Zielsetzungen die Motivation entscheidend bestimmt. War die Ergründung der Motivation in der Second Language Acquisition Research (SLAR) bisher eher von der Makroperspektive geprägt, in der es um eine ganze Sprachgemeinschaft ging, wurde nunmehr auf die einzelnen Lerner_innen und die aktuellen Lernsituationen wie Klassenzimmer, Lehrende und Curricula geschaut (vgl. Dörnyei 2005: 74–82). Besonders einflussreich ist in diesem Zusammenhang die *Selbstbestimmungstheorie der Motivation* (Deci & Ryan 1985, 2002), die auf verschiedene Typen extrinsischer und intrinsischer Motivation sowie auf deren Zusammenspiel bezüglich der Autonomie der Lerner_innen fokussiert (vgl. Riemer 2006: 38–39). Intrinsische Motivation kennzeichnet Lernen, wenn es aus dem eigenen Interesse am Lerngegenstand heraus erfolgt oder an sich lustbetont ist. Extrinsische Motivation bezieht sich auf ein Lernen, das v. a. aufgrund von äußeren Belohnungen oder Bestrafungen erfolgt wie bspw. guten Noten, eine Prüfung bestehen oder nicht bestehen, die nichts mit dem Lerngegenstand an sich zu tun haben (vgl. Krause & Stark 2010).

In der *Attributionstheorie* (Weiner 1992) geht es um den Einfluss der kausalen Attributionen zwischen vergangenen Erfahrungen und zukünftig erstrebten Zielen. So spielt es bspw. eine Rolle, ob die Lerner_innen die Ursache für ihre Erfolge oder Misserfolge eher in sich selbst oder in ihrer Umwelt suchen (vgl. Riemer 2004: 39). Schließlich ist in diesem Zusammenhang auch die *Aufgabenmotivation* zu nennen (*Task Motivation*), die für die SLAR gerade deswegen attraktiv ist, weil der komplexe

²⁷ Dörnyei (2005) weist darauf hin, dass in der *Second Language Acquisition Research* eine Verschiebung bzw. Fehlinterpretation stattgefunden hat, da Gardner und Lambert das Verhältnis von integrativer und instrumenteller Motivation ursprünglich nicht als Dichotomie formuliert hatten, sondern die instrumentelle Motivation lediglich ein Item der *Attitude/Motivation Test Battery* war (vgl. Dörnyei 2005: 70). Später wurde die instrumentelle Motivation von Gardner mit in die theoretische Fundierung aufgenommen (Gardner 2001: 7). In Gardner & Lambert (1959: 267) werden integrative und instrumentelle Motivation allerdings auch schon als die beiden Alternativen des Orientierungs-Indexes bezeichnet.

Prozess des Sprachenlernens in kleinere, abgrenzbare Einheiten aufgeteilt werden kann.

Die *prozessorientierte Periode* begann in den 1990er Jahren und nahm die bisher vernachlässigten Aspekte *temporale Variation* und *dynamischer Charakter der Motivation* ins Visier. In diesem Zusammenhang wird die Gardner'sche *Integrativeness* (2001) neu formuliert und Dörnyeis *L2 Motivational Self System* verortet (vgl. Dörnyei 2005: 83–117; Dörnyei 2011). Gingen frühere Untersuchungen davon aus, dass höhere Motivation größeren Erfolg nach sich zieht, wird das Verhältnis von Erfolg und Motivation nun als Wechselbeziehung gesehen: Eine hohe Motivation kann den Lernerfolg hervorrufen und umgekehrt können Lernerfolge die Motivation erhöhen bzw. aufrechterhalten (vgl. Ellis 1994: 515; Larsen-Freeman & Long 1991: 183).

Nach diesem Überblick werden nun im Folgenden das Modell des alterstypischen Antriebs von Klein (1996, 2001), dessen Weiterentwicklung bei Pagonis (2009b) sowie Dörnyeis *L2 Motivational Self System* (Dörnyei 2009a) ausführlicher dargestellt und diskutiert, weil sie bezüglich der Frage nach dem Einfluss des Alters und anderer individueller Variablen produktiv sind. Klein stellt drei Komponenten fest, die den Erstspracherwerb wie auch den gesteuerten und ungesteuerten L2-Erwerb bestimmen: der *Zugang* (Menge und Qualität des Inputs), der *Sprachverarbeiter* (biologische und kognitive Ausstattung, Weltwissen, Wissen über L1, L2 etc., Sprachlernereignung) und der *Antrieb* (alterstypische Bereitschaft).²⁸ Klein zufolge sind die beiden wesentlichsten Gründe, eine Sprache zu lernen, *soziale Integration* und *kommunikative Bedürfnisse*. Soziale Integration bezieht sich v. a. auf den Erstspracherwerb und den kindlichen L2-Erwerb:

„Ein Kind muss eine bestimmte soziale Identität erwerben, und dazu muss es die Sprache seiner sozialen Umgebung exakt reproduzieren. [...] Ein Kind, das das /s/ in allen Positionen gleich ausspräche, konsequent bei allen Objekten den Artikel „das“ verwendete und „laufte“ und „schwimmte“ statt „lief“ und „schwamm“ sagen würde, wäre im Grunde viel logischer und vernünftiger als seine soziale Umgebung – aber es wäre in dieser Umwelt ein Außenseiter.“ (Klein 2001: 609–610)

Soziale Integration ist wesentlich auf die Erfüllung der kommunikativen Bedürfnisse angewiesen. Die kommunikativen Bedürfnisse können jedoch auch ohne integrative Absichten verwirklicht werden: „umgekehrt will nicht jeder, der spanische Handelskorrespondenz lernt, auch ein Spanier werden“ (ebd.). Dies wirkt sich auf den End-

²⁸ Vgl. die Zusammenfassung des Modells in Czinglar (2014: 19 f.).

stand des L2-Erwerbs aus, wobei die Skala von der Ununterscheidbarkeit der Lerner_innen vom zielsprachlichen Umfeld bis zum Erwerb der Basisvarietät reicht, die den einfachsten kommunikativen Bedürfnissen angepasst ist (vgl. Klein 1996: 95f.). Auch wenn diese Begriffe an Gardner und Lamberts integrative und instrumentelle Motivation erinnern, ist es an dieser Stelle wichtig, zwischen Motivation und Antrieb zu unterscheiden. Während Motivationen mehr oder weniger bewusst und beeinflussbar sind, ist der alterstypische Erwerbsantrieb unbewusst und von entwicklungspsychologischen Faktoren determiniert. Deswegen bezeichnet Klein den Antrieb als Bereitschaft (*propensity*), eine Sprache zu lernen, in Abgrenzung zur Motivation (vgl. ebd.: 5).

Pagonis (2009b) greift in produktiver Weise Kleins Modell und Gardners und Lamberts Unterscheidung in instrumentelle vs. integrative Motivation auf. Er untersucht in einer longitudinalen Fallstudie die Unterschiede in den Erwerbsverläufen zweier russischer Lernerinnen. Die beiden Schwestern Nastja (AO = 8) und Dascha (AO = 14) werden über eineinhalb Jahre hinweg wöchentlich eine Stunde lang interviewt, wobei meistens freie, spontane Gespräche mit erwachsenen Interviewpartnern oder kindlichen/jugendlichen Freunden aufgezeichnet werden (vgl. auch Kap. 2.2 *Studien zum Faktor Alter*). Der Beginn der Interviews ist drei Wochen nach Nastjas und Daschas Ankunft in Deutschland. Beide sprechen Russisch als Erstsprache und hatten vor ihrer Ankunft in Deutschland keinen nennenswerten Kontakt mit der deutschen Sprache. Pagonis untersucht in seiner Erwerbsverlaufsstudie v. a. folgende Bereiche: Verbalflexion (Subjekt-Verb-Kongruenz, Tempora), Nominalflexion (Deklination des attributiven Adjektivs, ‚schwache‘ Deklination), Verbstellung, Abtönungspartikel und das Lexikon (vgl. Pagonis 2009b: 121–128). Nach ausführlicher Diskussion der CPH stellt Pagonis das Modell des altersabhängigen Erwerbsantriebs vor, worunter die Verknüpfung aus altersbedingter überindividueller Identifikationsbereitschaft und individueller Erwerbsmotivation versteht. Ihm zufolge verändert das Alter den Sprachverarbeiter, den Zugang und den Antrieb und somit den Endzustand des Spracherwerbs (vgl. das Schaubild in Pagonis 2009b: 64).

Der Vorteil von Pagonis' Modell gegenüber der CPH ist die Offenheit für individuelle Variation. Es lässt sowohl den überdurchschnittlichen Lernerfolg erwachsener Lerner_innen zu als auch unterdurchschnittliche Erfolge bei Kindern. Tab. 2 stellt das Modell im Überblick dar.

Tab. 2: Modell des alterstypischen Erwerbsantriebs (Tab. aus Pagonis 2009a).

	Erwerbs- antrieb	=	altersbedingte überindividuelle Identifikationsbereitschaft	+	individuelle Erwerbsmotivation
Kinder			hoch		hoch oder niedrig
Erwachsene			niedrig		hoch oder niedrig

Pagonis zeigt, dass die präpubertäre, frühe Lernerin Nastja mit Ausnahme des Lexikons alle untersuchten Bereiche zielsprachlicher als die postpubertäre, späte Lernerin Dascha erwirbt bzw. gebraucht. Ein genauerer Blick auf die Analyseergebnisse zeigt, dass sich die Lernerfolge von Nastja und Dascha etwa im Bereich der Adjektivdeklinaton besonders stark unterscheiden, der sich durch strukturelle Komplexität und hohen Lernaufwand auszeichnet. Während Nastja die Form-Funktions-Konvention nach 16 Kontaktmonaten zielsprachlich meistert, zeigt ihre ältere Schwester keine Annäherung an die zielsprachliche Norm mit Ausnahme der pauschalen Verwendung der e-Endung und dem teilweisen Gebrauch der Markierung -er (vgl. Pagonis 2009b: 256–257). Unterschiedliche altersbedingte Erwerbsantriebe erklären Pagonis zufolge diese Befunde:

„Warum sollte der hohe Lernaufwand zum zielsprachlichen Erwerb der Adjektivdeklinaton geleistet werden, so könnte der unbewusste Lernantrieb der postpubertären Lernerin möglicherweise paraphrasiert werden, wenn die Lerneräußerungen auch ohne formale Differenzierung der attributiv verwendeten Adjektive nach den Kategorien „Kasus“, „Genus“, „Numerus“ und „Artikelgebrauch“ die Befriedigung der kommunikativen Bedürfnisse gewährleistete und dabei der ungefährdete Transport konkreter kommunikativer Inhalte als altersbedingtes Erwerbsziel zugrunde liegt, nicht aber die u. a. sprachliche Ununterscheidbarkeit von den (zumeist muttersprachlichen) Mitgliedern der sprachlichen Umgebung?“ (Pagonis 2009b: 259)

Für Pagonis sind die sozial- und entwicklungspsychologischen Faktoren entscheidend bei der Erklärung tendenzieller Vorteile jüngerer Lerner_innen. Nastjas Integrationsbereitschaft steht dem fehlenden Integrationsbedürfnis der adoleszenten Lernerin Dascha gegenüber, deren Lernantrieb zwischen dem kindlichen identitätsstiftenden und dem erwachsenen identitätswahrenden Lernantrieb steht. Trotz der festgestellten Unterschiede bzgl. der Zielsprachlichkeit des erreichten Sprachniveaus der beiden Lernerinnen stellt Pagonis fest, dass Daschas formreduzierter Spracherwerb eigentlich genau so erfolgreich ist, wenn „das Maß an Verwirklichung der zugrundeliegenden „Erwerbsabsicht“ angenommen wird“ (Pagonis 2009b: 341).

Mit Gardners und Lamberts integrativem Motivationsmodell und dem von Klein und später Pagonis daraus entwickelten altersspezifischen Erwerbsantrieb lassen sich die gefundenen Unterschiede in Bezug auf den erreichten Sprachstand erklären. Geht es jedoch um Fremdsprachen-Lernen, wird deutlich, dass das Modell der *Integrativeness* nicht in dem Maße fundamental für die Motivation ist wie angenommen: Es ist auf bestimmte sozio-kulturelle Kontexte beschränkt, bspw. das multikulturelle und mehrsprachige Milieu Kanadas, das Gardner und Lambert beforchten, oder die Situation bezogen auf die Fragestellungen von Migration und Spracherwerb (Noels, Pelletier, Clément & Vallerand 2000: 60).

Dörnyei spricht sich vor diesem Hintergrund dafür aus, Gardners Theorie der *Integrativeness* zu reformulieren, weil es in vielen Kontexten des Fremdsprachlernens nicht um direkte Integration geht (vgl. Dörnyei 2005: 98). Auf der Grundlage einer groß angelegten Studie in Ungarn, die 1993, 1999 und 2004 Daten von insgesamt 13.000 Fremdsprachenlerner_innen erhob, die Englisch, Deutsch, Französisch, Italienisch oder Russisch lernten (vgl. Csizér & Dörnyei 2005a;b), ermittelt Dörnyei ein Strukturgleichungsmodell, das bis auf geringfügige Unterschiede bei allen Sprachen und in allen Erhebungsjahren zu demselben Ergebnis führt (vgl. die Abb. in Dörnyei 2009a: 27 für eine schematische Darstellung). Das wichtigste Ergebnis ist, dass *Integrativeness* bei der L2-Motivation die Schlüsselrolle spiele, wobei die *Attitudes toward L2 speakers/community* und *Instrumentality* der *Integrativeness* unmittelbar vorausgehen. Dörnyei entwickelt daraus das *L2 Motivational Self System*, indem er *Integrativeness* aus der *Self*-Perspektive als L2-spezifische Facette des idealen Selbst betrachtet:

„[...] if our ideal self is associated with mastery of an L2, that is, if the person that we would like to become is proficient in the L2, we can be described in Gardner's (1985) terminology as having integrative disposition. Thus, the central theme of the emerging new theory was the equation of the motivational dimension that has traditionally been interpreted as 'integrativeness / integrative motivation' with the *Ideal L2 Self*.“ (ebd.: 27)

Die *Self*-Perspektive erklärt Dörnyei zufolge auch die beiden eigentlich gegensätzlichen Vorstufen der *Integrativeness*: *Attitudes toward L2 speakers/community* und *Instrumentality*. Da die L2-Sprecher_innen die Vergleichsgröße für das ideale L2-Selbst sind, liegt es nahe zu vermuten, dass mit zunehmend positiver Einschätzung der L2-Sprecher_innen das idealisierte L2-Selbst an Attraktivität gewinnt. Somit sind

die *Attitudes toward L2 speakers/community* mit dem *Ideal L2 Self Image* verbunden. Bezüglich der *Instrumentality* stellt Dörnyei fest, dass zu unserem idealisierten Selbstbild auch der berufliche Erfolg gehört. Deswegen sind instrumentelle Motive, die förderlich für die Karriere sind, mit dem *Ideal L2 Self* verbunden (vgl. Dörnyei 2009a: 28).

Aus diesen Ergebnissen der empirischen Motivationsforschung und in enger Anknüpfung an die Modelle von Noels (2003) und Ushioda (2001) entwickelt Dörnyei das *L2 Motivational Self System*, mit dem sich gemäß der prozessorientierten Motivationstheorie die dynamischen und zeitlichen Aspekte der Motivation besser abbilden lassen. Es besteht aus drei Komponenten (vgl. Dörnyei 2005: 105–108; 2011: 29):

1. Das *Ideal L2 Self*: Das ideale L2-Selbst spricht die zu lernende Sprache und ist eine starke motivationale Kraft. Es beinhaltet Noels und Ushiodas integrative Aspekte, also persönliche Ziele, angestrebte Sprachkompetenz, akademisches Interesse usw.
2. Das *Ought-to L2 Self*: Dieses Selbst bezieht sich auf die Aspekte, von denen die Lerner_innen annehmen, dass sie sie besitzen müssen. Gemeint sind die extrinsischen Anteile der instrumentellen Motive.
3. Die *L2 Learning Experience* umfasst die situationspezifischen Motive, die in Bezug zum unmittelbaren Lernumfeld stehen.

Inzwischen liegen einige Studien zu Dörnyeis Modell vor, die in Japan, China, Ungarn und dem Iran in der Sekundarstufe, der Universität und der Erwachsenenbildung das *L2 Motivational Self System* untersucht haben (vgl. die Artikel in Dörnyei & Ushioda 2009).

Bezüglich dieses Modells ist zu bedenken, dass es eine Herausforderung ist, die Diversität und Dynamik der Selbstkonzepte mit quantitativen Mitteln zu erfassen und zu beschreiben. MacIntyre, Mackinnon & Clément (2009), die auf der Grundlage von qualitativen Erhebungen eine Skala entwickelten, benennen die Schwierigkeiten dieser Operationalisierung: Selbstkonzepte werden in verschiedenen Bereichen (akademisch, sozial, körperlich) entwickelt und müssen nicht notwendigerweise miteinander zusammenhängen. Bezogen auf die SLAR ist von Bedeutung, dass sich die

Selbstkonzepte beim Sprachenlernen verändern können, was sich wiederum auf die Motivation auswirkt. Des Weiteren sind interkulturelle Unterschiede bezüglich der allgemeinen Selbstkonzepte und der *L2 Self*-Konzepte im Besonderen zu berücksichtigen sowie die Frage nach dem Bezugszeitpunkt der Skala (vgl. ebd.: 194–196): Erfragt sie ein potentielles zukünftiges Selbst und/oder ein gegenwärtiges Selbst? Chen, Warden & Chang (2005) stellen auf der Basis einer Untersuchung von 567 Englischlerner_innen in Taiwan fest, dass die im Westen entwickelten Motivationsmodelle kaum auf die ostasiatischen L2-Lerner_innen übertragbar sind. Die starke Motivation der Lerner_innen in Taiwan, die Erwartungen der Gesellschaft, ihrer Eltern und des Bildungssystems zu erfüllen – wobei Erfüllung v. a. gute Noten bedeutet –, unterscheidet sich ihnen zufolge von westlichen Konzepten der extrinsischen Motivation. Dieses von ihnen als ‚chinesischer Imperativ‘ bezeichnete Konstrukt könnte spezifisch für die chinesische, konfuzianistisch geprägte Kultur sein (vgl. Chen et al. 623). Ushioda (2009) vertritt statt einer empirischen eine ontologische Position. Motivation entsteht für sie aus den Beziehungen zwischen realen Personen (*person-in-context relational view*), die sich nicht auf lineare Ursache-Wirkungs-Beziehungen reduzieren lassen. Vielmehr sind die motivationalen Reaktionen einer Person auf bestimmte Ereignisse und Lebenserfahrungen komplex und idiosynkratisch (vgl. Ushioda 2009: 219).

In diesem Unterkapitel wurde ein Überblick über verschiedene Motivationstheorien und den Erwerbsantrieb gegeben sowie auf einige der einschlägigen Studien und ihre zugrundeliegenden Instrumente Bezug genommen. Im nächsten Unterkapitel der vorliegenden Studie wird das Konzept der Sprachlerneignung präsentiert und aufgezeigt, welche Veränderungen es in den letzten 50 Jahren erfahren hat.

2.4.3 Sprachlerneignung

Neben der Motivation ist die Sprachlerneignung der wichtigste Faktor, der im Zusammenhang der Fragen nach dem Einfluss des Alters und anderer individueller Faktoren auf den Fremdsprachenlernerfolg diskutiert wird (vgl. Ellis 2004: 531). Unter (Fremd-)Sprachlerneignung (*foreign language aptitude*), wird die Fähigkeit verstanden, Sprachen schnell zu lernen und/oder ein höheres Niveau der Sprachkompetenz zu erreichen (vgl. Robinson 2013: 1). Gemeinhin als ‚Sprachtalent‘ oder ‚Sprachbegabung‘ verstanden, kann der Begriff unter Umständen ein diskriminieren-

des Potential entfalten und Assoziationen wie ‚undemokratisch‘ und ‚elitär‘ erwecken (vgl. Schlak 2008: 5). In der einschlägigen Literatur findet sich sowohl die Auffassung, dass sich Sprachlerneignung potentiell trainieren lässt (vgl. ebd.: 4), als auch die Ansicht, dass sie ein „largely innate, relatively fixed talent for language learning“ (Abrahamsson & Hyltenstam 2008: 485) ist. Wird Sprachlerneignung als Fähigkeit verstanden, eine Sprache schnell zu lernen, wird aus einem ausschließenden Begriff ein inkludierender: „Jeder Lernende kann eine Sprache lernen, einige benötigen dafür nur etwas weniger bzw. etwas mehr Zeit“ (Schlak 2008: 5). Kein anderer Faktor korreliert höher mit dem Sprachlernerfolg: Positive Korrelationen zwischen .4 und .65 wurden zumindest im Kontext gesteuerten Fremdsprachenlernens an US-Schulen in den 60er Jahren gemessen (vgl. Carroll 1962). Gleichzeitig verdeutlichen die Ergebnisse jedoch auch, dass die Sprachlerneignung nicht der einzige Faktor ist, von dem der Sprachlernerfolg abhängt: „Es handelt sich also um einen wichtigen Faktor, aber um keinen Faktor, der nicht durch Motivation, guten Unterricht, hohe Lernintensität und -dauer usw. auszugleichen wäre“ (Schlak 2010: 258).

Einigkeit besteht in der wissenschaftlichen Debatte darüber, dass Sprachlerneignung ein Konstrukt ist, das sich aus mehreren Komponenten zusammensetzt. Der von Carroll und Sapon (1959) in den 50er Jahren entwickelte *Modern Language Aptitude Test* (MLAT) galt lange Zeit als der Standardtest. Er untergliedert die Sprachlerneignung in vier verschiedene kognitive Fähigkeiten, die mit den fünf Subtests des MLAT gemessen werden. Bei der *Phonetic Coding Ability* geht es darum, Laute voneinander unterscheiden und sie sich in Verbindung mit ihrer schriftlichen Symbolisierung merken zu können. *Grammatical Sensitivity* ist die Fähigkeit, die grammatischen Funktionen von Wörtern innerhalb eines Satzes identifizieren zu können. *Inductive Language Learning Ability* ist die Fähigkeit, syntaktische und morphologische Muster in einem Korpus sprachlichen Materials erkennen und damit neue Sätze bilden zu können (vgl. induktives Lernen). Unter *Rote Learning Ability* ist die Fähigkeit zum Auswendiglernen bzw. des zügigen Etablierens von Verbindungen zwischen Laut und Bedeutung zu verstehen und/oder die Fähigkeit, zwischen dem Wortschatz der L1 und der L2 assoziative Verbindungen herstellen zu können (vgl. Carroll 1981: 105).

Neben dem MLAT ist die *Pimsleur Language Aptitude Battery* (PLAB) von Pimsleur (1966) weit verbreitet. Dieser Test wurde entwickelt, um den Fremdsprachenlerner-

folg von Schüler_innen in der Sekundarstufe vorhersagen zu können. Er ähnelt dem MLAT sehr, legt allerdings weniger Gewicht auf das Gedächtnis, sondern betont stärker auditive Elemente (vgl. Dörnyei & Skehan 2003: 594).²⁹

MLAT und PLAB weisen einige methodische Schwächen auf, bspw. das Messen sprachlichen Wissens der L1 oder die teilweise unklare Zuordnung der Subtests zu den Fähigkeiten, die der Spracherlernung zugeordnet werden (vgl. Aguado 2012: 56–57). Aus heutiger Perspektive wird an Carrolls Konzipierung der Spracherlernung und den verschiedenen Tests v. a. kritisiert, dass sie dem dynamischen Prozess des Fremdsprachenlernens nur wenig gerecht werden. Das Erlernen einer Fremdsprache erfordere spezifische Fähigkeiten in Abhängigkeit von dem Niveau, auf dem gelernt werde (vgl. Robinson 2013: 2). Dörnyei & Skehan (2003) referieren die Ergebnisse einer Studie, die nachweist, dass sich die Spracherlernung mit dem Alter verändert: Während Kinder v. a. in Bezug auf ihre Gedächtnisleistung gut abschneiden, erreichen ältere Lerner_innen mit ihren analytischen Fähigkeiten gute Ergebnisse (vgl. ebd.: 596). Ein anderer Kritikpunkt betrifft den Umstand, dass Fremdsprachenlernen in vielen verschiedenen Situationen und Unterrichtskontexten geschieht, wobei v. a. die Unterscheidung von explizit, implizit und inzidentell von Bedeutung ist. Robinson zeigt, dass der MLAT und andere Tests lediglich das explizite, intentionale Lernen messen (vgl. Robinson 2013: 3). Andererseits liegen einzelne Studien vor, die belegen, dass der MLAT auch in einem kommunikativ orientierten Fremdsprachenunterricht sowie beim ungesteuerten Zweitspracherwerb über Prognosefähigkeit verfügt (vgl. Schlak 2008: 15). Die von Meara (2005) entwickelten LLAMA *Language Aptitude Tests* basieren auf dem MLAT und messen mithilfe von vier Subtests die Fähigkeiten, Vokabeln zu lernen, Laute und Muster in gesprochener Sprache zu erkennen, Laute und Symbole einander zuzuordnen, sowie das Vermögen, grammatische Regeln induktiv abzuleiten.³⁰

Der *Cognitive Ability for Novelty in Acquisition of Language – Foreign-Test* (CANAL-F) wurde von Grigorenko, Sternberg & Ehrman (2000) als Alternative zum MLAT entwickelt und basiert auf der Analyse der Spracherwerbsprozesse. Dabei wird die Fähigkeit, mit Neuheit und Ambiguität umgehen zu können, als eine der zentralen

²⁹ Vgl. Dörnyei (2005: 37 f.) und Ellis (2008: 654) für einen detaillierteren Vergleich des MLATs und PLABs.

³⁰ Eine ausführlichere Darstellung des LLAMA findet sich in Kap. 3.4.2.6 *LLAMA-Tests*.

Kompetenzen für den Fremdsprachenerwerb betrachtet. Außerdem möchte der Test die damit verbundenen dynamischen Prozesse (enkodieren, speichern und wieder abrufen) erfassen und wiederholt deswegen vier der fünf Sektionen:

„1. Learning meanings of neologisms from context (immediate and delayed recall) 2. Understanding the meaning of passages (immediate and delayed-recall) 3. Continuous paired-associate learning (immediate and delayed recall) 4. Sentential inference (immediate and delayed recall) 5. Learning language rules. This section was presented last, and so it included only immediate recall questions.“ (vgl. ebd.: 394)

Ellis (2008) zufolge unterscheiden sich die Resultate des CANAL-F wenig von denen des MLATs. Allerdings ermöglicht der CANAL-F eine engere Verbindung zwischen einzelnen Fähigkeiten und bestimmten psycholinguistischen Prozessen (vgl. ebd.: 655–656).

Die *High-Level Language Aptitude Battery* (Hi-LAB) wird an der Universität Maryland entwickelt mit dem Ziel, diejenigen L2-Lerner_innen zu finden, die ein besonders hohes Niveau an Sprachkompetenz erreichen können (vgl. Doughty, Campbell, Mislevy, Bunting, Bowles & Koeth 2010). Der komplexe Test integriert neueste Erkenntnisse aus den Kognitionswissenschaften und wird derzeit bezüglich der Testgütekriterien überprüft (vgl. Linck, Hughes, Campbell, Silbert, Tare, Jackson, Smith, Bunting, Doughty 2013). Er umfasst in der vorliegenden Version 34 Subtests, die zur Bemessung der folgenden Komponenten entwickelt wurden (vgl. Doughty et al. 2010: 17)³¹:

- *Working Memory: Short Term Memory Capacity* (7), *Executive Capacity and Control* (13), *Rote Memory* (2)
- *Perceptual Acuity: Auditory Perceptual Acuity* (6)
- *Processing Speed: Discounting* (1), *Decision Speed* (1)
- *Tolerance of Ambiguity: Tolerance of Ambiguity* (3)
- *Primability: Priming* (2)
- *Induction: Explicit Induction* (1), *Implicit Induction* (1)
- *Pragmatic Sensitivity* (wird gerade entwickelt)
- *Fluency* (wird gerade entwickelt)

³¹ Die Zahl in Klammern gibt jeweils die Zahl der Subtests an.

Im Zusammenhang mit dem Forschungsansatz der *Aptitude-Treatment Interaction* (ATI) modellierte Skehan (1998) verschiedene Fremdsprachenlerneignungsprofile.³² Diese lassen sich folgendermaßen zusammenfassen: Lerner_innen verfügen in der Regel über ein gutes Gedächtnis oder über gute Analysefähigkeiten. Seltener ist beides stark ausgeprägt oder beide Fähigkeiten unterdurchschnittlich entwickelt. Skehan beobachtet differenzierte Wechselwirkungen zwischen den jeweiligen Lerneignungsprofilen und Unterricht. Je nach Profil reagieren Lerner_innen unterschiedlich auf den angebotenen Unterricht:

„The assumption of such research would be that there is no single best method but that it is the combination of specific method type with specific aptitude profile which creates optimal learning conditions.“ (ebd.: 199)

Wird entsprechend des ermittelten Sprachlerneignungsprofils unterrichtet – bspw. analytisch ausgerichteter Unterricht für Lerner_innen mit großer Analysefähigkeit – fällt der Lernerfolg größer aus und die Zufriedenheit der Lerner_innen mit dem Unterricht steigt ebenfalls an.

Robinson (2007) befindet die Differenzierung in gedächtnisorientierte und analytisch orientierte Lerner_innen als nicht ausreichend. In der von ihm vertretenen *Aptitude Complex Hypothesis* bezieht er auch die Lernbedingungen mit ein:

„Consequently there is a need to develop more contextually sensitive measures of aptitude if progress is to be made in linking IDs in cognitive abilities to the daily conditions of classroom learning and practice in a useful way.“ (ebd.: 270)

Sprachlerneignung – bzw. Sprachlerneignungen (*aptitudes*), wie in einigen Texten neueren Datums zu lesen ist, um den verschiedenen Teilkomponenten gerecht zu werden – wird vermehrt im Zusammenhang mit Intelligenztests und Arbeitsgedächtnis diskutiert bzw. miteinander korreliert. Verschiedene Studien zeigen sowohl Gemeinsamkeiten zwischen Sprachlerneignung und Intelligenz als auch deutliche Unterschiede. Im Vergleich zur Intelligenz ist Sprachlerneignung der erfolgreichere Prädiktor für L2-Lernerfolg (vgl. Schlak 2008: 6–7). In den letzten Jahren gab es zahlreiche Arbeiten, die den Zusammenhang des Arbeitsgedächtnisses mit der L2-Performanz und der Entwicklung zeigen. Diese Studien legen es nahe, das Arbeitsgedächtnis als einen der zentralen Bestandteile der Sprachlerneignung zu begreifen (vgl. den Überblick in Aguado 2012: 59). DeKeyser und Koeth (2011) kommen nach

³² Vgl. die Zusammenfassung von Skehans und Robinsons Modellierung der Sprachlerneignung in Aguado 2012: 62f.

Auswertung der einschlägigen Studien zu dem Ergebnis, dass das Arbeitsgedächtnis eine kritische Rolle bei dem Konstrukt Sprachlernneigung für den L2-Erwerb spielt, und zwar unabhängig von Alter, Sprachkompetenz, Sprachfertigkeiten und Lernkontext (vgl. ebd.: 402). Aguado (2012) schätzt das Arbeitsgedächtnis als Kernbestandteil der Sprachlernneigung ein, schließt sich aber auch Ortega (2009) an, die eine voreilige Simplifizierung befürchtet:

„[...] the ways in which memory facilitates differential rate and success of L2 learning may be more complex than a simple correlation between memory tasks and proficiency measures can capture.“ (Ortega 2009: 156)

Im Untersuchungszusammenhang des Einflusses von Alter und Sprachlernneigung auf den Fremdsprachenlernerfolg ist auf die Studien von DeKeyser (2000) und Abrahamsson & Hyltenstam (2008) zu verweisen, wobei im Folgenden DeKeyser (2000) dargestellt wird.

DeKeyser (2000) überprüft in seiner Studie „The Robustness of Critical Period Effects on Second Language Learning“ (vgl. ausführlich in Kap. 2.2 *Studien zum Faktor Alter*) Bley-Vromans (1988, 1990) *Fundamental-Difference-Hypothesis*. Im Zusammenhang der Sprachlernneigung ist der Umstand bemerkenswert, dass DeKeyser nur den *Words in Sentences*-Subtest des MLAT bzw. eine ungarische Version einsetzt, mit dem die grammatische Sensitivität gemessen wird. DeKeyser zufolge sagt dieser Subtest das grammatische Lernen am besten voraus (vgl. DeKeyser 2000: 509), was insofern folgerichtig erscheint, als die Bemessung der Sprachkompetenz der Proband_innen auf der Grundlage eines GJT erfolgt. Dennoch bleibt festzuhalten, dass somit nur eine Teilkompetenz der Sprachlernneigung überprüft wird und die Ergebnisse dementsprechend nicht auf das gesamte Konzept bezogen werden sollten. Von den sechs der 42 Späterwerber_innen (AO > 16), die vergleichbar gute Ergebnisse im GJT erzielen wie die jüngeren Lerner_innen (AO < 16; n = 15), verfügen fünf auch über eine hohe grammatische Sensitivität. Bei den jüngeren Lerner_innen ließ sich dagegen keine Korrelation zwischen gutem GJT-Ergebnis und hoher grammatischer Sensitivität feststellen. Abrahamsson & Hyltenstam (2008) zufolge kann auch die fehlende Varianz bei den Ergebnissen des GJTs für dieses Ergebnis verantwortlich sein: Wenn alle jüngeren Lerner_innen ein ähnlich hohes Ergebnis erzielen (Deckeneffekt), ist es nicht mehr trennscharf (vgl. ebd.: 500).

Nach diesem Überblick über die verschiedenen Ansätze zur Sprachlernung wird im nächsten Unterkapitel die Natur-Kultur-Debatte zusammengefasst, da die Einsicht in den damit verbundenen Paradigmenwechsel die Einordnung einiger Studien zum Faktor Alter erleichtert.

2.4.4 Die Natur-Kultur-Debatte

Wahrscheinlich so alt wie die Menschheit selbst ist die Debatte darüber, was der Mensch erlernen kann und was ihm angeboren ist. Schon in Platons Ideenlehre findet sich die Vorstellung, dass der Mensch mit einer unsterblichen Seele ausgestattet ist, die während ihrer körperlosen Zeit die vollkommenen Ideen schaut. Jede irdische Erkenntnis und jedes Lernen ist ihm zufolge ein Wiedererinnern (*anamnesis*) dieser Ideen (vgl. Menon 81a-84d, Phaidon 72e-77a und 91c-92e). Diese später als Nativismus bezeichnete Position steht dem Empirismus gegenüber, der, verallgemeinert ausgedrückt, davon ausgeht, dass wir alles durch Erfahrung lernen.

Im Bereich der Sprachwissenschaften wird die nativistische Position v. a. mit Noam Chomsky in Verbindung gebracht. Chomsky entwickelte seine Theorie einer Universalgrammatik (UG) v. a. in der Auseinandersetzung mit Skinners behavioristischem Modell des Spracherwerbs (vgl. Singleton & Ryan 2004: 186). Gemäß der Theorie der UG werden alle menschlichen Sprachen mithilfe derselben abstrakten Prinzipien konstruiert. Die UG stellt einen Versuch dar, diese universalen Strukturen mit dem Ansatz der Prinzipien und Parameter zu beschreiben. In diesem Zusammenhang geht Chomsky von einem universellen, kognitiven Spracherwerbsmechanismus aus, der Teil des menschlichen Genoms sei (vgl. Chomsky 1965, 1980, 2007). Demzufolge müsse Sprache nicht gelernt werden, vielmehr werde sie durch Input aus der UG abgeleitet:

„This Principles and Parameters (P & P) approach proposed a sharp distinction between process of acquisition and the format of the internal theory of language. The basic idea is that the genetic endowment provides a fixed system of principles that interact in the generation of the expressions of a language, and a fixed set of parameters that are assigned a value by exposure to external data. A language is determined by setting the values and selecting lexical items from a fixed and highly structured store, with properties of the kind I have mentioned.“ (Chomsky 2007: 13)

Zunächst von einer Phrasenstrukturgrammatik ausgehend, schränkt Chomsky in diesem Artikel Art und Umfang der Prinzipien der UG auf die Rekursivität ein (Chomsky 2007: 15). Begründet wird die UG durch die Beobachtung, dass der L1-

Erwerb universell und uniform abläuft und dabei die erworbene Sprachkompetenz der Kinder die des erfahrenen Inputs übersteigt. Kinder erwerben die Sprache vollständig und korrekt, obwohl sie nur wenig negative Evidenz erfahren (*Poverty of Stimulus*).

Ein anderer populärer Vertreter der nativistischen Position ist Steven Pinker, der einen menschlichen Sprachinstinkt annimmt, der im Gehirn verankert sein müsse (Pinker 1994, 2002: 60). Grammatikgene steuern dabei das Grammatikorgan, womit sich Pinker explizit auf Chomsky bezieht (vgl. Pinker 1996: 347). Weil dieses Organ auf Dauer überflüssig sei und den Metabolismus belaste, würde es nach erfolgreichem Erwerb der Erstsprache von den neurologischen Schaltkreisen entkoppelt (Pinker 1994: 294).

Chomskys UG, die der Erklärung der sprachlichen Kompetenzen von ‚Muttersprachler_innen‘ dient, wurde auch auf den L2-Erwerb übertragen. Für L2-Erwerber_innen gelte das *Poverty-of-Stimulus*-Argument ebenfalls, denn auch ihre Möglichkeit, Konstruktionen in der L2 zu produzieren, übersteige den Input:

„That is, if it can be shown that L2ers acquire abstract and subtle properties that are underdetermined by the L2 input, this suggests that interlanguage competence must be subject to the same constraints as native competence.“ (White 2015: 36)

Gerade auch in der Diskussion um die Erklärung altersbedingter Unterschiede beim Fremdspracherwerb spielt der nativistische Ansatz eine Rolle. Dabei ist strittig, ob und wie weitgehend erwachsene Lerner_innen Zugang zu dem angeborenen Spracherwerbsmechanismus haben. Die folgenden fünf Hypothesen zeigen die verschiedenen Positionen:

1. *Fundamental Difference Hypothesis* (Bley-Vroman 1990): Die UG und ihre Prinzipien sind älteren Lerner_innen nicht mehr verfügbar, sie müssen auf andere Problemlösestrategien setzen, um eine Sprache zu lernen. Die starke Auslegung der These wird durch eine Untersuchung von Bley-Vroman (2009: 26) selbst in Frage gestellt.
2. *Full Access Hypothesis* (Flynn 1987): Auf die UG kann beim frühen wie auch beim späten L2-Erwerb zugegriffen werden. Beobachtete Unterschiede erklären sich durch die Reifung der kognitiven Fähigkeiten und die verschiedenen Bedürfnisse der Lerner_innen.

3. *Partial Access Hypothesis* (Schachter 1988: 232): Die Prinzipien der UG sind verfügbar, nicht aber ihre Parameter.
4. *Indirect Access Hypothesis* (White 1986): Eine alternative Formulierung der *Partial Access Hypothesis* besagt, dass nur die Prinzipien wirksam werden können, die während des L1-Erwerbs aktiviert wurden.
5. Konkurrierende kognitive Systeme (Felix 1985): In dieser Kombination von Chomskys Theorien mit den von Jean Piaget beschriebenen Stadien der kognitiven Entwicklung haben auch erwachsene Lerner_innen Zugang zur UG. Dieser Zugang wird jedoch durch die ausgereiften Problemlösestrategien des Erwachsenen behindert, die nicht bewusst unterdrückt werden können.

White (2015) zufolge ist das verbindende Element dieser unterschiedlichen generativen Ansätze die Annahme, dass die *Interlanguage*-Grammatiken genauso wie die Grammatiken von ‚Muttersprachler_innen‘ unbewusste mentale Repräsentationen enthalten. Uneinigkeit besteht jedoch bezüglich der präzisen Beschreibung dieser Repräsentationen, beispielsweise bei der Art und dem Ausmaß des Einflusses der L1 und dem Status der Einschränkungen durch die UG (vgl. White 2015: 36).

Chomskys generative Grammatik, seine Formulierung universaler Prinzipien und die Vorstellung eines idealisierten Sprechers/Hörers einer Sprache, prägte die Linguistik und damit auch die SLAR. Folgende Argumente werden gegen den nativistischen Ansatz angeführt: Es fehle die neurobiologische Validität (vgl. Paradis 2004: 36), außerdem sei es nicht konstruktiv, mit angeborenem Wissen zu argumentieren angesichts alternativer Erklärungsmöglichkeiten (vgl. Dörnyei 2009b: 85). Das Tempo des Spracherwerbs spreche ebenfalls gegen die UG, die von einem schnellen Erwerb ausginge: Zwischenzeitlich wurde gezeigt, dass der Erstspracherwerb viel langsamer und mühsamer vonstatten geht als ursprünglich angenommen (vgl. Dimroth & Haberzettl 2008: 236–237). Des Weiteren wird kritisiert, dass die UG selektiv sei und sich v. a. auf Grammatik konzentriere, während andere Aspekte der Sprache so gut wie unberücksichtigt bleiben (vgl. MacWhinney 2006). Die tatsächliche Diversität der Sprachen, in denen sich viele Gegenbeispiele zu den als universal postulierten Prinzipien finden lassen (vgl. Evans & Levinson 2009: 430–431), spreche ebenfalls gegen den Nativismus. Schließlich bietet Tomasellos emergentistischer

Ansatz Argumente für den Erwerb der Sprache aus der Erfahrung: „language is a bio-cultural hybrid“ (ebd.: S. 431).

Als Protagonist aktueller emergentistischer Theoriebildung wird im Folgenden Tomasello Ansatz vorgestellt. Sein Konzept geht auf umfangreiche empirische Untersuchungen der Kommunikation und des Sozialverhaltens von Menschenaffen, des Spracherwerbs und Sozialverhaltens von Kleinkindern und der Gehörlosen-Kommunikation hervor. Tomasello zufolge unterscheiden sich die Menschen in erster Linie dadurch von Menschenaffen, dass sie kooperieren können und über eine geteilte Intentionalität verfügen (vgl. Tomasello 2009: 83–112 und 168–182).

„Was in diesen toleranten, friedlich zusammen fressenden Individuen selektiert worden sein könnte, ist die Fähigkeit gemeinsame Ziele und gemeinsame Aufmerksamkeit zu erzeugen. [...] Mutualistische Zusammenarbeit ist also die Geburtsstätte des gemeinsamen begrifflichen Hintergrunds, der notwendig ist für die inferentiell reichhaltige kooperative Kommunikation menschlichen Stils.“ (Tomasello 2009: 209–210)

Arbiträre stimmliche Konventionen, also die gesprochene Sprache, entwickelten sich aus Zeigegesten und dem Gebärdenspiel, da sie nicht entstanden sein können, „ohne irgendwie auf dem Rücken bereits bedeutungshaltiger Kommunikationshandlungen transportiert zu werden“ (ebd.: 259). Werkzeuge und Wörter ermöglichen es dem Menschen, seine Erfahrungen zu konservieren. Dieser „Wagenhebereffekt“ besteht Tomasello zufolge darin, dass es den Menschen gegeben sei, durch Imitation und andere Formen kulturellen Lernens, Neues zu erwerben (vgl. ebd.: S. 309). Tomasello hält die Evolution einer UG zwar nicht für unmöglich, sieht aber keine Notwendigkeit dafür, wenn Sprache in seinem Sinne verstanden wird (vgl. ebd.: 333–334). Stattdessen nimmt er eine „koevolutionäre Spiralbewegung“ an:

„Die kooperative Infrastruktur menschlicher Kommunikation, einschließlich der konventionellen sprachlichen Kommunikation, entsteht daher nicht nur aus der einzigartigen kooperativen, kulturellen Weise des Lebens und Denkens der Menschen, sondern trägt auch zu ihr bei.“ (ebd.: 364–365)

In der SLAR findet sich der emergentistische Ansatz innerhalb der *usage-based* Theorien zum L2-Erwerb. Ellis & Wulff (2015: 75) zufolge teilen die verschiedenen Ansätze die beiden folgenden Hypothesen:

1. Sprachenlernen basiert in erster Linie auf dem sprachlichen Input, den die Lerner_innen in der L2 erfahren.

2. Lerner_innen erzeugen die Regeln ihrer L2 aus diesem Input unter Verwendung allgemeiner kognitiver Mechanismen, die bei jedem Lernprozess involviert und nicht exklusiv für den Spracherwerb sind.

Die Ausgangsbasis ist dabei ein System, das bereits durch die L1 geprägt ist. Deswegen brauchen L2-Lerner_innen eine große und repräsentative Menge an Input, um ein neues rationales Modell abstrahieren zu können, das zum L2-Input passt (vgl. ebd.: 86). Slabakova (2013: 62) stellt fest, dass es in der SLAR etliche empirische Studien gibt, in denen der generative Ansatz überprüft und ausdifferenziert wird. Dagegen wird der *usage-based* Ansatz v. a. theoretisch modelliert und es gibt bislang nur wenige Studien, die emergentistische Vorhersagen überprüfen.

Die Diskussion um nativistische und emergentistische linguistische Beschreibung wurde dargestellt, um einen Einblick in diese divergierenden Ansätze zu geben und weil die CPH Teil des nativistischen Paradigmas ist. Im nächsten Kapitel dieser Arbeit wird das Konzept der ‚muttersprachlichen Kompetenz‘ diskutiert.

2.5 ‚Muttersprachliche Kompetenz‘

In diesem Kapitel wird gezeigt, dass die ‚muttersprachliche Kompetenz‘ lange Zeit als Ziel des fremdsprachlichen Lernens fungierte und nicht hinterfragt wurde, weder in Bezug auf die Angemessenheit dieser Vergleichsgröße noch darauf, was ‚muttersprachliche Kompetenz‘ eigentlich meint. Im nächsten Schritt wird ‚Muttersprachlichkeit‘ als solche thematisiert und genauer analysiert. Nach diesem Unterkapitel wird die geschichtliche Entwicklung des Begriffs ‚Muttersprachlichkeit‘ nachvollzogen und aufgezeigt, in welchen Zusammenhängen daraus ein ethnonationalistisches Konstrukt werden konnte. Welche praktischen Auswirkungen das haben kann, wird anhand der Situation von ‚nicht-muttersprachlichen‘ Fremdsprachenlehrer_innen beleuchtet. Schließlich wird in einem Exkurs die Situation an deutschen Schulen dargestellt und die Wichtigkeit aufgezeigt, die ein ‚ressourcenorientierter‘ Blick auf Mehrsprachigkeit für alle Beteiligten hat.

Die ausführliche Diskussion des Konzepts ‚muttersprachlichen Kompetenz‘ begründet sich dadurch, dass die Erhebung einer Vergleichsgruppe mit Personen mit doppeltem Erstspracherwerb ein Novum dieser Studie darstellt. Darüber hinaus lässt sich das

Thema bisher in komprimierter Form nicht nachlesen, während es eine umfangreiche Auswahl an Überblicksdarstellungen zum Faktor Alter und zu individuellen Variablen gibt.

2.5.1 Die ‚muttersprachliche Kompetenz‘ als Messlatte für den L2-Lernerfolg

In den Untersuchungen zum Faktor Alter – v. a. in den Untersuchungen, die sich mit Belegen oder Gegenbelegen zur CPH auseinandersetzen – geht es meist darum herauszufinden, welches Sprachniveau die Lerner_innen erreichen können bzw. ob es Späterwerber_innen gibt, die eine sogenannte ‚muttersprachliche Kompetenz‘ erreichen. So zeigen etwa Abrahamsson & Hyltenstam (2009: 287) in ihrer großangelegten Studie, die insgesamt 195 Bilinguale mit den Sprachen Schwedisch und Spanisch untersucht, Folgendes: Nur drei von 31 Frühlerner_innen (AO = 1–11), die in der ersten Untersuchungsphase als ‚Muttersprachler_innen‘ bewertet wurden, erreichen bei noch genauerer Untersuchung (*scrutinized nativeness*)³³ das Kompetenzprofil der muttersprachlichen Vergleichsgruppe (vgl. die Darstellung der Studie in Kap. 2.2 *Studien zum Faktor Alter*). Damit scheinen sie genau das auszuführen, was Davies bereits sechs Jahre früher konstatierte: „when nonnative speakers have been shown to perform as well as a native speaker on a test, the cry goes up for yet another test“ (Davies 2003: 213).³⁴

Davies (2011) hinterfragt die ‚muttersprachliche Kompetenz‘ u. a. im Hinblick auf Sprachprüfungen und Testformate. Ihm zufolge handelt es sich beim *native speaker* genau so um eine Idealisierung wie bei der *standard language* (ebd.: 306). Davies schlägt vor, bei der Leistungsbewertung zukünftig vom Kompetenzprofil idealer L2-Lerner_innen auszugehen und nicht mehr vom lange Zeit unhinterfragt gebliebenen

³³ Zu den 10 hochkomplexen und kognitiv anspruchsvollen Instrumenten, die eingesetzt und ausgewertet wurden, zählen: *production and perception of voice onset time, speech perception in noise, grammaticality judgement, grammatical, lexical, and semantic inferencing, formulaic language* (vgl. Abrahamsson & Hyltenstam 2009: 276–279).

³⁴ In diesem Zusammenhang ist es aufschlussreich, auf den Furor in Longs Ausdrucksweise zu achten, wenn er Studien zusammenfassend beschreibt, die ‚muttersprachliches Niveau‘ bei Späterwerbern festgestellt haben: „Some researchers claiming to find adult starters with native-like abilities in the L2 have really done no more (or less) than produce individuals able to fool (some) NS judges forced to make judgments of nativeness on the basis of *inadequate samples* of behavior and/or cases of highly educated, highly analytic learners able to score very well on paper-and-pencil tests of language-like behavior“ (Long 2005: 297, Hervorhebungen von SH).

Konzept ‚Muttersprachler_in‘ (ebd.: 296). In einer Rezension zu Davies wird dies folgendermaßen formuliert: „One can let the (ideal) native speaker [...] function as a goal or an inspiration for L2 learning, but let the ideal L2 user serve as a measure of L2 learning“ (Han 2004: 184).

Desgleichen Hulstijn (2011, 2012, 2015), der die idealisierte ‚muttersprachliche Kompetenz‘ generell in Frage stellt. Stattdessen nimmt er eine Unterscheidung in *Basic Language Cognition* (BLC) und *Higher Language Cognition* (HLC) vor, um die jeweilige Sprachkompetenz exakter zu beschreiben. Während die BLC v. a. Standardwortschatz und -phrasen umfasst und einem hohen Grad der Automatisierung unterliegt, beinhaltet die HLC Wortschatz und Phrasen niedriger Frequenz und bezieht sich auf gesprochene wie auch geschriebene Äußerungen (vgl. Hulstijn 2011: 230–231). Diese Unterscheidung in ein basales und ein höheres sprachliches Register wurde von Cummins 1979 in die Debatte eingeführt und dort als *Basic Interpersonal Communicative Skills* (BICS) und *Cognitive Academic Language Proficiency* (CALP) bezeichnet (vgl. Cummins 2008: 71). Unter BICS wird die Flüssigkeit der Kommunikationsfähigkeit der Lerner_innen verstanden, unter CALP die Fähigkeit, mündlich und schriftlich für den Schulerfolg relevante Konzepte zu verstehen und auszudrücken (vgl. ebd.). Cummins weist darauf hin, dass er mit den Begriffen keine ausdifferenzierte Konzepte vorlegt. Vielmehr nimmt er mit BICS und CALP eine konzeptionelle Unterscheidung vor, um über Klassenzimmer-Realitäten zu sprechen und auf diskriminierende Begutachtungen und Unterrichtspraktiken hinzuweisen, die viele bilinguale Schüler_innen betreffen.³⁵

Hulstijn stellt fest, dass man von ‚Muttersprachlichkeit‘ nur in Bezug auf die BLC sprechen kann. Dagegen sei die HLC auch von L2-Lerner_innen zu erreichen, wenn sie ein entsprechendes Bildungsniveau haben. Umgekehrt verfügen nicht alle ‚Muttersprachler_innen‘ über HLC (Hulstijn 2011, 242). Neben der Unterscheidung in BLC und HLC führt Hulstijn eine weitere Dimension ein, um Sprachkompetenzunterschiede zu beschreiben: Kernkomponenten vs. periphere Komponenten (*Core ver-*

³⁵ „The children seem to have much greater English proficiency than they actually do because their spoken English has no accent and they are able to converse on a few everyday, frequently discussed subjects. Academic language is frequently lacking. Teachers actually spend very little time talking with individual children and tend to interpret a small sample of speech as evidence of full English proficiency“ (Vincent 1996: 195, zit. nach Cummins 2008: 74).

Periphery). Kernkomponenten gehören zur linguistischen Kognition in den Bereichen Phonetik/Phonologie, Morphologie, Morphosyntax und Lexik. Periphere Komponenten sind metakognitive Kompetenzen wie bspw. metalinguistisches Wissen über verschiedene mündliche und schriftliche Diskurstypen und -strategien (vgl. Hulstijn 2015: 41f.). Hulstijn zufolge fällt die BLC in den Bereich der Kernkomponenten, während die HLC periphere und Kernkomponenten umfasst (vgl. ebd.: 44f.).

Die theoretische Modellierung einer Unterscheidung in eine basale und eine höhere Sprachkompetenz – wobei die Aneignung der letzteren v. a. auch in Zusammenhang mit dem Bildungshintergrund steht – lässt sich mit empirischen Studien unterstützen. In ihrem Review-Artikel fasst Dabrowska (2012) die Ergebnisse einiger Studien zusammen, die individuelle Unterschiede bezüglich der morphologischen und syntaktischen Grammatikperformanz bei erwachsenen ‚Muttersprachler_innen‘ untersuchen:

„As we will see, some of these constructions [Polish genitive and dative inflections, English passives, the universal quantifier *every* and several types of sentences containing subordinate clauses] are not fully mastered by all native speakers; in other cases “the same” knowledge is represented differently by different speakers.“ (Dabrowska 2012: 221)

Neben dem Nachweis individueller Unterschiede bei den ‚Muttersprachler_innen‘ ist es besonders bedeutungsvoll, dass Dabrowski anhand mehrerer Studien zeigt, dass es einen deutlichen Zusammenhang zwischen Bildungshintergrund und sprachlicher Kompetenz gibt (vgl. Dabrowska 2012: 227ff.; Korrelationskoeffizienten, die hier wiedergegeben werden könnten, finden sich in diesem Abschnitt nicht).

Zur Frage, warum es eine Beziehung zwischen Bildung und unterschiedlichem grammatischen Wissen geben sollte, führt Dabrowska drei verschiedene Erklärungsmöglichkeiten an. So könnte bspw. die größere Menge an Erfahrungen mit geschriebener Sprache dafür verantwortlich sein oder die größere Menge an sprachlicher Erfahrung generell. Dabrowska zufolge spiele Sprache im Arbeitsalltag von Akademiker_innen eine größere Rolle als bei Nicht-Akademiker_innen: „they are likely to have more exposure to all the structures tested than the non-graduates“ (Dabrowska 2012: 242). Als dritten möglichen Erklärungsansatz führt Dabrowska an, dass sich die bisherige sprachliche Erfahrung der Proband_innen qualitativ danach unterscheidet, ob expliziter Sprachunterricht bezogen auf die L1 stattgefunden habe oder nicht (vgl. Dabrowska 2012: 243). Dabrowska zufolge ist die Beziehung zwischen Bildungshintergrund und sprachlicher Kompetenz in beide Richtungen vorstellbar:

„This raises the possibility that the HAA [High Academic Achievers] participants are simply better language learners – in other words, that they are better at inferring meaning from context, noticing patterns in the input, and generalizing those patterns – and that their academic success is a result of their better linguistic skills.“ (Dabrowska 2012: 243)

Die gefundenen Unterschiede unter ‚Muttersprachler_innen‘ zeigen, dass die Annahme einer einheitlichen ‚muttersprachlichen Kompetenz‘ nicht haltbar ist. Diese Ergebnisse sind ein weiterer Beleg gegen die *Fundamental Difference Hypothesis* und ihre theoretische Voraussetzung, dass L1-Sprecher_innen alle über dieselbe Grammatik verfügen, während L2-Lerner_innen sich diesbezüglich unterscheiden (vgl. Dabrowska 2012: 220 sowie die Darstellung in Kap. 2.4.4 *Die Natur-Kultur-Debatte*). Der Zusammenhang zwischen Bildung und grammatischem Wissen unterstützt die Modellierungen von Cummins und Hulstijn (s. o.), die eine basale, alltägliche Kompetenz von einer höheren, akademischen Sprachkompetenz unterscheiden.

Zusammenfassend lässt sich festhalten, dass mehr Differenzierung bezüglich dessen, was ‚muttersprachliche Kompetenz‘ meint, von Nöten ist sowie mehr Fairness bei der Bewertung der L2-Lernerfolge. ‚Muttersprachliche Kompetenz‘ darf nicht unhinterfragt als Ziel und Maß des L2-Lernerfolgs angesetzt werden. Vor diesem Hintergrund ist der Gemeinsame Europäische Referenzrahmen für Sprachen (GER) zur Bewertung der L2-Kenntnisse zu begrüßen, der in seinen Kompetenzbeschreibungen gänzlich ohne ‚muttersprachliche Kompetenzen‘ auskommt (vgl. URL 3; 2000).

2.5.2 Konstruktion und Dekonstruktion eines Begriffs

Die Rede von der ‚Muttersprache‘ stellt zwar eine Selbstverständlichkeit dar, es ist aber nicht klar, wovon dabei eigentlich gesprochen wird. Im strengen Sinne kann es keine ‚Muttersprachler_innen‘ geben, sondern nur Sprecher_innen eines bestimmten Dialektes, Soziolektes oder, bezogen auf die individuelle Ebene, eines Ideolektes. Ein weiterer Grund dafür, den Maßstab der ‚Muttersprachlichkeit‘ zu hinterfragen, liegt darin, dass dieser Terminus diskriminierende und rassistische Bedeutungsaspekte transportieren kann. Bonfiglio (2010) hat in seiner Untersuchung „Mother tongues and nations: the invention of the native speaker“ Folgendes gezeigt: Bei dem Begriff *native* handelt es sich um ein Konstrukt, das sowohl rassistisch motiviert ist als auch dazu dient, Macht und Einfluss einer dominierenden Gruppe abzusichern, die den Standard definiert: „These terms are no innocuous intuitions: They are the divisive implements of ethnolinguistic nationalism“ (ebd.: 222). Aus diesem Grund legt er die historischen Bauplätze der Begriffe *native* und *mother tongue* frei,

um zu zeigen, wie die Begriffe ihre tendentiöse Ladung erhielten, und sie dadurch zu dekonstruieren.

In einem ausführlichen Gang durch die Geschichte zeigt Bonfiglio, wie *native* und *mother tongue* ihre ausgrenzenden und rassistischen Bedeutungsfacetten erhielten. In der Antike ist der Begriff von Sprache fluide. Das Beherrschen mehrerer Sprachen galt als politisch, ökonomisch und militärisch sinnvoll. Es war unerheblich, ob ein_e Sprecher_in eine Sprache als L1 oder L2 spricht. Mehrsprachigkeit wurde ganz pragmatisch gelebt, und Bonfiglio zufolge wurde nach folgenden Kriterien entschieden, wann welche Sprache zum Einsatz kam:

„The language used should be the one that: (1) you know best for the topic concerned; (2) you believe your interlocutor knows best for the topic concerned; (3) you used the last time you addressed this person; (4) includes or excludes a third party; (5) asserts the most advantageous social group membership in the interaction.“ (Langslow 2002, zit. nach Bonfiglio 2010: 55)

Im Mittelalter wird der Begriff der ‚Muttersprache‘ geprägt und die organistischen Metaphern für Sprache in der Frühmoderne haben hier ihren Ursprung. Allerdings kann von einer außerordentlichen Rolle der ‚Muttersprache‘ resp. Nationalsprache keine Rede sein. Vielmehr verhinderte das Primat des Lateinischen die Herausbildung einer schriftsprachlichen Kultur der Volkssprachen.

Erst das Auftauchen der Nationalstaaten in der frühen Moderne generierte den ethnolinguistischen Nationalismus (Bonfiglios Terminus für die Vorstellung, eine Nation zeichne sich durch eine gemeinsame Sprache und eine gemeinsame ethnische Herkunft aus), der unserem zeitgenössischen ähnelt (vgl. ebd.: 219). Welche gedanklichen Konstrukte diesem ethnolinguistischen Nationalismus zugrunde liegen, lässt sich am Beispiel des Sprachphilosophen Johann Gottfried Herder nachvollziehen. Herder gebraucht das Wort ‚Muttersprache‘ häufig und synonym mit ‚deutsche Sprache‘, ‚Landessprache‘ und ‚Sprache des Vaterlandes‘. Er bezeichnet damit sowohl die Sprache der individuellen Sozialisation und des gesellschaftlichen Verkehrs als auch die Schriftsprache (vgl. Ahlzeit 1994: 110). Herder identifiziert Sprache mit Nation resp. Volk:

„Wer in derselben Sprache erzogen ward, wer sein Herz in sie schütten, seine Seele in ihr ausdrücken lernte, der gehört zum *Volk dieser Sprache*.“ (Herder 1795: 287)

Diese organische Einheit impliziert, dass ein Mensch nur eine ‚Muttersprache‘ haben kann, dass ihm andere Sprachen fremd und etwaige Transplantationen schwierig sind. Und umgekehrt ist und bleibt die jeweilige Sprache für die Außenseiter_innen letztendlich unzugänglich (vgl. Bonfiglio 2010: 133). Der Terminus ‚Sprachgefühl‘,

über das nur ‚Muttersprachler_innen‘ verfügen sollen, geht auf Herder zurück (vgl. Ahlzweig 1994: 117).³⁶ Sprache und Nation bilden in diesem Diskurs eine Einheit – ‚Muttersprachler_in‘ kann nur sein, wer auch der ethnischen Gruppe angehört, um deren Sprache es sich handelt. Dass gerade im deutschen Sprachraum ein großes Interesse an der Identifikation von Sprache und Nation bestand, erklärt Ahlzweig (1994: 54ff.) durch den Umstand, dass Deutschland erst spät zu einer politischen Entität wurde: Mangels eines klar durch äußere Grenzen definierten territorialen Bezugsrahmens habe sich ein „Kulturpatriotismus“ (ebd.: 54) herausgebildet, dessen zentraler Dreh- und Angelpunkt die Sprache als Trägerin „deutscher Kultur“ gewesen sei. Diese Vorstellung habe bis heute nichts an Wirkmächtigkeit verloren und zeige sich immer dann, wenn Personen die Anerkennung als ‚genuine Muttersprachler_innen‘ versagt wird, weil sie anders aussehen und andere Namen tragen als von L1-Sprecher_innen der betreffenden Sprache erwartet.³⁷ Anhand des Umgangs der Deutschen mit den „türkischen Gastarbeitern“ – die immer noch so heißen, auch wenn sie bereits in zweiter oder dritter Generation in Deutschland leben und vielleicht monolingual Deutsch aufwachsen – zeigt Bonfiglio: „In the perception of nativity, identity tends to be judged first and language later“ (Bonfiglio 2010: 12).

Bonfiglio zeigt, wie sich die exklusive Bedeutung der Termini ‚Muttersprache‘ bzw. „*native*“ auch in anderen europäischen und nordamerikanischen Ländern etablierte (vgl. ebd.: 122 ff.). Er konstatiert, dass die Standardisierung der Sprache zeitgleich mit der Verbreitung der Alphabetisierung stattfindet. Die schriftsprachliche Kompetenz sei etwas, das erlernt werden müsse, und nichts, in das man hineingeboren werde (vgl. Bonfiglio 2010: 8). Standardisierung komme wertneutral daher, sei es aber in keiner Weise. Im amerikanischen Englisch würden *standard* und „*proper speech and pronunciation*“ seit Franklins Zeiten in rassistischen und ethnischen Begriffen gefasst. Franklin verstand Standard als die Form, die von den Engländer_innen

³⁶ An dieser Stelle muss Herder auch verteidigt werden: Sprache und Kultur sind bei ihm miteinander verkoppelt und durch seine Sammlung an Volksliedern (Stimmen der Völker in Liedern 1778/79) trägt er zu einer Demokratisierung des Kulturbegriffs bei. Außerdem zeigt seine Sammlung, dass er im Gegensatz zu den kolonialistischen Diskursen seiner Zeit in keiner Weise verschiedene Kulturen hierarchisiert.

³⁷ Vgl. hierzu Noah Sows „Deutschland Schwarz Weiß. Der alltägliche Rassismus“ über die Diskriminierung trotz bester Absichten: „Sie sprechen aber gut Deutsch“ (ebd.: 252 ff.).

in Nordamerika benutzt werde, auf jeden Fall nicht die Formen, die Deutsche oder Schwarze gebrauchen (vgl. ebd.).

Anhand des Beispiels einer Stellenanzeige in einer singapurischen Zeitung illustriert Bonfiglio (2010) die heutige Virulenz dieser rassistischen Bedeutungsschichten:

„On Thursday, July 12, 1990, the Singapore newspaper *The Straits Times* listed the following advertisement: „Established private school urgently requires native speaking expatriate English teachers for foreign students.“ By Saturday, July 14, the advertisement had been changed as such: „Established private school urgently requires native speaking *Caucasian* English teachers for foreign students“ [...] The belated addition of the word „Caucasian,“ however, indicates that the semantic field of the term „native“ in the original advertisement extends well beyond purely linguistic criteria; it clearly contains notions of race and ethnicity.“ (Bonfiglio 2010: 1)

Schon Paikeday (1985) findet hierfür klare Worte: „[...] sometimes you begin to wonder, when people start recruiting “native speakers,” of English, for example, whether they don’t really mean, “White Anglo-Saxon protestants; Scots maybe, but no Irish need apply” “ (ebd.: 46–47). Holliday (2008) erweitert den Problemkreis noch, indem er zahlreiche Beispiele für die Diskriminierung nicht-nativer Fremdsprachenlehrender auf dem Arbeitsmarkt gibt, die trotz besserer Qualifikationen gegenüber den ‚Muttersprachler_innen‘ das Nachsehen haben (vgl. ebd.: 121).

Unter *native-speakerism* versteht Holliday die weit verbreitete Ideologie innerhalb des Fachdiskurses zum *English Learning Teaching* (ELT), die folgendermaßen charakterisiert wird: „‘native-speaker’ teachers represent a ‘Western culture’ from which spring the ideals both of the English language and of English language teaching methodology“ (Holliday 2006: 385). Der Einfluss dieser Ideologie mache sich in vielen Bereichen von der Beschäftigungspolitik bis zu sprachlichen Repräsentationen bemerkbar. Das zugrundeliegende Thema ist dabei das *othering*³⁸ von Studierenden und Kolleg_innen außerhalb des englischsprachigen Westens, hervorgerufen durch die Orientierung an essentialistischen regionalen oder religiös-kulturellen Stereotypisierungen (vgl. ebd.). Für Holliday ist das *othering* im Bereich ELT v. a. eine Kompensationsstrategie, wenn es sich erweist, dass die Lerngruppe Schwierigkeiten hat mit den kommunikativen, aktivierenden, lernerzentrierten Methoden, die im westlichen Diskurs als allen herkömmlichen Methoden überlegen gehandelt werden. Die von Holliday in der Fachliteratur gefundenen Kennzeichnungen zeigen, dass es ge-

³⁸ Differenzierung und Distanzierung der Gruppe, der man sich zugehörig fühlt, von anderen Gruppen (vgl. URL 4; 2014: o. S.).

rechtfertigt ist, in diesem Zusammenhang von neokolonialistischen Stereotypisierungen zu sprechen:

„Such a perspective is native-speakerist because it negatively and confiningly labels what are in effect 'non-native speaker' 'cultures' as 'dependent', 'hierarchical', 'collectivist', 'reticent', 'indirect', 'passive', 'docile', 'lacking in self esteem', 'reluctant to challenge authority', 'easily dominated', 'undemocratic', or 'traditional' and, in effect, uncritical and unthinking.“ (ebd.: 385–386)

Tran (2009) zeigt die Wirkmächtigkeit dieses kolonialen Erbes am Beispiel des DaF-Unterrichts in Vietnam im Spannungsfeld von Globalisierung und Rückbesinnung auf Traditionen bzw. kommunikationsorientiertem DaF-Unterricht und Konfuzianismus.

2.5.3 Der ‚monolinguale Habitus‘ und seine Folgen

Seit Gogolins (1994) Studie, die dem deutschen Bildungswesen und seinen professionellen Akteurinnen und Akteuren einen ‚monolingualen Habitus‘ nachwies, ist viel geschehen: Das Bekenntnis zum Einwanderungsland Deutschland wird zwar nach wie vor nicht von allen politischen Akteur_innen geteilt. Die Reform des Staatsangehörigkeitsrechts und die Neuregelung der Einbürgerung im Jahr 2000 (vgl. Storz & Wilmes 2007: 1) sind aber klare Zeichen für die Einsicht in den Handlungsbedarf angesichts einer Bevölkerungsquote mit Migrationshintergrund von 20 %, von denen knapp die Hälfte Ausländer_innen sind (vgl. URL 5; 2012: 7). Auch auf der bildungspolitischen Ebene wird mittlerweile gehandelt, etwa mittels der Einführung obligatorischer DaZ-Veranstaltungen in der Lehrer_innenausbildung oder der Etablierung bilingualer Schulen.³⁹ Dennoch ist zu befürchten, dass Gogolins Beurteilung der grundsätzlichen Orientierung des Bildungswesens an der Einsprachigkeit und des gleichzeitigen Desinteresses an mehrsprachiger Kompetenz oder sogar defizitärer Wahrnehmung von Mehrsprachigkeit auch heute noch Gültigkeit hat.

Sind monolinguale ‚Muttersprachler_innen‘ das Ideal, dann kann dieser Ideologie zufolge eine Sprache als L2 nicht auf dem gleichen hohen Kompetenzniveau erlernt bzw. beherrscht werden. Mehrsprachigen Kindern werden sprachliche Defizite der Zielsprache unterstellt und ihre vermuteten anderen Sprachkenntnisse werden in der

³⁹ Im Land Berlin ist seit WS 2007/2008 ein verpflichtendes DaZ-Modul Bestandteil der Lehrer_innenausbildung. In Nordrhein-Westfalen wurde dies zum WS 2010/2011 eingeführt. In den anderen Bundesländern werden DaZ-Inhalte ebenfalls in die Lehrer_innenausbildung integriert, sind aber kein eigenes Modul (Quelle: Gespräch mit Anke Börsel, Fachgebiet Deutsch als Zweitsprache, TU Berlin, 19.08.2013).

Regel nicht aktiv bei der Entwicklung sprachlicher Ausdrucksfähigkeit genutzt: „Vielsprachigkeit in der Schule [...] wird weder gefördert, noch werden die Chancen genutzt, die sich daraus [...] ergeben.“ (Hoodgarzadeh 2010: 37). Vor allem werden sie dann nicht genutzt, wenn es sich um die Sprache einer „ethnolinguistisch minorisierten“ (Vermees 1998: 12) Gruppe handelt. Vermees erfasst mit dem Begriff die Hierarchisierung von Herkunftsländern und -kulturen, die unterschiedlich wertgeschätzt werden:

„In der Tat hat die Erforschung unterschiedlicher Konstellationen von Mehrsprachigkeit aufgezeigt, wie sehr die Minorisierung bestimmter Sprachen durch Machtbeziehungen zwischen den beteiligten Gruppen bestimmt wird.“ (ebd.: 13)⁴⁰

Dem entgegenzusetzen wäre ein ressourcenorientierter Blick auf Mehrsprachigkeit, der sie als Bereicherung und Chance sowohl für die Schüler_innen als auch für die Lehrenden wahrnimmt. Tatsächlich vorhandene sprachliche Schwierigkeiten von Kindern und Jugendlichen in der L2 dürfen dabei natürlich nicht negiert werden. Wichtig ist aber, sie nicht verkoppelt mit schulischem Misserfolg auf die Mehrsprachigkeit zurückzuführen, sondern andere Faktoren in den Blick zu nehmen, bspw. soziale Hintergründe und Bildungstraditionen und -ideale der Herkunftsfamilien. Schließlich zeigen Sprachstandserhebungen zum Zeitpunkt der Einschulung und in der gymnasialen Oberstufe sprachliche Defizite auch bei Kindern ohne Migrationshintergrund (vgl. Petersen 2013). So wurde 2012 in Brandenburg bei 18 % der Schulanfänger_innen sprachlicher Förderbedarf festgestellt (vgl. Rockmann, Rehkämper & Leerhoff 2013: 38). In Berlin – wo bezeichnenderweise die Sprachkenntnisse ausschließlich bei Einschulungskindern mit Migrationshintergrund getestet werden – ergab sich ein nach Aufenthaltsdauer und elterlichem Bildungsniveau differenziertes Bild. Bei hohem Bildungsniveau der Eltern lag der Anteil guter bis sehr guter Ergebnisse im Sprachtest bei knapp 80 %, bei einem bildungsfernen Elternhaus schnitten nur 59 % gut oder sehr gut ab (vgl. ebd.: 2013: 42).

Dass die Einflussfaktoren auf den Schulerfolg nicht auf Mehrsprachigkeit bzw. eine abweichende ‚Muttersprache‘ zurückzuführen sind, zeigt auch ein Blick auf die vietnamesischen Einwanderer_innen: Im Vergleich zu anderen Gruppen mit Migrations-

⁴⁰ Zu denken ist hierbei etwa an die vergleichsweise geringe Anerkennung von Türkisch- und Arabischkenntnissen gegenüber Englisch, Französisch oder Spanisch. Vgl. auch Meißner (2013: 76), der in diesem Zusammenhang von „Sprachenstolz“ und „Sprachenscham“ spricht. Darüber hinaus wurde Mehrsprachigkeit lange Zeit defizitär behandelt, wenn etwa von „doppelseitiger Halbsprachigkeit“ (Stötting 1980) die Rede war.

hintergrund weisen sie einen überdurchschnittlichen Schulerfolg auf, denn zwei Drittel von ihnen machen Abitur (Mai 2011). Dass die Eltern dieser Kinder – die sogenannten Boatpeople in der BRD sowie die ehemaligen Vertragsarbeiter_innen der DDR – zumeist nur über geringe Deutschkenntnisse verfügen, verweist eindeutig darauf, dass andere Faktoren wie sozialer Hintergrund, Bildungstraditionen und Werte eine größere Rolle spielen als die Sprache des Elternhauses (vgl. Schöningh 2013: 144).

Angesichts der Verabschiedung der EU-Rahmenstrategie 2005 zur Förderung von Mehrsprachigkeit und lebenslangem (Sprachen-)Lernen (vgl. URL 6; 2005) wäre es erstrebenswert, wenn eine ressourcenorientierte Perspektive auf Mehrsprachigkeit den defizitären Blick ablösen würde, der Bildungsmisserfolge vorschnell ethnisiert und es dadurch versäumt, die sozialen Hintergründe zu beleuchten.

2.5.4 Alternativen zum Begriff der ‚Muttersprache‘

Hoodgarzadeh (2010: 39) befragte 86 Gymnasiast_innen mit Einwanderungsgeschichte im Alter von 16 bis 20 Jahren aus 23 Herkunftsländern mit einer Fragebogenenerhebung u. a. danach, welche Sprache(n) sie als ihre ‚Muttersprache‘ anerkennen. Dabei ergaben sich drei Dimensionen mit fließenden Grenzen: Die Herkunftssprache der Eltern wird als ‚Muttersprache‘ anerkannt, wenn rationale Aspekte für diese Wahl zugunsten der emotionalen Perspektive in den Hintergrund treten. Häufig benannt werden dabei der Bezug zur Familie und der Stolz auf die Herkunftskultur. In vielen Fällen sind Abgrenzungstendenzen beobachtbar, die einhergehen mit der Ausgrenzung und Ablehnung durch die deutsche Mehrheitsgesellschaft und schulischen oder beruflichen Misserfolgen (vgl. ebd.: 39–40).

In der zweiten Dimension wird Deutsch die Rolle der ‚Muttersprache‘ zugewiesen, wobei die Gründe rationale und emotionale Aspekte haben: Einerseits geht es um die sprachliche Ausdrucksfähigkeit und Sicherheit im Deutschen und die Motivation zur Integration. Andererseits wird angeführt, auf Deutsch zu sprechen, zu denken und zu träumen. Ausschlaggebend hierfür können ein überwiegend deutscher Freundeskreis sein, ein Elternhaus mit hohem Engagement für die Schule und/oder eigenes Interesse am Schulerfolg (vgl. ebd.: 41). Diese Fälle verdeutlichen eindrucksvoll, dass die „[...] vielfach automatisch vorgenommene Synonymisierung der Begriffe „Muttersprache“ und „Erstsprache“ nicht berechtigt ist“ (ebd.: 42).

In der dritten Dimension werden sowohl die Herkunftssprache als auch die deutsche Sprache als ‚Muttersprache(n)‘ empfunden, wenn eine tiefe emotionale Verbundenheit mit den Eltern bei gleichzeitiger kommunikativer Sicherheit im Deutschen vorhanden ist. Die Autorin hält fest, dass bei keiner der Antworten eine Zerrissenheit zwischen den Kulturen nachweisbar war, vielmehr wird aus Herkunftskultur und deutscher Kultur ein eigenes System generiert (vgl. ebd.: 42–43).⁴¹ Hoodgarzadeh zieht aus diesen Dimensionen das Resümee, dass die befragten Schüler_innen eine „Zuschreibungsemanzipation“ entwickeln, in deren Folge ‚Muttersprache‘ zu einem flexiblen, frei wählbaren Konstrukt wird, „das jeweils vor dem Hintergrund individueller Erfahrungen und sozialisationsbedingter Bildungsbiografien verstanden werden muss“ (ebd.: 44).

Darüber hinaus wurden im Rahmen der Projektwoche BEN (Türkisch: ich) – die die Kommunikation über die sprachlichen Ressourcen einer heterogenen achten Hauptschulklasse zum Gegenstand hatte – neben vielfältigen anderen Instrumenten Aufsätze zum Thema ‚Muttersprache‘ geschrieben und ausgewertet. Es ging darum, nicht nur die Reflexionen von sog. „bildungserfolgreichen Migrant_innen“ zu erfassen (vgl. ebd.: 44). Die Aussagen der Hauptschüler_innen zeigen einerseits einen pragmatischen Umgang mit der Frage, wenn etwa die ‚Muttersprache‘ als Sprache der Mutter verstanden wird. Andererseits lässt sich ein großes Sprachpotential mit teilweise drei bis vier gesprochenen Sprachen in der Familie und einem hohen Reflexionsniveau bezüglich der persönlichen ‚Muttersprache‘ nachweisen (ebd.: 44). Hoodgarzadeh zieht aus den Untersuchungen das Fazit, dass Begrifflichkeiten wie ‚Muttersprache‘, Deutsch als ‚Muttersprache‘, Deutsch als Zweitsprache – so wie sie derzeit in der Literatur verwendet werden – nicht (mehr) ausreichen, um den Realitäten einer mehrsprachigen Gesellschaft gerecht zu werden (vgl. ebd.: 46).

Davies (2011) schlägt den Terminus *native user* vor, um „proficient nonnative speakers“ (ebd.: 296) zu bezeichnen und/oder „proficient in ELF (English as Lingua Franca, SH)“ (ebd.: 294). Es bleibt unklar, ob Davies damit das Konzept *native speaker* ablösen will oder *native user* und *native speaker* parallel verwendet werden sollen. Für Davies gibt es funktional keine Unterschiede zwischen *native users* und *native*

⁴¹ „Dies wird dann deutlich, wenn auf die Frage: Sehen Sie sich eher als Iraner oder eher als Deutscher? wie folgt geantwortet wird: Weder noch. Ich sehe mich eher als norddeutscher Perser.“ (Hoodgarzadeh 2010: 38).

speakers. Auch bezogen auf die Sprachkompetenz sind die Gruppen teilweise deckungsgleich. Allerdings bleibt ein Unterschied bestehen: „being a native speaker is ascribed, whereas being a native user is attained“ (ebd.).

Dieser Unterschied wird in den von Hyltenstam & Abrahamsson (2012) aktuell verwendeten Definitionen der Begriffe deutlich: Unter „native speaker“ verstehen sie den „archetypical L1 speaker, who is not necessarily monolingual but who has learned that language and continued to use it regularly throughout the life span“ (ebd.: 182). Ein „nativelike speaker“ ist ihnen zufolge eine Person, die in jeder Hinsicht wie ein „native“ spricht, obwohl die Sprache nicht seine L1 ist (vgl. ebd.). Ein „near-native speaker“ ist eine Person, die sich „native“ anhört, sich aber bei genauerer linguistischer Untersuchung unterscheidet (vgl. ebd.). Demgegenüber ist ein „advanced L2 learner“ jemand, dessen L2 fast „native“ ist, dessen L2-Sprachgebrauch aber in der alltäglichen mündlichen und schriftlichen Kommunikation deutlich zu merken ist.

Im deutschen sprachwissenschaftlichen Diskurs wird zumeist von ‚Erstsprache‘ gesprochen. Begriffe wie Erstsprachler_in finden sich jedoch nur selten (vgl. Ahrenholz 2008: 5). Der Begriff ‚Muttersprache‘ verzerrt Ahrenholz zufolge die tatsächliche Erwerbssituation, an der mehr Personen und nicht ausschließlich die Mütter beteiligt sind (ebd.: 3). Ein anderer Grund, auf den Begriff ‚Muttersprache‘ in der Fachliteratur zu verzichten, besteht darin, dass mit ‚Erstsprache‘ die wissenschaftsgeschichtlichen Konnotationen des Begriffs ‚Muttersprache‘ vermieden werden, der in der älteren Fachliteratur nicht selten mit einer negativen Bewertung früher Mehrsprachigkeit einhergegangen ist (vgl. Andresen 2010: 184). Siebert-Ott (2003) verwendet die Begriffe Erstsprache, Zweitsprache usw., wenn es um die Erwerbsreihenfolge geht und ‚Muttersprache‘, „wenn damit die „Sprache des kulturellen Erbes“ (ebd.: 31) einer Sprachgruppe bezeichnet werden soll. Dies geschieht unabhängig davon, ob diese Sprache tatsächlich im Alltag dieser Gruppe die ausschließliche oder überwiegend verwendete Sprache ist“ (ebd.). Ahrenholz weist darauf hin, dass dem Begriff Erstsprache die emotionale Dimension fehle, die mit ‚Muttersprache‘ mitgemeint wird (Ahrenholz 2008: 4).

Die optimale Auflösung der Arbeit an den Begrifflichkeiten wäre der Umgang im Sinne der von Hoodgarzadeh befragten Jugendlichen: ‚Muttersprachlichkeit‘ als flexibles Konstrukt, das je nach eigenem Bedürfnis und eigener Verortung verwendet wird.

Solange aber die gesellschaftliche Konstruktion der Begriffe eine andere ist und ‚Muttersprachlichkeit‘ als Begriff dazu benutzt wird, nach nicht ausschließlich sprachlichen Kriterien Menschen ein- und auszuschließen (vgl. Kap. 2.5.2 *Wer bestimmt den Standard?*) ist es wichtig, alternative Begriffe zu benutzen.

2.6 Einfluss der Ausgangssprache

Zum Abschluss des theoretischen Teils dieser Arbeit soll der Einfluss der Ausgangssprache auf den Lernerfolg einer L2 betrachtet werden. In diesem Zusammenhang wird auf die Frage eingegangen, welche Rolle strukturelle bzw. typologische Unterschiede zwischen der Ausgangs- und der Zielsprache spielen. Danach erfolgt ein kurzer sprachtypologischer Vergleich der Sprachen Mandarin und Deutsch. Darüber hinaus wird ein Überblick zur Situation des DaF-Unterrichts in China gegeben.

Bezüglich des Einflusses der Ausgangssprache auf den L2-Lernerfolg werden innerhalb der Spracherwerbsforschung verschiedene Hypothesen diskutiert. Vertreter_innen der im Zusammenhang mit behavioristischen Lerntheorien entwickelten kontrastiven Erwerbtheorien wie bspw. Lado (1967) und Fries (1962, beide zit. nach Kleppin 1997: 31) gehen davon aus, dass Lernende die sprachlichen Strukturen ihrer L1 auf die L2 übertragen. Die starke Version der Kontrastiv-Hypothese besagt, dass mithilfe der kontrastiven Beschreibung von Ausgangs- und Zielsprache potentielle Schwierigkeiten vorhergesagt werden können. Die schwache Version der Hypothese erklärt Lernschwierigkeiten mittels kontrastiver Analysen (vgl. Königs 2010: 756): Positiver Transfer zwischen den Sprachen führt zu einem leichteren Erwerb der zielsprachlichen Struktur, negativer Transfer bzw. Interferenz zu einem erschweren bzw. fehlerhaften Erwerb. Kritisiert wird an dieser Hypothese, dass mit Transfer ein komplexer Prozess wie der L2-Erwerb nur teilweise erklärt werden kann. Außerdem konnten bisher keine empirischen Nachweise für vorhergesagten Transfer sprachlicher Strukturen erbracht werden (vgl. Boeckmann 2010: 169).

Der Identitätshypothese zufolge, die auf Dulay & Burt (1974, zit. nach Aguado 2010: 123) zurückgeht, sind der Erwerb einer L1 und einer L2 isomorphe Prozesse. Die Hypothese beruht auf den Annahmen der Generativen Grammatik und nativistischen Spracherwerbtheorien (vgl. Kap. 2.4.4 *Die Natur-Kultur-Debatte*) und besagt im Gegensatz zur Kontrastivhypothese, dass die L1 keinen Einfluss auf den Erwerb der

L2 hat. Stattdessen sind die Strukturen der Zielsprache maßgeblich für den Verlauf des L2-Erwerbs. Bezogen auf einzelne linguistische Phänomene konnte dies empirisch belegt werden (vgl. Huneke & Steinig 2000: 25). Gegen die Hypothese sprechen die mittlerweile durch Studien belegten Unterschiede des L1- und L2-Erwerbs, die bspw. von der unterschiedlichen kognitiven und sozialen Entwicklung der L1- und L2-Erwerbenden bedingt werden. Trotzdem kann für einige morphosyntaktische Phänomene von universellen Erwerbsequenzen ausgegangen werden, da diese bei L1- und L2-Erwerbenden gleich verlaufen (vgl. Aguado 2010: 123).

Die auf Selinker (1972) zurückgehende *Interlanguage*-Hypothese geht von der Annahme aus, dass Lerner_innen während ihres L2-Erwerbs ein für sie spezifisches Sprachsystem herausbilden. Die *Interlanguage*, die im Bereich DaF/DaZ auch als Interimsprache, Lernervarietät oder approximatives System bezeichnet wird, ist sowohl von Elementen der Ausgangssprache als auch der Zielsprache geprägt und umfasst darüber hinaus auch davon unabhängige Merkmale. Sie ist ein variables, dynamisches System, das sich mit den Hypothesen der Lernenden über die Zielsprache verändert (vgl. Kleppin 1997: 39). Dabei durchläuft sie teilweise vorhersagbare und von der Ausgangssprache unabhängige Stadien, wobei sie zunehmend zielsprachlicher und komplexer wird (vgl. Aguado 2010: 142). Mit der *Interlanguage*-Hypothese wird der L2-Erwerb als kreativer, kognitiver Prozess modelliert.

Jede dieser ‚klassischen‘ Hypothesen zum L2-Erwerb beschreibt für den Einfluss der Ausgangssprache auf den L2-Erwerbsprozess wesentliche Aspekte, und dennoch ist jede einzelne dieser Hypothesen angesichts der gegenwärtigen Erkenntnisse über die Komplexität des L2-Erwerbs nicht hinreichend als Erklärungsmodell.

Aktuell üben die Schwellen- und die Interdependenzhypothese großen Einfluss auf die fachliche Diskussion aus, die beide auf Cummins zurückgehen. Dieser formulierte auf Grundlage einer Synthese sich widersprechender empirischer Studien zum Schulerfolg bilingualer Kinder die zweistufige Schwellenhypothese:

„This [der Vergleich von bilingualen Kindern in Immersionsprogrammen und bilingualen Programmen] raises the possibility that there may be not one but two thresholds. The attainment of the first threshold would be sufficient to avoid cognitive retardation but the attainment of a second, higher, level of bilingual competence might be necessary to lead to accelerated cognitive growth.“ (Cummins 1976: 24)

Kritisiert wurde diese Hypothese aufgrund der nicht näher bestimmten Definition des Schwellenniveaus und wegen des unklaren Verhältnisses von kognitiven und sprachlichen Fähigkeiten. Cummins entwickelte seinen Ansatz weiter und stellte die Interdependenzhypothese auf. Diese formuliert einen positiven wechselseitigen Zu-

sammenhang der erreichten Sprachkompetenzen in der L1 und der L2, vorausgesetzt, dass beide Sprachen gefördert werden und eine angemessene Motivation für den Erwerb von beiden Sprachen besteht. Cummins Metapher des Eisbergs mit zwei Gipfeln verdeutlicht die Hypothese: Während wir an der Oberfläche zwei getrennte bzw. unabhängige Systeme, L1 und L2, zu sehen glauben, erblicken wir unter der Wasseroberfläche ein gemeinsames System, das als *Common Underlying Proficiency* (CUP) bezeichnet wird (vgl. Cummins 1981: 22ff.). In Cummins (2000) werden empirische Studien präsentiert, die seine Hypothesen unterstützen. So konnte bspw. ein positiver Zusammenhang von Bilingualismus und metakognitiven Fähigkeiten nachgewiesen werden. Die Annahme einer CUP und wechselseitiger Beziehungen zwischen den Sprachen bedeutet, dass nicht nur die Ausgangssprache die Zielsprache beeinflusst, sondern auch die Zielsprache die Ausgangssprache (vgl. hierzu auch Cooks Konzept der *multi-competence*; Cook 1991; 2003).

In der vorliegenden Studie wird vor dem Hintergrund der oben skizzierten Hypothesen der Fokus auf Mandarin als Ausgangssprache gelegt, um zur Klärung des Zusammenhangs zwischen Sprachähnlichkeit und UA beizutragen. Bialystok (1997: 130) zeigte, dass Aspekte der Zweitsprache, die sich strukturell von der Erstsprache unterscheiden, für Lerner_innen schwerer zu erwerben sind. Diese Schwierigkeiten zeigten sich bei allen von ihr untersuchten Lerner_innen unabhängig von ihrem AO. In diese Richtung deuten auch die Ergebnisse von Bongaerts, Mennen & van der Silk (2000) und Bongaerts, van Summeren, Planken & Schils (1997), wonach die erfolgreichsten Lerner_innen über der Zielsprache ähnelnde Ausgangssprachen verfügen.

Birdsong & Molis (2001) replizierten die Johnson- und Newport-Studie und erhielten andere Ergebnisse als Johnson und Newport: Ihre Proband_innen mit der Ausgangssprache Spanisch schneiden in der Altersgruppe der Späterwerber_innen durchschnittlich besser ab als die Testpersonen von Johnson und Newport, die die Ausgangssprachen Chinesisch und Koreanisch haben (vgl. Abb. 4 in Kap. 2.2 *Studien zum Faktor Alter*). Die Ergebnisse von McDonald (2000) weisen ebenfalls in diese Richtung, denn die von ihm untersuchten Englischerwerbenden mit der L1 Spanisch erzielen beim GJT bessere Ergebnisse als diejenigen mit der L1 Vietnamesisch. Die Gruppe der Früherwerbenden (*early acquirers* im Original) mit L1 Spanisch (AO < 5) erzielt dasselbe Ergebnis wie die ‚muttersprachliche‘ Kontrollgruppe, während die Gruppe der Früherwerbenden mit der L1 Vietnamesisch schlechter ab-

schneidet. McDonald zeigte, dass die Fehler v. a. bezüglich Flexion, Pronomen, Artikel und Subjekt-Verb-Kongruenz gemacht wurden, womit es um grammatische Merkmale geht, die sich aufgrund der verschiedenen typologischen Profile des Englischen und des Vietnamesischen unterscheiden (vgl. ebd.: 413). An dieser Stelle lohnt jedoch ein genauer Blick in die Daten, mit dem sich feststellen lässt, dass sich die Früherwerbenden mit der L1 Vietnamesisch untereinander noch einmal erheblich unterscheiden. Denn im Bereich $AO < 2$ gibt es einige, die über 90 Punkte erreichen und somit im Bereich der Ergebnisse der ‚Muttersprachler_innen‘ liegen.⁴² Die Proband_innen, die deutlich unterhalb der Ergebnisse der ‚Muttersprachler_innen‘ liegen, haben alle einen $AO = 5$. Es ist die Frage, inwiefern hier noch von ‚Frühbilingualen‘ gesprochen werden kann. Auch fehlen Informationen darüber, wie viel Sprachkontakt die Proband_innen hatten. Da die Späterwerber_innen mit der L1 Spanisch ebenfalls in den oben benannten Bereichen die meisten Fehler machen und Spanisch und Englisch sich diesbezüglich nicht so stark unterscheiden wie Vietnamesisch und Englisch, sollte mit der Interpretation dieser Ergebnisse vorsichtig umgegangen werden.

Aguado, Grotjahn & Schlak (2007: 140) halten fest:

„Nur weitere Untersuchungen können zeigen, ob auch Sprecher weniger ähnlicher Sprachen ein mit Muttersprachlern vergleichbares Sprachniveau erreichen können und wie stark der Einfluss der Erstsprache auf die Zweitsprache hinsichtlich des erreichbaren Sprachstands ist.“

Deutsch und Mandarin sind typologisch verschieden. Mandarin, eine sinotibetische Sprache, wird zu den analytischen Sprachen gezählt, während Deutsch als indoeuropäische Sprache zu den synthetischen Sprachen gehört (Zhao 2010: 627). Mandarin ist eine flexionslose Sprache. Im Unterschied zum Deutschen gibt es keine Präfixe oder Suffixe, die Tempus, Genus, Kasus oder Numerus angeben. Satzgliedfunktionen und syntaktische Beziehungen werden mit lexikalischen Mitteln, der Wortstellung und semantischem Kontext realisiert (vgl. ebd.: 629). Zhao leitet daraus ab, dass „Wortstellungsfreiheit, die morphologischen Markierungen sowie die Verwendung von Konjunktionen Lernprobleme für chinesische Deutschlerner sein können.“

⁴² Wie viele es genau sind, ist nicht ersichtlich, weil die Werte in einem Streudiagramm wiedergegeben werden und es keine tabellarischen Angaben gibt (vgl. ebd.: 410).

(ebd.: 631). Die Sprachen unterscheiden sich strukturell enorm, was Zhao zufolge eine Rückwirkung auf komparatistische Studien hat:

„Obwohl diese beiden Sprachen formal identische Spracheinheiten wie Morphem, Wort, Wortgruppe, Satz und Text auf hierarchischen Ebenen aufweisen, werden unter den gleichen Bezeichnungen häufig ganz unterschiedliche Sprachkomponenten bzw. Darstellungsweisen verstanden. Aufgrund der oben genannten Problematik betreffen die deutsch-chinesischen kontrastiven Arbeiten überwiegend Einzelprobleme und erzielen punktuelle Ergebnisse, eine systematische Gesamtdarstellung ist aber bis heute nicht vorhanden.“ (ebd.)

Unterschiede bestehen jedoch nicht nur im lexikologischen, morphologischen und syntaktischen Bereich, denn im Gebiet der Phonetik – die bezüglich des Einfluss des Faktors Alter häufig diskutiert wird – ist ebenfalls eine beträchtliche Differenz vorhanden.⁴³ Mandarin ist eine Tonsprache, d. h. dass die unterschiedlichen Stimmverläufe der Silben bedeutungsunterscheidend sein können. Ein weiterer hauptsächlicher phonetischer Unterschied besteht darin, dass das Chinesische eine Silbensprache ist. Zwischen den Silben wird eine Junktur artikuliert, die die Silbengrenze kennzeichnet, während im Deutschen mehrere Silben ohne Junktur vorkommen können. Darüber hinaus gibt es weitere artikulatorische Herausforderungen für Lerner_innen mit chinesischen Ausgangssprachen zu bewältigen, bspw. die Differenzierung von [l] und [r] bzw. [ʀ] und das Vorkommen von Konsonantenkombinationen sowie Konsonanten in Endstellung (vgl. ebd.: 629).

Ein weiterer Grund, den Fokus auf Mandarin zu legen, besteht darin, dass asiatische Sprachen im Vergleich zu europäischen Sprachen in der L2-Erwerbsforschung unterrepräsentiert sind. Dabei steigt die Zahl der Deutschlerner_innen in China rasant: Laut den statistischen Erhebungen des Netzwerks Deutsch, einer Initiative des Auswärtigen Amts, des Deutschen Akademischen Austauschdiensts, des Goethe Instituts und der Zentralstelle für Auslandsschulwesen, lernten 2005 4.990 chinesische Studierende Deutsch, während es 2010 bereits 35.000 waren (vgl. URL 7; 2010: 5). 2015 wird ein weiterer Anstieg um knapp ein Viertel auf 44.945 Lernende verzeichnet (vgl. URL 8; 2015: 10). Gleichzeitig veränderte sich die chinesische Hochschullandschaft rasant und zeichnet sich durch Dynamik und Leistungsstärke aus (vgl. Steinmüller 2010). Deutsch wird in China in erster Linie an den Hochschulen gelernt.

⁴³ Ein Blick auf Singletons Zusammenstellung (vgl. Tab. 1, Kap. 2.1 *Kritische Periode(n)*) zeigt, dass die meisten Untersuchungen sich auf die Phonetik/Phonologie beziehen.

Aber auch im schulischen Bereich ist die Zahl der Deutschlerner_innen steigend: 2009 zählte die Zentralstelle für das Auslandsschulwesen (ZfA) 2500 Deutschlerner_innen an 28 Fremdsprachenschwerpunktschulen, die das Fach Deutsch anbieten (vgl. Hernig 2010: 1640). Die 25.564 Studierenden aus China, die in Deutschland studieren, machen 12,5 % der ausländischen Studierenden (Bildungsausländer_innen) in Deutschland insgesamt aus (vgl. URL 9; 2014: o. S.). Ein Blick auf die mittelfristige Entwicklung zeigt, dass sich die Zahl in den letzten 15 Jahren verfünffacht hat, denn 1999 waren es 5.054 Studierende aus China (vgl. URL 10; 2014: o. S.).

Der fremdsprachliche Unterricht in China hat sich in den letzten Jahrzehnten gewandelt. Steinmüller zufolge unterlagen die Ziele des Fremdsprachenunterrichts früher dem „Primat der Nützlichkeit“. Gleichzeitig wurde befürchtet, dass fremde Sprachen auch fremdkulturelle Einflüsse bedeuten, die eine eventuell schädigende Wirkung haben können (vgl. Steinmüller 2007: 17). Heute werden die Ziele in erster Linie funktional und instrumentell definiert:

„[...] im Zusammenhang von Technologietransfer und Qualifikation für internationale Kooperationen treffen sich die wirtschaftlichen Interessen des Staates mit den Vorstellungen von einem modernen Fremdsprachenunterricht [...]“ (Steinmüller 2013: 25)

Aktuell gewinnt ihm zufolge mit den wachsenden Kommunikationsbedürfnissen einer globalisierten Welt die interkulturelle Kommunikations- und Handlungsfähigkeit als weitere Zweckbestimmung des Fremdsprachenunterrichts an Bedeutung (vgl. ebd.). Die Situation im Bereich Deutsch als Fremdsprache an den Universitäten ist vielfältig: An den renommierten Deutschabteilungen der landesweit bekannten Universitäten, bspw. in Beijing, Shanghai, Chongqing oder Xian, findet ‚moderner‘ kommunikativ-orientierter DaF-Unterricht statt mit Lehrkräften, die i. d. R. in Deutschland studiert haben, sowie deutschen Lektor_innen. Dagegen ist die Situation an den kleinen, später gegründeten Deutschabteilungen vieler Universitäten im ländlichen Raum anders. Hier gibt es viele Lehrkräfte, die direkt nach dem Abschluss ihres vierjährigen Bachelorstudiums ohne weitere pädagogische Ausbildung zu unterrichten begonnen haben, bisher keinen längeren Deutschlandaufenthalt absolvieren konnten und de-

ren Lehrmethoden sich eher an der traditionellen Grammatik-Übersetzungs-Methode orientieren.⁴⁴

⁴⁴ Diese Informationen entstammen einer E-Mail-Korrespondenz mit Johanna Varuzza, die als Bosch-Lektorin in Chongqing (China) war.

3. Die Studie

Im vorangehenden theoretischen Teil dieser Arbeit wurde der Forschungsstand zum Faktor Alter durch eine Einführung in die Diskussion der CPH und die Darstellung einschlägiger Studien präsentiert. Zur Beantwortung der Frage, was Alter ist und was ältere von jüngeren Lerner_innen unterscheidet, wurden neurophysiologische Grundlagen der Gehirnreifungsprozesse diskutiert und die individuellen Variablen Motivation und Sprachlerneignung dargestellt. Eine kurze Einführung in die Natur-Kultur-Debatte, bei der nativistische und emergentistische theoretische Ansätze vorgestellt wurden, rundete dieses Unterkapitel ab, um den wissenschaftsparadigmatischen Zusammenhang aufzuzeigen, in den die Debatte um den Faktor Alter eingebettet ist. Die sich daran anschließende ausführliche Auseinandersetzung mit dem Begriff ‚Muttersprache‘ unterstrich die Relevanz einer Vergleichsgruppe, die nicht allein aus monolingualen ‚Muttersprachler_innen‘ besteht, sondern aus Personen mit doppeltem Erstspracherwerb. Zum Abschluss des theoretischen Teils dieser Arbeit wurde der Einfluss der Ausgangssprache im Zusammenhang von Veröffentlichungen zum Faktor Alter dargestellt.

Im nun folgenden empirischen Teil der Arbeit wird zunächst der Forschungsansatz aufgezeigt und beschrieben, inwiefern auf Daten aus dem Forschungsprojekt „Alter und Spracherwerb“ von Aguado, Grotjahn und Schlak (2005, 2007) zurückgegriffen werden kann. Nach der Darstellung der Ziele und Forschungsfragen werden das Forschungsdesign und die Testpersonen beschrieben. Bei den verwendeten Instrumenten wird entweder ihre Herkunft dokumentiert oder gezeigt, wie sie entwickelt wurden. Die Dokumentation findet ihren Abschluss in der Schilderung der Rahmenbedingungen der Studie und der Proband_innen-Suche.

3.1 Forschungsansatz

Das im Folgenden skizzierte Forschungsprojekt ermittelt mit einem Mehr-Methoden-Design, das in zwei Phasen abläuft, die Profile erfolgreicher erwachsener DaF-Lerner_innen. In diesem Rahmen werden in einer ersten Phase die Instrumente *Grammatical Judgement Test* (GJT), zeitlimitierte und nicht zeitlimitierte C-Tests (S-C-Tests und C-Tests), ein Phonetik-Test, Skalen zur Motivation, Introversion und

Fremdsprachenverwendungsangst (TESTATT-Skalen) sowie ein Fragebogen zur Sprachlernbiografie eingesetzt. Diese Instrumente werden in einer zweiten Phase ergänzt durch Sprachlerneignungstests (LLAMA-Tests) und einen zweiten Fragebogen zur Motivation (*L2 Motivational Self System*) und weitere biografische Hintergründe ausgewählter Proband_innen (vgl. die Übersichtsdarstellung in Abb. 8).

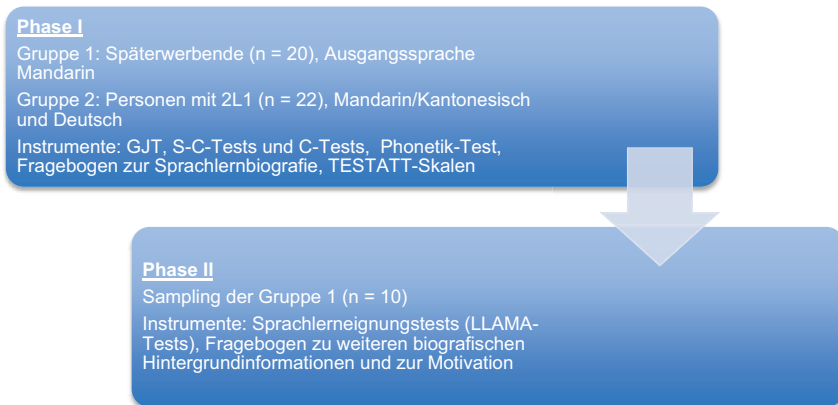


Abb. 8 Übersichtsdarstellung Phase I und II.

Getestet werden 20 Späterwerber_innen, die im Alter von 16 Jahren oder älter begonnen haben, Deutsch zu lernen ($AO \geq 16$), seit mindestens fünf Jahren in Deutschland leben ($LOR \geq 5$) und über einen akademischen Bildungshintergrund verfügen. Die Vergleichsgruppe besteht aus 22 Personen mit 2L1 ($AO = 0-3$), die mit den Sprachen Mandarin bzw. Kantonesisch und Deutsch aufgewachsen sind.

Die Erhebung einer mehrsprachigen Vergleichsgruppe geschieht vor dem Hintergrund der Diskussion um den Begriff ‚muttersprachliche Kompetenz‘. Davies (2011: 296) kritisiert den Begriff v. a. im Hinblick auf Sprachprüfungen und Testformate. Hulstijn (2011; 2012) zeigt, inwiefern es sich bei der ‚muttersprachlichen Sprachkompetenz‘ um eine Idealisierung handelt, und unterscheidet deswegen in *Basic Language Cognition* (BLC) und *Higher Language Cognition* (HLC) (vgl. die ausführliche Darstellung der Diskussion in Kap. 2.5 ‚Muttersprachliche Kompetenz‘). Der Vergleich der Kompetenzprofile von Späterwerber_innen und Personen mit 2L1 ist ein Vergleich von mehrsprachigen Gehirnen mit mehrsprachigen Gehirnen. Er ist dementsprechend angemessener und aussagekräftiger als der Vergleich mit monolingualen ‚Muttersprach-

ler_innen' (vgl. Muñoz & Singleton 2011: 4–5; Grosjean 1989). Auch DeKeyser (2012) zufolge ist der Vergleich von L2-Lerner_innen mit Personen mit doppeltem Erstspracherwerb zu bevorzugen:

„Using bilinguals with a very young age of acquisition avoids both problems of comparing either with an idealized non-empirical norm or comparing with empirical data from monolinguals (in themselves perhaps already an idealized norm from the point of view of bilingualism research).“ (DeKeyser 2012: 261)

In diesem Zusammenhang muss angemerkt werden, dass bei der Erhebung der ‚Muttersprachler_innen‘ nicht abgefragt wurde, inwiefern die Personen Dialektsprecher_innen und somit ebenfalls bilingual bzw. mehrsprachig sein könnten. Den Testleiter_innen zufolge wurde darauf geachtet, dass die Proband_innen dieser Gruppe einheitlich entweder Hochdeutsch oder lediglich mit einem schwachen ruhrdeutschen Einschlag sprechen. Die Diskussion um Bilingualität im Zusammenhang mit Dialekt- und Sprachgrenzen wird im Rahmen dieser Arbeit nicht geführt.

3.2 Vorarbeiten

Das Dissertationsprojekt schließt an das Forschungsprojekt „Erwerbsalter und Sprachlernerfolg“ von Aguado, Grotjahn & Schlak (2005, 2007) an. Ihr Projekt besteht aus zwei Teilstudien. In der einen wird angestrebt, außergewöhnlich gute Späterwerber_innen mit verschiedenen Ausgangssprachen zu finden, die trotz ihres späten Lernbeginns ($AO \geq 16$) ‚muttersprachliche Kompetenz‘ bezogen auf diverse sprachliche Fertigkeiten erreichen. In der zweiten Teilstudie geht es um die Frage, wie häufig ‚muttersprachliche Kompetenz‘ von Lerner_innen mit verschiedenen Ausgangssprachen erreicht werden kann (anvisiert waren jeweils 20 Testpersonen mit den L1 Arabisch, Englisch, Polnisch, Spanisch und Türkisch). Als Vergleichsgruppe wurden zunächst Muttersprachler_innen getestet. Diese Vergleichsdaten wurden später ergänzt durch die Testung von Personen mit 2L1.

Die Bedingungen für die Testpersonen der vorliegenden Studie entsprechen denen des Projekts „Erwerbsalter und Sprachlernerfolg“ ($AO \geq 16$, $LOR \geq 5$ und akademischer Bildungshintergrund). In Phase I wird auf das Standardinstrumentarium zurückgegriffen, das im Projekt von Aguado, Grotjahn und Schlak verwendet und teilweise dafür entwickelt wurde (vgl. dazu die Ausführungen in Kapitel 3.4.2 *Instrumente*).

Die im Rahmen des Projekts „Erwerbsalter und Sprachlernerfolg“ erhobenen 232 Datensätze, für die Werte des GJT, der S-C-Tests und C-Tests, der TESTATT-Skalen sowie der standardisierten Interviews zum Sprachenprofil vorliegen, stehen für die Nutzung in dem hier beschriebenen Projekt zu Verfügung. Diese Datensätze wurden bisher noch nicht analysiert und veröffentlicht. Die 232 Datensätzen reduzieren sich um 8, bei denen keine Daten zum biografischen Hintergrund resp. der Sprachlernbiografie vorliegen, so dass somit keine Zuordnung vorgenommen werden kann zwischen ‚Muttersprachler_innen‘, Personen mit 2L1 und Späterwerber_innen. 45 Datensätze sind von Personen mit diversen Ausgangssprachen, deren AO = 4–15 ist. Sie haben keine Relevanz für die Beantwortung der Forschungsfragen, werden aber im ersten Schritt der Datensichtung und -aufbereitung miteinbezogen.

Um den Datenbestand für vergleichende und ergänzende Analysen nutzen zu können, erfolgt zunächst eine deskriptive Auswertung (vgl. Tab. 3 zur Verteilung der L1). Von besonderem Interesse sind hierbei die Datensätze der 69 L2-Lerner_innen mit AO \geq 16, für die Daten aus den GJTs, den S-C-Tests, den C-Tests, den TESTATT-Skalen und den standardisierten Interviews vorliegen:

Tab. 3: Späterwerber_innen und ihre Ausgangssprachen. Daten aus dem Forschungsprojekt von Aguado, Grotjahn & Schlak (2005, 2007).

n	L1	n	L1
21	Mandarin	2	Thai
10	Russisch	1	Englisch
5	Arabisch	1	Griechisch
5	Japanisch	1	Kiswahili
5	Spanisch	1	Mongolisch
3	Aserbaidzhanisch	1	Persisch
3	Vietnamesisch	1	Polnisch
2	Armenisch	1	Portugiesisch
2	Dari	1	Ukrainisch
2	Indonesisch	1	Usbekisch

Als Vergleichsgruppe wurden Daten von deutschen Erstsprachler_innen (n = 80) gesammelt sowie von 28 Personen mit 2L1 (AO = 0–3), die neben Deutsch mit verschiedenen Sprachen aufgewachsen sind (vgl. Tab. 4).⁴⁵

Tab 4: Personen mit 2L1. Darstellung der Sprachen der Testpersonen. Nutzung der Daten aus dem Forschungsprojekt von Aguado, Grotjahn & Schlak (2005, 2007).

n	2L1 (Deutsch +)	n	2L1 (Deutsch+)
5	Englisch	1	Italienisch
4	Polnisch	1	Koreanisch
4	Türkisch	1	Kurdisch
3	Arabisch	1	Norwegisch
2	Portugiesisch	1	Rumänisch
2	Spanisch	1	Russisch
1	Französisch	1	Vietnamesisch

3.3 Ziele und Forschungsfragen

Die Zielstellung der vorliegenden Studie ist die Identifizierung der Faktoren, die individuelle Leistungsunterschiede beim L2-Lernen von Späterwerber_innen (Erwerbsbeginn ≥ 16 Jahre) beeinflussen. Dabei wird die Sprachkompetenz von Späterwerber_innen (AO ≥ 16) gemessen und mit der von Personen mit doppeltem Erstspracherwerb (AO = 0–3 Jahre) und monolingualen Erstsprachler_innen verglichen. Auf der Basis dieser Leistungsmessung werden die Profile erfolgreicher Späterwerber_innen unter Berücksichtigung von kognitiven, affektiv-motivationalen und sozio-kulturellen Einflussgrößen spezifiziert.

Bei der Darstellung des Forschungsansatzes in Kap. 3.1 wurde mit Bezug auf Grosjean (1989) darauf verwiesen, dass ein Vergleich mehrsprachiger Gehirne miteinander angemessener und aussagekräftiger ist, weil ein_e Bilinguale_r nicht das gleiche

⁴⁵ Die Daten der Personen mit 2L1 wurden im Rahmen des Projekts „Erwerbsalter und Sprachlernerfolg aus bilingualer Perspektive“ unter Prof. Dr. Torsten Schlak (†) am Fachgebiet Deutsch als Fremdsprache der TU Berlin erhoben.

ist wie zwei Monolinguale in einer Person. Aus diesem Grund werden die Daten der monolingualen Erstsprachler_innen aus dem Projekt von Aguado, Grotjahn und Schlak (2005; 2007) zwar ausgewertet und zu Vergleichszwecken herangezogen, sie sind aber nicht Teil der Datenbasis zur Beantwortung der Forschungsfragen.

Forschungsfragen:

1. Welches Lerner_innenprofil hinsichtlich Motivation und Persönlichkeit weisen die erfolgreichsten, die durchschnittlichen und die weniger erfolgreichen Späterwerber_innen (AO \geq 16) auf?
2. Wie häufig kommt eine Sprachkompetenz in der Gruppe der Späterwerber_innen vor, die der Sprachkompetenz der Personen mit doppeltem Erstspracherwerb entspricht?
3. Welchen Einfluss hat die Sprachähnlichkeit auf die erreichte Sprachkompetenz?

Die Beantwortung der Fragen zielt darauf ab, einen entscheidenden Beitrag zur Frage der Erfolgsbedingungen für L2-Späterwerber_innen zu leisten und Profile erfolgreicher Lerner_innen zu erstellen. Besonders der Erkenntnisgewinn bezüglich des Vergleichs der sprachlichen Kompetenzen von Personen mit 2L1 und Späterwerber_innen wird grundlegend für weitere Forschungen sein, da er eine Innovation darstellt.

3.4 Forschungsdesign

3.4.1 Testpersonen

Die Testpersonen unterteilen sich in drei Gruppen: Späterwerber_innen (AO \geq 16), Personen mit doppeltem Erstspracherwerb und deutsche ‚Muttersprachler_innen‘. Alle Testpersonen verfügen über einen akademischen Bildungshintergrund, knapp die Hälfte der Späterwerber_innen (n = 34) lebt schon seit mindestens fünf Jahren in Deutschland (LOR \geq 5). Abb. 9 zeigt eine Gesamtdarstellung der untersuchten Gruppen und weist die Erhebungsorte und Zeiträume aus.

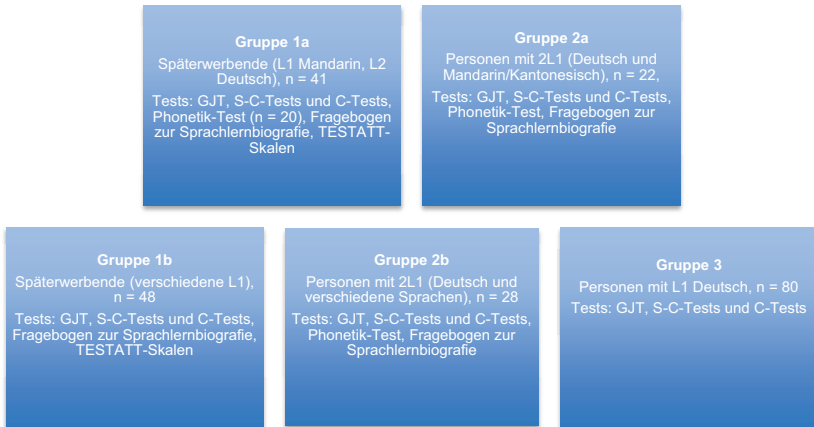


Abb. 9: Gesamtdarstellung der untersuchten Gruppen und durchgeführten Tests der Phase I. 20 Fälle der Gruppe 1a und alle Fälle der Gruppe 2a wurden im Rahmen des Dissertationsprojekts erhoben. Gruppe 2b wurde von 2010 bis 2011 an der TU Berlin erhoben, 21 Fälle der Gruppe 1a und die Gruppen 1b und 3 wurden von 2006 bis 2011 an der Ruhr Universität Bochum, dem CDC Dortmund, der Universität Kassel und der TU Berlin erhoben.

Statistische Angaben zu den Testpersonen

63 Frauen und 26 Männer stellten sich als Testpersonen in der Gruppe der Späterwerber_innen zu Verfügung. Das durchschnittliche Alter zum Testzeitpunkt war 29,03 Jahre (Min: 20 und Max: 59). Das AO lag im Durchschnitt bei 21,87 Jahren (Min: 16 und Max: 42). Die LOR lag im Durchschnitt bei 56,9 Monaten (Min: 1 und Max: 240). In der Gruppe der Personen mit 2L1 waren 35 Frauen und 15 Männer. Das Alter zum Testzeitpunkt lag im Durchschnitt bei 23,42 Jahren (Min: 15 und Max: 47). 71 Frauen und 9 Männer sind in der Gruppe der ‚Muttersprachler_innen‘ getestet worden. Das durchschnittliche Alter bei Testung war 26,20 Jahre (Min: 21 und Max: 48).

Im Rahmen der Diskussion um die CPH ist ein $AO \geq 16$ eine konservative Grenze, die in der Regel für den Erwerb semantischen und pragmatischen Wissens angesetzt wird (vgl. Tab. 1, Kap. 2.1 *Kritische Periode(n)*). Bei aller Uneinheitlichkeit und Uneinigkeit in der Forschung besteht Konsens über Folgendes: Wenn es kritische oder sensible Perioden gibt, haben sie bis zum Alter von 16 Jahren ihren *Offset* erreicht. Bezogen auf neurophysiologische Erkenntnisse ist 16 eine sehr frühe Grenze, denn verschiedene Untersuchungen deuten daraufhin, dass ein menschliches Ge-

hirn erst ab 20 oder 21 Jahren als erwachsen gelten kann (vgl. bspw. Keulers, Stiers & Jolles 2011).

AO kann verschieden operationalisiert werden. In vielen Studien zum Faktor Alter wird darunter das *Age of Arrival (AOA)* oder *Age of Immersion (AOI)* verstanden. Flege (1999) argumentiert für das *Age of Arrival*, das er seiner Untersuchung zugrunde legt:

„The participants were selected on the basis of AOA rather than the age at which they first began to study English at school in Korea. This is because, of the two variables, AOA is the more potent predictor of L2 performance (e.g., Johnson & Newport, 1989).“ (ebd.: 79)

Und auch Singleton & Muñoz (2011: 413) weisen daraufhin, dass es besser sei, getrennte Analysen für den Lernbeginn in der Heimat und den Beginn der Immersion vorzunehmen.

In der vorliegenden Untersuchung geht es teilweise um Testpersonen, die einen BA in einem germanistischen Fach in ihrem Heimatland erworben haben. Das ist eine Einflussgröße, die nicht vernachlässigt werden darf. Deswegen ist hier beim AO der Lernbeginn gemeint, unabhängig davon, ob er im Heimatland oder einem deutschsprachigen Land stattgefunden hat. Darüber hinaus werden Analysen vorgenommen, die AO noch weiter ausdifferenzieren nach dem Lernbeginn in der Heimat und in Deutschland.

Das Interesse der hier vorliegenden Untersuchung richtet sich nicht auf das Lerntempo (*rate of acquisition*), sondern auf den erreichten Sprachstand (UA) der Testpersonen (vgl. Kap. 2.1 *Kritische Periode(n)*). In der Diskussion gibt es verschiedene Positionen dazu, wann dieses UA erreicht ist. In Untersuchungen zum Faktor Alter wird das UA in der Regel durch LOR zwischen 5 und 10 Jahren operationalisiert (vgl. Abrahamsson & Hyltenstam 2009: 252; Birdsong 2009; Muñoz & Singleton 2011: 17).

Oyama (1976) und Johnson & Newport (1989) setzen das UA nach fünf Jahren an. Bohn & Flege (1990) bestätigen diesen Wert und stellen ein Abflachen der L2-Lernkurve nach fünf Jahren fest (zitiert nach Bohn 1998: 6). DeKeyser (2000) weitet die Zeitspanne auf zehn Jahre aus; erst dann habe die LOR keinen signifikanten Einfluss mehr. Wichtig ist, sich darüber bewusst zu sein, dass LOR eine sehr unspezifische Kategorie ist, was den tatsächlichen zweitsprachlichen Input anbelangt, den

eine Person erfahren hat. Sowohl dessen Qualität als auch Quantität können immens differieren:

„Allerdings gibt es bislang keine einfache Antwort auf die Frage, wie qualitativ und quantitativ unterschiedlicher L2-Input zu gewichten ist, wenn ein L2-Lerner z. B. am Arbeitsplatz verschiedenen Dialekten der L2 ausgesetzt ist und zu Hause fremdsprachlich akzentuierten Input erhält.“ (Bohn 1998: 5)

Auch Moyer (2004) betont, dass Input nicht nur eine quantitative Frage ist, sondern eine entscheidendere Rolle spielt „why learners use the target language as they do, that is, for what functions and to what social and psychological effects“ (ebd.: 168). Im Rahmen des Projekts liegen 34 Datensätze von Späterwerbenden vor mit einer LOR ≥ 5 .

Fast die Hälfte der Gruppe der Späterwerber_innen spricht die Ausgangssprache Mandarin (n = 41). Auch in der Gruppe 2 liegt der Anteil der neben Deutsch mit Mandarin bzw. Kantonesisch aufgewachsenen Personen (n = 22) bei fast 50 %. Der Fokus auf diese Sprachen begründet sich durch die typologische Verschiedenheit zwischen Mandarin bzw. Kantonesisch und Deutsch (vgl. Kap. 2.6 *Einfluss der Ausgangssprache*). Gleichzeitig spielen pragmatische Überlegungen eine Rolle, da bereits 21 Datensätze von Späterwerber_innen mit der Ausgangssprache Mandarin vorlagen, die am besten zu einer Gruppenstärke ergänzt werden konnten, die statistisch gesehen zu relevanten Aussagen führen kann.

Neben den ‚Muttersprachler_innen‘ wurde auch eine Gruppe von Personen mit doppeltem (oder bilingualem) Erstspracherwerb (2L1) als Vergleichsgruppe erhoben (AO = 0–3 Jahre). Meisel (2007, 2011) argumentiert dafür, dass der simultane Erwerb der Mehrsprachigkeit als eine Form des doppelten Erstspracherwerbs verstanden werden sollte:

„1) Bilingual aufwachsende Kinder sind in der Lage, das internalisierte Wissen von ihren Sprachen von früh an zu trennen. 2) Sie durchlaufen dieselben grammatischen Erwerbssequenzen wie monolinguale Kinder, die die gleiche Sprache erwerben. 3) Sie erreichen eine grammatische Kompetenz, die sich qualitativ nicht von der vergleichbarer Monolingualer unterscheidet.“ (Meisel 2007: 9)

Dagegen gibt es die Position, die bspw. Herschensohn (2007) referiert, nach der Personen mit 2L1 nicht den Monolingualen entsprechen, weil sie andere Zeitfenster brauchen, um dieselben Meilensteile in der grammatikalischen Entwicklung zu erreichen, und sich ihr grammatischer Endstand unterscheidet (vgl. ebd.: 230–231).

Dass nur beim Erwerb von zwei Erstsprachen im Alter von null bis drei Jahren von doppeltem Erstspracherwerb gesprochen werden kann, zeigen Untersuchungen zur syntaktischen und semantischen Wohlgeformtheit von Sätzen. Wurde die L2 nach dem dritten Lebensjahr erworben, zeigen sich die EKPs verzögert oder in ihrer Amplitude stark verändert. Ein extrem hohes Leistungsniveau in der später gelernten Sprache gleicht diese Unterschiede allerdings aus (vgl. Wartenburger 2012: 183).

3.4.2. Instrumente

Nach der ausführlichen Beschreibung des Forschungsdesigns, der Forschungsfragen, der Vorarbeiten und der Testpersonen werden im Folgenden die Instrumente vorgestellt, die in dem Mehrmethoden-Design der vorliegenden Studie Verwendung finden. In der ersten quantitativen Phase kommen GJTs, S-C-Tests, C-Tests, die TESTATT-Skalen, ein Fragebogen zur Sprachlernbiographie und ein Phonetik-Test zum Einsatz. Nach der Auswertung und Analyse der Ergebnisse wird in einer zweiten Phase eine Vertiefung mit einzelnen Testpersonen angestrebt, um die in der ersten Phase gefundenen Zusammenhänge besser deuten zu können. In der zweiten Phase werden Sprachlerneignugstests und ein weiterer Fragebogen zu biografischen Hintergrundinformationen und zur Motivation eingesetzt.

3.4.2.1 Grammatikalitätssurteilstests

Grammatikalitätssurteilstests (*Grammatical Judgement Tests*) kommen allgemein in der Fremdsprachenforschung und im Besonderen in Studien zum Faktor Alter häufig vor, um sprachliches Wissen zu testen. Typischerweise besteht ein GJT aus einer bestimmten Anzahl von grammatisch richtigen und falschen Sätzen. Die Testpersonen werden dazu aufgefordert zu entscheiden, ob die Sätze korrekt oder nicht korrekt sind und ob sie sich bei ihrem Urteil sicher sind. In manchen Versionen sollen sie darüber hinaus bei den falschen Sätzen angeben, was falsch ist, eine korrekte Version schreiben und/oder die grammatische Regel benennen, die von dem Satz nicht erfüllt wird (vgl. Ellis 1991: 162).

Der hier verwendete GJT ist Aldona Sopatas Dissertation „Universalgrammatik und Fremdsprachendidaktik“ (2004) entnommen, in der sie den Erwerb der Null-Expletiva und der Pronominabindung in Infinitivkonstruktionen von DaF-Lerner_innen mit der

Ausgangssprache Polnisch untersucht.⁴⁶ Die Untersuchung dieser grammatischen Phänomene begründet sich darin, dass sich Polnisch und Deutsch bezüglich dieser Parameter unterscheiden und entsprechende Lernschwierigkeiten bei polnischen Deutschlerner_innen zu erwarten sind. Sopatas Testpersonen werden in vier Gruppen eingeteilt, die sich danach unterscheiden, wie lange sie bereits Deutsch lernen.⁴⁷ Die Kenntnisse der Proband_innen über Pronomina, Infinitivkonstruktionen, Pronominabindung und Expletivumgebrauch werden durch verschiedene Tests gemessen, darunter Ersetzungstests, *Multiple Choice*, gelenkte Produktion und GJT. Die Auswertung zeigt, dass bereits Lerner_innen der zweiten Gruppe die Pronomina zum größten Teil fehlerfrei interpretieren. Dagegen stellt das Erlernen des Expletivumsgebrauchs auch noch in der Gruppe 4, die am längsten Deutsch gelernt hat, eine Schwierigkeit dar (vgl. Sopata 2004: 141f). Sopata zufolge bestätigen diese Ergebnisse die Hypothese, dass die UG L2-Lerner_innen zugänglich sei, wenn es sich dabei um unparametrisierte UG-Prinzipien handelt (vgl. ebd.: 142).

Für das Forschungsprojekt „Alter und Sprachlernerfolg“ wurde dieser Test gewählt, weil UG-basierte Phänomene der Kerngrammatik getestet werden sollten. Die ursprünglich geplante zusätzliche Testung peripherer – d. h. nicht mit Hilfe der UG beschreibbarer – sprachspezifischer Probleme des Deutschen wie bspw. Genuszuweisung und Verbflexion unterblieb aus pragmatischen Gründen. Da die Bearbeitungszeit für die Testbatterie der ersten Phase bereits zwischen 70 und 100 Minuten dauert, sollte das Testset nicht noch weiter ausgedehnt werden, um einer möglichen Proband_innen-Mortalität entgegenzuwirken. Von besonderem Interesse für das Forschungsprojekt von Aguado, Grotjahn und Schlak ist die Ausrichtung des GJT auf den Null-Subjekt-Parameter, mit dem die Verwendung von Expletiva einhergeht: „Im Unterschied zu Null-Subjekt-Sprachen verfügen das Deutsche und alle weiteren Nicht-Null-Subjekt-Sprachen über solche Expletiva.“ (Aguado, Grotjahn & Schlak 2005: 282). Sopatas Untersuchungsergebnisse zeigen, dass selbst fortgeschrittenen Lerner_innen der Gebrauch der Expletiva schwerfällt (s. o.). Aguado, Grotjahn und Schlak merken an, dass bisher nicht geklärt ist, ob diese Schwierigkeit auch für Ler-

⁴⁶ Alle verwendeten Tests, Skalen und Fragebögen, die in Papierform eingesetzt wurden, finden sich im Anhang.

⁴⁷ Gruppe 1 hat i. d. R. seit 7 Jahren Deutsch an der Schule gelernt (n = 49). Gruppe 2 sind Germanistik-Studierende des ersten Jahres (n = 44), Gruppe 3 des zweiten Jahres (n = 32) und Gruppe 4 des dritten Jahres (n = 39).

ner_innen mit Ausgangssprachen gelten, die nicht zu den Nullsubjekt-Sprachen zählen (ebd.: 282).

Im Hinblick auf Mandarin und Kantonesisch ist es sinnvoll, diesen Parameter zu testen, da Sätze auch dann grammatisch wohlgeformt sind, wenn sie kein Subjekt haben. Es handelt sich also um Sprachen, die sich bezüglich des Null-Subjekt-Parameters vom Deutschen unterscheiden. Item 1, das auf die Adjektivkongruenz abzielt, ist ebenfalls relevant für den Kontrast von Mandarin und Deutsch, da Mandarin keine Flexion kennt.

Der GJT besteht aus 32 Test-Items, wovon 20 für das Null-Expletiva-Phänomen relevant sind und von Sopata in vier Untergruppen unterteilt werden: a) einfache und b) zusammengesetzte Sätze, die den grammatischen und ungrammatischen Gebrauch von referentiellen Subjekten zeigen; c) einfache Sätze, die den grammatischen und ungrammatischen Gebrauch des Expletivums präsentieren; sowie d) zusammengesetzte Sätze, in denen das Expletivum in einer passivischen Konstruktion grammatisch und ungrammatisch ist (vgl. 104f). Für eine bessere Übersicht habe ich die Items in Tab. 5 den Gruppen zugewiesen, wobei die im Zusammenhang der Forschungsfragen von Sopata als Distraktoren dienenden Items einem grammatischen Phänomen zugeordnet werden bzw. in die Rubrik korrekte Sätze fallen.⁴⁸

Tab. 5: Zuordnung der GJT-Items in Gruppen bzw. Themen.

Gruppe/Phänomen	Item-Nr.
a) einfache Sätze mit referentiellem Subjekt	2+19, 6+25,
b) zusammengesetzte Sätze mit referentiellem Subjekt	3+9, 8+22,
c) einfache Sätze mit Expletivum	7+27, 10+17, 14+21
d) zusammengesetzte Sätze mit Expletivum im Passiv	11+15, 12+18, 24+29,
d') wie d), aber aktivische Konstruktion	4+16, 28+32,
Adjektivkongruenz	1
Tempus bzw. Infinitivkonstruktion	31
korrekte Sätze	5, 13, 20, 23, 26, 30

⁴⁸ Nach der hier vorgenommenen Einstufung sind es acht Distraktoren und nicht zwölf. Da es bei Sopata keine vollständige Übersicht über ihre Zuordnungen gibt, lässt sich nicht herausfinden, an welchen ItemItems bzw. Zuordnungen es liegt.

24 Items sind gepaart und wiederholen sich jeweils in einer korrekten und nicht-korrekten Version, die Tabelle gibt dies durch die Pluszeichen zwischen den Zahlen wieder. Die Präsentation von solchen Satzpaaren wird in der Regel gewählt, um herauszufinden, welchen der Sätze die Proband_innen bevorzugen bzw. als wohlgeformter empfinden (vgl. Ellis 1991: 171).⁴⁹ Sopata präsentiert ihre GJT-Items randomisiert, um nicht auf die Ermittlung einer Vorzugsentscheidung (*preference choice*) der Proband_innen abzielen. Dies führt zu der Frage, warum die Items gepaart und randomisiert sind und welcher Effekt damit erzielt wird. Möglicherweise ist in Betracht zu ziehen, dass dieses Vorgehen die eigentliche Item-Zahl verringert.

Die Proband_innen in der Versuchsanordnung bei Sopata entscheiden in einem ersten Schritt, ob sie den jeweiligen Satz grammatisch finden oder nicht. Sie geben dann auf einer sechsstufigen Skala an, ob sie sich bei diesem Urteil sicher sind. Darüberhinaus werden sie dazu aufgefordert, eine korrigierte Version des Satzes zu notieren. Bei Aguado, Grotjahn und Schlak und auch in diesem Projekt entfällt dies und Korrektheit und Sicherheit des Urteils werden in einem Schritt erhoben (vgl. Tab. 6).

Tab. 6: Bewertungsraster für den GJT.

	korrekt		nicht korrekt	
	sicher	unsicher	sicher	unsicher
Bsp. I: Ich verstehe nicht.				
Bsp. II: Ich verstehen nicht.				

Die Testpersonen sollen die Grammatikalitätsurteile zügig fällen, jeden Satz nur einmal lesen und ihre getroffenen Urteile auch nicht mehr ändern. Diese Vorgaben zielen auf den Erhalt „intuitive[r] Antworten“ ab (Sopata 2004: 105). Eine streng gehandhabte Zeitbegrenzung wurde nicht gesetzt, allerdings darauf hingewiesen, dass die Bearbeitung des GJT nicht länger als 5 Minuten dauern soll.

Die Items 7, 13 und 21 sind ambivalent, weil sie sowohl korrekt als auch unkorrekt bewertet werden können, je nachdem welcher semantische Kontext dazu gedacht

⁴⁹ Ellis zufolge galten GJT-Konstruktionen gemäß dem Prinzip *preference choice* zum damaligen Zeitpunkt als vergleichsweise reliabel und valide (vgl. 171f).

wird. Item 7 (*Auf der Autobahn wird es sehr schnell gefahren*) ist der Lösungsschablone zufolge falsch. Diese Lösung trifft zu, wenn ‚es‘ als Expletivum betrachtet wird. Es wäre aber auch vorstellbar, dass Testpersonen ‚es‘ als Pronomen für ein Fahrzeug interpretieren und deswegen folgerichtig den Satz als korrekt bewerten. Item 13 (*Er setzte sich zwischen seinen Schülern*) ist gemäß dem Lösungsbogen falsch, allerdings kann dieser Satz auch als korrekt interpretiert werden. Falsch ist der Satz, wenn ‚setzte sich‘ als zielgerichtete Bewegung verstanden wird, weil die der Wechselpräposition ‚zwischen‘ folgende lokale Angabe dann im Akkusativ stehen muss. Bedingt durch die Präteritumform des Verbs kann der Satz semantisch so interpretiert werden, dass vom Vorhandensein des Subjekts ‚zwischen seinen Schülern‘ ausgegangen werden kann und folglich der Dativ ebenfalls grammatisch ist. Bei Item 21 (*Während des Krieges wird es gelitten*) ist die Lösungsambivalenz ähnlich begründet wie bei Item 7, denn das ‚es‘ muss nicht zwangsläufig als überflüssiges Expletivum interpretiert werden. Wenn ‚gelitten‘ nicht als Partizip II von ‚leiden‘ im Sinne von ‚Leid ertragen‘, sondern in der homonymen Wortbedeutung ‚etwas akzeptieren, dulden‘ interpretiert wird, kann das ‚es‘ als Pronomen fungieren, das für eine Sache steht, die zu Friedenszeiten nicht akzeptabel wäre. Item 26 ist ebenfalls problematisch, denn die Nomen-Verb-Verbindung „Gesundheit“ + „untergraben“ (*Der Alkohol begann, seine Gesundheit zu untergraben*) ist ungewöhnlich. Darüber hinaus sollten Themen, die mit Drogen, Sexualität oder Krieg (also auch Item 21) zu tun haben, möglichst nicht in Fremdsprachen-Tests vorkommen, um die eventuell durch Irritation verursachte Benachteiligung einzelner Testpersonen zu vermeiden (Schlak, Zimmermann, Molnár 2010: 17). Bei der Auswertung werden die hier diskutierten Items einzeln analysiert, um eventuelle Auffälligkeiten zu erkennen und gegebenenfalls einzelne Items aus der Analyse herauszunehmen.

3.4.2.2 C-Tests und S-C-Tests

C-Tests sind ein gängiges Mittel zur Messung der allgemeinen Sprachkompetenz und werden v. a. als Einstufungstests sowie in der Forschung eingesetzt (vgl. Grotjahn 1995: 44f; Grotjahn 2011: 131f). Ein C-Test-Set besteht aus mehreren kurzen Lückentexten, die jeweils thematisch in sich abgeschlossen sind. In den Lückentexten wird die hintere Hälfte jedes zweiten Worts gelöscht, wobei der erste und letzte Satz ganz erhalten bleiben, um dem Text einen Rahmen zu geben (vgl. Klein-Braley & Raatz 1984: 136, die den ersten C-Test entwickelten und validierten). C-Tests stel-

len eine Weiterentwicklung der Cloze-Tests dar und beruhen auf dem Prinzip der reduzierten Redundanz: Sprache enthält eine Vielzahl redundanter Informationen, die durch das Weglassen der Endungen verringert wird (vgl. Klein-Braley 1997: 63). Hinter dem Prinzip steht die Annahme, dass ein_e L1-Sprecher_in diese Texte wieder herstellen kann. L2-Sprecher_innen haben je nach erreichtem sprachlichen Niveau mehr oder weniger Schwierigkeiten, die Lücken auszufüllen. C-Tests weisen signifikant positive Korrelationen mit allen sprachlichen Teilfertigkeiten auf (vgl. die Übersicht in Eckes & Grotjahn 2006: 295f).

Im hier vorliegenden Forschungsprojekt wird die allgemeine Sprachkompetenz mit den Schwerpunkten Lexik und Morphologie mittels zeitlimitierter (S-C-Tests) und nicht zeitlimitierter C-Tests erhoben. Der S-C-Test, den Aguado, Grotjahn und Schlak für das Forschungsprojekt entwickelt haben, ermöglicht es, sehr hohe sprachliche Kompetenzprofile differenziert zu messen und sogenannte Deckeneffekte zu vermeiden (Grotjahn, Schlak & Aguado 2010: 298). C-Tests sind ein reliables und valides Mittel, um den Sprachstand bis zu einem Niveau von C1 (GER) zu messen, etwa zu Einstufungszwecken in einen passenden Sprachkurs.⁵⁰ Durch S-C-Tests wird es möglich, auch im Bereich C2 Leistungsniveaus voneinander unterscheiden zu können (vgl. Aguado, Grotjahn & Schlak 2005: 284).

Der gemeinsame Einsatz von C-Tests und S-C-Tests ermöglicht neben der Messung der allgemeinen lexikalischen und morphosyntaktischen Kompetenz auch die Bestimmung der prozeduralen und deklarativen Wissensanteile (vgl. Grotjahn, Schlak & Aguado 2010: 303). Aguado, Grotjahn und Schlak gehen davon aus, dass die „kompetente Verwendung von Sprache unter Echtzeitbedingungen [...] sowohl eine breite deklarative (sprachliche) Wissensbasis als auch eine ausreichende Prozeduralisierung (Automatisierung) des Wissens“ (Aguado, Grotjahn & Schlak 2007: 145) voraussetzt. Ihre Annahme ist dabei, dass Erstsprachler_innen über große deklarative sprachliche Wissensbestände verfügen, die zumeist auch ausreichend prozeduralisiert sind. L2-Lerner_innen werden dagegen häufig auch bei hohem deklarativen Wissensanteil nur auf eine unzureichende Prozeduralisierung zurückgreifen können (vgl. ebd.: 145).

⁵⁰ Beispielsweise mit dem onSET-Deutsch, einem Online-Einstufungstest, der im Bereich A2–C1 (GER) einstuft (vgl. URL 11; o. J.: o. S.).

Gemäß Andersons ACT-Theorie (ACT = *adaptive control of thought*) bedeutet der Erwerb von Fähigkeiten und/oder Fertigkeiten einen Übergang von deklarativem zu prozeduralem Wissen (Anderson & Lebiere 1998). Dieser Prozess wird als Prozeduralisierung bezeichnet, während unter Automatisierung der Endpunkt dieser Entwicklung verstanden wird (vgl. auch Segalowitz & Hulstijn 2005: 372). Deklaratives Wissen bedeutet bewusstes, für die Fähigkeiten relevantes Wissen, das eine Person beschreiben kann. Prozedurales Wissen zeigt sich im Verhalten einer Person, ist ihr aber nicht bewusst und kann von ihr nicht beschrieben werden (vgl. Anderson & Lebiere 1998: 5). Segalowitz (2003) führt in diesem Zusammenhang das Beispiel an, dass deklaratives Wissen mit dem expliziten Wissen über grammatische Konstruktionen von L2-Lerner_innen vergleichbar ist. Prozedurales Wissen meint dagegen das Wissen der L1-Sprecher_innen über die Wohlgeformtheit von Sätzen, die aber oft nicht erklären können, warum etwas falsch oder richtig ist (vgl. ebd.: 395).

Dabei ist eine schnelle von einer automatischen Verarbeitung (*fast vs. automatic processing*) zu unterscheiden (vgl. Segalowitz 2003: 385f.): Schnelle Verarbeitung ist die Beschleunigung aller Prozesse, die mit dem anfänglichen Stadium des Fähigkeitserwerbs zusammenhängen, während automatisierte Verarbeitung eine grundsätzliche Umorganisation und Abkürzung einiger serieller Prozesse bedeutet. Paradis (2009) zufolge können L2-Lerner_innen zwar sukzessive vom ausschließlichen Gebrauch explizit gelernter und bewusst vorhandener Wissensseinheiten zur Anwendung impliziter Wissensformen gelangen. Es ist aber schwer zu beurteilen bzw. zu messen, ob sie dabei ihre Verarbeitungsprozesse graduell beschleunigen oder tatsächlich den Schritt zur Automatisierung vollziehen (vgl. ebd.: XI).

Der Entwicklungsprozess der S-C-Tests ist ausführlich in Grotjahn, Schlak & Aguado (2010: 303f) beschrieben; weitere Überlegungen zur Konstruktion der S-C-Tests und Deutungsmöglichkeiten ihrer Ergebnisse finden sich in Grotjahn (2010). Das im Forschungsprojekt eingesetzte Testset besteht aus sechs Texten à 25 Lücken mit jeweils variierender zeitlicher Limitierung zwischen 1:05 min und 1:55 min. Das Testset setzt keinen Eisbrecher-Text ein. Die Verwendung eines Eisbrecher-Textes hat den Vorteil, dass sich die Teilnehmer_innen (TN) an das Format gewöhnen können und eventuelle anfängliche Schwierigkeiten nicht in die Wertung eingehen. Auch im Hinblick darauf, dass die S-C-Tests vor den nicht zeitlimitierten Tests bearbeitet werden, wäre der Einsatz eines Eisbrecher-Tests wünschenswert.

Einige der Texte sind aus inhaltlichen Gründen problematisch: Sowohl der erste Text „Krieg“ als auch der sechste Text „Risiken der Technik“ könnte TN in Bedrängnis bringen und dadurch benachteiligen, etwa falls sie persönliche, vielleicht sogar traumatisierende Erfahrungen, die in Zusammenhang mit diesen Themen stehen, mitbringen. Die *Association of Language Testers in Europe* (ALTE) führt in ihrer Handreichung für Test-Autor_innen an, dass Texte aufgrund unangemessener Themen ungeeignet sind, „such as war, death, politics and religious beliefs, which may offend or distress some candidates“ (URL 12; 2005: 55) (vgl. auch die Anmerkung zur *test bias* in Kap. 3.4.2.1 *Grammatical Judgement Test*). Bei dem Text „Lenin-Denkmal“, in dem es um den Abbau des Lenin-Denkmal in Dresden geht, gefährdet meiner Ansicht nach die Tagesaktualität und Spezifik des Themas die Chancengleichheit. TN, die den Test mit weniger Abstand zum Ereignis machen, haben Vorteile gegenüber denjenigen mit mehr Abstand (Schlak, Zimmermann & Molnár 2010: 17). Auch ist zu bedenken, dass generell bestimmte Personengruppen mit einem Thema mehr anzufangen wissen als andere oder durch die Themenwahl benachteiligt werden (Schön, Zimmermann & Johnson 2012: 66). Vor diesem Hintergrund könnte es heikel sein, dass zwei von sechs Texten (Text 4 „Lenin-Denkmal“ und Text 5 „Kunstleben“) Ereignisse behandeln, die mit der Wende resp. dem Zerfall der Sowjetunion zu tun haben. Da die S-C-Tests als Testset für das Projekt „Erwerbsalter und Sprachlernerfolg“ entwickelt und von 100 Testpersonen pilotiert wurden, werden trotz der oben angeführten Bedenken sämtliche Tests in die Auswertung miteinbezogen, um die Validität des Sets nicht zu gefährden.

3.4.2.3 Fragebogen zum Sprachenprofil

Mit einem Fragebogen zum Sprachenprofil werden allgemeine Informationen, also Alter, Geschlecht, beruflicher Hintergrund etc., wie auch spezifische Daten über den L2-Erwerb erhoben: AO, LOR, Lernkontext, DaF-Unterricht in der Heimat und in Deutschland, Daten zum gegenwärtigen Gebrauch der L1, des Deutschen und weiterer Fremdsprachen etc. Der Fragebogen liegt in zwei Fassungen vor: eine, die für das Projekt „Erwerbsalter und Sprachlernerfolg“ entwickelt wurde, und eine, die für das Projekt „Erwerbsalter und Sprachlernerfolg aus bilingualer Perspektive“ modifiziert und ergänzt wurde. In der auf Personen mit 2L1 zugeschnittenen Version gibt es neben den spezifischen Fragen zur doppelten Erstsprachigkeit auch eine Ausdifferenzierung der Fragen nach dem Gebrauch des Deutschen:

- Item 17: „Wie häufig nutzen Sie deutschsprachige Medien?“
- Item 18: „Welche Sprachen werden in Ihrer Wohnung gesprochen?“
- Item 19: „Welche Sprachen werden da, wo sie derzeit arbeiten/studieren gesprochen?“

Das Item 11: „In welchen Sprachen sind Sie alphabetisiert worden?“ ist im Kontext von Personen mit 2L1 Deutsch und Mandarin bzw. Kantonesisch ungünstig formuliert, um den Erwerb der damit verbundenen logographischen Schriftsysteme zu erfragen. Item 20 „Schätzen Sie bitte Ihre Deutschkenntnisse auf einer Skala von 1 (sehr niedrig) bis 6 (fast muttersprachlich) ein“ ist angesichts des Adressatenkreises (Personen mit 2L1) nicht angemessen. Zumindest sollte „fast“ vor „muttersprachlich“ gestrichen werden oder eine andere Bezeichnung für die Erfragung sehr hoher Sprachkompetenz gewählt werden.

3.4.2.4 TESTATT-Skalen

Die TESTATT-Skalen wurden im Rahmen des Projektes *Tests and Attitude Scales for the Year Abroad* (1998–2001) entwickelt, einem Nachfolgeprojekt des *European Language Proficiency Surveys* (ELPS, 1993–1996). Eins der Ergebnisse der ELPS ist, dass der wichtigste Faktor für den Fremdsprachenlernerfolg während des Studiums ein entsprechender Auslandsaufenthalt ist. Allerdings wird festgestellt, dass nicht alle Studierende im gleichen Maß davon profitieren. Um diese Unterschiede genauer bestimmen zu können, wurden die TESTATT-Instrumente als Kombination aus verschiedenen C-Tests und den Skalen entwickelt (vgl. Grotjahn 2004: 1).

Bezüglich des Lerner_innenprofils sind v. a. die individuellen Faktoren Motivation, Extraversion/Introversion und Fremdsprachenverwendungsangst von Interesse, die mit Hilfe der Skalen des TESTATT ermittelt werden. Der Einfluss der Motivation auf den L2-Lernerfolg wurde schon oft nachgewiesen (vgl. Kap. 2.4.2 *Motivation und Erwerbsantrieb*). Ähnlich eindeutig ist der Befund bei der Fremdsprachenverwendungsangst: Zumindest im schulischen Kontext ist sie einer der besten Prädiktoren für den L2-Lernerfolg: „Fremdsprachenverwendungsangst kann die Verarbeitung, das Behalten und die Produktion von Sprache auf allen Ebenen negativ beeinflussen [...]“ (Grotjahn 2004: 5). Bezüglich der Extraversion sind die Ergebnisse weniger eindeutig: Ausgehend von der Annahme, dass extravertierte Persönlichkeiten mehr mit Sprecher_innen der Zielsprache interagieren und demzufolge mehr zielsprachli-

chen Input erhalten, wird den extravertierten L2-Lerner_innen gemeinhin mehr Erfolg prognostiziert. Signifikante Korrelationen wurden bisher jedoch ausschließlich für die Flüssigkeit einer Äußerung im mündlichen Ausdruck nachgewiesen, während es bei schriftlichen Aufgaben keine relevanten Unterschiede zwischen introvertierten und extravertierten Lerner_innen gibt (vgl. ebd.: 4).

Die TESTATT-Skalen sind mehrfach erprobt worden und ökonomisch zu handhaben. Bei den 16 Items der Extraversionsskala und der Fremdsprachenverwendungsangst-Skala wurden von Aguado, Grotjahn und Schlak nur minimale Veränderungen für das Projekt „Erwerbssalter und Sprachlernerfolg“ vorgenommen. Der Fragebogen zur Motivation wurde hingegen stärker modifiziert: Vier der zehn TESTATT-Items wurden gestrichen und neun neue hinzugefügt. Die nun insgesamt 15 Items enthalten je fünf Items zur instrumentellen und fünf zur integrativen Motivation, je ein Item zu den Motivationstypen *travel orientation*, *friendship orientation* und *knowledge orientation* (vgl. Clément & Kruidenier 1983) sowie zwei Items zur intrinsischen Motivation (vgl. Aguado, Grotjahn & Schlak 2005: 287–288). Die Items beginnen immer mit „Ich habe Deutsch gelernt ... bzw. ich lerne Deutsch ...“

Items zur instrumentellen Motivation:

- ... für meinen zukünftigen Beruf (1)⁵¹
- ... um ein höheres Einkommen zu erzielen (6)
- ... weil es für mein Studium wichtig ist (9)
- ... weil ich im Studium erfolgreich sein möchte (13)
- ... weil ich im Beruf erfolgreich sein möchte (15)

Items zur integrativen Motivation:

- ... weil ich die Deutschen mag (3)
- ... um die deutsche Lebensweise und das Land, in dem diese Sprache gesprochen wird, besser zu verstehen (4)
- ... weil ich diese Sprache mag (7)
- ... weil mir Deutschland gefällt (12)
- ... weil mich die deutsche Kultur und Sprache interessieren (14)

⁵¹ Die Zahl in Klammern gibt jeweils die Nummer auf dem Fragebogen an.

Item zur *travel orientation*:

- ... um in deutschsprachige Länder reisen zu können (2)

Item zur *friendship orientation*:

- ... um in meinem Leben eine größere Anzahl deutschsprachiger Menschen kennen zu lernen (8)

Item zur *knowledge orientation*:

- ... um eine höhere Bildung zu erlangen (11)

Item zur intrinsischen Motivation:

- ... weil es mir Spaß macht (5)
- ... weil ich gerne Sprachen lerne (10)

Das Item (2), das hier zur *travel orientation* gezählt wird, könnte in anderen Zusammenhängen auch unter die instrumentelle Motivation fallen. In der Auswertung wird die instrumentelle Motivation gesondert berechnet, wozu die fünf oben aufgeführten Items sowie das Item (2) zusammengezählt werden.

Die Fremdsprachenverwendungsangst-Skala ist für den Kontext Klassenzimmer bestimmt. Acht von elf Items nehmen direkt auf Situationen im Klassenzimmer Bezug, da sie der *Foreign Language Classroom Anxiety Scale* entnommen wurden (vgl. Grotjahn 2004: 5). Es stellt sich die Frage, inwiefern sie etwas über die Fremdsprachenverwendungsangst bei Proband_innen aussagen kann, die schon seit Jahren im Land der Zielsprache leben und deren letzter Fremdsprachenunterricht eventuell schon viele Jahre zurückliegt. Hinzu kommt der Umstand, dass die Skala durchgängig negativ formuliert ist – bspw. „Ich zittere, wenn ich weiß, dass ich im Fremdsprachenunterricht gleich aufgerufen werde“ (1). Tatsächlich wird eine der Probandinnen die Rückmeldung geben, dass für sie der Fremdsprachenunterricht immer eher spaßig und spielerisch war. Dagegen seien Beiträge auf Deutsch in einer Arbeitsbesprechung oder eine Präsentation auf Deutsch nach wie vor für sie angstbesetzt.

Trotz dieser berechtigten Kritik wurden in der vorliegenden Studie die TESTATT-Skalen genau so übernommen, wie sie im Projekt „Erwerbsalter und Sprachlernerfolg“ eingesetzt wurden. Der Rückgriff auf die in Aguados, Grotjahns und Schlaks Projekt erhobenen Daten und die Möglichkeit einer gemeinsamen Auswertung sollte nicht durch die Veränderung der eingesetzten Instrumente beschränkt werden.

3.4.2.5 Phonetik-Test

Der Bereich der Phonetik wird in Studien zum Faktor Alter am häufigsten untersucht und diskutiert. Um die Aussprache zu testen, sprechen die Proband_innen in der vorliegenden Untersuchung zwei Minuten frei. Das Sprechen wird durch folgende Impulse angeregt:

- Was machen Sie gerne in Ihrer Freizeit?
- Reisen Sie gerne? Können Sie eine Ihrer letzten Reisen beschreiben?
- Erzählen Sie den Inhalt eines Filmes, den Sie in letzter Zeit gesehen haben?
- Erzählen Sie den Inhalt eines Buches, das Sie in letzter Zeit gelesen haben?

Im Vergleich zu der im Projekt „Erwerbsalter und Sprachlernerfolg“ vorgesehenen Vorgehensweise stellt dieses Verfahren eine starke Modifikation dar (vgl. Aguado, Grotjahn & Schlak 2005: 281–282 und die Darstellung der Pilotstudie zur Phonetik-Auswertung des Projekts von Molnár 2010).

Aufgrund der von Molnár geschilderten Erfahrungen veränderte ich den Frageimpuls für die freie Rede. „Wie läuft bei Ihnen normalerweise ein Wochenende ab?“ erwies sich als eine für die Mehrzahl der Testpersonen zu persönliche Frage (vgl. ebd.: 54). Aus diesem Grund bot ich den Proband_innen vier verschiedene Impulse an, so dass sie wählen konnten. Die Impulse orientieren sich an den von Moyer eingesetzten Materialien:

„A. Beschreiben Sie eine bedeutsame Erfahrung in Ihrem Leben. Was ist passiert, wie alt waren Sie, wer war wichtig in dieser Erfahrung, wie hat Sie diese Erfahrung beeinflusst usw. ...?“

B. Beschreiben Sie eine sehr wichtige Person in Ihrem Leben. Wie haben Sie ihn/sie kennengelernt? Wie hat diese Person Ihr Leben verändert oder Ihre Meinung beeinflusst? (das darf auch ein beliebtes Haustier sein)

C. Beschreiben Sie ein Problem oder eine schwierige Situation, mit der Sie kürzlich konfrontiert wurden. Wie haben Sie eine Lösung gefunden? Haben Sie etwas Wichtiges dabei gelernt? Wie würden Sie solche Probleme in der Zukunft vermeiden? usw. ...

D. Besprechen Sie ein Thema Ihrer Wahl. Warum ist dieses Thema/dieses Prinzip wichtig? Wann und unter welchen Umständen haben Sie sich dafür erstmals interessiert?“ (Moyer 2004: 153)

Bei der Durchführung der Datenerhebung erwies sich, dass die zweite von mir getestete Person diese Fragen ebenfalls als viel zu persönlich empfand. Der Redefluss wurde dadurch nicht nur gehemmt, sondern gestoppt. Um dennoch zu einer Sprechprobe zu gelangen, stellte ich einige offene, möglichst unverfängliche Impulsfragen.

Diese Impulse behielt ich dann für die folgenden Befragungen bei; mit der ersten Testperson habe ich das kurze Interview mit den veränderten Impulsen wiederholt, um gleiche Bedingungen für alle getesteten Personen zu gewährleisten.

Die ursprünglich für das Projekt „Erwerbsalter und Sprachlernerfolg“ vorgesehenen Vorlesetexte wurden nicht eingesetzt, weil im Fokus der hier vorliegenden Forschungsarbeit die Bewertung von Akzent im Alltagssprachlichen Kontext stand. Das Vorlesen eines Textes spielt dabei eine untergeordnete Rolle. In der Pilotuntersuchung von Molnár sehen wir darüber hinaus, dass in fünf von neun Fällen einer der vorgelesenen Texte genau gleich wie das freie Sprechen bewertet wurde; in den vier anderen Fällen beträgt die höchste Abweichung .34⁵² (vgl. Molnár 2010: 50).

Aus den im Rahmen der hier vorliegenden Studie erhobenen zweiminütigen Sprechproben wurde ein Ausschnitt von jeweils 30 Sekunden extrahiert, der dem Bewertungsteam vorgelegt wurde. Diese Kürzung der Interviews geschah, um die zeitlichen Ressourcen der Rater zu schonen, die insgesamt 50 Sprechproben bewerteten. Darüber hinaus ging es darum, Hinweise auf die Herkunft der Testpersonen zu tilgen. Dabei stellte es sich heraus, dass es bei einigen Interviews schwierig war, Passagen auszuwählen, die keine Hinweise auf die Herkunft enthielten, etwa wenn die Testpersonen antworteten, dass sie ausschließlich auf Chinesisch lesen oder die meisten Reisen nach China machen.

In den einschlägigen Studien finden sich unterschiedlich zusammengestellte Rater-Teams, die in der Regel die zu bewertende Sprache als L1 sprechen. Sie unterscheiden sich v. a. dadurch, dass sie einen philologischen bzw. fremdsprachendidaktischen professionellen Hintergrund haben oder nicht. Bei Patkowski (1980: 452) sind es zwei Rater, die seit mindestens fünf Jahren Englisch als Fremdsprache unterrichten. Bei Flege et al. (1999: 83) kommen zehn Rater zum Einsatz, deren Geschlecht, Alter und regionale Herkunft angegeben wird, während die beruflichen Hintergründe nicht dokumentiert sind. Molnár (2010: 46) setzte vier Rater mit Deutsch als L1 ein, wovon einer professionelle Erfahrung hat, was das beurteilen phonetischer Kompetenz angeht. Munro, Derwing & Morton (2006) setzten bezüglich der Ausgangssprachen ein heterogenes Rater-Team ein. Sie untersuchen, inwiefern die

⁵² Die Sprechproben wurden auf einer Skala bewertet: 1 = sehr starker ausländischer Akzent: definitiv kein deutscher Muttersprachler bis 5 = kein ausländischer Akzent: definitiv deutscher Muttersprachler.

L1 der Rater (Englisch, Japanisch, Kantonesisch und Mandarin) die Beurteilung der Testpersonen (L1 Japanisch, Kantonesisch, Spanisch und Polnisch) beeinflusst.

Die für das hier vorliegende Forschungsprojekt erhobenen Sprechproben wurden von einem heterogenen Team bewertet: eine Person, die mit der L1 Deutsch (aktiv) und Plattdeutsch (passiv) aufgewachsen ist; zwei Personen, die Deutsch als Späterwerber_innen gelernt haben; eine mit der L1 Rumänisch (AO = 7), eine mit der L1 Mandarin (AO = 25); zwei Personen, die mit 2L1 Deutsch und Polnisch bzw. Italienisch aufgewachsen sind. Alle Bewerber_innen verfügen über einen Hochschulabschluss; vier sind einschlägig im Fach Deutsch als Fremdsprache ausgebildet. Vor dem Hintergrund der Argumentation in Kapitel 2.5 ‚*Muttersprachliche Kompetenz*‘ und auch im Hinblick auf den Erkenntnisgewinn, inwiefern sich die Beurteilungen dieser heterogenen Rater unterscheiden, fiel die Entscheidung für diese Zusammensetzung anstelle eines Teams, das ausschließlich aus Personen mit der L1 Deutsch besteht.

Die Rater bewerten die 30-sekündigen Sprechproben der hier vorliegenden Studie zunächst auf einer fünfstufigen Skala⁵³:

- 1 = sehr starker Akzent
- 2 = starker Akzent
- 3 = weniger starker Akzent
- 4 = leichter Akzent
- 5 = kein Akzent (akzentfrei)

Dies ist eine Abwandlung der Skala von Molnár (2010: 47), die eine in der Fremdsprachenerwerbsforschung häufig verwendete fünfstufige Skala ins Deutsche übertrug, um die Vergleichbarkeit mit diesen Studien zu gewährleisten (für die Originalskala vgl. Bongaerts, Planken & Schils 1995: 39). Im Unterschied zum Original wird in der hier vorliegenden Skala nicht „sehr starker ausländischer Akzent“ etc. verwendet, sondern lediglich „sehr starker Akzent“. Akzent kann auch ohne spezifische Attribuierung bewertet werden und die Kontrollgruppe (n = 10) besteht aus Dia-

⁵³ Das Informationsblatt, in dem die Rater über das Prozedere informiert werden, befindet sich im Anhang. Auf dem Informationsblatt wird auf eine Definition von Akzent zurückgegriffen, die aus Wikipedia stammt. Mir ist bewusst, dass Wikipedia als Ressource problematisch ist. Im Zusammenhang der Begriffsklärung für Rater, die nicht alle einen fachwissenschaftlichen Bezug haben, war es angemessen, auf eine Definition zurückzugreifen, die allgemein verständlich ist (vgl. URL 13; 2013: o. S.).

lekt sprecher_innen, deren Akzent somit auf der gleichen Skala erfasst werden kann. Als mögliches Problem könnte es sich hier ergeben, dass eine oder mehrere Testpersonen L2-Lerner_innen sind und eine dialektale Varietät sprechen. Durch die Gelegenheit, die abgegebenen Bewertungen zu begründen bzw. zu kommentieren, können die Rater eventuelle Uneindeutigkeiten benennen.

Abrahamsson & Hyltenstam (2009: 267) lassen ihre Rater eine binäre Entscheidung treffen, denn sie verstehen ‚Muttersprachlichkeit‘ als binäres Merkmal analog zu Eigenschaften wie *verheiratet*, *tot* etc. Die Rater, die alle ‚Muttersprachler_innen‘ waren, sollten beurteilen, ob die Sprechproben von Stockholmer Sprecher_innen stammen, nach einem anderen Dialekt klingen oder einen ausländischen Akzent haben. Dadurch wollten sie die Rater davon abhalten, sich ausschließlich auf die Wahrnehmung von ausländischem Akzent zu konzentrieren und jede Abweichung vom Standard sofort als fremdsprachlich wahrzunehmen. Darüber hinaus wird die Aufgabe dadurch vergleichbar mit einer „authentic, everyday speaker-judgement situation, in which many sources of phonetic variation may come to people’s minds“ (ebd.: 266).

Im Rahmen der hier vorliegenden Untersuchung wird die Entscheidung über dialektalen oder fremdsprachlichen Akzent mit denselben Intentionen aufgegriffen, die Abrahamsson & Hyltenstam formulieren. Im zweiten Schritt des Phonetik-Ratings sollen die Bewerter_innen angeben, ob Sie den Akzent als fremdsprachlich oder dialektbedingt wahrnehmen. In einem dritten Schritt beurteilt das Rater-Team auf einer fünfstufigen Skala die Verständlichkeit der Äußerung:⁵⁴

- 1 = Die Person ist nicht zu verstehen
- 2 = Die Person ist schwer zu verstehen, ich musste mich beim Hören sehr anstrengen
- 3 = Die Person ist stellenweise schwer zu verstehen, ich musste mich beim Hören anstrengen
- 4 = Die Person ist zu verstehen, ich musste mich beim Hören aber etwas anstrengen
- 5 = Die Person spricht klar und deutlich und ist gut zu verstehen

⁵⁴ Diese für das vorliegende Forschungsprojekt entwickelte Skala für die Verständlichkeit einer Äußerung wird auch in einer empirischen Untersuchung von Lahti (eingereicht) eingesetzt.

Schließlich hatten die Rater optional die Gelegenheit, ihre Meinung zu begründen oder zu kommentieren.

Die Unterscheidung zwischen Akzent und Verständlichkeit muss bezüglich des Lernerfolgs von L2-Lerner_innen eingebracht werden. Zahlreiche Studien zeigen den komplexen Zusammenhang von Akzent- und Verständlichkeitsbewertungen, wenn L2-Sprecher_innen bewertet werden. Verständliches Sprechen ist ein rationales und besser erreichbares Ziel als akzentfreies Sprechen (vgl. Jun & Li 2010: 53). Trotzdem ist es schwer, beide Kategorien klar voneinander abzugrenzen. So zeigen Munro & Derwing (1995), dass die Leichtigkeit, mit der mündliche Äußerungen verstanden werden, von der Stärke des fremdsprachlichen Akzents beeinflusst wird.

3.4.2.6 LLAMA-Tests

Die LLAMA-Tests bestehen aus vier Subtests und wurden in Anlehnung an („based loosely on“) Carroll und Sapon (1959) entwickelt (vgl. Meara 2005: 2):

- LLAMA B – Vokabeln lernen
- LLAMA D – Laute und Muster in gesprochener Sprache erkennen
- LLAMA E – Laute und Symbole zuordnen
- LLAMA F – Grammatische Regeln induktiv ableiten

Die LLAMA-Tests haben mehrere Vorteile: sie sind unabhängig von der L1 und sonstigen Sprachkenntnissen der Testperson, lassen sich am PC durchführen und sind Freeware (URL 14). Ersteres wird dadurch erreicht, dass v. a. mit Bildern gearbeitet wird und die phonetischen Stimuli auf einem westkanadischen indigenen Dialekt beruhen, die in verfremdender Aussprache präsentiert werden (vgl. Meara 2005). Kritisch anzumerken ist, dass sich die theoretischen Konzepte zur Sprachlernung seit Carroll & Sapon (1959) weiter entwickelt haben und neuere Ansätze deutlich komplexere Modellierungen der Sprachlernung(en) vornehmen, wobei gerade auch dem Arbeitsgedächtnis und der Ambiguitätstoleranz eine zentrale Rolle zukommt (vgl. Kap. 2.4.3 *Sprachlernung*).

Meara weist darauf hin, dass die Tests nicht standardisiert seien und deswegen nur mit Vorsicht eingesetzt werden dürfen (vgl. Meara 2005: 21). An diesem Punkt setzt Granenas (2013a) explorative Validierungsstudie der LLAMA-Tests an, die die Beurteilung der Reliabilität der Tests ermöglicht und ihre zugrundeliegende Struktur er-

forscht (vgl. ebd.: 106). Unter Reliabilität versteht Granena in diesem Zusammenhang die interne Konsistenz und Stabilität über eine Zeitspanne hinweg (vgl. ebd.: 113). Granena ließ insgesamt 186 Testpersonen mit drei verschiedenen Ausgangssprachen (Chinesisch, Spanisch und Englisch) die LLAMA-Tests sowie zwei verschiedene Test-Batterien mit Intelligenz- und Arbeitsgedächtnistests absolvieren.⁵⁵ Die Varianzanalysen ergaben keine signifikanten Differenzen zwischen den drei verschiedenen L1 – weder bezüglich des Gesamtergebnisses noch bezüglich der einzelnen Subtests LLAMA B, D, E und F (vgl. ebd.: 117–118). Die interne Konsistenz und zeitliche Stabilität der Tests wurde mit Cronbachs Alpha und der Test-Retest-Reliabilität nach zwei Jahren gemessen. Die interne Konsistenz der LLAMA-Tests ist insgesamt gut ($\alpha = .77$), für die einzelnen Subtests jedoch nur zufriedenstellend: LLAMA D, E und F liegen unter $.70$ (vgl. ebd.: 118). 20 der 186 Testpersonen haben den Test nach zwei Jahren wiederholt, wobei sich moderate und moderat starke Korrelationen ergaben (vgl. ebd.: 118–119).

Die faktorielle Validität des Tests überprüft Granena mit einer *Principal Component Analyses* (PCA) sowohl der einzelnen Subtests untereinander als auch in Bezug auf die eingesetzten Intelligenz- und Kurzzeitgedächtnistests. Dabei stellt sich heraus, dass sich LLAMA D (Laute und Muster in gesprochener Sprache erkennen) anders verhält als die anderen Subtests (vgl. ebd.: 119–120). LLAMA B, E und F haben gemeinsam, dass sie eine Studierphase beinhalten und die Möglichkeit bieten, Strategien und Problemlösungstechniken anzuwenden. Granena zufolge spielen Auswendiglernen, explizit assoziatives Lernen und analytische Fähigkeiten in allen drei Tests eine Rolle. LLAMA B ist ein Vokabellernstest, man kann aber auch Beziehungen zwischen Bezeichnungen und Objekten herausarbeiten. Die stärksten analytischen Fähigkeiten werden beim LLAMA F gebraucht. LLAMA D hat wenig mit den anderen Tests zu tun, weil er ganz andere Fähigkeiten misst:

„This uncorrelated variation could be used to look at aptitude abilities through a componential approach by identifying individual aptitude profiles, for example L2 learners with strong sound sequence learning ability, but low grammatical inferencing skills,

⁵⁵ Die eine Batterie, die von 110 Test-TN absolviert wurde, umfasst einen *general intelligence test* (GAMA) und einen *probabilistic serial reaction time task* (SRT). Die andere Batterie, die den anderen 76 TN vorgelegt wurde, besteht aus folgenden Tests: dem *operation span test* (OSPAN), der das Arbeitsgedächtnis testet, dem *letter span test*, der das Kurzzeitgedächtnis misst, dem *digit-symbol correspondence test*, der die Verarbeitungsgeschwindigkeit erfasst, und dem *Simon task*, einem Test der Aufmerksamkeitskontrolle (vgl. Granena 2013a: 114–117).

and vice-versa – a situation with obvious potential for aptitude-treatment interaction (ATI) studies [...].“ (Granena 2013: 122–123)

Diese Perspektive auf die Komponenten bestätigt sich durch die PCAs mit den anderen Testformaten. LLAMA D stimmte dabei mit dem SRT-Task überein, der das implizite Sequenzlernen misst. Außerdem ergeben die PCAs mit den Testformaten, die *Working Memory*, *Storage Capacity*, *Processing Speed* und *Attention Control* messen, dass sich die von den Tests LLAMA B, E und F gemessenen Sprachlernerkomponenten von den mittels der anderen Tests gemessenen Gedächtnisleistungen unterscheiden (vgl. ebd.: 124). Granena folgert aus den Analysen, dass die LLAMA-Tests zwei verschiedene Sprachlernerkomponenten messen: eine analytische Fähigkeit, die v. a. für das explizite Sprachenlernen relevant ist, und eine auditive Sequenz-Lernfähigkeit, die für das implizite Sprachenlernen eine Rolle spielt (vgl. ebd.: 126).

Die Ergebnisse bei der Validierung der Subtests sind nicht überraschend, weil Laute zu diskriminieren und wiederzuerkennen (LLAMA D) eine grundsätzlich andere Fähigkeit ist, als bspw. Satzstrukturen zu analysieren oder Vokabeln zu lernen. Unerwartet ist dagegen die Korrelation der analytischen Tests (LLAMA E und F) mit dem Vokabellerntest (LLAMA B), da sich analytische Fähigkeiten und Memorisierungsleistungen voneinander unterscheiden.

Auch wenn bei den LLAMA-Tests konzeptionelle Schwächen zu attestieren sind, werden sie im Rahmen dieser Forschungsarbeit dazu genutzt, die Sprachlernerkomponenten ausgewählter Teilnehmer_innen zu messen. Denn für einen Einsatz der LLAMA-Tests sprechen die positive Validierungsstudie von Granena (2013) und die Praktikabilität des Testsets durch seine Sprachenunabhängigkeit, computerbasierte Erhebung und Auswertung sowie der Zugänglichkeit als Freeware.

3.4.2.7 Fragebogen biografische Hintergrundinformationen und Motivation

Der in der zweiten Phase eingesetzte Fragebogen besteht aus zwei Teilen. Der erste Teil erfragt weitere Hintergrundinformationen über die Testpersonen. Der zweite Teil besteht aus Items, die die Einstellungen und Motivationen bezogen auf das Deutschlernen messen.

Die Fragebogen-Items des ersten Teils (Item 1–15) erheben zusätzliche Informationen über die Proband_innen und ergänzen somit den Fragebogen zur Sprachlernbi-

ografie der ersten Phase. Einige der Items, die sich mit dem Gebrauch der Zielsprache und dem Input beschäftigen, stammen aus dem Fragebogen, der im Projekt „Erwerbssalter und Sprachlernerfolg aus bilingualer Perspektive“⁵⁶ eingesetzt wurde:

Item 12 erfragt die Häufigkeit der Mediennutzung von TV, Radio, deutschsprachiger Musik, Zeitungen, Zeitschriften, Büchern und Internet.⁵⁷ Die Proband_innen kreuzen hierzu jeweils einen Wert von 1 (sehr selten) bis 4 (sehr häufig) an. Auch die anderen übernommenen Items erfassen den Gebrauch der deutschen Sprache. Generell muss bei Selbstberichten bedacht werden, dass sich die Selbstwahrnehmung der Proband_innen nicht immer mit den tatsächlichen Gegebenheiten decken muss. Dies gilt auch für Auskünfte zum Sprachgebrauch:

„One problem with this approach is that self-reported language use is notorious for its lack of reliability. John Gumperz, who studies language use in real-life settings, reports: “Attempts to elicit self-report information on bilingual usage regularly showed significant discrepancies between speakers’ descriptions of their own usage and empirical studies of tape recorded texts.” (Gumperz 1982: 62, zit. nach Bialystok & Hakuta 1994: 71)

Bei den hier eingesetzten Items geht es darum, eine Aussage darüber zu treffen, ob die L1, die L2 oder andere Sprachen in verschiedenen Kontexten gesprochen werden – nicht darum einzuschätzen, in welcher Quantität welche Sprache gesprochen wird.

Item 13: Welche Sprachen werden in Ihrer Wohnung gesprochen? (Mehrfachantworten möglich)

-) Deutsch
-) die Sprache, mit der ich aufgewachsen bin
-) andere Sprachen
-) Ich wohne allein.

Item 14: Welche Sprachen werden da gesprochen, wo Sie derzeit arbeiten/studieren? (Mehrfachantworten möglich)

-) Deutsch

⁵⁶ Ein Projekt an der TU Berlin, das auf dem Projekt „Alter und Sprachlernerfolg“ basiert und eine Gruppe von Personen mit 2L1 testet unter Verwendung der Instrumente aus dem Projekt „Alter und Sprachlernerfolg“ (GJT, S-C-Tests, C-Tests, TESATATT-Skalen). Der Fragebogen zur Sprachlernbiografie wurde an die Personen mit 2L1 angepasst und durch einige zusätzliche Items erweitert.

⁵⁷ Für den Einsatz des Fragebogens in der vorliegenden Studie wurde die Wahlmöglichkeit „Internet“ ergänzt.

die andere(n) Sprache(n), mit der (denen) ich aufgewachsen bin

andere Sprachen

Item 15: Welche Sprachen werden da gesprochen, wo Sie Ihre Freizeit verbringen?
(Mehrfachantworten möglich)

Deutsch

die Sprache, mit der ich aufgewachsen bin

andere Sprachen

Andere Items wurden neu kreiert, wobei sie sich inhaltlich v. a. an dem von Loup et al. (1994) präsentierten Profil und den Ergebnissen der Studien von Moyer (2004) orientieren (vgl. Kap. 2.2 *Studien zum Faktor Alter*). Dies betrifft die Fragen zu einer/einem deutschsprachigen Lebenspartner_in (Item 3), zum Kontakt zu deutschsprachigen Freund_innen (Item 4), nach dem selbstgesteckten Ziel des Deutschlernens (Item 6), der Selbsteinschätzung der Fähigkeit, Sprachen, Dialekte oder Akzente nachzuahmen (Item 7), sowie die Fragen zum Feedback und dem Umgang damit (Items 8–10). Item 11 fragt ausführlich ab, welche Bereiche (Sprechen, Hören, Lesen, Schreiben, Grammatik, Wortschatz, Aussprache) beim Deutschlernen im Vordergrund standen und ob es um alltägliche, berufliche und/oder akademische Kommunikation ging.

Der zweite Teil des Fragebogens (Items 16–55) wurde aus verschiedenen Studien aus dem Band *Motivation, Language Identity and the L2 Self* (Dörnyei & Ushioda 2009) zusammengestellt. Dörnyeis *L2 Motivational Self System* (vgl. Dörnyei 2009a) und die sich daran orientierenden Untersuchungen nehmen in der Regel L2-Lerner_innen während des Prozesses des Fremdsprachenlernens in den Blick (vgl. bspw. Taguchi, Magid & Papi 2009, Ryan 2009). Die für den Fragebogen ausgewählten Items wurden aus dem Englischen übersetzt, entsprechend dem vorliegenden Kontext modifiziert – bspw. wurde *Englisch lernen* durch *Deutsch lernen* ersetzt – und in die Vergangenheitsform gebracht, da es bei den hier untersuchten Proband_innen nicht um den aktuellen Spracherwerbsprozess geht, sondern um eine Retrospektive.

Im Folgenden werden die Items thematisch gegliedert dargestellt und mit Nachweisen versehen. Für den Fragebogen wurden die Items mithilfe von *Microsoft Excel für Mac 2011, Version 14.2.2* randomisiert. Dabei stand den Testpersonen für die Beantwortung eine vierstufige Likert-Skala (1 = trifft sicher nicht zu bis 4 = trifft sicher

zu) zur Verfügung, im Unterschied zu den Studien, denen die Skalen entnommen sind (fünf- bzw. sechsstufige Skalen). Die Reduzierung der fünfstufigen Likert-Skalen wurde vorgenommen, damit die Testpersonen eine Entscheidung treffen müssen (*forced choice*). Die Reduzierung von sechs auf vier geschah, um die Testpersonen nicht mit sehr feinen Abstufungen zu konfrontieren, die eventuell nur schwer zu unterscheiden sind.⁵⁸

Items auf der Grundlage der *Integrativeness*

- War/ist es Ihnen wichtig, deutsche Filme zu verstehen? (55)⁵⁹
- War/ist es Ihnen wichtig, deutsche Literatur zu lesen? (28)
- Gefällt es Ihnen, Deutsch zu sprechen? (23)
- Fühlen Sie sich mit Ihren deutschen Sprachkenntnissen respektiert? (44)⁶⁰

Items auf der Grundlage des *Ideal L2 Self*

- Als ich damals Deutsch-Unterricht hatte, konnte ich mir vorstellen, in der Zukunft Deutsch zu verwenden. (46)
- Als ich damals Deutsch-Unterricht hatte, stellte ich mir oft vor, dass ich jemand bin, der Deutsch fließend verwenden kann. (35)
- Als ich damals Deutsch-Unterricht hatte, konnte ich mir vorstellen, es mit deutschen Freunden zu sprechen. (52)
- Als ich damals Deutsch-Unterricht hatte und über meine Zukunft nachdachte, wusste ich, dass Deutschkenntnisse wichtig sind. (21)
- Als ich damals Deutsch-Unterricht hatte, wusste ich, dass ich Deutsch für meine Zukunft brauchen würde. (50)⁶¹
- Mein Ziel war es, so gut Deutsch zu beherrschen wie eine Person, die mit Deutsch aufgewachsen ist. (18)⁶²

⁵⁸ Vgl. auch die Diskussion in Grotjahn (2004: 3f.) zur Ökonomie einer vierstufigen Skala und zur Vermeidung der Tendenz zur Mitte.

⁵⁹ Die Zahlen in Klammern nach den Items geben die tatsächliche Stelle im Fragebogen an.

⁶⁰ Die Fragen sind dem Anhang zur Untersuchung von MacIntyre et al. (2009: 213–214) entnommen.

⁶¹ Item 46, 35, 21, 50 und 52: vgl. Ryan (2009: 143). Bei Item 21 und 52 sind die Übergänge zur *Instrumentality (Promotion)* fließend.

⁶² Vgl. Taguchi et al. (2009: 91).

Items auf der Grundlage von *Intended Learning Effort/Motivated Behaviour and Effort*

- Ich habe hart gearbeitet, um Deutsch zu lernen. (22)
- Es war mir extrem wichtig, Deutsch zu lernen. (30)
- Als ich damals Deutsch-Unterricht hatte, war das Deutsch-Lernen einer der wichtigsten Aspekte meines Lebens. (32)
- Als ich damals Deutsch-Unterricht hatte, habe ich immer die Hausaufgaben gemacht. (17)
- Als ich damals Deutsch-Unterricht hatte, habe ich versucht auch durch deutsche Musik, deutsches Fernsehen etc. Deutsch zu lernen. (41)
- Als ich damals Deutsch-Unterricht (in Deutschland) hatte, habe ich auch außerhalb des Klassenzimmers versucht, viel Deutsch zu sprechen. (16)
- Wenn es einen Deutschkurs gäbe, der mich auf meinem Niveau noch verbessern könnte, würde ich daran teilnehmen. (43)
- Wenn ich ein deutsches Lied im Radio höre, höre ich oft sorgfältig zu und versuche jedes Wort zu verstehen. (36)
- Ich kann ehrlich sagen, dass ich mein bestes gegeben habe, um Deutsch zu lernen. (24)
- Ich sehe oft deutsches Fernsehen, um meine Sprachkenntnisse zu verbessern. (33)
- Ich bin ein Typ, der sich sehr anstrengt, um Deutsch zu lernen. (5)⁶³

Items auf der Grundlage *L2 Self-Confidence*

- Als ich damals Deutsch-Unterricht hatte, war ich mir sicher, dass ich eine Fremdsprache lernen kann. (26)⁶⁴
- Als ich damals Deutsch-Unterricht hatte, befürchtete ich, dass die anderen Schülerinnen und/oder Schüler über mich lachen könnten, wenn ich Deutsch sprach. (51)
- Eine Fremdsprache zu lernen ist eine schwierige Aufgabe für mich. (48)

⁶³ Vgl. Al-Shehri (2009): 170–171 und Ryan (2009): 143.

⁶⁴ Dieses Item ist im Unterschied zu den anderen ItemItems dieser Rubrik positiv, bei der Auswertung muss dies entsprechend berücksichtigt werden und die angegebenen Werte werden umgekehrt.

- Ich glaube, dass ich der Typ bin, der sich schnell ängstlich und unwohl fühlt, wenn ich mit jemandem in der Fremdsprache sprechen muss. (27)
- Als ich damals Deutsch-Unterricht hatte, fühlte ich immer, dass meine Klassenkameradinnen und/oder Klassenkameraden besser waren als ich. (40)⁶⁵

Items auf der Grundlage *Ought-to L2 Self*

- Ich habe Deutsch gelernt, weil enge Freundinnen und/oder Freunde von mir dachten, dass das wichtig ist. (19)
- Ich musste Deutsch lernen, weil ich sonst meine Eltern enttäuscht hätte. (29)
- Deutsch lernen war wichtig, weil mein Umfeld das von mir erwartet hat. (38)
- Meine Eltern glauben, dass Fremdsprachenkenntnisse der Schlüssel zum Erfolg sind. (53)
- Deutsch lernen war wichtig für mich, um die Anerkennung meiner Freundinnen und/oder Freunde zu gewinnen. (42)
- Deutsch lernen war wichtig für mich, um die Anerkennung meiner Lehrer_innen zu gewinnen. (45)
- Deutsch lernen war wichtig für mich, um die Anerkennung meiner Familie zu gewinnen. (25)
- Deutsch lernen war wichtig für mich, um die Anerkennung meiner Chefin und/oder meines Chefs zu gewinnen. (47)⁶⁶

Items auf der Grundlage *Instrumentality (Promotion)*

- Deutsch lernen war wichtig für mich, weil ich glaubte, dadurch eines Tages einen guten Job zu bekommen. (37)
- Deutsch lernen war wichtig für mich, weil ich dachte, dass mit einem hohen Sprachniveau viel Geld zu verdienen ist. (31)
- Deutsch lernen war wichtig für mich, weil ich dachte, dass es eines Tages nützlich wäre, um einen guten Job zu bekommen oder viel Geld zu verdienen. (20)

Items auf der Grundlage *Instrumentality (Prevention)*

⁶⁵ Vgl. Ryan (2009): 143.

⁶⁶ Vgl. Taguchi et al. (2009): 92.

- Ich lernte Deutsch, weil ich sonst meinen Universitäts-Abschluss nicht bekommen hätte. (39)
- Ich lernte Deutsch, weil ich keine schlechten Noten bekommen wollte. (49)
- Deutsch lernen war mir wichtig, weil ich ohne gute Noten in Deutsch als schwache_r Student_in gegolten hätte. (34)⁶⁷

Der Fragebogen wurde von fünf Personen pilotiert, vier davon sind weiblich, eine männlich. Vier sprechen Deutsch als Fremdsprache mit AO = 6, 18 (n = 2) oder 19 und den Ausgangssprachen Arabisch, Chinesisch und Russisch (n = 2). Eine Person ist mit den Sprachen Deutsch und Griechisch aufgewachsen (2L1). Die Pilotierung ergab verschiedene Änderungen wie bspw. die Tilgung von schwierigem Wortschatz, die Zusammenfassung von Items und die ersatzlose Streichung eines Items.⁶⁸

3.5 Durchführung der Untersuchung

Nach der Darstellung der verwendeten Instrumente im vorangehenden Unterkapitel folgt nun die Beschreibung des zeitlichen und örtlichen Rahmens des Forschungsprojekts, weiterer Bedingungen und der Vorgehensweise bei der Proband_innen-Suche.

3.5.1 Rahmenbedingungen

Die Datenerhebungen der ersten Phase des Forschungsprojektes fanden im Zeitraum vom 23.01.2013 bis 23.11.2013 statt. Die Länge des Zeitraums erklärt sich durch die Herausforderungen bei der Suche nach Testpersonen (vgl. Kap. 3.5.3 *Proband_innen-Suche*). Die Datenerhebungen dauerten zwischen 60 und 90 Minuten, je nachdem wie schnell die Proband_innen sich durch die Tests und Skalen durcharbeiteten.⁶⁹ Die Datenerhebungen fanden in einem Büroraum der TU Berlin

⁶⁷ Alle ItemItems zu *Instrumentality (Promotion und Prevention)* vgl. Taguchi et al. (2009: 93–95).

⁶⁸ Das Item „Verhalten Sie sich „deutsch“?“ wurde bei der Pilotierung kritisiert. Es ist unklar, was deutsches Verhalten in diesem Zusammenhang sein soll: deutsche Feste feiern, deutsche Verhaltensregeln befolgen? Deutsches Essen essen, Bier trinken? etc.

⁶⁹ Die zeitliche Differenz erklärt sich v. a. durch das Bearbeiten der sechs C-Tests ohne Zeitlimitierung. Es sind dafür 30 Minuten vorgesehen, es gab aber TN, die das gesamte Set in zwölf Minuten bearbeiteten.

statt, den Räumlichkeiten von chinesischen Samstagsschulen und in zwei Fällen auf Wunsch der Testpersonen auch in ihren jeweiligen Büroräumen. Die meisten Datenerhebungen bei den Späterwerber_innen wurden in Einzelsitzungen durchgeführt; bei den Personen mit 2L1 konnten die schriftlichen Tests in Gruppen und die Audioaufzeichnungen in Dyaden durchgeführt werden.

Die Datenerhebungen der zweiten Phase fanden vom 01.03.2014 bis zum 25.03.2014 statt. Dabei gab es ausschließlich Einzeltermine. Die Datenerhebungen dauerten zwischen 45 und 60 Minuten und fanden entweder in einem Büroraum der TU Berlin, in einer der Samstagsschulen oder, in einem Fall, im Büro der Testperson statt. Die Durchführung der LLAMA-Tests musste intensiv angeleitet und begleitet werden.

In den Feedback-Gesprächen mit den Proband_innen nach den Datenerhebungen bekundeten die meisten, dass ihnen die Teilnahme an der Untersuchung trotz des hohen Zeitaufwands Freude gemacht habe.

Die Daten wurden mit informierter Einwilligung (*informed consent*) der Proband_innen erhoben. Das bedeutet, dass die Ziele der Untersuchung, Dauer, Belastung und Risiko sowie der Umgang mit den Daten und Ergebnissen offengelegt werden und die TN auf die Anonymisierung ihrer Daten und die Freiwilligkeit ihrer Teilnahme hingewiesen werden (Miethe 2010: 929). Die TN hätten ihre Teilnahme jederzeit verwehren bzw. abbrechen können.

3.5.2 Proband_innen-Suche

Eine Herausforderung dieses Dissertationsprojektes stellte die Proband_innen-Suche dar. Es galt jeweils 20 Personen zu finden, die Mandarin und Deutsch als 2L1 sprechen oder Mandarin als L1 sprechen, ab dem Alter von 16 Jahren Deutsch gelernt haben, über einen akademischen Bildungshintergrund verfügen und mindestens schon fünf Jahre im deutschsprachigen Raum leben. Die Suche nach den zehn Personen, die als Kontrollgruppe beim Audio-Rating fungierten, war demgegenüber vergleichsweise einfach.

Der erste Schritt war die Veröffentlichung des Proband_innen-Gesuchs (s. Anhang) in einem Forschungskolloquium, an dem auch Doktorand_innen aus China teilnah-

men.⁷⁰ Auf diese Weise wurden die ersten Testpersonen gefunden, die das Forschungsprojekt nicht nur durch ihre Bereitschaft zur Datenerhebung unterstützten, sondern auch in ihrem Bekanntenkreis für eine Teilnahme an der Untersuchung warben und den Kontakt zu weiteren Testpersonen anbahnten. Weitere Schritte bei der Proband_innen-Suche waren die Kontaktaufnahme mit chinesischen Vereinen und Institutionen, die Veröffentlichung der Suche unter Kolleg_innen und Studierenden der TU Berlin und Aushänge an den drei Berliner Universitäten. Per E-Mail oder Postbrief wurden von mir die folgenden Vereine und Institutionen angeschrieben: Gesellschaft für chinesische Akademiker e. V. – Vereinigung an der TU Berlin, Studentischer Verein zum deutsch-chinesischen Kulturaustausch, Gesellschaft für deutsch-chinesische Freundschaft e. V., Das China Büro e. V., Berlin China Cultural Bridges e. V., Chinesisches Kulturzentrum Berlin, Konfuzius Institut an der Freien Universität Berlin sowie die Bildungsabteilung der Chinesischen Botschaft Berlin. Auf diese schriftlichen Anfragen gab es in einem Fall den Rücklauf, dass man aufgrund des Aufgabenprofils keinen Kontakt zu den gesuchten Personengruppen habe; in allen anderen Fällen fand gar kein Rücklauf statt. Die Reaktion auf die Aushänge war gering, denn darauf meldete sich nur eine Testperson. Es zeichnete sich ab, dass ohne persönliche Begegnung oder Beziehung die Proband_innen-Suche nicht gelingen würde. Ein Grund dafür könnte der relativ komplexe Ausschreibungstext sein. Mehrere Testpersonen meldeten zurück, dass sie erst nach der mündlichen Einführung verstanden hätten, um was es gehe.

Zielführend war schließlich die Kontaktaufnahme mit drei verschiedenen chinesischen Samstagsschulen. Diese Ergänzungsschulen sind in der Regel von chinesischen Einwander_innen gegründet worden mit dem Ziel, ihren in Deutschland aufwachsenden Kindern Mandarin resp. Kantonesisch in Wort und Schrift zu vermitteln. Im Rahmen dieses Forschungsprojekts wurde mit der HuaDe China Schule e. V., der Yi-Zhi-Chinesische-Kulturschule (Projekt des Deutsch-Chinesischen Vereins für Kultur und Bildung e. V.) und der Chinesischen Schule Berlin e. V. zusammengearbeitet. Nach der Kontaktaufnahme mit der Leitung und/oder einzelnen Lehrer_innen konnte das Forschungsprojekt in den entsprechenden Klassen vorgestellt werden. Neben der Möglichkeit, an diesen Schulen Personen mit 2L1 zu finden, konnten dort

⁷⁰ In diesem Zusammenhang möchte ich Prof. Dr. Steinmüller danken, der die Proband_innen-Suche unterstützt hat.

auch deren Eltern und Lehrer_innen angesprochen werden, die größtenteils als Erwachsene Deutsch lernten.

Ursprünglich sollten in der Gruppe der Personen mit 2L1 ausschließlich Personen getestet werden, die mit Deutsch und Mandarin aufgewachsen waren. Aus pragmatischen Gründen wurde das Profil um Kantonesisch erweitert, um das angestrebte Minimum von 20 TN zu erreichen. Diese Erweiterung ist legitim, weil Mandarin und Kantonesisch, das zu den südchinesischen Yuè-Dialekten gehört, über dieselben typologischen Merkmale verfügen (vgl. die ausführliche Beschreibung in Norman 1988: 210–221). Norman zufolge stellen die Yuè-Dialekte eine Subvarietät des Südchinesischen dar mit einer starken Überlagerung von Elementen aus dem Nordchinesischen (vgl. ebd.: 221). Kantonesisch, Mandarin und andere chinesische Sprachen unterscheiden sich nicht stärker voneinander als etwa die einzelnen romanischen Sprachen.⁷¹

Auch wurde das Mindestalter von 18 auf 16 Jahre herabgesetzt; damit einhergehend wurde statt eines akademischen Bildungshintergrundes der Besuch eines Gymnasiums bzw. der Sekundarstufe II erwartet. Dies war nötig durch den Umstand, dass die Datenerhebungen v. a. an den chinesischen Samstagsschulen durchgeführt wurden, die in der Regel bis zum Abitur besucht werden. Dieses pragmatische Vorgehen legitimiert sich durch die Auswertung der ersten Daten von 16-Jährigen, deren Leistungsspektrum sich statistisch nicht von dem der über 18-Jährigen unterschieden.

Das *Sampling* für die zweite Phase sollte ursprünglich als gezielte Auswahl (*purposeful sampling*) durchgeführt werden (Patton 2002: 230): Extremfälle – oberes und unteres Leistungsspektrum – und Fälle, die bezüglich der erhobenen Kompetenzen in der L2 den Durchschnitt repräsentieren (je zwei Fälle) sowie je zwei Fälle mit maximaler Variation bzgl. der Motivations- und Extraversionsprofile und der standardisierten Befragung zu den Lehr- und Lernkontexten. Da der Rücklauf bei diesem gezielten *Sampling* nur bei 50 % lag, wurden schließlich alle 20 TN aus der ersten Phase angeschrieben und alle getestet, die sich zurückmeldeten. Dabei ergab sich, dass sich das Leistungsspektrum gut in dieser Auswahl repräsentierte. Diese Art von

⁷¹ Diesen Vergleich prägte Dr. Andreas Guder, Leiter des Studienbereichs Chinesische Sprache am ostasiatischen Seminar der FU Berlin, in einem Gespräch am 24.06.2013.

selbstselektiver Stichprobe deutet auf eine hohe Motivation der Teilnehmer_innen hin, die an der zweiten Phase teilnahmen.

4. Ergebnisse

Nach der Vorstellung des Forschungsdesigns, der Instrumente, der Testpersonen und dem Einblick in den tatsächlichen Forschungsprozess werden nun die Ergebnisse präsentiert. Zunächst werden die Testformate der Phase I nach Instrumenten getrennt einer deskriptiven Analyse unterzogen. Um weitergehende Aussagen treffen zu können, werden Korrelationen berechnet. Das Phonetik-Rating wird zuletzt analysiert, weil die Daten nur für eine Teilgruppe der Proband_innen vorliegen (vgl. Kap. 3.4 *Forschungsdesign*). Danach erfolgt die Auswertung der Phase II, also des Fragebogens zu persönlichen Hintergrundinformationen, der Motivationskalen und der LLAMA-Tests. Auf Grundlage dieser Analysen werden die Forschungsfragen beantwortet, in einem abschließenden Schritt werden die gefundenen Ergebnisse diskutiert.

Inferentielle Statistik erlaubt es, Ergebnisse, die aus der Analyse der erhobenen Daten der Stichprobe resultieren, auf eine größere Population zu beziehen. Sie ist eigentlich nur zulässig, wenn die Stichprobe zufällig ausgewählt worden ist. Dies ist jedoch im Rahmen der Studien zum Lehren und Lernen von Fremdsprachen fast nie der Fall. Um zu vermeiden, dass man nur wenig aus den erhobenen Daten folgern kann und bei der deskriptiven Statistik bleiben muss, dürfen Generalisierungen nur mit äußerster Vorsicht formuliert werden (vgl. Larson-Hall 2010: 60).

Laut Vanhove (2013) bieten Regressionsmodelle eine valide Möglichkeit, um Aussagen über die Beziehung zwischen AO und UA treffen zu können. Er argumentiert, dass Regressionsmodelle den Berechnungen von Korrelationen und Mittelwerten überlegen sind:

„If we want to find out more about the relationship between AOA and UA, why throw away most of the AOA information and effectively reduce the UA data to group means and the variance in those groups?“ (Vanhove 2013: 6)

In der hier vorliegenden Arbeit war ursprünglich geplant, multiple Regressionen zu berechnen, allerdings ist die Anwendung parametrischer Verfahren aus mehreren Gründen problematisch: Die hier vorliegenden Daten sind zunächst einmal nicht normalverteilt. Bezüglich der Überprüfung der Normalverteilung der Daten wird empfohlen, von Testverfahren wie bspw. Kolmogorov-Smirnov- und Shapiro-Wilk-Test

Abstand zu nehmen, weil sie nicht immer exakt bestimmen können, ob die Daten normalverteilt sind oder nicht (vgl. Larson-Hall 2010: 74), oder die Testverfahren mit der Betrachtung von Streu-, Boxplot- oder Balkendiagrammen zu kombinieren.

Führen wir den Kolmogorov-Smirnov-Test durch, erhalten wir für die C-Tests und GJTs der ‚Muttersprachler_innen‘ und der Personen mit doppeltem Erstspracherwerb Werte von $p = .000$. Das bedeutet, dass die Daten nicht normalverteilt sind. Für die S-C-Tests werden für beide Gruppen p-Werte über $.05$ berechnet. Bei den Späterwerber_innen werden für alle drei Tests Werte über $.05$ berechnet. Das heißt, dass hier normalverteilte Daten vorliegen könnten. Diese Ergebnisse lassen sich durch eine Augenscheinanalyse der Boxplot-Grafiken bestätigen: Bei den Boxplots der laut Berechnung nicht normalverteilten Testergebnisse sind die Boxen und *Whiskers* nicht symmetrisch, d. h. die Daten sind schief verteilt (vgl. Abb. 11, 14, 15). Die nicht normalverteilten Daten der Gruppen der ‚Muttersprachler_innen‘ und Personen mit 2L1 bei den C-Tests und dem GJT sind als Deckeneffekte zu bewerten. Sie zeigen, dass bei äußerst kompetenten Sprecher_innen allein die S-C-Tests ein Testformat darstellen, das genügend differenzieren kann (zur Konstruktion der S-C-Tests vgl. Kap. 3.4.2.2 *C-Tests*). Gegen die Anwendung parametrischer Verfahren spricht des Weiteren, dass die Daten nicht unabhängig gesammelt wurden, die Einteilung in Gruppen vorher festgelegt wurde und die Gruppengrößen relativ klein sind. Darüber hinaus handelt es sich bei etlichen der erhobenen Daten um ordinalskalierte Daten.

Aus diesen Gründen wurden Spearman-Rho-Ranglistenkorrelationen und nicht Pearson-Korrelationen berechnet und deswegen wurden keine multiplen Regressionsanalysen vorgenommen, die sich eigentlich angeboten hätten, da es um mehrere Einflussvariablen geht.

Mit R anstelle von SPSS hätten sich aussagekräftigere Verfahren berechnen lassen. Vanhove stellt sein R-Script zur Verfügung, mit dem er Reanalysen der Studie von DeKeyser, Alfi-Shabatay und Ravit (2010) berechnet. Den Autor_innen zufolge sprechen ihre Ergebnisse aus zwei parallelen Studien für die CPH, und zwar mit Proband_innen, die mit der Ausgangssprache Russisch entweder Englisch als L2 in Nordamerika oder Hebräisch in Israel erwarben. Beide Gruppen weisen jeweils einen Kurvenverlauf von AOA und UA auf, der nicht linear ist, sondern ab einem be-

stimmten Punkt zwischen Pubertät und Erwachsenenalter flach ausläuft. Des Weiteren ist die Beziehung zwischen Fremdsprachenlerneignung und UA bei den jüngeren und älteren Einwander_innen unterschiedlich und nur bei den älteren signifikant. Diese Ergebnisse halten jedoch Vanhoves Reanalyse nicht stand. Er stellt fest, dass die AOA-UA-Kurve auch nach dem vermeintlichen Ende der CP stark abfällt und dass die beiden Proband_innen-Gruppen erheblich divergierende „Umschlagpunkte“ (AOA = 16 und AOA = 6) in ihren Kurvenverläufen haben (vgl. Vanhove 2013: 13), was ihm zufolge beides gegen die CPH spricht.

Der Hinweis auf die Möglichkeiten der statistischen Auswertung mit R erfolgte zu einem Zeitpunkt, an dem die Datenaufbereitung mit SPSS bereits fertiggestellt und die Datenanalyse schon weit fortgeschritten war. Die ausführliche Dokumentation der Streudiagramme und Boxplots geschieht mit der Absicht, einer Reduktion der Daten entgegenzuwirken.

4.1 Auswertung Phase I

Die erhobenen Daten wurden mit *IBM SPSS Statistics, Version 22* ausgewertet.

4.1.1 Grammatical Judgement Tests

Die Testpersonen bewerten die Sätze in zwei Schritten: Zuerst entscheiden sie, ob der Satz korrekt ist oder nicht, danach geben sie an, ob sie sich sicher oder unsicher sind bei ihrem Urteil. Für die Auswertung wurden diese Urteile zusammengezogen, da es Studien gibt, die zeigen, dass sich die Proband_innen bei der Angabe über die Sicherheit ihres Urteils und auch beim Grammatikalitätsurteil oft selbst widersprechen. Ellis (1991: 177) zeigt: Wenn man mit einer Proband_innen-Gruppe einen GJT nach einer Woche wiederholt und die Proband_innen dabei laut denken lässt, stimmen die schriftlichen Testergebnisse aus der 1. Woche mit den Laut-Denken-Protokollen in 22,5 % der Fälle nicht überein. Dabei bezieht sich jeweils die Hälfte der Fälle auf die Entscheidung korrekt vs. nicht korrekt und auf die Sicherheit, mit der das Urteil gefällt wurde. Der Genauigkeit wegen muss erwähnt werden, dass die von Ellis eingesetzte Skala im selben Schritt die drei Möglichkeiten grammatisch, ungrammatisch, unsicher offerierte.

Bevor Entscheidungen über die zur Anwendung kommenden Testverfahren getroffen werden, sollen die erhobenen Daten deskriptiv ausgewertet werden. Dabei ist v. a. die grafische Aufbereitung aufschlussreich (vgl. Larson-Hall 2010: 75).

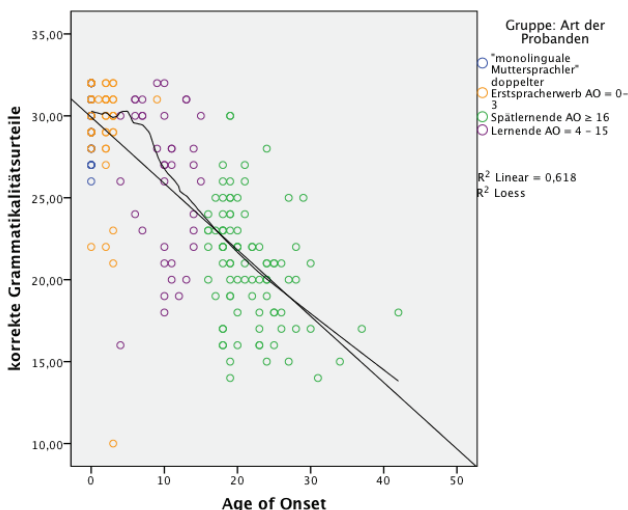


Abb. 10: Streudiagramm AO und korrekte Grammatikalitätsurteile (theoretisches Maximum = 32). In das Streudiagramm wurden ebenfalls die Fälle mit AO = 4–15 ($n = 53$) einbezogen, die vorliegen, aber für die Forschungsfragen dieser Arbeit keine Rolle spielen.

Ein Blick auf das Streudiagramm (vgl. Abb. 10) zeigt, dass es einen Alterseffekt bezüglich der GJT-Ergebnisse gibt: Die geometrische Regressionslinie – hier zu deskriptiven Zwecken berechnet – verläuft von links oben nach rechts unten, was einen negativen Zusammenhang von AO und korrekten Grammatikalitätsurteilen bedeutet. Die Loess-Kurve (*local polynomial regression*), die nicht auf einen linearen Verlauf festgelegt ist, passt sich den tatsächlichen Variablen-Werten besser an und ermöglicht zu untersuchen, ob die Daten in einem linearen Zusammenhang stehen (vgl. Larson-Hall 2016: 193). In der hier vorliegenden Analyse weicht sie von der linearen Regressionslinie bis zu AO = 10 deutlich nach oben ab, um sich ihr später fast anzugleichen. Gleichzeitig sehen wir, dass die Daten weit streuen, d. h., dass sich die Ergebnisse individuell stark unterscheiden, und zwar bezogen auf jedes Anfangsalter (AO).

Tab. 7 zeigt die GJT-Ergebnisse (Anzahl der korrekten Urteile) differenziert nach den Gruppen ‚Muttersprachler_innen‘, Personen mit 2L1 und Spätler_innen.

Tab. 7: Deskriptive Auswertung der GJT-Ergebnisse.

Gruppe	N	Min	Max (max. 32)	Median	M	SD
‚Muttersprachler_innen‘	80	26	32	30	30,1500	1,55978
Personen mit 2L1	49	21	32	30	29,6122	2,65217
Spätler_innen	89	14	30	21	20,8652	3,60615

Anmerkung: M = arithmetisches Mittel; SD = Standardabweichung

Die Darstellung der GJT-Ergebnisse als Boxplot-Grafik aufgeteilt nach Gruppen gibt uns weiteren Aufschluss (vgl. Abb. 11). Schon jetzt sehen wir, dass sich die Personengruppe mit doppeltem Erstspracherwerb aller Wahrscheinlichkeit nach nicht signifikant von der Gruppe ‚Muttersprachler_innen‘ unterscheidet. In beiden Gruppen liegt der Median bei 30 Punkten und die Streuweite der Daten ist ähnlich. Allerdings gibt es in der Gruppe der Personen mit doppeltem Erstspracherwerb einen Ausreißer und vier Extremwerte.

Für den Umgang mit Ausreißern und Extremwerten gibt es keine allgemeingültige Vorgehensweise. Ausreißer (in Abb. 11 dargestellt als Kreis) liegen zwischen dem 1,5-fachen und dem 3-fachen des Interquartilsabstands.⁷² Extremwerte (in Abb. 11 dargestellt als Stern) befinden sich außerhalb des 3-fachen Interquartilsabstands und liegen somit weit außerhalb des Bereichs einer unauffälligen Streuung.

Das Vorgehen im Rahmen dieser Arbeit ist, bei den Extremwerten und Ausreißern jeweils in den Datensätzen und vorliegenden Unterlagen nachzuschauen, ob etwas auffällig ist oder ein Messfehler vorliegt, weswegen der Fall eventuell nicht mit ausgewertet werden sollte. Sollte diese nähere Betrachtung aber ein unauffälliges Ergebnis erzielen, werden die Ausreißer und Extremwerte in die Analyse miteinbezogen.

⁷² „Interquartilsabstand ist definiert als Q3-Q1“ (Gültekin-Karakoç & Feldmayer 2014: 191). Das entspricht 50 % der Daten.

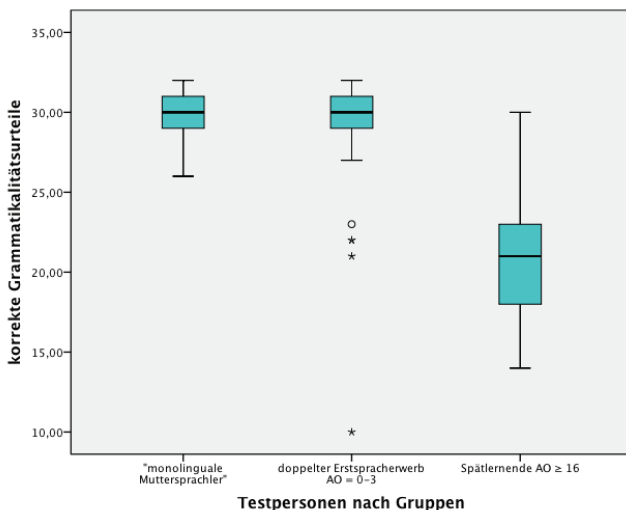


Abb. 11: Boxplot Grammatikalitätsurteile nach Gruppen.

Der auffälligste Extremwert aus der Gruppe der Personen mit 2L1 mit nur 10 korrekten Antworten sollte tatsächlich aus den weiteren Analysen ausgeschlossen werden, da ein Blick in die Unterlagen ergab, dass der GJT von dieser Person nur zur Hälfte ausgefüllt wurde (Items 1–16). Warum der Test danach abgebrochen wurde, ist nicht klar, denn es gibt diesbezüglich keine Anmerkung in den Unterlagen. Bei den anderen drei Extremwerten sowie dem Ausreißer zeigen sich keine Auffälligkeiten; folglich werden sie in die GJT-Auswertungen miteinbezogen.⁷³

Die deskriptive Analyse ergibt, dass die Ergebnisse der Spätlerner_innen viel weiter streuen als die der anderen Gruppen. Zudem können wir sehen, dass ca. 50 % der Späterwerber_innen im Bereich der Personen mit 2L1 abgeschlossen haben, wenn wir die Extremwerte (ohne den ausgeschlossenen Fall) berücksichtigen.

Ein Nebeneffekt dieser GJT-Auswertung ist, dass sie die Wichtigkeit der Reproduktion von Studien unterstreicht. Bei Sopata (2004), deren Studie der Test entnommen ist, gibt es eine Kontrollgruppe ‚Muttersprachler‘ (n = 25), die die Items des GJT 100 % korrekt löst. In der vorliegenden Studie sind es nur 23,7 % der ‚Muttersprach-

⁷³ Es gibt zwei Testpersonen, die 22 korrekte Antworten hatten; auf der Boxplot-Grafik ist aber nur ein Stern zu sehen, der etwas dunkler dargestellt ist.

ler_innen' und 18,37 % der Personen mit 2L1, die den Test vollständig korrekt lösen im Sinne der Lösungsschablone. Angesichts der Kritikwürdigkeit mancher Items (vgl. Kap. 3.4.2.1 *Grammatical Judgement Test*) ist dies nicht überraschend bzw. sind eher die 100 % bei Sopata erstaunlich.

Von allen ‚Muttersprachler_innen‘ wurden die Items 10, 19, 20, 21, 23, und 24 korrekt gelöst. Die vier in Kap. 3.4.2.1 *Grammatical Judgement Test* diskutierten Items schneiden sehr unterschiedlich ab: Item 7 „Auf der Autobahn wird es sehr schnell gefahren“ wurde von 96,3 % zutreffend als nicht korrekt eingestuft, Item 21 „Während des Krieges wird es gelitten“ von 100 %. Die angeführten Bedenken scheinen also nicht für Irritationen gesorgt zu haben. Item 13 musste aus der Wertung herausgenommen werden: Ein Teil der Daten, die aus Bochum und Berlin übernommen wurden, wurde von wissenschaftlichen Hilfskräften mit der Original-Schablone ausgewertet, die den Satz als falsch wertet. Ein anderer Teil arbeitete mit einer Schablone, die den Satz als richtig und falsch akzeptiert. Item 26 („Der Alkohol begann, seine Gesundheit zu untergraben“) wurde nur von 66,3 % zutreffend als „korrekt“ beurteilt. Weitere fünf Items wurden unter 90 % zutreffend als jeweils „korrekt“ oder „nicht korrekt“ eingestuft.⁷⁴

Von den Personen mit 2L1 wurden die Items 1, 6, 9, 10, 19, 20 und 30 zu 100 % richtig eingestuft. Die im GJT-Kapitel problematisierten Items wurden folgendermaßen eingestuft: Item 7 (82 %), Item 21 (92 %), Item 26 (76 %). Insgesamt neun Items wurden von unter 90 % nicht zutreffend eingestuft.⁷⁵

Bei den Späterwerber_innen gibt es kein einziges Item, das von allen richtig beurteilt wurde. Um sich einen Eindruck zu verschaffen, welche Sätze für die Späterwerber_innen besonders schwer einzuordnen waren, wurden in Tab. 8 alle Items eingetragen, die von unter 60 % zutreffend eingestuft wurden. Die Tabelle zeigt deutlich, dass der korrekte Gebrauch von Expletiva für die Späterwerber_innen eine Schwierigkeit darstellt. Dieses Ergebnis entspricht der Annahme, dass es sich dabei um ei-

⁷⁴ Item 4 (87,5 %), Item 16 (68,8 %), Item 18 (88,5 %), Item 25 (88,8 %) und Item 29 (88,8 %). An dieser Stelle wird die jeweilige Problematik dieser Items nicht weiter diskutiert, weil der Focus dieser Arbeit ein anderer ist. Auf jeden Fall sollten diese Items aber bei einer neuerlichen Verwendung des GJTs geändert werden.

⁷⁵ Item 4 (86 %), Item 11 (88 %), Item 12 (84 %), Item 14 (88 %), Item 15 (84 %), Item 16 (82 %), Item 18 (82 %), Item 25 (80 %), Item 29 (86 %).

ne besondere Lernschwierigkeit der deutschen Sprache handelt und schwer implizit zu erwerben ist.

Tab. 8 Aufteilung der Items des GJT nach grammatischen Phänomenen, die von den Späterwerber_innen unter 60 % richtig als korrekt oder nicht korrekt eingestuft wurden.

Gruppe/Phänomen	Item -Nr. (Prozentsatz der Testpersonen, die Satz richtig als „korrekt“ bzw. „nicht korrekt“ eingestuft haben)
a) einfache Sätze mit referentiellem Subjekt	/
b) zusammengesetzte Sätze mit referentiellem Subjekt	/
c) einfache Sätze mit Expletivum	21 (19,9 %), 7 (34,8 %), 14 (49,5 %), 27 (50,6 %)
d) zusammengesetzte Sätze mit Expletivum im Passiv	24 (19,1 %), 29 (24,7 %), 18 (33,7 %), 12 (34,8 %), 15 (39,5 %), 11 (50,6 %)
d') wie d) aber aktivische Konstruktion	16 (42,7 %), 32 (46,1 %), 4 (58,4 %)
Adjektiv-Kongruenz	/
Tempus bzw. Infinitivkonstruktion	31 (55 %)
korrekte Sätze	23 (49,4 %), 26 (55 %)

Die Reliabilität des GJT, die mit Cronbachs Alpha geschätzt wurde, ergibt für sämtliche erhobenen Datensätze auf Basis der Lückenwerte ($n = 266^{76}$) $\alpha = .735$. Dieser Wert kann noch als gut bezeichnet werden. Die für die Gruppe der Späterwerber_innen ($n = 86$) geschätzte Testreliabilität fällt mit $\alpha = .649$ akzeptabel aus. Für die monolingualen Erstsprachler_innen ($n = 80$) liegt die geschätzte Reliabilität für die GJT bei $\alpha = .549$. Bei den Personen mit 2L1 ($n = 49$) ist $\alpha = .507$. Die Testreliabilität fällt für diese beiden Gruppen fragwürdig aus.

⁷⁶ Für diese Berechnung wurden sämtliche erhobenen Daten ausgewertet incl. der Fälle mit AO = 4–15.

4.1.2 C-Tests und S-C-Tests

Die C-Tests wurden differenziert ausgewertet, wobei sieben verschiedene Bewertungskategorien verwendet wurden:

- 0 = falsch
- 1 = Originallösung
- 2 = akzeptable Variante ohne Orthografiefehler
- 3 = Originallösung mit Orthografiefehler
- 4 = akzeptable Variante mit Orthografiefehler
- 5 = unklare Klassifizierung/Zweifelsfall
- 6 = nicht lesbar
- 7 = Kongruenz nicht erfüllt (ohne und mit Orthografiefehler)

Für die vorliegende Auswertung wurden alle Ergebnisse als richtig gewertet, die mit 1 oder 2 codiert wurden. Orthografiefehler wurden nicht zu den richtigen Ergebnissen gezählt, da sich bei der Pilotierung der S-C-Tests Folgendes gezeigt hat: Es werden nur wenige Orthografiefehler gemacht und die Zahl der Orthografiefehler korreliert negativ mit der Zahl der korrekten Lösungen, d. h., das leistungsstärkere Proband_innen weniger Orthografiefehler machen (vgl. Grotjahn, Schlak & Aguado 2010: 307).

In derselben Weise wie die GJT-Ergebnisse werden die S-C-Test- und C-Test-Ergebnisse aufbereitet, um eine erste deskriptive Analyse vorzunehmen. Die Streudiagramme zeigen ebenfalls einen klaren Alterseffekt, wobei wir zugleich eine enorme individuelle Ausdifferenzierung der Ergebnisse sehen (vgl. Abb. 12 und 13). In Abb. 13 deutet die Diskrepanz von linearer Regression und Loess-Kurve darauf hin, dass der Zusammenhang der Variablen eventuell nicht linear ist.

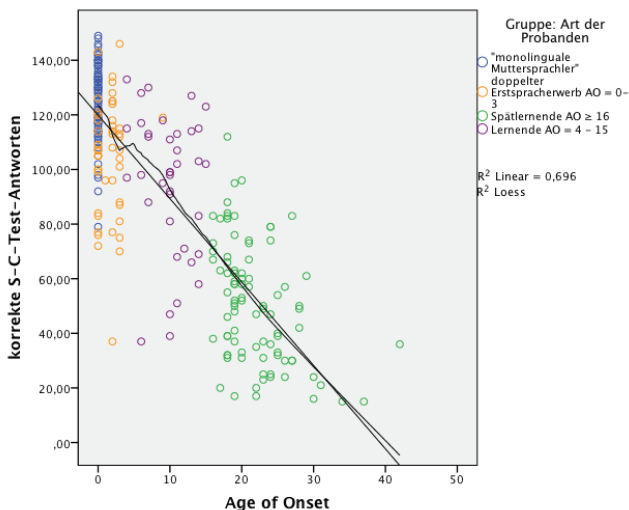


Abb. 12: Streudiagramm AO und korrekte S-C-Test-Antworten (theoretisches Maximum = 150). In das Streudiagramm wurden ebenfalls die Fälle mit AO = 4–15 ($n = 53$) einbezogen, die vorliegen, aber für die Forschungsfragen dieser Arbeit keine Rolle spielen.

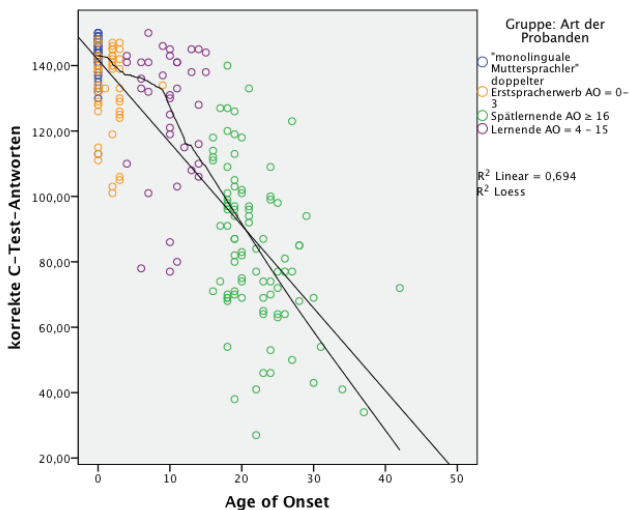


Abb. 13: Streudiagramm AO und korrekte C-Test-Antworten (theoretisches Maximum = 150). In das Streudiagramm wurden ebenfalls die Fälle mit AO = 4–15 ($n = 53$) einbezogen, die vorliegen, aber für die Forschungsfragen dieser Arbeit keine Rolle spielen.

Tab. 9 zeigt die S-C-Test-Ergebnisse (Anzahl der korrekten Urteile) differenziert nach den Gruppen ‚Muttersprachler_innen‘, Personen mit 2L1 und Späterwerber_innen.

Tab. 9: Deskriptive Auswertung der S-C-Test-Ergebnisse.

Gruppe	N	Min	Max (max.=150)	Median	M	SD
‚Muttersprachler_innen‘	80	79	149	129	126,2875	13,27336
Personen mit 2L1	50	37	146	107	104,9600	20,74329
Spätler_innen	89	15	112	50	50,7865	22,00024

Anmerkung: M = arithmetisches Mittel; SD = Standardabweichung

Tab. 10 zeigt die C-Test-Ergebnisse (Anzahl der korrekten Urteile) differenziert nach den Gruppen.

Tab. 10: Deskriptive Auswertung der C-Test-Ergebnisse.

Gruppe	N	Min	Max (max.=150)	Median	M	SD
‚Muttersprachler_innen‘	79	130	150	144	143,5316	5,07845
Personen mit 2L1	50	101	148	139	133,0000	12,68013
Spätler_innen	89	27	140	83	84,1124	24,23860

Anmerkung: M = arithmetisches Mittel; SD = Standardabweichung

Wieder gibt uns die Darstellung als *Boxplot*-Grafik Aufschluss über die Verteilung der Daten, ihre Streuweite, Ausreißer und Extremwerte (vgl. Abb. 14 und 15).

Bei dem Ausreißer in der Gruppe der ‚monolingualen Muttersprachler_innen‘ handelt es sich um eine Person, die v. a. mit der Zeitbeschränkung Schwierigkeiten hatte, denn in allen sechs Tests bleiben die letzten Lücken leer, in einzelnen Tests bis zu 50 % der Lücken. Für ein Zeitproblem spricht auch der Umstand, dass die Versuchsperson den C-Test ohne Zeitlimitierung mit 146 (von 150) korrekten Antworten sehr gut löste. Der Ausreißer unter den Personen mit 2L1 scheint ein Zeitdruck- und ein sprachliches Problem zu haben, denn das C-Test-Ergebnis ohne Zeitlimitierung

ist mit 101 Punkten ebenfalls ein Ausreißer. Da in beiden Fällen die Daten jedoch sonst keine Auffälligkeiten aufweisen, spricht nichts dagegen, die Ausreißer in die Analyse miteinzubeziehen.

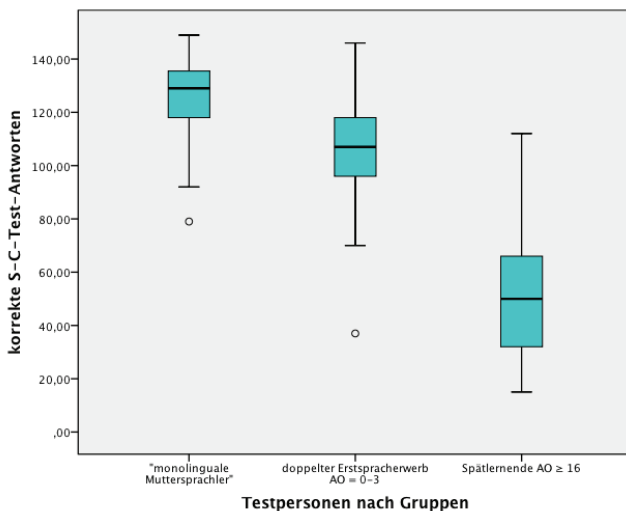


Abb. 14: Boxplot korrekte S-C-Test-Antworten nach Gruppen.

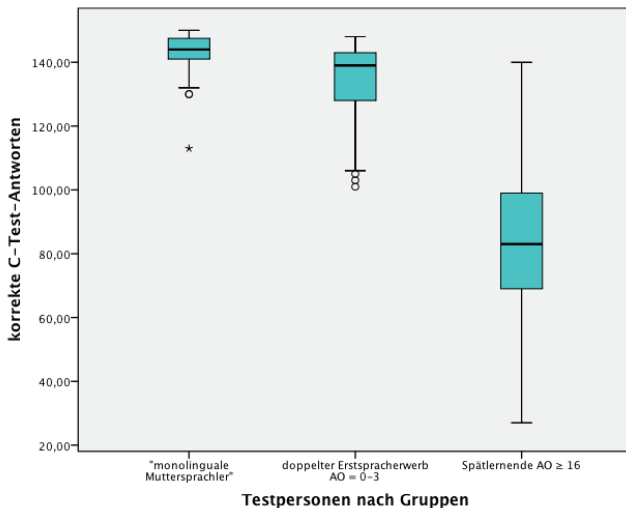


Abb. 15: Boxplot korrekte C-Test-Antworten nach Gruppen.

In der Gruppe der ‚monolingualen Muttersprachler_innen‘ gibt es zwei Ausreißer, die nur knapp außerhalb des unteren *Whiskers* liegen, und einen Extremwert. Die Datensätze der punktgleichen Ausreißer sind unauffällig, die beiden erzielten beim S-C-Test ebenfalls die gleiche Punktzahl und liegen mit 117 Punkten am oberen Ende des unteren *Whiskers*. Ein genauer Blick auf den Datensatz des Extremwerts ergibt, dass der sechste C-Test komplett nicht bearbeitet wurde. Der Grund dafür lässt sich im Nachhinein nicht mehr klären. Dieser Fall wird von den weiteren C-Test-Analysen ausgeschlossen. Bei den Personen mit 2L1 gibt es drei Ausreißer, wobei einer davon die Person ist, die auch beim S-C-Test Ausreißer ist, während die beiden anderen im Bereich des unteren *Whiskers* abgeschnitten haben. Dennoch gilt für alle drei, dass die Datensätze vollständig und unauffällig sind, sodass sie folglich in die C-Test-Analysen miteinbezogen werden.

Die Boxplot-Grafiken für die S-C-Tests und C-Tests lassen beide vermuten, dass hier die Unterschiede zwischen allen Gruppen signifikant sind. Des Weiteren können wir sehen, dass bei den S-C-Tests fast 75 % der Späterwerber_innen im Bereich der Personen mit 2L1 abgeschnitten haben, wenn wir den Ausreißer als Minimalwert akzeptieren. Wenn wir den Ausreißer bei den Personen mit 2L1 nicht berücksichtigen, sind es fast 25 %. Keine_r der Späterwerbenden schneidet im Bereich des ersten Quartils der Personen mit 2L1 ab, d. h. niemand aus dieser Gruppe erreicht die Maximalleistungen der Personen mit 2L1. Bei den C-Tests sind es mit oder ohne Ausreißer knapp 25 % der Späterwerber_innen, die im Bereich der Personen mit 2L1 abschneiden (vgl. Abb. 14 und 15).

Von besonderem Interesse ist, dass die Daten der C-Tests auch bei den ‚Muttersprachler_innen‘ variieren, wobei die Spannweite (Varianz) unter der Zeitdruckbedingung erheblich größer ist (S-C-Test Min: 79, Max: 149; C-Test Min: 130, Max: 150). In der Gruppe mit doppeltem Erstspracherwerb und auch bei den Späterwerber_innen ist die Spannweite größer. Dies lässt sich bei den Späterwerber_innen nicht am Wert der Standardabweichung ablesen, die sich zwischen den C-Tests und S-C-Tests nur um 2,3 Punkte unterscheidet. Aguado, Grotjahn und Schlak erwarten bei weit fortgeschrittenen Lerner_innen beim S-C-Test eine erhöhte Varianz und eine deutlich erhöhte Reliabilität. Tatsächlich kann an den Daten abgelesen werden, dass die Varianz in der Gruppe der 2L1 am höchsten ist. Von besonderem Interesse ist der Umstand, dass die Personen mit 2L1 zwar generell schlechter abschneiden

als die ‚Muttersprachler_innen‘, der Unterschied ist aber für die zeitlimitierte Testbedingung deutlich ausgeprägter als im C-Test ohne Zeitlimitierung. Es kann davon ausgegangen werden, dass für die ‚Muttersprachler_innen‘ wie auch für die Personen mit 2L1 Deckeneffekte beim C-Test ohne Zeitlimitierung bestehen. Die Schätzung der Reliabilität mit Cronbachs Alpha ergibt für sämtliche erhobenen Datensätze auf Basis der Lückenwerte für die S-C-Tests ($n = 269^{77}$) $\alpha = .986$ und für die C-Tests ($n = 271$) $\alpha = .971$. Für die monolingualen Erstsprachler_innen ($n = 80$) liegt die geschätzte Reliabilität für die S-C-Tests bei $\alpha = .924$ und für die C-Tests bei $\alpha = .851$. Bei den Personen mit 2L1 ($n = 50$) ist $\alpha = .950$ für die S-C-Tests und $\alpha = .585^{78}$ für die C-Tests. In der Gruppe der Späterwerber_innen ($n = 89$) ist $\alpha = .986$ bei den S-C-Tests und $\alpha = .959$ bei den C-Tests. Außer bei den C-Tests der Personen mit 2L1 kann die interne Konsistenz der S-C-Tests und C-Tests als sehr gut bezeichnet werden.

Die Person mit 149 S-C-Test- und 150 C-Test-Punkten in der Gruppe der ‚Muttersprachler_innen‘ ist dieselbe. Insgesamt gibt es fünf ‚Muttersprachler_innen‘, die beim C-Test ohne Zeitlimitierung alle Lücken mit der korrekten Lösung oder einer Lösungsvariante ausfüllen können.

4.1.3 Gruppenvergleiche

Der Mann-Whitney-U-Test ist eine nonparametrische Alternative zum t-Test. Er wird im Folgenden aufgrund der nicht normal verteilten Daten eingesetzt, um zu ermitteln, ob sich die Gruppen der ‚Muttersprachler_innen‘, Personen mit 2L1 und Späterwerber_innen signifikant voneinander unterscheiden. Weil dieser Test jeweils zwei Gruppen miteinander vergleicht, wurde der Test gepaart für die Gruppen Muttersprachler_innen vs. Doppelter Erstspracherwerb, Doppelter Erstspracherwerb vs. Späterwerber_innen und Späterwerber_innen vs. Muttersprachler_innen durchgeführt.

⁷⁷ Für diese Berechnung wurden sämtliche erhobenen Daten ausgewertet incl. der Fälle mit AO = 4–15.

⁷⁸ Entgegen der Erwartung fällt dieser Wert im Vergleich zu den anderen deutlich kleiner aus. Bei der nach der Berechnung des Wertes stattgefundenen Überprüfung der zugrundeliegenden Daten konnten keine offensichtlichen Fehler im Datenset gefunden werden.

Beim GJT unterscheiden sich, wie bereits vermutet, die ‚Muttersprachler_innen‘ nicht von den Personen mit 2L1; beide unterscheiden sich aber signifikant von den Späterwerber_innen (vgl. Tab. 11). Für sämtliche Mann-Whitney-U-Tests ist das Signifikanzniveau $\alpha = 5\%$, einseitig. Das Problem des massiv erhöhten Alpha-Fehlers durch multiple U-Tests konnte hier vernachlässigt werden, da es nur einen p-Wert $> .000$ gibt.

Tab. 11: Gruppenunterschiede GJT.

Gruppen	Mann-Whitney-U	p	Z	Φ
‚Muttersprachler_innen‘ und Personen mit 2L1 (n = 130)	1855.000	.241	- .708	- .06
Personen mit 2L1 und Späterwerber_innen (n = 139)	175.000	.000	- 8.625	- .73
‚Muttersprachler_innen‘ und Späterwerber_innen (n = 169)	86.000	.000	- 10.974	- .84

Bei den S-C-Tests unterscheiden sich alle drei Gruppen signifikant (vgl. Tab. 12).

Tab. 12: Gruppenunterschiede S-C-Tests.

Gruppen	Mann-Whitney-U	p	Z	Φ
‚Muttersprachler_innen‘ und Personen mit 2L1 (n = 130)	711.000	.000	- 6.17	- .54
Personen mit 2L1 und Späterwerber_innen (n = 139)	198.000	.000	- 8.895	- .75
‚Muttersprachler_innen‘ und Späterwerber_innen (n = 169)	22.000	.000	- 11.141	- .86

Bei den C-Tests unterscheiden sich alle drei Gruppen signifikant (vgl. Tab. 13).

Tab. 13: Gruppenunterschiede C-Tests.

Gruppen	Mann-Whitney-U	p	Z	Φ
„Muttersprachler_innen“ und Personen mit 2L1 (n = 130)	885.000	.000	- 5.344	- .47
Personen mit 2L1 und Späterwerber_innen (n = 139)	158.500	.000	- 9.071	- .77
„Muttersprachler_innen“ und Späterwerber_innen (n = 169)	31.500	.000	- 11.114	- .85

Unterteilt man die Gruppe der Späterwerber_innen in eine DaF-Gruppe (Lernbeginn in der Heimat, n = 71) und eine DaZ-Gruppe (Lernbeginn in Deutschland, n = 177⁹⁹) ergeben sich keine signifikanten Unterschiede beim Mann-Whitney-U-Test (vgl. Tab. 14).

Tab. 14: Gruppenunterschiede DaF- und DaZ- Lerner_innen.

Gruppen DaF und DaZ	Mann-Whitney-U	p	Z	Φ
GJT	486.000	.108	- 1.246	- .13
S-C-Tests	525.000	.206	- .830	- .09
C-Tests	454.000	.114	- 1.58	- .17

4.1.4 Korrelationen

In der statistischen Auswertung kommt das Spearman-Rho-Verfahren zum Einsatz, und zwar aus folgenden Gründen: Die Daten der Testergebnisse sind nicht alle normalverteilt, nicht alle ausgewerteten Daten sind metrisch skaliert und schließlich liegen bei den Analysen der einzelnen Gruppen teilweise kleine Fallzahlen (n < 50) vor.

Im Folgenden werden die Spearman-Rho-Ranglistenkorrelationen (vgl. Tab. 15) berichtet, die zwischen den Sprachstandsmessungen und den psycho-sozialen und bi-

⁹⁹ Die Summe ergibt hier nicht 89, weil bei einer Testperson die Werte fehlen.

ografischen Einflussgrößen bestehen. Letztere wurden mit den TESTATT-Skalen und Fragebögen erhoben.

Tab. 15: Spearman-Rho-Ranglistenkorrelationen für verschiedene Variablen individueller Differenz. Korrelation ist bei Niveau .05 signifikant (zweiseitig). Signifikante Korrelationen sind farblich gekennzeichnet.

Variable	n	S-C-Tests Rho (Sig.)	C-Tests Rho (Sig.)	GJT Rho (Sig.)
S-C-Tests	89	/	.870 (.000)	.490 (.000)
C-Tests	89	/	/	.584 (.000)
AO	89	-.444 (.000)	-.447 (.000)	-.379 (.000)
Motivation für Deutsch auf hohem Niveau	89	.388 (.000)	.418 (.000)	.261 (.000)
LOR	87	.337 (.001)	.449 (.000)	.366 (.000)
Fremdsprachenverwendungsangst	80	-.343 (.002)	-.307 (.006)	-.145 (.198)
DaF-Unterricht in der Heimat	69	.334 (.005)	.341 (.004)	.137 (.261)
Deutschgebrauch im Alltag	89	.279 (.008)	.210 (.049)	.155 (.147)
DaF-Unterricht insgesamt	84	.275 (.012)	.354 (.001)	.244 (.025)
Motivation (instrumentell)	79	.263 (.019)	.255 (.023)	.16 (.159)
Gebrauch weiterer Sprachen im Alltag	89	-.247 (.020)	-.267 (.011)	-.139 (.193)
DaF-Unterricht in Deutschland	68	-.202 (.099)	-.137 (.265)	.032 (.797)
Motivation (insgesamt)	78	.129 (.261)	.147 (.198)	.08 (.486)
L1-Gebrauch im Alltag	89	.096 (.373)	.143 (.182)	.207 (.051)
Extraversion	77	.089 (.443)	.089 (.44)	.035 (.765)
Alter zum Zeitpunkt des Tests	88	-.009 (.931)	.136 (.207)	.009 (.931)

Die Interpretation der Werte als starke und schwache Korrelationen ist vom Kontext abhängig. Wenn zwei Variablen fast dasselbe messen, dann ist Rho = .7 eventuell als schwacher Zusammenhang zu bewerten. Werden sehr verschiedene Größen gemessen, dann ist .5 oder .3 möglicherweise außergewöhnlich hoch. Beispielswei-

se fällt die Korrelation zwischen den unlimitierten und den zeitlimitierten C-Tests mit $Rho = .870$ erwartungsgemäß hoch aus. Die Ausprägung der anderen Korrelationen ist an dieser Stelle nicht überraschend, weil die Variablen sehr verschiedene Sachverhalte messen.

Die Interkorrelation zwischen den C-Tests und S-C-Tests fällt am höchsten aus, zwischen den S-C-Tests und GJTs am niedrigsten. Angesichts der Verwendung der gleichen Texte für die C-Tests und S-C-Tests ist die hohe Korrelation erwartbar: Wer unter der Zeitdruckbedingung den Test gut löst, dem gelingt dies auch ohne Zeitdruck. Die schwächer ausgeprägten Korrelationen zwischen C-Tests, S-C-Tests und GJT lassen sich durch die Unterschiedlichkeit der Testformate erklären, wobei v. a. folgender Punkt hervorzuheben ist: Der GJT hebt auf spezifische Lernschwierigkeiten der deutschen Sprache ab, während die C-Tests und S-C-Tests die allgemeine Sprachkompetenz messen (vgl. Kap. 3.4.2.1 *Grammatical Judgement Test* und 3.4.2.2 *C-Tests und S-C-Tests*).

Die höchste Korrelation bezüglich der Variablen individueller Differenz konnte im Zusammenhang dieser Untersuchung erwartungsgemäß für den Lernbeginn (AO) ermittelt werden. Die Variable AO bezieht sich entweder auf das Alter des Lernbeginns im Herkunftsland oder den Umzug in ein deutschsprachiges Land, je nachdem, was von den Testpersonen im Fragebogen angegeben wurde. Wenn man diese Variablen nicht in AO zusammenfasst, sondern einzeln auswertet, ergeben sich die folgenden Korrelationen (vgl. Tab. 16).

Tab. 16: Spearman-Rho-Ranglistenkorrelationen für AO und seine Differenzierungen. Korrelation ist bei Niveau .05 signifikant (zweiseitig). Signifikante Korrelationen sind farblich gekennzeichnet.

Variable	n	S-C-Tests Rho (Sig.)	C-Tests Rho (Sig.)	GJT Rho (Sig.)
Alter zu Beginn des Deutschlernens im Herkunftsland	70	-549 (.000)	-508 (.000)	-335 (.005)
AO	89	-444 (.000)	-447 (.000)	-379 (.000)
Alter beim Umzug in ein deutschsprachiges Land	80	-314 (.005)	-246 (.028)	-2 (.076)
Alter beim ersten Besuch eines deutschsprachigen Landes	88	-295 (.005)	-328 (.002)	-258 (.015)

Die Zahlen zeigen, dass das Alter beim ersten Besuch bzw. tatsächlichen Umzug nach Deutschland weniger Einfluss hat als das Alter des Lernbeginns in der Heimat. Dieses Ergebnis ist unerwartet, denn in der Literatur wird dafür argumentiert, dass das Alter des Umzugs in das Land der Zielsprache und der damit einsetzenden Immersion stärker den L2-Lernerfolg vorhersagen kann (vgl. Kap. 3.4.1 *Testpersonen*). Die etwas höhere Korrelation beim Alter zu Beginn des Deutschlernens im Herkunftsland könnte dadurch zu erklären sein, dass die Testpersonen durchschnittlich jünger waren, wenn sie schon im Herkunftsland mit dem Deutschlernen begannen.

Natürlich ist es bei AO von besonderem Interesse, die Korrelation nicht nur für die Gruppe der Späterwerber_innen zu berechnen. Bezieht man die Fälle mit AO = 4–15 mit ein, fällt das Ergebnis bei einem Signifikanzniveau von $p < .001$ deutlich aus: Spearmans Rho: $-.762$ (GJT und AO), $-.814$ (S-C-Tests und AO) und $-.805$ (C-Tests und AO).

Das zweitstärkste Ergebnis betrifft den Zusammenhang des Items Nr. 21 aus dem Fragebogen zur Sprachlernbiografie:

Wie wichtig ist es Ihnen, Deutsch auf einem sehr hohen Niveau zu sprechen?			
Bitte tragen Sie es auf einer Skala von 1 (völlig unwichtig) bis 4 (sehr wichtig) ein. Markieren Sie die entsprechende Zahl mit einem Kreis.			
völlig unwichtig	unwichtig	wichtig	sehr wichtig
1	2	3	4

Dieses Ergebnis entspricht ebenfalls der aus der Diskussion der Forschungsliteratur gewonnenen Hypothese, dass die festgesetzten Ziele bzw. Zweckbestimmungen des L2-Lernens und deren Einfluss auf die Motivation bzw. den Arbeitseinsatz für die erwachsenen Lerner_innen eine entscheidende Rolle spielen.

Tab. 17 zeigt die deskriptive Auswertung der LOR. LOR ist ebenfalls ein starker Prädiktor. Gegenläufig zur Erwartung ist jedoch, dass der Wert für die C-Tests (Rho = $.449$ ($p = .000$)) höher ausfällt als für die S-C-Tests (Rho = $.337$ ($p = .001$)). Eine höhere LOR und die damit verbundene Immersion lässt es naheliegend erscheinen, dass der Grad der Automatisierung des sprachlichen Wissens gesteigert wird. Dies kann mit den vorliegenden Daten jedoch nicht bestätigt werden.

Tab. 17: Deskriptive Auswertung der LOR.

Variable	N	Min	Max	Median	M	SD
LOR (in Monaten)	87	1	240	20	49.87	56.905

Anmerkung: M = arithmetisches Mittel; SD = Standardabweichung

Der negative Einfluss der Angst vor der Fremdsprachenverwendung auf den L2-Lernerfolg fällt ebenfalls deutlich aus, obwohl es sich um erwachsene Testpersonen handelt, die ein sehr hohes oder hohes Niveau der Zielsprache erreicht haben. Es wäre zu vermuten, dass das Ergebnis eventuell noch stärker sein könnte, wenn die Skalen-Items sich nicht nur auf den schulischen Kontext bezögen.

Tab. 18: Deskriptive Auswertung der Unterrichtslänge.

Variable	N	Min	Max	Median	M	SD
Länge des DaF-Unterrichts in der Heimat (in h)	89	0	2520	80	306.58	464.315
Länge des DaF-Unterrichts in Deutschland (in h)	89	0	1020	48	101.68	161.601
Länge des DaF-Unterrichts insgesamt (in h)	89	0	2520	250	408.26	440.063

Anmerkung: M = arithmetisches Mittel; SD = Standardabweichung

Tatsächlich ist der Gesamtumfang des DaF-Unterrichts in der Heimat wirkmächtiger als der des Unterrichts in Deutschland. Dazu muss angemerkt werden, dass die Dauer des DaF-Unterrichts mit 0–2520 Stunden enorm variiert (vgl. Tab. 18). Es wurde wesentlich mehr DaF-Unterricht im Ausland angegeben. Hierzu ist anzumerken, dass explizit nach DaF-Unterricht gefragt wurde. Beispielsweise wird ein Studium in Deutschland nicht erfasst, das viele der Testpersonen absolviert haben und zur sprachlichen Weiterentwicklung auf hohem Niveau sicherlich beiträgt. Es liegen keine Daten zu Unterrichtsformen, methodischen Ansätzen der Lehrkräfte und Lernstrategien der Späterwerber_innen vor; Aussagen über die Lehr- und Lernkontexte können deswegen nicht getroffen werden.

4.1.5 Phonetik-Rating

Das Phonetik-Rating ist ebenfalls Bestandteil der ersten Phase. Allerdings liegen Phonetik-Proben nur von den Testpersonen vor, die im Rahmen des hier vorliegen-

den Forschungsprojekts erhoben wurden: Personen mit 2L1 (n = 21⁸⁰) und Späterwerber_innen (n = 20), sowie von zehn ‚Muttersprachler_innen‘, die als Kontrollgruppe fungieren. Die Sprechproben wurden mit einem Diktiergerät (*Olympus Digital Voice Recorder VN-713 PC*) aufgenommen und im MP3-Format gespeichert. Geschnitten wurden die Sprechproben mit *Audacity 2.0.5* (Freeware/Opensource, URL 15).

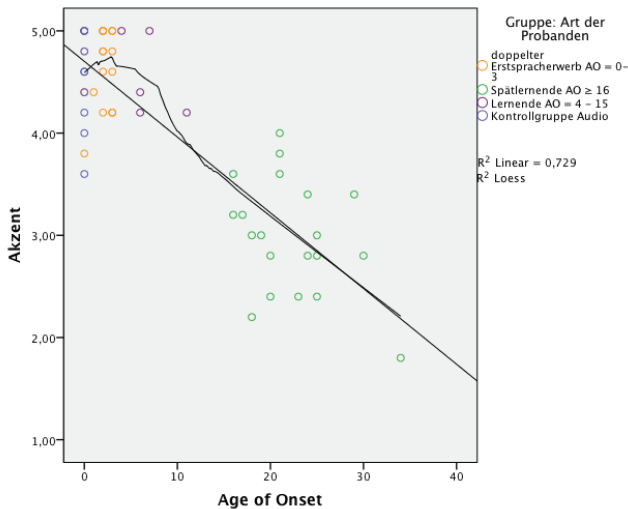


Abb. 16: Streudiagramm AO und Ergebnis des Akzent-Ratings. In das Streudiagramm wurden ebenfalls die Fälle mit AO = 4–15 (n = 6) einbezogen.

Trotz der heterogenen Zusammenstellung des Ratingteams⁸¹ und der niedrigen Item-Zahl (jeweils 1–5) fällt die Inter-Rater-Reliabilität sehr hoch aus (Cronbachs Alpha): $\alpha = .89$ für das Akzent-Rating und $\alpha = .828$ für die Verständlichkeit. Auch beim Phonetik-Rating veranschaulichen wir uns die Daten zunächst durch ein Streudiagramm (Abb. 16 für Akzent und Abb. 17 für Verständlichkeit).

⁸⁰ Trotz größter Sorgfalt bezüglich des Umgangs mit den Daten, ihrer Archivierung sowie des Einsatzes mehrerer Backupsysteme ist ein Audio einer Probandin verloren gegangen.

⁸¹ Zusammensetzung des Ratingteams: 1 x L1 Deutsch, 2 x 2L1 Deutsch und Italienisch bzw. Polnisch, 1 x L1 Mandarin und L2 Deutsch und 1 x L1 Rumänisch und L2 Deutsch (vgl. ausführlich in Kap. 3.4.2.5 *Phonetik-Test*).

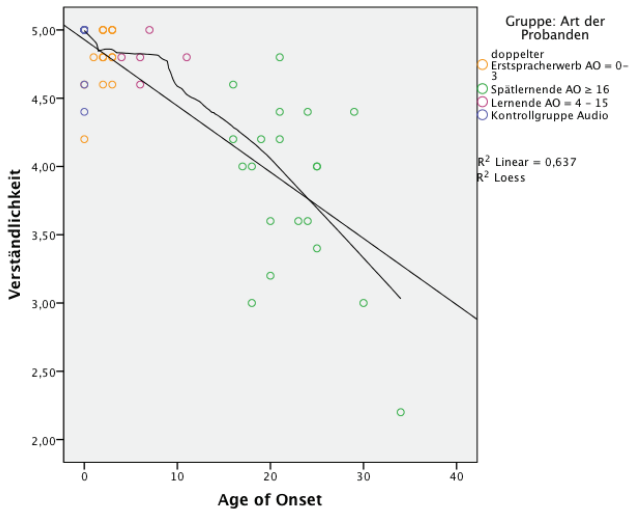


Abb. 17: Streudiagramm AO und Ergebnis des Verständlichkeits-Ratings. In das Streudiagramm wurden ebenfalls die Fälle mit AO = 4–15 (n = 6) einbezogen.

In Tab. 19 sehen wir, dass die Phonetik-Proben-Bewertung der Spätlerner_innen am weitesten streut.

Tab. 19: Deskriptive Statistik des Phonetik-Ratings.

Gruppe	N	Min	Max	Median	M	SD
Kontrollgruppe Akzent	10	3.6	5.0	4.600	4.5200	.47329
Kontrollgruppe Verständlichkeit	10	4.4	5.0	5.000	4.9000	.21602
Personen mit 2L1 Akzent	20	3.8	5.0	4.800	4.6500	.36056
Personen mit 2L1 Verständlichkeit	20	4.2	5.0	4.800	4.8200	.21423
Spätlerner_innen Akzent	20	1.8	4.0	3.000	2.9800	.56531
Spätlerner_innen Verständlichkeit	20	2.20	4.8	4.000	3.8400	.64105

Anmerkung: M = arithmetisches Mittel; SD = Standardabweichung

Die ebenfalls nicht unbeträchtliche Streuung beim Akzent-Rating der ‚Muttersprachler_innen‘ kann dadurch erklärt werden, dass als Kontrollgruppe größtenteils Dialekt-Sprecher_innen eingesetzt wurden und das Ratingteam sowohl fremdsprachlichen wie dialektalen Akzent berücksichtigen sollte.

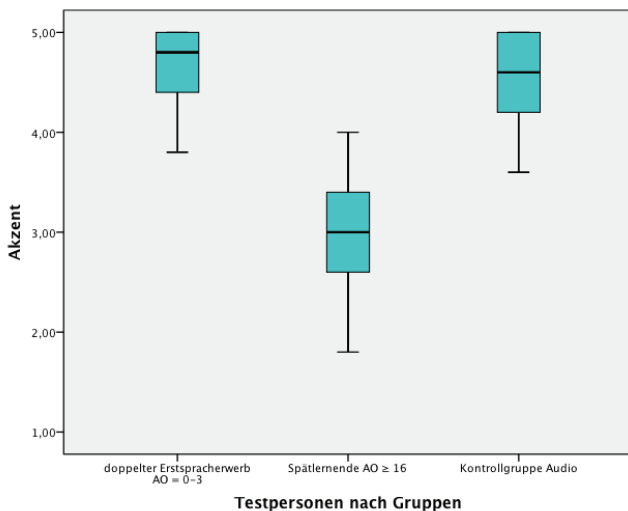


Abb. 18: Boxplot Akzentbewertung nach Gruppen.

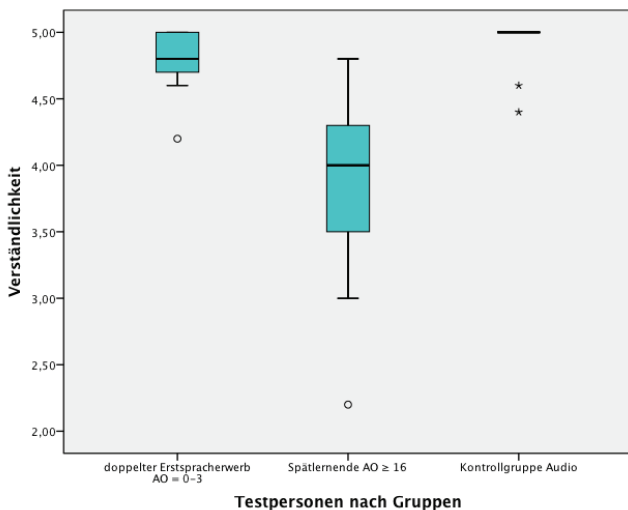


Abb. 19: Boxplot Verständlichkeitsbewertung nach Gruppen.

Beim Akzent- und Verständlichkeits-Rating unterscheiden sich, wie die grafische Darstellung vermuten lässt, die ‚Muttersprachler_innen‘ nicht von den Personen mit 2L1. Beide unterscheiden sich aber signifikant von den Späterwerber_innen (vgl. Abb. 18 und 19, Tab. 20 und 21).

Tab. 20: Gruppenunterschiede Akzent.

Gruppen	Mann-Whitney-U	p	Z	Φ
‚Muttersprachler_innen‘ und Personen mit 2L1 (n = 30)	85.500	.265	- .654	- .12
Personen mit 2L1 und Späterwerber_innen (n = 40)	1.500	.000	- 5.394	- .85
‚Muttersprachler_innen‘ und Späterwerber_innen (n = 30)	3.500	.000	- 4.26	- .78

Tab. 21: Gruppenunterschiede Verständlichkeit.

Gruppen	Mann-Whitney-U	p	Z	Φ
‚Muttersprachler_innen‘ und Personen mit 2L1 (n = 30)	72.000	.082	- 1.372	- .25
Personen mit 2L1 und Späterwerber_innen (n = 40)	15.500	.000	- 5.044	- .80
‚Muttersprachler_innen‘ und Späterwerber_innen (n = 30)	5.000	.000	- 4.233	- .77

Tab. 22 zeigt die Spearman-Rho-Korrelationen. Auffällig ist, dass die Korrelation zwischen Verständlichkeit und S-C-Tests und Verständlichkeit und C-Tests höher ausfällt. Dabei muss berücksichtigt werden, dass letztere nicht das Signifikanzniveau erreicht. Zu vermuten wäre, dass der Grad der Automatisierung gerade auch bei der spontanen mündlichen Rede eine große Rolle spielt. Ebenfalls auffällig ist die deutlich größere Rolle der Motivation. Es wurde keine einzige Korrelation signifikant, die AO, LOR, das Alter beim Umzug in ein deutschsprachiges Land oder die Länge des Deutschunterrichts im Inland oder Ausland betrifft. Dies könnte eventuell dadurch zu erklären sein, dass die Proband_innen dieser Gruppe alle einen AO ≥ 16 und einen LOR ≥ 5 haben. Darüber hinaus beeinflusst die kleine Stichprobengröße (n = 20)

das Erreichen des Signifikanzniveaus. Betrachten wir die Korrelation für alle Proband_innen – also auch mit der Kontrollgruppe (n = 10), den Personen mit 2L1 (n = 20) und den Lerner_innen mit einem AO = 4–15 (n = 6) – ergeben sich hohe Korrelationswerte: Akzent und AO: -.699 (.000); Verständlichkeit und AO: -.720 (.000).

Tab. 22: Spearman-Rho-Ranglisten-Korrelationen Phonetik-Rating. Signifikante Korrelationen sind farblich gekennzeichnet.

Variable	n	Akzent Rho (Sig.)	Verständlichkeit Rho (Sig.)
Akzent	20	/	.916 (.000)
Verständlichkeit	20	.916 (.000)	/
S-C-Tests	20	.469 (.037)	.535 (.015)
C-Tests	20	.436 (.054)	.401 (.08)
GJT	20	.472 (.036)	.332 (.152)
Selbsteinschätzung Deutschkenntnisse	20	.542 (.014)	.417 (.068)
Motivation für Deutsch auf hohem Niveau	20	.484 (.03)	.343 (.138)
Motivation (insgesamt)	20	.682 (.001)	.576 (.01)
Motivation (instrumentell)	20	.551 (.014)	.428 (.068)

4.2 Auswertung Phase II

Aufgrund der kleinen Anzahl der Testpersonen sind die Ergebnisse nur mit größter Vorsicht inferentiell auszuwerten. Im Folgenden werden einige Wechselbeziehungen zwischen den Skalen zur Motivation bzw. den LLAMA-Tests und den gemessenen L2-Kompetenzen präsentiert, die jedoch nur bedingt verallgemeinerbar sind.

4.2.1 Fragebogen biografische Hintergrundinformationen und Motivation

In diesem Abschnitt werden die Motivationsskalen ausgewertet. Die Ergebnisse der Erfassung der Hintergrundinformationen zu den Testpersonen werden in Kap. 4.3 *Beantwortung der Forschungsfragen* präsentiert.

Die Anzahl der Testpersonen in der zweiten Phase ist klein ($n = 10$) und infolgedessen dürfen die Daten nur vorsichtig ausgewertet werden. Trotzdem kann festgestellt werden, dass die Ergebnisse teilweise jenen entsprechen, die Dörnyeis *Motivational Self System*-Modell empirisch validieren (vgl. die Zusammenfassung verschiedener Studien in Dörnyei 2009a: 31 f.). Studien, die den Zusammenhang zwischen *Integrativeness* und *Ideal L2 Self* untersucht haben, ergeben eine durchschnittliche Korrelation von $.54^{82}$ zwischen den beiden Variablen – für Dörnyei ein klarer Hinweis darauf, dass die beiden Konzepte miteinander in enger Beziehung stehen. In der vorliegenden Studie erreichte wahrscheinlich aufgrund der kleinen Fallzahlen kaum eine der berechneten Korrelationen das Signifikanzniveau. Dafür entspricht die gefundene Spearman-Rho-Korrelation zwischen *Integrativeness* und *Ideal L2 Self* mit $Rho = .535$ ($p = .111$) fast genau Dörnyeis Durchschnittswert.

Der *Intended Learning Effort* wird in den Studien zum *Motivational Self System* in der Regel als *Criterion Measure* eingesetzt. Dörnyei zufolge erklärt das *Ideal L2 Self* 42 % der Varianz des *Intended Learning Effort* (vgl.: ebd.). Dieses Ergebnis fällt bei den zehn Testpersonen der zweiten Phase mit $Rho = .433$ ($p = .212$) schwächer aus. Dasselbe gilt für den Zusammenhang von *Intended Learning Effort* und *Integrativeness*: Bei Dörnyei erklären sie 32 % der Varianz, und in der hier vorliegenden Studie ist $Rho = .114$ ($p = .754$). *Instrumentality (Promotion)* und *Instrumentality (Prevention)* korrelieren bei den hier getesteten zehn Personen am stärksten mit dem *Intended Learning Effort*: $Rho = .619$ ($p = .056$) für *Promotion*, $Rho = .468$ ($p = .173$) für *Prevention*. Gemäß Dörnyeis Synopse ist der Zusammenhang zwischen *Ideal L2 Self* und *Instrumentality (Promotion)* höher als der zwischen *Ideal L2 Self* und *Instrumentality (Prevention)*, während die Beziehung zum *Ought-to L2 Self* umgekehrt ist: Für *Instrumentality (Prevention)* sehr stark, für *Instrumentality (Promotion)* deutlich schwächer, aber immer noch stark.⁸³ Die beiden Typen der instrumentellen Motivation teilen sich bei Dörnyei nur 12 % der Varianz. Dies gilt für ihn zusammen mit den zuvor präsentierten Ergebnissen als positive Evidenz dafür, dass instrumentelle Motivation in die beiden Bereiche *Prevention* und *Promotion* aufgeteilt

⁸² Dörnyei gibt an dieser Stelle keinen Korrelationskoeffizienten an. Da die meisten Studien, die der Zusammenschau zugrundeliegen, metrische Verfahren nutzen, ist von Pearson-Korrelationen auszugehen.

⁸³ Diese Zusammenhänge werden in Dörnyeis Darstellung nicht mit Zahlen wiedergegeben.

werden kann. Die Ergebnisse der hier vorliegenden Studie weichen davon ab: *Prevention* und *Promotion* korrelieren sehr stark mit $Rho = .856$ ($p = .002$), wobei auch bemerkenswert ist, dass diese Korrelation das Signifikanzniveau erreicht. In diesem Zusammenhang erscheint es wahrscheinlich, dass das Moment der Rückschau der Testpersonen in dieser Studie eine Rolle spielt. Bei den Fragen des Fragebogens der zweiten Phase könnte es für diese hohe Korrelation sorgen, da vermutlich aus der Perspektive der Rückschau nur schwer zwischen *Promotion* und *Prevention* unterschieden werden kann. Berichtenswert ist des Weiteren, dass für das *Ought-to L2 Self* starke Zusammenhänge gefunden wurden, die das Signifikanzniveau erreichen: zur *Promotion* ($Rho = .816$, $p = .004$), zur *Prevention* ($Rho = .687$, $p = .028$) sowie zum *L2 Self-Confidence* ($Rho = .752$, $p = .012$).

Wenn wir die Fragebogenergebnisse zur Motivation mit den Ergebnissen der Sprachtests korrelieren, erhalten wir die folgenden Werte (vgl. Tab. 23).

Tab.: 23 Spearman-Rho-Ranglistenkorrelationen für Motivation und Sprachkompetenz.

Motivation (retrospektiv)	n	S-C-Test Rho (Sig.)	C-Test Rho (Sig.)	GJT Rho (Sig.)	Akzent Rho (Sig.)	Verständ Rho (Sig.)
<i>Integrativeness</i>	10	.07 (.848)	.013 (.972)	.368 (.295)	.421 (.226)	.122 (.737)
<i>Ideal L2 Self</i>	10	-.255 (.478)	-.067 (.853)	.302 (.397)	.235 (.514)	-.087 (.811)
<i>L2 Self-Confidence</i>	10	-.189 (.601)	-.207 (.566)	-.117 (.748)	.288 (.42)	.01 (.979)
<i>Ought-to L2 Self</i>	10	-.09 (.806)	.185 (.61)	.155 (.669)	.388 (.268)	-.125 (.731)
<i>Instrumentality (Prevention)</i>	10	.055 (.879)	.313 (.379)	.102 (.779)	.328 (.355)	-.087 (.811)
<i>Instrumentality (Promotion)</i>	10	.056 (.878)	.284 (.426)	.075 (.838)	.474 (.167)	.147 (.685)
<i>Intended- Learning-Effort</i>	10	-.061 (.867)	.182 (.614)	.254 (.479)	.387 (.27)	.08 (.826)

Keine der errechneten Korrelationen erreicht das Signifikanzniveau. Auffällig ist allerdings die Reihe der gefundenen negativen Zusammenhänge. Auch hier müsste überprüft werden, inwiefern die Situation der Retrospektive auf den Lernprozess verantwortlich für diese Ergebnisse ist. Vergleichsweise hohe Werte liegen bei der instrumentellen Motivation und den nicht zeitlimitierten C-Tests vor. Dieses Ergebnis scheint mit jenem aus der ersten Phase zu korrespondieren, bei dem die instrumentelle Motivation signifikant wurde. Übertroffen wird dieser Wert von den Werten zwischen Akzent und den Komponenten des *Motivational Self Systems*. Für alle Werte gilt der Hinweis, dass sie aufgrund der kleinen Fallzahl und des nicht erreichten Signifikanzniveaus nicht verallgemeinerbar sind.

Tab. 24: Skalenwerte einzelner Testpersonen unterschieden nach Leistung im GJT, in den S-C-Tests und den C-Tests.

	Beste Testpersonen	Mittlere Testpersonen	Schwächste Testpersonen
<i>Integrativeness</i>	3.5	3.75	3.25
	3.5	3	3.75
<i>Ideal L2 Self</i>	3.5	4	2.75
	3.25	3.5	3.75
<i>L2 Self-Confidence</i>	1.8	2.4	1
	1	1.6	2
<i>Ought-to L2 Self</i>	3.75	2.63	1.25
	1.25	1.88	2.25
<i>Instrumentality (Prevention)</i>	4	4	2
	1	2.67	1.33
<i>Instrumentality (Promotion)</i>	4	4	2.8
	1.2	3	2.6
<i>Intended Learning Effort</i>	3.64	3.55	3.09
	2.27	3.18	2.82
Motivation (TESTATT)	3.53	3.73	2.93
	2.93	2.6	3.67
Extraversion	2.63	2.56	3.25
	2.94	2.81	2.5
Fremdsprachenverwendungsangst	1.09	1.73	1.27
	1	1.91	2.45

Die Werte der Motivationsskalen weisen große Unterschiede auf. Wenn wir uns auf der deskriptiven Ebene die Werte anschauen, die jeweils von den zwei besten, den zwei mittleren und den zwei schwächsten Testpersonen⁸⁴ angegeben wurden, sehen wir, dass die Skalenwerte stark individuell variieren und zumindest bei diesen sechs Testpersonen in keinem Zusammenhang mit dem Leistungsniveau stehen (vgl. Tab. 24).

4.2.2 LLAMA-Tests

Die Korrelationen für die LLAMA-Tests fallen gering aus und korrelieren teilweise sogar negativ. Sie erreichen bis auf einen Wert nicht das Signifikanzniveau (vgl. Tab. 25). Auch hier ist die kleine Fallzahl zu berücksichtigen, um nicht zu weitgehende Folgerungen zu ziehen. Für die zehn Testpersonen lässt sich ein Zusammenhang der LLAMA-Tests (außer LLAMA D) mit den C-Tests und den GJT feststellen. Insgesamt sind die berechneten Korrelationen für die GJT-Ergebnisse bei den einzelnen LLAMA-Tests ausgeprägter. Die Werte für die S-C-Tests sind bis auf LLAMA B alle im negativen Bereich. Diese Ergebnisse lassen sich mit den Funden der beiden Studien von Granena und Long (2012) und Granena (2013) erklären. Granena und Long (2012) finden einen Zusammenhang zwischen Lexik (*lexis and collocation*) und Sprachlerneignung sowie zwischen Phonetik (*pronunciation*) und Sprachlerneignung, aber nicht zwischen Morphosyntax und Sprachlerneignung (vgl. ebd.: 332). Granena (2013b) stellt fest, dass die Sprachlerneignung dann eine Rolle spielt, wenn sich die Testpersonen bei den Tests Zeit lassen und sich auf die Form konzentrieren können (vgl. ebd.: 197–198).

LLAMA D (phonetische Diskrimination) wirkt sich bei den hier getesteten zehn Personen kaum auf den Akzent aus. Dieses Ergebnis war nicht zu erwarten, denn in der Regel wird lautliches Diskriminierungsvermögen als Voraussetzung für eine gute Aussprache angesehen. Es wäre eventuell zu diskutieren, inwiefern der LLAMA D nicht nur das phonetische Diskriminierungsvermögen misst, sondern v. a. auch das Kurzzeitgedächtnis.⁸⁵

⁸⁴ Auf der Grundlage der S-C-Tests, C-Tests und des GJT wurden jeweils Rangplätze vergeben, die untereinander verrechnet wurden, um diese Zuordnung vornehmen zu können.

⁸⁵ Während des LLAMA D-Tests hören die Testpersonen zunächst zehn Wörter in einer unbekanntem Sprache. In der Testphase werden diese Wörter wieder abgespielt, allerdings gemischt mit anderen,

Tab. 25: Spearman-Rho-Ranglistenkorrelationen für LLAMA-Tests und Sprachkompetenz. Signifikante Korrelationen sind farblich gekennzeichnet.

LLAMA-Test	n	S-C-Test	C-Test	GJT	Akzent	Verständ
		Rho (Sig.)	Rho (Sig.)	Rho (Sig.)	Rho (Sig.)	Rho (Sig.)
LLAMA B (Vokabeln lernen)	10	.236 (.511)	.318 (.37)	.572 (.084)	.648 (.043)	.46 (.181)
LLAMA D (Phonetische Diskrimination)	10	-.04 (.913)	-.439 (.204)	-.104 (.774)	.08 (.826)	.229 (.524)
LLAMA E (Laute und Symbole)	10	-.115 (.753)	.383 (.275)	.46 (.181)	-.056 (.878)	-.433 (.212)
LLAMA F (Grammatik ableiten)	10	-.291 (.415)	.131 (.719)	.179 (.621)	-.142 (.697)	-.304 (.393)
Sprachlern-eignung insgesamt⁸⁶	10	-.189 (.601)	.207 (.567)	.321 (.366)	.083 (.82)	-.188 (.602)

Die Auswertung bestätigt die Vermutung einer Probandin: Sie gab nach dem Absolvieren der zweiten Phase zu Protokoll, dass sie auf die Ergebnisse der TN gespannt sei, die sich sicherlich nicht durch ein neues Alphabet überraschen ließen aufgrund des Umstands, bereits zwei verschiedene Schriftsysteme verinnerlicht zu haben. Tatsächlich erreichen die Proband_innen im LLAMA E die meisten Punkte (vgl. Tab. 26).

Tab. 26: Deskriptive Statistik der LLAMA-Test-Ergebnisse.

Test	N	Min	Max	Median	M	SD
LLAMA B	10	20	80	42.50	43	18.288
LLAMA D	10	10	60	37.50	35	17.159
LLAMA E	10	50	100	80.00	80	17.638
LLAMA F	10	0	80	60.00	51	29.981

Anmerkung: M = arithmetisches Mittel; SD = Standardabweichung

neuen Wörtern, die teilweise phonetische Ähnlichkeiten aufweisen. Die Testpersonen sollen nach jedem Wort angeben, ob sie es das zweite Mal hören oder nicht.

⁸⁶ Summe aus LLAMA B, LLAMA D, LLAMA E und LLAMA F.

Dies ist gerade auch vor dem Hintergrund bemerkenswert, dass Meara zur Testentwicklung der Vorläuferversion festhält:

„that it identified people who were easily fazed by an unfamiliar spelling system. Most of our testees produced scores which were close to chance levels. The test was also good at identifying analytical linguists – particularly those with formal training in phonetics.“ (Meara 2005: 14)

Dem ließe sich hinzufügen, dass der Test bei Testpersonen, die im Wechseln des Schriftsystems Erfahrung haben, ebenfalls besonders gut funktioniert. Das macht eine neuerliche Diskussion über die Unabhängigkeit des Formats von der L1 bzw. deren Schriftsystem erforderlich.

4.3 Beantwortung der Forschungsfragen

Nachdem nun die Daten der beiden Projektphasen statistisch ausgewertet sind, werden in diesem Kapitel die Forschungsfragen beantwortet. Dazu werden zusätzlich zu den statistischen Analysen die erhobenen biografischen Daten herangezogen, um prototypische Profile sehr erfolgreicher, erfolgreicher und vergleichsweise weniger erfolgreicher Lerner_innen zu erstellen.

Frage 1:

- Welches Lerner_innenprofil hinsichtlich Motivation und Persönlichkeit weisen die erfolgreichsten, die durchschnittlichen und die weniger erfolgreichen Späterwerber_innen (AO \geq 16) auf?

In Phase I des hier vorliegenden Forschungsprojekts wurden folgende Erkenntnisse gewonnen: Die Testpersonen mit den stärksten Leistungsprofilen weisen in der Regel eine ausgeprägte Motivation für Deutsch auf hohem Niveau auf. Ihnen ist es sehr wichtig, über sehr gute Deutschkenntnisse zu verfügen. Besonders hoch sind ihre Werte auf der Motivationsskala bei den Items zur instrumentellen Motivation. Ein weiteres deutliches Ergebnis ist der negative Einfluss der Fremdsprachenverwendungsangst auf den L2-Lernerfolg. Positiv sind dagegen der Deutschgebrauch im Alltag und auch der L1-Gebrauch, auch wenn dieser nicht signifikant wurde. Wird jedoch eine weitere Sprache benutzt, wirkt sich das negativ auf das Niveau des Deutschen aus. Angemerkt werden muss hier, dass AO, LOR und Länge des DaF-Unterrichts deutlich höhere Korrelationen zum Sprachniveau aufweisen als die anderen individuellen Variablen.

Aus den Daten von Phase II lassen sich die folgenden Profile ableiten: Für die Lerner_innen-Profile der besten (+), mittleren (+/-) und schwächsten (-) Lerner_innen wurden die zwei besten, die zwei mittleren und die zwei schwächsten Fälle ausgewertet.⁸⁷ Um dieses Ranking vornehmen zu können, wurden jeweils Rangplätze vergeben, die sich auf die Ergebnisse aus den S-C-Tests, C-Tests und GJTs beziehen, und miteinander verrechnet. Im Folgenden wird ein Überblick über die Daten gegeben, die aus den Fragebögen gewonnen wurden und die Grundlage für die Profile bilden. An dieser Stelle muss explizit darauf hingewiesen werden, dass es nunmehr lediglich um sechs Testpersonen geht und diese Profile keinesfalls verallgemeinert werden dürfen.

Aus Tab. 27 geht hervor, dass die schwächeren Lernerinnen Deutsch lernten, um im Alltag kommunizieren zu können. Die mittleren und erfolgreichsten Lernerinnen lernten Deutsch auch, um im Beruf kommunizieren zu können. Auch die eigene Zielstellung scheint ein Indikator für den Lernerfolg zu sein: Während die schwächeren TN gute oder sehr gute Deutsch-Kenntnisse anstreben, wollen die mittleren und eine der erfolgreichsten TN ein sehr gutes oder ‚muttersprachliches‘ Niveau erreichen. Zugleich können wir feststellen, dass eine erfolgreiche Lernerin sehr gute Deutschkenntnisse anstrebt.

Tab. 27: Selbstgesteckte Ziele und Zweck des Deutschlernens (✓ = ja; X = nein) der beiden besten (+), mittleren (+/-) und schwächsten (-) Lerner_innen.

	+	+	+/-	+/-	-	-
Deutsch für alltägliche Kommunikation (Zweck)	✓	✓	✓	✓	✓	✓
Deutsch für berufliche Kommunikation (Zweck)	✓	✓	✓	✓	X	X
Deutsch für akademische Kommunikation (Zweck)	✓	X	✓	✓	X	X
‚Muttersprachliches‘ Niveau (Ziel)	✓	X	✓	✓	X	X
Sehr gutes Deutsch (Ziel)	X	✓	X	X	X	✓
Gutes Deutsch (Ziel)	X	X	X	X	✓	X

Ein eindeutigeres Bild ergibt sich bezogen auf die erhobenen Daten zum Sprachgebrauch (vgl. Tab. 28). Die erfolgreichsten Lernerinnen unterscheiden sich eindeutig von allen anderen dadurch, dass Deutsch in ihrer Wohnung gesprochen wird. Im Un-

⁸⁷ Tatsächlich handelt es sich bei diesen sechs Fällen ausschließlich um Frauen, weswegen im Folgenden die femininen Genusformen verwendet werden.

terschied zu den anderen ist ihr Umgang mit der deutschen Sprache umfassend: bei der Arbeit, in der Freizeit und zu Hause.

Tab. 28: Gebrauch des Deutschen und der Ausgangssprache (✓ = ja; X = nein) der beiden besten (+), mittleren (+/-) und schwächsten (-) Lerner_innen.

	+	+	+/-	+/-	-	-
Deutsch auf der Arbeit	✓	✓	✓	✓	✓	✓
L1 auf der Arbeit	✓	✓	✓	✓	X	✓
Deutsch in der Freizeit	✓	✓	✓	✓	✓	X
L1 in der Freizeit	✓	✓	✓	✓	✓	✓
Deutsch in der Wohnung	✓	✓	X	X	X	X
L1 in der Wohnung	X	✓	X	X	✓	✓
1-Personen-Haushalt	X	X	✓	✓	✓	X

Tab. 29 zeigt ein eindeutiges Ergebnis: Die beiden schwächeren Lernerinnen haben oder hatten nie einen deutschen Partner, die beiden erfolgreichsten haben oder hatten einen. Bei den mittleren Lernerinnen gibt es beide Alternativen.

Tab. 29: Biografische und andere Angaben (✓ = ja; X = nein) der beiden besten (+), mittleren (+/-) und schwächsten (-) Lerner_innen.

	+	+	+/-	+/-	-	-
deutsche_r Partner_in	✓	✓	X	✓	X	X
deutschsprachige Freund_innen	✓	✓	✓	✓	✓	✓
Studium in Deutschland	✓	✓	✓	✓	✓	✓
Sprachliches Feedback erwünscht	✓	✓	✓	✓	✓	✓
aktiv zum Feedback auffordern	X	✓	X	✓	X	✓

Die schwächsten TN der Studie geben zwar an, Deutsch auf der Arbeit zu sprechen. Bei der Frage, für welche Zwecke sie Deutsch gelernt haben, erwähnen sie jedoch ausschließlich die alltägliche Kommunikation, nicht die berufliche oder akademische. Beide haben oder hatten nie eine_n deutsche_n Partner_in, wohl aber deutsche Freund_innen, mit denen sie regelmäßig Zeit verbringen und Deutsch sprechen. Beide haben in Deutschland studiert. Bei der schwächsten TN ist das selbstgesteckte Ziel, gut Deutsch zu sprechen, bei der zweitschwächsten, sehr gut Deutsch zu sprechen. Beide freuen sich über sprachliches Feedback und nehmen die Korrekturen i. d. R. an. Eine der TN fordert ihr Umfeld auch aktiv zu Korrekturen auf. Bei bei-

den wird in der Wohnung die Ausgangssprache gesprochen; bei der Arbeit in einem Fall die Ausgangssprache und Deutsch, im anderen Fall nur Deutsch; in der Freizeit wird Deutsch und die Ausgangssprache bzw. nur die Ausgangssprache gesprochen.

Die beiden TN, die im Mittelfeld liegen, haben ebenfalls in Deutschland studiert. Auf der Arbeit sprechen beide Deutsch und ihre Ausgangssprache. Bei den Zwecken, für die sie Deutsch gelernt haben, geben sie alltägliche, berufliche und akademische Kommunikation an. Beide haben deutschsprachige Freund_innen, mit denen sie regelmäßig Zeit verbringen und Deutsch sprechen. Eine hat auch einen deutschen Partner. Das selbstgesteckte Ziel ist bei beiden, Deutsch wie ein_e ‚Muttersprachler_in‘ zu beherrschen. Beide freuen sich über sprachliches Feedback und nehmen die Korrekturen i. d. R. an. Eine der beiden fordert ihr Umfeld auch aktiv dazu auf, ihr Rückmeldung zu geben. Beide wohnen allein und sprechen in der Freizeit jeweils Deutsch und die Ausgangssprache.

Die beiden besten TN gaben an, Deutsch und ihre Ausgangssprache auf der Arbeit zu sprechen. Die eine hat Deutsch für die alltägliche und berufliche Kommunikation gelernt, die andere für die alltägliche, berufliche und akademische. Beide haben oder hatten eine_n deutsche_n Partner_in und deutsche Freunde, mit denen sie regelmäßig Zeit verbringen und Deutsch sprechen. Beide haben in Deutschland studiert. Bei der besten TN ist das selbstgesteckte Ziel, Deutsch wie ein_e ‚Muttersprachler_in‘ zu sprechen, bei der zweitstärksten, sehr gut Deutsch zu sprechen. Beide reagieren positiv auf sprachliches Feedback und nehmen die Korrekturen i. d. R. an. Eine fordert ihr Umfeld auch dazu auf, sie zu korrigieren. Bei einer TN wird in der Wohnung nur Deutsch gesprochen, bei der anderen Deutsch und die Ausgangssprache. Beide sprechen in ihrer Freizeit Deutsch und die Ausgangssprache.

Frage 2:

- Wie häufig kommt eine Sprachkompetenz in der Gruppe der Späterwerbber_innen vor, die der Sprachkompetenz der Personen mit doppeltem Erstspracherwerb entspricht?

Für die Beantwortung dieser Frage wurde bei jedem Test das 10. Perzentil (P10) der Personen mit 2L1 als Grundlage genommen. Der Sprachkompetenz der Personen mit doppeltem Erstspracherwerb zu entsprechen, bedeutet ein Ergebnis zu erzielen,

das mindestens so hoch wie das P10 der Personen mit 2L1 ist. Die unten stehenden Prozentzahlen beinhalten die Fälle, die P10 erreicht haben oder höher liegen (Späterwerber_innen: n = 89; Personen mit 2L1: n = 49):

- S-C-Tests: 14,6 %
- C-Tests: 15,7 %
- GJT: 5,6 %

Bemerkenswert ist, dass ungefähr gleich viele Späterwerber_innen bei den S-C-Tests im Bereich der Personen mit 2L1 abschneiden wie bei den C-Tests. Das deutlich schlechtere Abschneiden beim GJT ist sicherlich auf die Items zurückzuführen, die konstruiert wurden, um spezifische Lernschwierigkeiten der deutschen Sprache abzubilden.

Beim Rating von Akzent und Verständlichkeit ist das zugrundeliegende n deutlich kleiner, weil die Daten nur für die Testpersonen vorliegen, die von mir erhoben wurden (Späterwerbende: n = 20; Personen mit 2L1: n = 21).

- Akzent: 5 %
- Verständlichkeit: 10 %

Dass immerhin 5 % der Späterwerber_innen beim Akzent-Rating im Bereich der Personen mit 2L1 abschneiden, ist aus der Perspektive der Vertreter_innen der CPH überraschend. Das Erreichen einer zielsprachlichen Phonetik gilt sonst als Hauptschwierigkeit des L2-Lernens von Erwachsenen (vgl. Kap. 2.1 *Kritische Periode(n)*). Zurückgeführt werden kann dieses Ergebnis auf die Anweisung in der Handreichung für das Rater-Team, sowohl dialektalen als auch fremdsprachlichen Akzent zu bewerten. Doppelt so viele der Späterwerbenden (10 %) erreichen bei der Beurteilung der Verständlichkeit Werte im Bereich der Personen mit 2L1. Auch hier ist der Hinweis wichtig, dass die Zahlen nicht ohne Weiteres verallgemeinert werden können, da das zugrundeliegende n sehr klein ist.

Um das Bild zu vervollständigen, werden an dieser Stelle auch die Zahlen für den Vergleich der Späterwerber_innen sowie der Personen mit 2L1 mit den monolingualen Erstsprachler_innen angegeben. Auch hier dient P10 der monolingualen Erstsprachler_innen als untere Grenze. 1,1 % der Späterwerber_innen (n = 89) schneiden bei den S-C-Tests wie die Erstsprachler_innen (n = 80) ab, bei den C-Tests sind

es ebenfalls 1,1 %. Beim GJT erreichen 3,4 % der Späterwerber_innen das Niveau der muttersprachlichen Vergleichsgruppe. Die Werte fallen bei diesem Vergleich deutlich niedriger aus und geben einen Hinweis darauf, wie wichtig es ist, Bilinguale mit Bilingualen zu vergleichen. Die folgenden Prozentzahlen beinhalten die Fälle der Personen mit 2L1, die mindestens P10 erreichen (monolinguale Erstsprachler_innen: n = 80; Personen mit 2L1: n = 49 für die S-C-Tests, C-Tests und GJT; monolinguale Erstsprachler_innen: n = 10; Personen mit 2L1: n = 21 für Akzent und Verständlichkeit):

- GJT: 88 %
- S-C-Tests: 42 %
- C-Tests: 58 %
- Akzent: 100 %
- Verständlichkeit: 95,2 %

Aus den vorliegenden Daten geht hervor, dass die monolingualen ‚Muttersprachler_innen‘ sowohl beim S-C-Test als auch beim C-Test im Durchschnitt ein besseres Ergebnis erreichen. Des Weiteren ist auffällig, dass ihre Daten beim C-Test (SD: 5) viel weniger streuen als beim S-C-Test (SD: 13). Bezüglich der Phonetik unterscheiden sich die beiden Gruppen kaum. Der Vergleich bzgl. der Phonetik beruht nur auf zehn ‚Muttersprachler_innen‘, die zur Kontrollgruppe beim Phonetik-Rating gehörten, sowie 21 Fällen mit 2L1.

Frage 3:

- Welchen Einfluss hat die Sprachähnlichkeit auf die erreichte Sprachkompetenz?

Im Rahmen dieser Arbeit ist es nicht möglich, die Lernerfolge der TN nach einzelnen Ausgangssprachen getrennt auszuwerten. Zwar gibt es eine große Gruppe von Testpersonen mit der L1 Mandarin (n = 41). Die anderen Ausgangssprachen sind jedoch in kleinen Zahlen oder nur einzeln vertreten (vgl. Tab. 30), so dass ein Gruppenvergleich nicht möglich ist.

Tab. 30: Späterwerber_innen und ihre Ausgangssprachen (gesamt).

L1	n	L1	n
Mandarin	41	Thai	2
Russisch	10	Englisch	1
Arabisch	5	Griechisch	1
Japanisch	5	Kiswahili	1
Spanisch	5	Mongolisch	1
Aserbaidshanisch	3	Persisch	1
Vietnamesisch	3	Polnisch	1
Armenisch	2	Portugiesisch	1
Dari	2	Ukrainisch	1
Indonesisch	2	Usbekisch	1

Möglich ist allerdings ein Vergleich nach sprachtypologischen Gesichtspunkten, wenn bspw. flektierende, agglutinierende und isolierende Sprachen jeweils als Gruppe gezählt werden. Dabei werden die flektierenden und agglutinierenden Sprachen aufgrund ihrer Zugehörigkeit zum synthetischen Sprachtyp als eine Gruppe gefasst. Sicherlich sind die Zuordnungen, die dabei getroffen werden müssen, teilweise problematisch, da es Sprachen gibt wie das Englische, das über agglutinierende, flektierende und isolierende Merkmale verfügt oder Indonesisch, das zwar überwiegend isolierende Merkmale besitzt, aber auch agglutinierende Elemente aufweist (vgl. Görlach 1993: 164). Gleichzeitig ist festzustellen, dass der sprachtypologische Vergleich auf morphologischer Ebene nur eine Vorgehensweise unter anderen ist. Typologische Untersuchungen gibt es auch bezüglich der Lexik, Syntax, Phonologie und Sprachwandeltypen. Darüber hinaus lassen sich Gruppen aufgrund von genetischer, geographischer und kultureller Nähe bilden (vgl. Moravcsik 2013). Da sprachtypologische Diskussionen nicht im Zentrum dieser Arbeit stehen, wurden für die vorliegenden grafischen Auswertungen die Sprachen mithilfe der ‚klassischen‘ Sprachtypologie eingeteilt. Mandarin, Thai, Vietnamesisch und Indonesisch wurden der Gruppe der isolierenden Sprachen zugeordnet (n = 45), alle anderen Sprachen

in die Gruppe der flektierenden und agglutinierenden Sprachen (n = 44) (vgl. Abb. 20, Abb. 21 und Abb. 22).

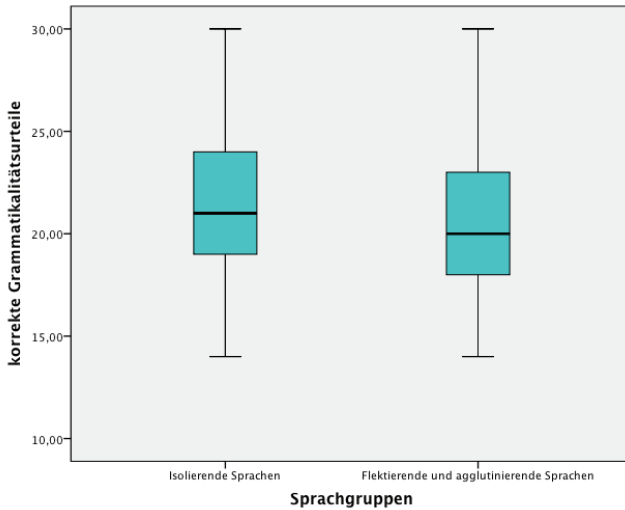


Abb. 20: Ergebnisse des GJTs nach sprachtypologischen Gruppen.

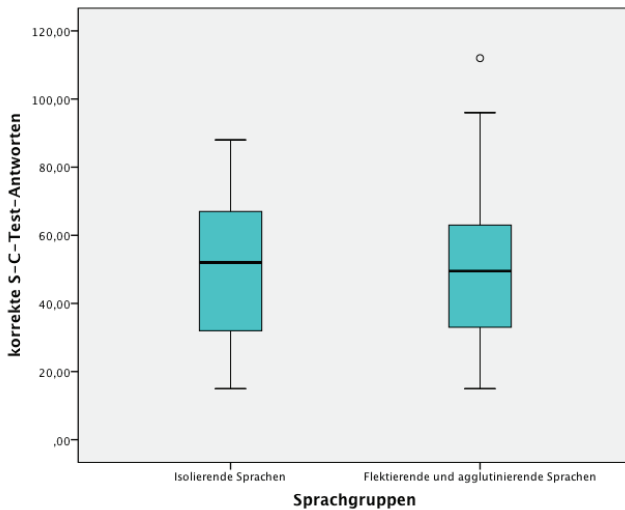


Abb. 21: Ergebnisse der S-C-Tests nach sprachtypologischen Gruppen.

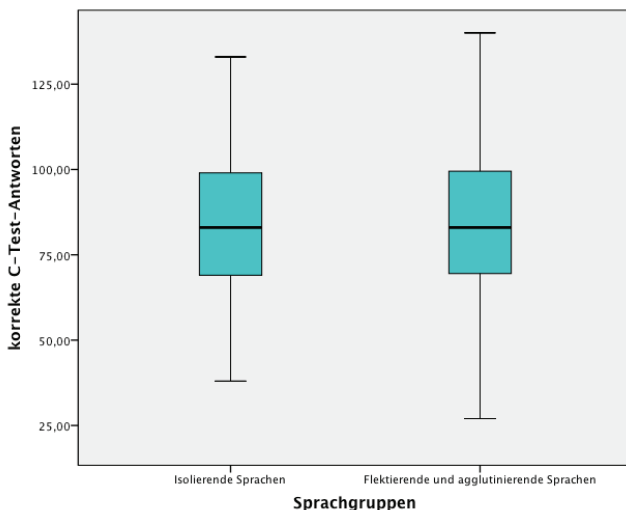


Abb. 22: Ergebnisse des C-Tests nach sprachtypologischen Gruppen.

Der Mann-Whitney-U-Test bestätigt den bloßen Augenschein: Es gibt keinen signifikanten Unterschied zwischen den beiden Gruppen bezogen auf die S-C-Test-, C-Test und GJT-Ergebnisse (vgl. Tab. 31).

Tab. 31: Gruppenunterschiede isolierende Sprachen vs. flektierende und agglutinierende Sprachen in den S-C-Tests, C-Tests und dem GJT.

Gruppen / Test	Mann-Whitney-U	p	Z	Φ
Isolierende Sprachen vs. flektierende und agglutinierende Sprachen S-C-Tests (n = 89)	988.500	.990	-.012	-.001
Isolierende Sprachen vs. Flektierende und agglutinierende Sprachen C-Tests (n = 89)	985.000	.967	-.041	-.004
Isolierende Sprachen vs. flektierende und agglutinierende Sprachen GJT (n = 89)	22.000	.000	-.634	-.067

Allerdings muss in diesem Zusammenhang auch die LOR der Testpersonen berücksichtigt werden, da sich LOR und Ausgangssprachen ungleich verteilen. Von den 44

Datensätzen – bei einem Fall liegt keine Angabe vor – mit einer isolierenden Ausgangssprache leben 25 Testpersonen seit mindestens fünf Jahren in einem deutschsprachigen Land. Von den 43 Datensätzen – auch hier liegt einmal keine Angabe vor – der Proband_innen mit flektierenden oder agglutinierenden Ausgangssprachen sind lediglich neun seit mindestens fünf Jahren im Land der Zielsprache.

Die grafische Auswertung zeigt ein deutlich besseres Abschneiden der neun Testpersonen mit flektierenden und agglutinierenden Ausgangssprachen bei den S-C-Tests und C-Tests (vgl. Abb. 24 und 25). Bei den GJTs hingegen sind die Testpersonen mit isolierenden Ausgangssprachen eindeutig stärker (vgl. Abb. 23).

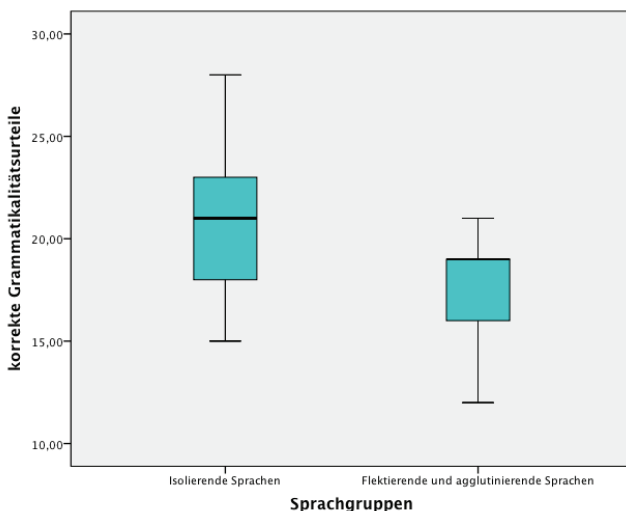


Abb. 23: Ergebnisse des GJTs nach sprachtypologischen Gruppen bei $LOR \geq 5$.

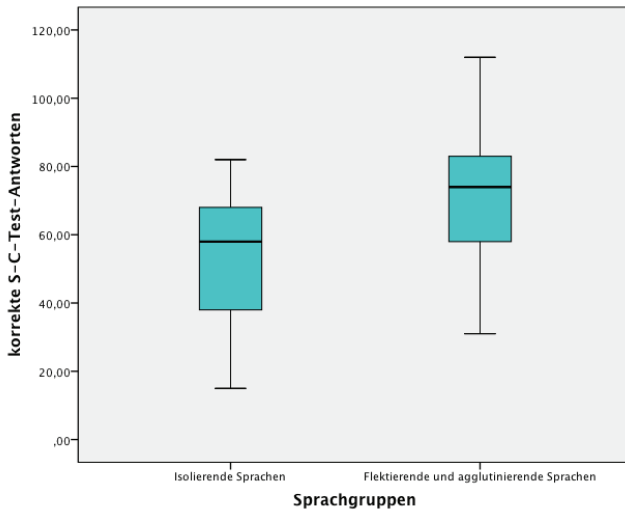


Abb. 24: Ergebnisse der S-C-Tests nach sprachtypologischen Gruppen bei $LOR \geq 5$.

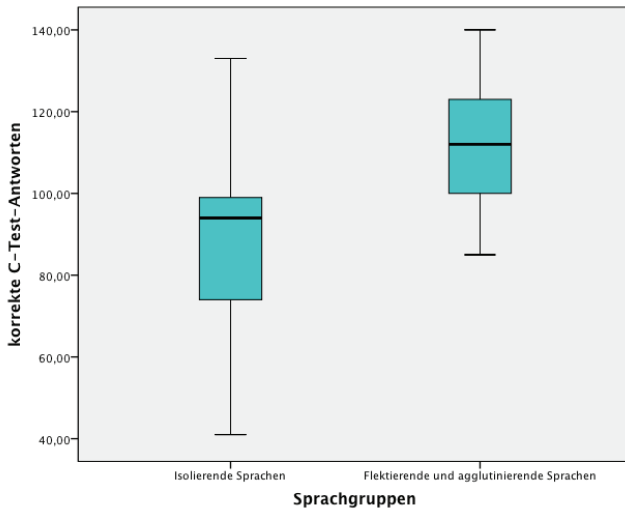


Abb. 25: Ergebnisse der C-Tests nach sprachtypologischen Gruppen bei $LOR \geq 5$.

Auch hier bestätigt der Mann-Whitney-U-Test den optischen Eindruck: Es gibt signifikante Unterschiede zwischen den beiden Gruppen bezogen auf die S-C-Test-, C-Test- und GJT-Ergebnisse (vgl. Tab. 32).

Tab. 32: Gruppenunterschiede isolierende Sprachen vs. flektierende und agglutinierende Sprachen mit $LOR \geq 5$ in den S-C-Tests, C-Tests und dem GJT.

Gruppen / Test	Mann-Whitney-U	p	Z	Φ
Isolierende Sprachen vs. flektierende und agglutinierende Sprachen S-C-Tests (n = 34)	62.000	.049	- 1.973	- .338
Isolierende Sprachen vs. flektierende und agglutinierende Sprachen C-Tests (n = 34)	38.500	.004	- 2.890	- .496
Isolierende Sprachen vs. flektierende und agglutinierende Sprachen GJT (n = 34)	59.000	.036	- 2.099	- .360

Für die Gruppe der Testpersonen mit einer LOR unter fünf Jahren (n = 19 mit isolierenden, n = 34 mit flektierenden und agglutinierenden Ausgangssprachen) fallen die Vergleiche wie folgt aus (vgl. Abb. 26, 27, 28).

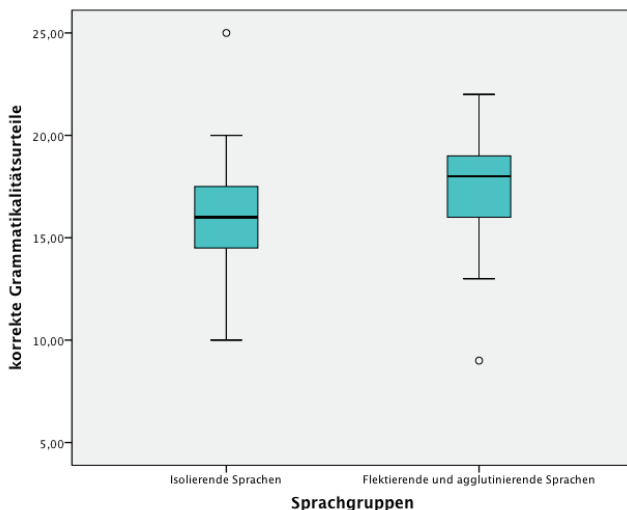


Abb. 26: Ergebnisse der GJTs nach sprachtypologischen Gruppen und einer $LOR < 5$.

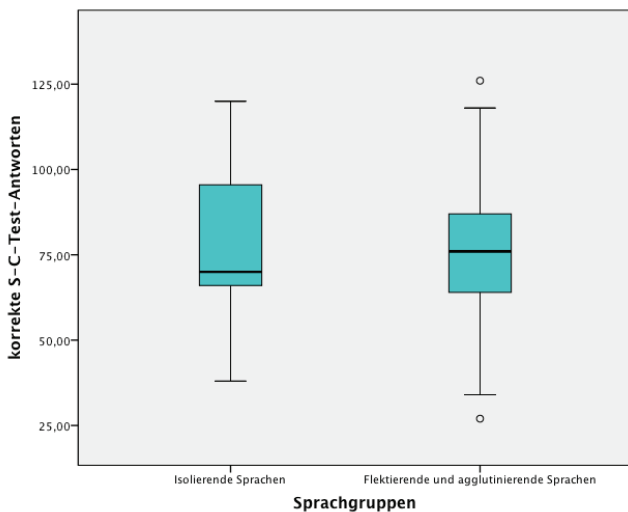


Abb. 27: Ergebnisse der S-C-Tests nach sprachtypologischen Gruppen und einer LOR < 5.

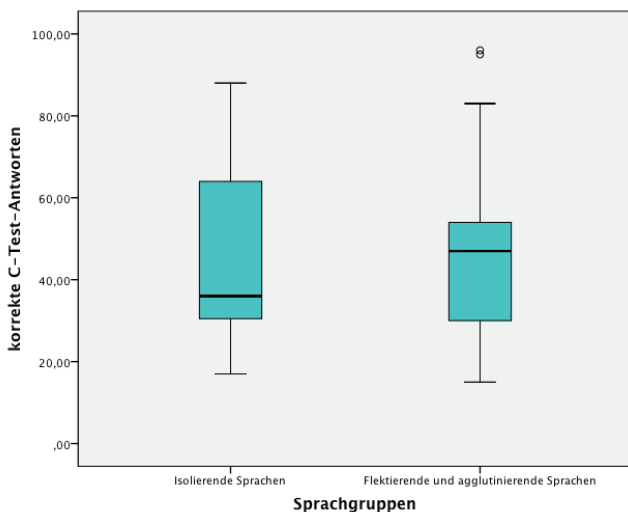


Abb. 28: Ergebnisse der C-Tests nach sprachtypologischen Gruppen und einer LOR < 5.

Bei den S-C- und C-Tests gibt es statistisch keine Unterschiede zwischen den Gruppen. Bei den GJTs schneidet die Gruppe der Testpersonen mit flektierenden und agglutinierenden Sprachen besser ab (vgl. Tab. 33).

Tab. 33: Gruppenunterschiede isolierende Sprachen vs. flektierende und agglutinierende Sprachen mit LOR < 5 in den S-C-Tests, C-Tests und dem GJT.

Gruppen / Test	Mann-Whitney-U	p	Z	Φ
Isolierende Sprachen vs. flektierende und agglutinierende Sprachen S-C-Tests (n = 34)	313.500	.860	- .175	- .024
Isolierende Sprachen vs. flektierende und agglutinierende Sprachen C-Tests (n = 34)	323.000	1.000	.000	0
Isolierende Sprachen vs. flektierende und agglutinierende Sprachen GJT (n = 34)	195.000	.017	- 2.391	- .328

Ein Blick auf die Mittelwerte der beiden Gruppen bezüglich des LORs zeigt jedoch, dass die Gruppen erheblich variieren. Zwar sind alle weniger als fünf Jahre im Land der Zielsprache. Bei der Gruppe mit isolierenden Ausgangssprachen ist der Mittelwert aber 18,32 Monate (Min: 3; Max: 52), während der Mittelwert bei der Gruppe der flektierenden und agglutinierenden Ausgangssprachen bei 8,21 Monaten liegt (Min: 1; Max: 54). Dieser Unterschied schränkt die gefundene Gruppengleichheit bei den S-C-Tests und C-Tests ein.

Der Einfluss der Ausgangssprache scheint sich abhängig von der LOR unterschiedlich auszuwirken. Allerdings können aufgrund der kleinen Fallzahl (n = 9) in der Gruppe der Testpersonen mit flektierenden und agglutinierenden Ausgangssprachen mit LOR ≥ 5 keine Aussagen getroffen werden, die die deskriptive Ebene überschreiten.

4.4 Diskussion der Ergebnisse

Die Auswertung ergibt, dass ein klarer Alterseffekt vorliegt. Die Regressionslinie verläuft bei allen Testformaten von links oben nach rechts unten (vgl. Abb. 10, Abb. 12, Abb. 13, Abb. 16 und Abb. 17), was auf einen klaren negativen Zusammenhang von AO und Testergebnissen verweist. Bezogen auf die Korrelationen auf Grundlage von allen Testsätzen ergeben sich deutliche Zahlen: AO korreliert am stärksten mit $Rho = -.444 (.000)$ bei den S-C-Tests, $Rho = -.447 (.000)$ bei den C-Tests und $Rho = -.379 (.000)$ beim GJT. Die Zahlen bedeuten jedoch auch, dass sich ein großer Teil des Lernerfolgs nicht durch das Alter erklären lässt, sondern von anderen Variablen abhängig ist. Infolgedessen sollte die Rolle des Alters weder ignoriert noch überschätzt werden (vgl. Moyer 2004: 89).

Die vorliegenden Ergebnisse werden im Folgenden in Bezug auf die beiden Ziele dieser Forschungsarbeit gebündelt und diskutiert.

Ziel 1:

- Messung und Vergleich der Sprachkompetenz von Personen mit doppeltem Erstspracherwerb (AO = 0–3 Jahre), Späterwerber_innen (AO \geq 16) und monolingualen Erstsprachler_innen.

Es zeigt sich, dass die Ergebnisse in allen Gruppen (‘Muttersprachler_innen’, Personen mit 2L1 und Späterwerber_innen) variieren. Wie zu erwarten, ist die Standardabweichung bei den Späterwerber_innen am höchsten und bei den ‘Muttersprachler_innen’ am niedrigsten (vgl. Tab. 7, Tab. 9, Tab. 10). Dennoch ist die individuelle Varianz der Ergebnisse ein Hinweis darauf, dass das erreichte Sprachniveau bei den ‘Muttersprachler_innen’ nicht so gleichförmig ist, wie das in der Literatur v. a. von Vertreter_innen der UG angenommen wird (vgl. Kap. 2.4.5 *Die Natur-Kultur-Debatte*).

Der Anteil der Späterwerber_innen, die im Bereich der Personen mit 2L1 abschneiden, fällt bei den schriftlichen Tests höher als erwartet aus: 14,6 % bei den S-C-Tests, 15,7 % bei den C-Tests, 5,6 % beim GJT, 5 % beim Akzent-Rating und 10 % beim Verständlichkeits-Rating. Hierzu ist zu bemerken, dass es sich bei den S-C-Tests, C-Tests und dem GJT um schriftliche Tests handelt. Bereits Johnson (1992)

zeigte, dass die Späterwerber_innen besser abschneiden, wenn sie die GJT-Items lesen anstatt sie zu hören (vgl. Abb. 3 in Kap. 2.2 *Studien zum Faktor Alter*). Das schwächere Ergebnis beim GJT könnte darauf beruhen, dass er mit dem Ziel konstruiert wurde, besondere Lernschwierigkeiten des Deutschen zu überprüfen, und daher mit entsprechenden Items geflutet wurde (vgl. Kap. 3.4.2.1 *Grammatical Judgement Test*). Insofern liegt die Vermutung nahe, dass die L2-Lerner_innen bei diesem Testformat mehr Schwierigkeiten haben als die beiden anderen Gruppen und bei den S-C-Tests und C-Tests, die zur Messung der allgemeinen Sprachkompetenz konstruiert wurden und nicht auf spezifische Lernschwierigkeiten des Deutschen ausgerichtet sind, öfter im Bereich der Personen mit 2L1 bzw. der ‚Muttersprachler_innen‘ abschneiden (vgl. Kap. 3.4.2.2 *C-Tests*).

Das Ergebnis beim Akzent-Rating ist mit 5 % erwartbar, denn Akzentfreiheit resp. Phonetik gilt als der Bereich der sprachlichen Kompetenz, der für die Späterwerber_innen – zumindest ohne explizites Training – schwer zu erreichen ist. Umgekehrt verhält es sich so, dass die Offsets für Phonetik besonders früh angenommen werden (vgl. Tab. 1, Kap. 2.1 *Kritische Periode(n)*). Ein genauer Blick auf die Ergebnisse der Phonetik-Ratings ergibt, dass nur drei von zehn Testpersonen aus der ‚muttersprachlichen‘ Kontrollgruppe mit 5 Punkten beim Akzent-Rating bewertet wurden. Bei den anderen wurde ein dialektal bedingter Akzent festgestellt. Dieses Ergebnis relativiert die hohe Wertigkeit, die der Akzentfreiheit bei L2-Lerner_innen zugeschrieben wurde, denn zumindest unter den hier getesteten Proband_innen wird sie von sieben von zehn nicht erreicht. Beim Verständlichkeits-Rating schneiden 10 % im Bereich der Personen mit 2L1 ab. Verständlichkeit ist, was die Erfüllung kommunikativer Bedürfnisse anbelangt, der wichtigere Wert.

Der Gruppenvergleich zwischen monolingualen ‚Muttersprachler_innen‘ und Personen mit 2L1 ergab signifikante Unterschiede bei den S-C- und C-Tests. Die ‚Muttersprachler_innen‘ schneiden besser ab und variieren weniger. Der größte Unterschied bei den Mittelwerten findet sich bei den S-C-Test-Ergebnissen. Dies könnte ein Hinweis darauf sein, dass der Automatisierungsgrad bei monolingualen ‚Muttersprachler_innen‘ höher ist als bei Personen mit 2L1. Bei den gefundenen Unterschieden zwischen monolingualen ‚Muttersprachler_innen‘ und Personen mit 2L1 bei S-C-Tests und C-Tests muss allerdings berücksichtigt werden, dass die Personen

mit 2L1 oftmals beträchtliche Lebensabschnitte in einem nicht deutschsprachigen Land verbracht wurden. So verbrachten 13 der 50 Personen ihre ersten Lebensjahre nicht in Deutschland; weitere 13 lebten in einem späteren Lebensabschnitt nicht in Deutschland.

Der Erkenntnisgewinn aus dem Vergleich der sprachlichen Kompetenzen von Personen mit 2L1, Späterwerber_innen und deutschen ‚Muttersprachler_innen‘ ist hoch einzuschätzen. Die ‚Muttersprachler_innen‘ erzielen die besten Ergebnisse und variieren am wenigsten. Bei den Personen mit 2L1 lassen sich zwar ebenfalls erheblich bessere Ergebnisse als bei den Späterwerber_innen feststellen, zugleich variieren die Ergebnisse in beiden Gruppen aber mehr als bei den ‚Muttersprachler_innen‘ (vgl. Tab. 9). Dabei ist die Standardabweichung in der Gruppe der Personen mit 2L1 bei den S-C-Tests (20.74) nur unwesentlich kleiner als in der Gruppe der Späterwerber_innen (22.00). Bei den C-Tests und S-C-Tests liegt der Anteil der Späterwerber_innen, die im Kompetenzbereich der Personen mit 2L1 abgeschnitten haben, bei 14,6 % bzw. 15,7 %. Diese hohen Prozentsätze deuten darauf hin, dass der Vergleich von bilingualen bzw. mehrsprachigen Testpersonen untereinander angemessener ist als der Vergleich mit monolingualen Erstsprachler_innen. Hinzugefügt werden muss, dass bei den Testpersonen keine Vorauswahl von besonders erfolgreichen Deutschler_innen stattfand.

Deutlich wird, dass die Nutzung eines Testformats, das eine sehr hohe allgemeine Sprachkompetenz differenziert misst (S-C-Test), geeigneter ist für die Bestimmung von Kompetenzunterschieden zwischen L1- und L2-Sprecher_innen als ein GJT, der auf besonders lernintensive grammatische Phänomene abzielt. Der Anteil der Späterwerber_innen, die wie Personen mit 2L1 abschneiden, liegt höchstwahrscheinlich aus diesem Grund bei den GJTs vergleichsweise niedrig bei 5,6 %. Dennoch liegen selbst diese Ergebnisse vergleichsweise hoch, wenn man sie mit den 5 % Späterwerber_innen vergleicht, die Birdsong (2005) zufolge im Bereich der „*nativelikeness*“ abschneiden sollten, um die CPH in Longs Formulierung ablehnen zu können:

„Moreover, a 5 % or greater incidence of *nativelikeness* would imply that there are substantial numbers of late learners who have attained near-nativelike proficiency. Age in SLA is not so constraining a factor that it prevents late learners from making remarkable strides toward *nativelikeness*, a fact recognized by Lenneberg [...]“ (Birdsong 2005: 121)

Der Einfluss der Ausgangssprache, hier verstanden als entweder isolierende Ausgangssprachen der Späterwerber_innen oder flektierende bzw. agglutinierende, wirkt sich in Abhängigkeit von der LOR unterschiedlich aus. Wenn alle Testpersonen getrennt nach dem Kriterium isolierend vs. flektierend bzw. agglutinierend ohne LOR zu berücksichtigen ausgewertet werden, können keine Gruppenunterschiede gefunden werden (vgl. Abb. 20, 21 und 22). Unter Berücksichtigung der LOR schneiden die neun Testpersonen mit flektierenden bzw. agglutinierenden Sprachen, die 5 Jahre und länger in Deutschland leben, bei den C-Tests und S-C-Tests deutlich besser ab (vgl. Abb. 24 und 25), während bei den GJTs die Proband_innen mit isolierenden Sprachen besser sind (vgl. Abb. 23). Aufgrund der kleinen Fallzahl ($n = 9$) in der Gruppe der Testpersonen mit flektierenden und agglutinierenden Ausgangssprachen mit $LOR \geq 5$ kann jedoch keine Aussage getroffen werden, die die deskriptive Ebene überschreitet.

Ziel 2:

- Spezifikation der Profile erfolgreicher Späterwerber_innen unter Berücksichtigung der ermittelten kognitiven, affektiv-motivationalen und sozio-kulturellen Einflussgrößen.

Neben dem großen Einfluss, den AO und LOR auf die erreichte Sprachkompetenz haben, konnten weitere Korrelationen festgestellt werden. Die einflussreichste Variable nach AO ist eine hohe Motivation für Deutsch. Darauf folgen die Fremdsprachenverwendungsangst (–), die Länge des DaF-Unterrichts in der Heimat, der Deutschgebrauch im Alltag, die Länge des DaF-Unterrichts insgesamt, die instrumentelle Motivation und der Gebrauch weiterer Sprachen im Alltag (–) (vgl. Tab. 15 in Kap. 4.1.4 *Korrelationen*). Bei folgenden Variablen wurde das Signifikanzniveau nicht erreicht: DaF-Unterricht in Deutschland in Stunden, L1-Gebrauch im Alltag, Motivation (insgesamt), Extraversion und Alter zum Zeitpunkt des Tests. Die gefundenen Korrelationen zeigen, dass Sprachlernerfolg multifaktoriell zu erklären ist. Es wäre erstaunlich gewesen, hier eine Variable zu finden, die deutlich höher liegt als alle anderen. Aus didaktischer Perspektive ist dieses Ergebnis positiv einzuschätzen: Während das Lebensalter der Lerner_innen nicht durch Unterricht zu beeinflussen ist, kann beispielsweise aktiv an der Motivation gearbeitet werden oder daran,

die Fremdsprachenverwendungsangst zu überwinden. Es können auch Wege gesucht werden, die Zielsprache im Alltag öfter anzuwenden.

Die Korrelationen für die LLAMA-Tests zur Messung der Sprachlerneignung erreichen bis auf eine nicht das Signifikanzniveau (vgl. Tab. 25), was angesichts der kleinen Proband_innenzahl ($n = 10$) in Phase II des hier vorliegenden Forschungsprojekts nicht irritierend ist. Die berechneten Korrelationen fallen gering aus und teilweise sogar negativ. Mit der Einschränkung, dass es sich möglicherweise um reine Zufallseffekte handelt, werden im Folgenden verschiedene Erwägungen präsentiert, um die geringen und negativen Werte zu erklären. Es könnte sein, dass das Testformat nicht angemessen ist und nicht das misst, was es vorgibt zu messen. Angesichts der Diskussion um Sprachlerneignung(en) – die gegenwärtig komplexer modelliert werden als zur Entstehungszeit des MLATs, auf dem die LLAMA-Tests basieren – wäre das nicht ausgeschlossen (vgl. Kap. 2.4.3 *Sprachlerneignung*). In den neueren Modellen spielt u. a. das Arbeitsgedächtnis eine wesentliche Rolle, das mit den LLAMA-Tests nicht erfasst werden kann. Ein anderer Erklärungsansatz bestünde in der Vermutung, dass die Auswahl der Testpersonen bezüglich des Sprachlernerfolgs zu homogen ist, um deutlichere Ergebnisse zu bekommen. Sicherlich variieren die Ergebnisse des GJTs, der S-C-Tests und der C-Tests, dennoch ist die Gruppe insofern einheitlich, als dass sie aus Personen besteht, die seit mindestens fünf Jahren in Deutschland leben und arbeiten und die größtenteils hier erfolgreich ein Studium absolviert haben. Würden wir diese Gruppe bspw. mit Studierenden vergleichen, die an der DSH- oder test-DaF-Prüfung trotz guter Vorbereitung scheitern, könnten die Ergebnisse vielleicht stärkere Zusammenhänge aufzeigen.

Die folgenden Anmerkungen zur Auswertung des Motivationsfragebogens der Phase II sind aufgrund der kleinen Zahl der Testpersonen genauso vorsichtig zu interpretieren wie die zu den LLAMA-Tests. Keine der berechneten Korrelationen erreicht das Signifikanzniveau. Die ermittelten Werte auf Grundlage des Fragebogens zur Motivation konnten die Skalenverhältnisse im Vergleich zu Dörnyei nur teilweise bestätigen. Allerdings muss dabei berücksichtigt werden, dass die Gruppe der untersuchten Späterwerber_innen zum Zeitpunkt der Erhebung die Zielsprache nicht mehr in einem expliziten Kontext lernte und der Fragebogen deswegen die meisten Merkmale in der Retrospektive erfasst. Werden die verschiedenen Motivationskomponenten auf die erreichte Sprachkompetenz bezogen, wie sie durch den GJT, die

S-C-Tests und C-Tests gemessen werden, ergeben sich keine nennenswerten Einflussgrößen, sondern die Werte variieren individuell stark. Dieser Befund weist in die Richtung von Riemers „Einzelgänger-Hypothese“, die besagt, dass der Fremdsprachenerwerb ein individueller Prozess ist, „bei dem u. a. bezüglich der Effektivität von Motivationen nur im Einzelfall entscheidbar ist, welche Motivationskomponenten tatsächlich zum Erfolg bzw. Misserfolg führen“ (Riemer 2005: 46). Bemerkenswerterweise ist Riemers Hypothese mit einem ganz anderen Forschungsdesign generiert worden, einer Longitudinalstudie über drei Fälle (vgl. Riemer 1997).

Die Auswertung der Hintergrundinformationen bestätigt die von Ioup et al. (1994) erfassten biografischen Daten der sehr erfolgreichen Fremdsprachenerwerberin, die sie untersuchten. Die besonders guten TN haben deutsche Partner_innen, einen deutschsprachigen Freundeskreis und sprechen Deutsch infolgedessen in den Lebensbereichen Wohnen und Freizeit. In den Bereichen Freizeit und Arbeit gebrauchen sie sowohl ihre Erstsprache als auch Deutsch. Sie lernen Deutsch für alltägliche sowie für berufliche und/oder akademische kommunikative Zwecke. Dabei war es ihr Ziel, ein sehr hohes bzw. ‚muttersprachliches‘ Niveau zu erreichen. Sie studieren in Deutschland und wünschen sich Feedback zu ihren zielsprachlichen Äußerungen.

Abschließend sollen Einsichten wiedergegeben werden, die sich im Laufe des Forschungsprojekts einstellten, sowie mögliche Themen und Fragen für weitere Forschungen benannt werden. In zukünftigen SLAR-Projekten, in denen quantitative Daten erhoben werden, sollten die Auswertungen mit den Verfahren der robusten Statistik vorgenommen werden. Diese Verfahren sind weniger voraussetzungsreich als die metrischen Verfahren. Beispielsweise müssen die Messwerte bzw. Residuen nicht normalverteilt sein, ordinalskalierte Variablen sind von der Analyse nicht ausgeschlossen und die Daten müssen nicht unabhängig voneinander gesammelt werden. Darüber hinaus ist es empfehlenswert über den ganzen Projektzeitraum hinweg das Einholen von statistischer Beratung und Expertise regelmäßig und fest zu verankern.

Bezogen auf die individuellen Differenzen bei den Späterwerber_innen wäre ein differenzierter Blick auf die verschiedenen Lehr- und Lernkontexte und die Lernstrategien erstrebenswert. Bezüglich der Sprachlerneignungstests muss festgehalten wer-

den, dass die Niveauunterschiede unter den getesteten Personen eventuell größer sein sollten. Auf jeden Fall ist zu empfehlen, wenn es darum geht spontansprachliche Äußerungen von Testpersonen für Phonetik-Ratings aufzunehmen, auf *elicited production tests* zurückzugreifen. Damit lässt sich die Intimsphäre der Proband_innen schützen und vermeiden, dass die Testpersonen Informationen über sich preisgeben, die die Rater beeinflussen könnten. Die Testpersonen über ein allgemeines und unverfängliches Thema sprechen zu lassen, ist ebenfalls eine Möglichkeit. So handhaben es Abrahamsson & Hyltenstam (2009: 264), es erfordert jedoch entsprechendes Vorwissen der Testpersonen und schließt die Erwähnung biografischer Daten nicht aus.

Von ganz besonderem Interesse wäre es, weit fortgeschrittene L2-Lerner_innen – so wie sie in der vorliegenden Arbeit in der Gruppe der Späterwerber_innen untersucht wurden – hinsichtlich aktueller Spracherwerbsprozesse genauer zu erforschen. Angeregt wurde dieser Gedanke durch die Nachfragen einzelner TN der Studie, was genau sie tun könnten, um ihr Deutsch weiter zu verbessern. Dies wird unterstützt von Hyltenstam und Abrahamsson (2012), die feststellen:

„As will be shown, the last phase of successful L2 acquisition is actually the one that is least well understood; however, the searchlight has begun to be directed toward the issues implicated at this level“ (ebd.: 177).

Eine weitergehende Untersuchung wäre an dieser Stelle sehr zu begrüßen. Dafür könnte es hilfreich sein, die Dimensionen BLC und HLC sowie Kernkomponenten und periphere Komponenten von Hulstijn entsprechend zu operationalisieren (2015, vgl. Kap. 2.5.1 *Die ‚Muttersprachliche Kompetenz‘ als Messlatte für den L2-Lernerfolg*). Studien mit dieser Perspektivierung könnten dazu beitragen, den „Bi/multilingual turn“ (Ortega 2013: 16) innerhalb der Spracherwerbsforschung zu befördern.

5. Zusammenfassung

Im Rahmen dieser Arbeit wurden Bedingungen erforscht, die Erwachsene erfolgreich lernen lassen, und Faktoren identifiziert, die individuelle Leistungsunterschiede beim L2-Lernen von sog. Späterwerber_innen ($AO \geq 16$) beeinflussen. Dabei wurden die anvisierten Ziele erreicht: Die Sprachkompetenz von Personen mit bilingualem Erstspracherwerb ($AO = 0-3$), Späterwerber_innen ($AO \geq 16$) und monolingualen Erstsprachler_innen wurden gemessen und miteinander verglichen. Dabei wurden die Profile erfolgreicher Späterwerber_innen spezifiziert.

Ausgehend von der umfangreichen Diskussion über den Einfluss des Faktors Alter auf das Fremdsprachenlernen wurden die Forschungsfragen und Ziele der Arbeit entwickelt. In der fachlichen Debatte wurde lange Zeit v. a. die CPH diskutiert. Anhand der Texte von Penfield (1959) und Lenneberg (1967) konnte gezeigt werden, inwiefern sie sich überhaupt auf den L2-Erwerb beziehen wollten und vor welchem Hintergrund und mit welcher Weitsicht sie dies getan haben. Danach wurden die zur CPH klassischen Studien von Patkowsky (1980) und Johnson & Newport (1989) ausführlich beschrieben und diskutiert. Die Darstellung aktuellerer Studien zum Thema (bspw. Abrahamsson & Hyltenstam 2009, 2012) sowie ein Überblick über die Forschungslage in Deutschland rundeten den Forschungsstand zum Faktor Alter ab.

Unter der Fragestellung „Was ist Alter? Was unterscheidet ältere von jüngeren Lerner_innen?“ wurden chronologisches Alter und die soziale Konstruktion von Alter gegenübergestellt. Des Weiteren wurde in diesem Kapitel der Forschungsstand zu den neurobiologischen Grundlagen der Reifungsprozesse zusammengefasst und diskutiert. Zusammen mit dem Thema Motivation wurden der altersspezifische Erwerbsantrieb (Klein 1996, 2001; Pagonis 2009b) und Dörnyeis (2009a) dynamisches Motivationsmodell dargestellt. Ebenso wurde die Entwicklung der Theorien zur Sprachlerneignung überblicksmäßig für die letzten 50 Jahre nachvollzogen. Neueste Entwicklungen bezüglich der Testformate (Hi-LAB) und Einblicke in die theoretische Diskussion (Robinsons ATI-Modell) vervollständigen dieses Unterkapitel. Seinen Abschluss findet das „Was ist Alter?“-Kapitel in einer kurzen Skizzierung der Naturkultur-Debatte. Diese ist insofern grundlegend für das Thema Alter, als die beiden Theorietypen die jeweilige Fundierung der Diskussionen über den Einfluss des Alters bilden. Während das nativistische Paradigma der Universalgrammatik seit Chomsky

die sprachwissenschaftliche Forschung bestimmt, zeichnet sich gegenwärtig ein Übergang zu emergentistischen bzw. interaktionistischen Ansätzen ab.

Im nächsten Kapitel wurde ausführlich das Thema ‚muttersprachliche Kompetenz‘ diskutiert. Dabei wurde hinterfragt, was darunter verstanden werden kann und ob die ‚Muttersprachlichkeit‘ ein angemessenes Ziel für den Fremdsprachenlernerfolg sein kann. In diesem Kapitel wurde ebenfalls erörtert, inwiefern die Überwindung des monolingualen Habitus (Gogolin 1994) dazu beiträgt, Chancengleichheit in einer mehrsprachigen Gesellschaft zu fördern. Im letzten Kapitel des Theorieteils wurde die Diskussion um den Einfluss der Ausgangssprache auf den Fremdsprachenlernerfolg zusammengefasst und beleuchtet, welche Rolle die typologische Ähnlichkeit bzw. Differenz von Ausgangssprache und Zielsprache möglicherweise spielt.

Im empirischen Teil wurde das Forschungsprojekt ausführlich vorgestellt. Ein besonderes Anliegen war dabei, den Bezug zum Forschungsprojekt ‚Erwerbsalter und Sprachlernerfolg‘ von Aguado, Grotjahn und Schlak darzustellen (2005, 2007), aus dem das hier vorliegende Projekt hervorgegangen ist. Zum einen entstammen die zum Einsatz gekommenen Instrumente der Phase I (GJT, S-C-Tests, C-Tests, TESTATT-Skalen und Fragebogen zur Sprachlernbiographie) dem Projekt. Zum anderen konnte für die Auswertung auf die von Aguado, Grotjahn und Schlak erhobenen Daten zurückgegriffen werden. Im Rahmen der vorliegenden Arbeit wurden zusätzlich Daten von 20 Späterwerber_innen des Deutschen mit der Ausgangssprache Mandarin sowie von 22 Personen mit 2L1 und den Sprachen Deutsch und Mandarin bzw. Kantonesisch erhoben. Die oben benannten Instrumente der ersten Phase werden mit einem Phonetiktest ergänzt. In der zweiten Phase absolvierten zehn ausgewählte Proband_innen aus der Gruppe der Späterwerber_innen Sprachlerneignungstests (LLAMA-Tests), einen weiteren Fragebogen zu persönlichen Hintergrundinformationen sowie eine zusätzliche Motivationskala.

Bei der statistischen Auswertung der Daten wurden sowohl die Ergebnisse der einzelnen Sprachkompetenztests präsentiert als auch Korrelationen zwischen den Sprachtestergebnissen und individuellen Variablen berechnet. Besonders bedeutsam ist, dass hierbei Späterwerber_innen mit Personen mit bilinguaem Erstspracherwerb und nicht ausschließlich mit monolingualen Erstsprachler_innen verglichen werden. Die Ergebnisse der Personen mit 2L1 variieren stärker als die der monolingualen Erstsprachler_innen. Einerseits gibt es einen Anteil an bilingualen Erst-

sprachler_innen, die wie die monolingualen abschneiden. Andererseits streuen die Daten in der Gruppe mit 2L1 weiter, was zur Folge hat, dass es mehr Späterwerber_innen gibt, die innerhalb des Bereichs der Personen mit 2L1 liegen. Zirka ein Sechstel der hier untersuchten Späterwerber_innen schneidet bei den C-Tests und S-C-Tests im Bereich der Personen mit 2L1 ab. Der Vergleich mit den bilingualen Erstsprachler_innen ist aussagekräftiger, wenn es darum geht, den Lernerfolg von Späterwerber_innen zu bemessen, weil Mehrsprachige mit Mehrsprachigen verglichen werden. Darüber hinaus ist das Ergebnis dieses Vergleichs motivierender für erwachsene L2-Lerner_innen.

Es wurde ein deutlicher Alterseffekt festgestellt. AO ist der Faktor, der am stärksten mit der gemessenen Sprachkompetenz korreliert. Dennoch bleibt festzuhalten, dass andere Faktoren ebenfalls eine wichtige Rolle spielen: Nach AO weisen die Motivation der Lerner_innen, Deutsch auf hohem Niveau zu beherrschen, und die LOR den größten Zusammenhang auf. Es besteht ebenfalls ein Zusammenhang mit der Fremdsprachenverwendungsangst, der Länge des DaF-Unterrichts in der Heimat, dem Deutschgebrauch im Alltag, der Länge des DaF-Unterrichts insgesamt und der instrumentellen Motivation. Weitere Sprachen im Alltag zu benutzen, hat einen negativen Effekt. Für den Gebrauch der L1 im Alltag kann ein geringer positiver Zusammenhang auf das zielsprachliche Kompetenzprofil gezeigt werden.

Zwischen den Späterwerber_innen mit isolierenden Ausgangssprachen und denjenigen mit flektierenden oder agglutinierenden wurden in Abhängigkeit von der LOR Gruppenunterschiede gefunden. Wird die LOR der Testpersonen bei der Auswertung nicht berücksichtigt, ergeben sich bei den hier untersuchten Personen keine Gruppenunterschiede.

In der zweiten Phase konnten Interpretationen auf Grundlage der Auswertungen der LLAMA-Tests und des Motivationsfragebogens wegen der geringen Datenbasis (n = 10) nur mit großen Einschränkungen vorgenommen werden. Die hier gewonnenen Hinweise können in die Entwicklung zukünftiger Forschungsprojekte eingehen.

Zum Schluss wurden biografische Profile erarbeitet, die erfolgreiche, mittlere und schwächere L2-Lerner_innen beschreiben. Erfolgreiche L2-Lerner_innen zeichnen sich v. a. dadurch aus, dass sie Deutsch in allen Bereichen (Arbeit, Wohnen, Freizeit) sprechen, sowohl beruflich wie auch privat Deutsch und ihre Erstsprache spre-

chen, einen deutschen Freundeskreis und eine_n deutsche_n Partner_in haben. Diese Profile dürfen aufgrund der kleinen Datengrundlage nicht verallgemeinert werden und sind explorativ zu verstehen.

6. Abkürzungsverzeichnis

ACT = Adaptive Control of Thought

ALTE = Association of Language Testers in Europe

AO = Age of Onset

AOA = Age of Arrival

AOI = Age of Immersion

BICS = Basic Interpersonal Communicative Skills

BLC = Basic Language Cognition

CALP = Cognitive Academic Language Proficiency

CANAL-F = Cognitive Ability for Novelty in Acquisition of Language – Foreign

CP = Critical Period

CPH = Critical Period Hypothesis

CUP = Common Underlying Proficiency

DaF = Deutsch als Fremdsprache

DaZ = Deutsch als Zweitsprache

DaZ-AF-Korpus = Korpus Deutsch als Fremdsprache – Altersfaktor

DSH = Deutsche Sprachprüfung für den Hochschulzugang

EEG = Elektroenzephalogramm

EKP = Ereigniskorrelierte Potentiale

ELPS = European Language Proficiency Surveys

ELF = English as Lingua Franca

ELT = English Learning Teaching

FDH = Fundamental Difference Hypothesis

fMRT = Funktionelle Magnetresonanztomographie

GAMA = General Intelligence Test

GJT = Grammatical Judgement Test
Hi-LAB = High-Level Language Aptitude Battery
HLC = Higher Language Cognition
L1 = Erstsprache
L2 = Zweit- oder Fremdsprache
2L1 = bilingualer Erstspracherwerb
LLAMA = Language Aptitude Tests
LOR = Length of Residency
MLAT = Modern Language Aptitude Test
onSET-Deutsch = Online Einstufungstest Deutsch als Fremdsprache
OSPAN = Operation Span Test
P10 = 10. Perzentil
PCA = Principal Component Analysis
PET = Positronen-Emissions-Tomographie
PLAB = Pimsleur Language Aptitude Battery
S-C-Tests = Speeded C-Tests (zeitlimitierte C-Tests)
SLAR = Second Language Acquisition Research
SRT = Probabilistic Serial Reaction Time Task
TESTATT = Tests and Attitude Scales for the Year Abroad
TN = Teilnehmer_in
UA = Ultimate Attainment
UG = Universalgrammatik
V2-Stellung = Verb-Zweit-Stellung
VE-Stellung = Verb-End-Stellung
XV-Stellung = Distanzstellung des finiten und infiniten Verbteils in der Satzklammer

7. Literatur

- Abrahamsson, Niclas (2012), Age of onset and nativelike attainment of morphosyntactic and phonetic intuition. *Studies in Second Language Acquisition*, 34: 2, 187–214.
- Abrahamsson, Niclas & Hyltenstam, Kenneth (2008), The Robustness of Aptitude Effects in Near-Native Second Language Acquisition. *Studies in Second Language Acquisition*, 30: 4, 481–509.
- Abrahamsson, Niclas & Hyltenstam, Kenneth (2009), Age of Onset and Nativelikeness in a Second Language: Listener Perception Versus Linguistic Scrutiny. *Language Learning. A Journal of Research in Language Studies*, 59: 2, 249–306.
- Aguado, Karin (2010), Critical Period. In: Hans Barkowski & Hans-Jürgen Krumm (Hrsg.), *Fachlexikon Deutsch als Fremd- und Zweitsprache*. Tübingen [u. a.]: Francke, 39.
- Aguado, Karin (2010), Identitätshypothese. In: Hans Barkowski & Hans-Jürgen Krumm (Hrsg.), *Fachlexikon Deutsch als Fremd- und Zweitsprache*. Tübingen [u. a.]: Francke, 123.
- Aguado, Karin (2010), Interlanguage-Hypothese. In: Hans Barkowski & Hans-Jürgen Krumm (Hrsg.), *Fachlexikon Deutsch als Fremd- und Zweitsprache*. Tübingen [u. a.]: Francke, 142.
- Aguado, Karin (2012), Language learning aptitude and foreign language learning. In: Wai Meng Chan, Kwee Nyet Chin, Sunil Kumar Bhatt & Izumi Walker (Hrsg.), *Perspectives on Individual Characteristics and Foreign Language Education*. Boston, Berlin: De Gruyter Mouton, 51–69.
- Aguado, Karin (2013), Sind L2-Chunks lebenslang erlernbar oder gibt es eine sensible Phase für ihren erfolgreichen Erwerb? In: Annette Berndt (Hrsg.), *Fremdsprachen in der Perspektive lebenslangen Lernens*. Frankfurt a. M. [u. a.]: Peter Lang, 129–142.
- Aguado, Karin; Grotjahn, Rüdiger & Schlak, Torsten (2005), Erwerbssalter und Sprachlernerfolg. Theoretische und methodologische Grundlagen eines empirischen Forschungsprojekts. *Zeitschrift für Fremdsprachenforschung*, 16: 2, 275–293.

- Aguado, Karin; Grotjahn, Rüdiger & Schlak, Torsten (2007), Erwerbsalter und Sprachlernerfolg: Zeitlimitierte C-Tests als Instrument zur Messung prozeduralen sprachlichen Wissens. In: Helmut Johannes Vollmer (Hrsg.), Synergieeffekte in der Fremdsprachenforschung. Empirische Zugänge, Probleme, Ergebnisse. Frankfurt a. M. [u. a.]: Peter Lang, 137–149.
- Ahlzweig, Claus (1994), Muttersprache – Vaterland. Die deutsche Nation und ihre Sprache. Opladen: Westdeutscher Verlag.
- Ahrenholz, Bernt (2008), Erstsprache – Zweitsprache – Fremdsprache. In: Bernt Ahrenholz & Ingelore Oomen-Welke (Hrsg.), Deutsch als Zweitsprache. Baltmannsweiler: Schneider, 3–17.
- Ahrenholz, Bernt (Hrsg.) (2006), Kinder mit Migrationshintergrund. Spracherwerb und Förderungsmöglichkeiten. Freiburg im Breisgau: Fillibach.
- Al-Shehri, Abdullah S. (2009), Motivation and Vision: The Relation Between the Ideal L2 Self, Imagination and Visual Style. In: Zoltán Dörnyei & Ema Ushioda (Hrsg.), Motivation, Language Identity and the L2 Self. Bristol [u. a.]: Multilingual Matters, 164–171.
- Anderson, John R. & Lebiere, Christian (1998), The atomic components of thought. Mahwah, NJ: Erlbaum.
- Andresen, Helga (2010), Erstsprache. In: Helmut Glück (Hrsg.), Metzler Lexikon Sprache (4. Aufl.). Stuttgart: Metzler, 184.
- Andrew, Patricia (2012), The Social Construction of Age. Adult Foreign Language Learners. Bristol [u. a.]: Multilingual Matters.
- Andrews, Edna; Frigau, Luca; Voyvodic-Casabo, Clara; Voyvodic, James & Wright, John (2013), Multilingualism and fMRI: Longitudinal Study of Second Language Acquisition. *Brain Sciences*, 3, 849–876.
- Bennett, Craig M.; Baird, Abigail A.; Miller, Michael B. & Wolford, George L. (2009), Neural Correlates of Interspecies Perspective Taking in the Post-Mortem Atlantic Salmon: An Argument For Multiple Comparisons Correction. *Proceedings from Human Brain Mapping*.
- Bennett, Craig M.; Baird, Abigail A.; Miller, Michael B. & Wolford, George L. (2010), Neural Correlates of Interspecies Perspective Taking in the Post-Mortem Atlantic Salmon: An Argument For Multiple Comparisons Correction. *Journal of Serendipitous and Unexpected Results*, 1–5.

- Bennett, Craig M. & Miller, Michael B. (2010), How reliable are the results from functional magnetic resonance imaging? *Annals of the New York Academy of Sciences*, 1191, 133–155.
- Berndt, Annette (2003), *Sprachenlernen im Alter. Eine empirische Studie zur Fremdspracherwerbsforschung*. München: Iudicium.
- Bialystok, Ellen (1997), The structure of age: in search of barriers to second language acquisition. *Second Language Research*, 13: 2, 116–137.
- Bialystok, Ellen & Hakuta, Kenji (1994), *In Other Words. The Science and Psychology of Second-Language Acquisition*. New York: BasicBooks.
- Birdsong, David (2005), Interpreting Age Effects in Second Language Acquisition. In: Judith F. Kroll & Annette M. B. de Groot (Hrsg.), *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford: Oxford University Press, 109–127.
- Birdsong, David (2009), Age and the end state of second language acquisition. In: William C. Ritchie & Thejk Bhatia (Hrsg.), *The New Handbook of Second Language Acquisition*. Amsterdam: Elsevier, 401–424.
- Birdsong, David & Molis, Michelle (2001), On the Evidence for maturational constraints in second language acquisition. *Journal of Memory and Language*, 44, 235–249.
- Bley-Vroman, Robert W. (1990), The logical problem of foreign language learning. *Linguistic Analysis*, 20: 1–2, 3–49.
- Bley-Vroman, Robert W. (2009), The evolving context of the fundamental difference hypothesis. *Studies in Second Language Acquisition*, 31, 175–198.
- Bley-Vroman, Robert W.; Felix, Sascha W. & Ioup, Georgette L. (1988), The accessibility of Universal Grammar in adult language learning. *Second Language Research*, 4, 1–32.
- Boeckmann, Klaus-Börge (2010), Kontrastivhypothese. In: Hans Barkowski & Hans-Jürgen Krumm (Hrsg.), *Fachlexikon Deutsch als Fremd- und Zweitsprache*. Tübingen: Francke, 169.
- Bohn, Ocke-Schwen (1998), Wahrnehmung fremdsprachlicher Laute: Wo ist das Problem? In: Heide Wegener (Hrsg.), *Eine zweite Sprache lernen: empirische Untersuchungen zum Zweitspracherwerb*. Tübingen: Narr, 1–20.
- Bonfiglio, Thomas Paul (2010), *Mother tongues and nations: the invention of the native speaker*. New York: De Gruyter Mouton.

- Bongaerts, Theo; Mennen, Susan & van der Silk, Frans (2000), Authenticity of pronunciation in naturalistic second language acquisition: The case of very advanced late learners of Dutch as a second language. *Studia Linguistica*, 54: 2, 298–308.
- Bongaerts, Theo; Planken, Brigitte & Schils, Erik (1995), Can late starters attain a native accent in a foreign language? A test of the critical period hypothesis. In: David Singleton & Zsolt Lengyl (Hrsg.), *The Age Factor in Second Language Acquisition*. Clevedon: Multilingual Matters, 30–50.
- Bongaerts, Theo; Planken, Brigitte; Schils, Erik & van Summeren, Chantal (1997), Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19, 447–465.
- van Boxtel, Sonja; Bongaerts, Theo & Coppen, Peter-Arno (2005), Native-like attainment of dummy subjects in Dutch and the role of the L1. *International Review of Applied Linguistics*, 43, 355–380.
- Button, Katherine S.; Ioannidis, John P. A.; Mokrysz, Claire; Nosek, Brian A.; Flint, Jonathan; Robinson, Emma S. J. & Munafò, Marcus R. (2013), Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14, 365–376.
- Carroll, John B. (1962), The prediction of success in intensive foreign language training. In: Robert Glaser (Hrsg.), *Training research and education*. New York: John Wiley & Sons, 87–136.
- Carroll, John B. (1981), Twenty-five years of research on foreign language aptitude. In: Karl C. Diller (Hrsg.), *Individual Differences and Universals in Language Learning Aptitude*. Rowley: Newbury House, 83–118.
- Carroll, John B. & Sapon, Stanley M. (1959), *Modern Language Aptitude Test*. Washington: Second Language Testing Inc.
- Chen, Judy F.; Warden, Clyde A.; Chang, Huo-Tsan (2005): Motivators That Do Not Motivate: The Case of Chinese EFL Learners and the Influence of Culture on Motivation. *TESOL Quarterly*, 39: 4, 609–633.
- Chomsky, Noam (1965), *Aspects of the theory of syntax*. Cambridge: M.I.T. Press.
- Chomsky, Noam (1980), *Rules and representations*. New York: Columbia University Press.
- Chomsky, Noam (2007), Bilingualistic explorations: Design, development, evolution. *International Journal of Philosophical Studies*, 15, 1–12.

- Clément, Richard & Kruidenier, Bastian G. (1983), Orientations in second language acquisition: The effects of ethnicity, milieu, and target language on their emergence. *Language Learning*, 33, 273–291.
- Cook, Vivian & Singleton, David (2014), Is There a Best Age for Learning a Second Language? In: Vivian Cook & David Singleton (Hrsg.), *Key Topics in Second Language Acquisition*. Bristol [u. a.]: *Multilingual Matters*, 17–36.
- Cummins, Jim (1976), The influence of bilingualism on cognitive growth: a synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism*, 9, 1–43.
- Cummins, Jim (1981), The role of primary language development in promoting educational success for language minority students. In: *Schooling and Language Minority Students: A Theoretical Framework*. Los Angeles: Evaluation, Dissemination and Assessment Center California State University, 3–49.
- Cummins, Jim (2000), *Language, Power and Pedagogy. Bilingual Children in the Crossfire*. Clevedon: *Multilingual Matters*.
- Cummins, Jim (2008), BICS and CALP: Empirical and theoretical status of the distinction. In: Brian Street & Nancy H. Hornberger (Hrsg.), *Encyclopedia of Language and Education*. Band 2: Literacy (2. Aufl.). New York: Springer, 71–83.
- Czinger, Christine (2014), *Grammatikerwerb vor und nach der Pubertät. Eine Fallstudie zur Verbstellung im Deutschen als Zweitsprache*. Berlin [u. a.]: De Gruyter.
- Dabrowska, Ewa (2012), Different speakers, different grammars. Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, 2(3), 219–253.
- Davies, Alan (2003), *The Native Speaker: Myth and Reality* (2. Aufl.). Clevedon; Buffalo: *Multilingual Matters*.
- Davies, Alan (2011), Does Language Testing Need the Native Speaker? *Language Assessment Quarterly*, 8: 3, 291–308.
- De Bot, Kees (2008), Review article: The imaging of what in the multilingual mind? *Second Language Research*, 24: 1, 111–133.
- Deci, Edward L. & Ryan, Richard M. (1985), *Intrinsic motivation and self-determination in human behavior (Perspectives in social psychology)*. New York: Plenum.

- DeKeyser, Robert (2000), The Robustness of Critical Period Effects on Second Language Learning. *Studies in Second Language Acquisition*, 22: 4, 499–533.
- DeKeyser, Robert (2012), Individual differences in native language attainment and their implications for research on second language acquisition. *Linguistic Approaches to Bilingualism*, 2(3), 260–263.
- DeKeyser, Robert; Alfi-Shabatay, Iris & Ravit, Dorit (2010), Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics*, 31, 413–438.
- DeKeyser, Robert & Koeth, Joel (2011), Cognitive Aptitudes for Second Language Learning. In: Eli Hinkel (Hrsg.), *Handbook of Research in Second Language Teaching and Learning*. New York [u. a.]: Routledge, 395–406.
- DeKeyser, Robert & Larson-Hall, Jenifer (2005), What does the Critical Period really mean? In: Judith F. Kroll & Annette M. B. DeGroot (Hrsg.), *Handbook of Bilingualism*. Oxford: Oxford University Press, 88–108.
- Dimroth, Christine (2007), Zweitspracherwerb bei Kindern und Jugendlichen: Gemeinsamkeiten und Unterschiede. In: Tanja Anstatt (Hrsg.), *Mehrsprachigkeit bei Kindern und Erwachsenen: Erwerb, Formen, Förderung*. Tübingen: Narr, 115–138.
- Dimroth, Christine & Haberzettl, Stefanie (2008), Je älter desto besser: der Erwerb der Verbflexion im Kindesalter. In: Bernt Ahrenholz, Ursula Bredel, Wolfgang Klein, Martina Rost-Roth & Romuald Skiba (Hrsg.), *Empirische Forschung und Theoriebildung. Beiträge aus der Soziolinguistik, Gesprochene-Sprache- und Zweitspracherwerbsforschung*. Frankfurt a. M.: Peter Lang, 227–239.
- Dimroth, Christine & Klein, Wolfgang (2003), Der ungesteuerte Zweitspracherwerb Erwachsener: Ein Überblick über den Forschungsstand. In: Utz Maas & Ulrich Mehlem (Hrsg.), *Qualitätsanforderungen für die Sprachförderung im Rahmen der Integration von Zuwanderern*. Osnabrück: Institut für Migrationsforschung und Interkulturelle Studien, 127–161.
- Dörnyei, Zoltán (2005), *The Psychology of the Language Learner. Individual Differences in Second Language Acquisition*. New York, London: Lawrence Erlbaum Associates.
- Dörnyei, Zoltán (2009a), The L2 Motivational Self System. In: Zoltán Dörnyei & Ema Ushioda (Hrsg.), *Motivation, Language Identity and the L2 Self*. Bristol [u. a.]: Multilingual Matters, 9–42.

- Dörnyei, Zoltán (2009b), *The Psychology of Second Language Acquisition*. Oxford, New York: Oxford University Press.
- Dörnyei, Zoltán (2011), The L2 Motivational Self System. In: Zoltán Dörnyei & Ema Ushioda (Hrsg.), *Motivation, Language, Identity and the L2 Self*. Bristol: *Multilingual Matters*, 9–42.
- Dörnyei, Zoltán & Skehan, Peter (2003), Individual differences in second language learning. In: Catherine J. Doughty & Michael H. Long (Hrsg.), *The Handbook of Second Language Acquisition*. Oxford: Blackwell, 589–630.
- Dörnyei, Zoltán & Ushioda, Ema (Hrsg.) (2009), *Motivation, Language Identity and the L2 Self*. Bristol [u. a.]: *Multilingual Matters*.
- Doughty, Catherine J.; Campbell, Susan G.; Mislevy, Meredith A.; Bunting, Michael F.; Bowles, Anita R. & Koeth, Joel T. (2010), Predicting near-native L2 ability. *Proceedings of the 2008 Second Language Research Form*. Somerville (MA): Cascadilla Press.
- Eckes, Thomas & Grotjahn, Rüdiger (2006), A closer look at the construct validity of C-tests. *Language Testing*, 23: 3, 290–325.
- Ellis, Nick C. & Wulff, Stefanie (2015), Usage-Based Approaches to SLA. In: Bill VanPatten & Jessica Williams (Hrsg.), *Theories in Second Language Acquisition. An Introduction* (2. Aufl.). New York; London: Routledge, 75–93.
- Ellis, Rod (1991), Grammaticality judgements and second language acquisition. *Studies in Second Language Acquisition*, 13: 2, 161–186.
- Ellis, Rod (1994), *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Ellis, Rod (2004), Individual differences in second language learning. In: Alan Davies & Catherine Elder (Hrsg.), *The Handbook of Applied Linguistics*. Oxford: Wiley-Blackwell, 525–551.
- Ellis, Rod (2008), *The Study of Second Language Acquisition* (2. Aufl.). Oxford: Oxford University Press.
- Evans, Nicholas & Levinson, Stephen C. (2009), The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32, 429–492.
- Felix, Sascha W. (1985), More evidence on competing cognitive systems. *Second Language Research*, 1, 47–72.

- Flege, James Emil; Yeni-Komshian, Grace H. & Liu, Serena (1999), Age Constraints on Second-Language Acquisition. *Journal of Memory and Language*, 41, 78–104.
- Flynn, Suzanne (1987), *A Parameter-Setting Model of L2 Acquisition: Experimental Studies in Anaphora*. Dordrecht: Reidel.
- Gardner, Robert C. (2001), Integrative Motivation and Second Language Acquisition. In: Zoltán Dörnyei & Richard W. Schmidt (Hrsg.), *Motivation and second language acquisition*. Honolulu: University of Hawaii Press, 1–20.
- Gardner, Robert C. & Lambert, Wallace E. (1959), Motivational Variables in Second Language Acquisition. *Canadian Journal of Psychology*, 13: 4, 266–280.
- Glück, Martin (2010a), Brocasches Zentrum. In: Helmut Glück (Hrsg.), *Metzler Lexikon Sprache* (4. Aufl.). Stuttgart: Metzler, 103.
- Glück, Martin (2010b), Wernickesches Feld. In: Helmut Glück (Hrsg.), *Metzler Lexikon Sprache* (4. Aufl.). Stuttgart: Metzler, 689.
- Gogolin, Ingrid (1994), *Der monolinguale Habitus der multilingualen Schule*. Münster, New York: Waxmann.
- Görlach, Manfred (1993), Englisch. In: Helmut Glück (Hrsg.), *Metzler Lexikon Sprache*. Stuttgart: Metzler, 164–165.
- Granena, Gisela (2013a), Cognitive aptitudes for second language learning and the LLAMA Language Aptitude Test. In: Gisela Granena & Mike Long (Hrsg.), *Sensitive periods, language aptitude, and ultimate L2 attainment*. Amsterdam: Benjamins, 105–130.
- Granena, Gisela (2013b), Reexamining the robustness of aptitude in second language acquisition. In: Gisela Granena & Mike Long (Hrsg.), *Sensitive periods, language aptitude, and ultimate L2 attainment*. Amsterdam: Benjamins, 179–204.
- Granena, Gisela & Long, Michael H. (2012), Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, 29(3), 311–343.
- Grigorenko, Elene L.; Sternberg, Robert J. & Ehrman, Madeline E. (2000), A Theory-Based Approach to the Measurement of Foreign Language Learning Ability: The Canal-F Theory and Test. *The Modern Language Journal*, 84: 3, 390–405.
- Grosjean, François (1989), Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36(1), 3–15.

- Grotjahn, Rüdiger (1995), Der C-Test: State of the Art. Zeitschrift für Fremdsprachenforschung, 6: 2, 37–60.
- Grotjahn, Rüdiger (2003), Der Faktor „Alter“ beim Fremdsprachenlernen. Mythen, Fakten, didaktisch-methodische Implikationen. Deutsch als Fremdsprache, 40: 1, 32–41.
- Grotjahn, Rüdiger (2004), „Tests and Attitude Scales for the Year Abroad“ (TESTATT): Sprachlernmotivation und Einstellungen gegenüber Sprechern der eigenen und der fremden Sprache. Zeitschrift für Interkulturellen Fremdsprachenunterricht, 9: 2, 1–23.
- Grotjahn, Rüdiger (2005), Je früher, desto besser? – Neuere Befunde zum Einfluss des Faktors „Alter“ auf das Fremdsprachenlernen. In: Heiner Pürschel & Thomas Tinnefeld (Hrsg.), Moderner Fremdspracherwerb zwischen Interkulturalität und Multimedia: Reflexionen und Anregungen aus Wissenschaft und Praxis. Bochum: AKS-Verlag, 186–202.
- Grotjahn, Rüdiger (2010), Gesamtdarbietung, Einzeltextdarbietung, Zeitbegrenzung und Zeitdruck: Auswirkungen auf Item- und Testkennwerte und C-Test-Konstrukt. In: Rüdiger Grotjahn (Hrsg.), Der C-Test: Beiträge aus der aktuellen Forschung/The C-Test: Contributions from current research. Frankfurt a. M.: Lang, 265–296.
- Grotjahn, Rüdiger (2011), C-Tests – Aspekte der Validität. Deutsch als Fremdsprache, 48: 3, 131–137.
- Grotjahn, Rüdiger & Schlak, Torsten (2010), Alter. In: Hans-Jürgen Krumm, Christian Fandrych, Britta Hufeisen & Claudia Riemer (Hrsg.), Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch (2. Aufl.). Berlin, New York: De Gruyter Mouton: 867–876.
- Grotjahn, Rüdiger; Schlak, Torsten & Aguado, Karin (2010), S-C-Tests: Messung automatisierter sprachlicher Kompetenzen anhand von C-Tests mit massiver textspezifischer Zeitlimitierung. In: Rüdiger Grotjahn (Hrsg.), Der C-Test: Beiträge aus der aktuellen Forschung. Frankfurt a. M. [u. a.]: Peter Lang, 297–319.
- Grotjahn, Rüdiger & Schlak, Torsten (2013), Alter und Fremdsprachenlernen – Ein Forschungsüberblick. In: Annette Berndt (Hrsg.), Fremdsprachen in der Perspektive lebenslangen Lernens. Frankfurt a. M. [u. a.]: Peter Lang, 7–23.
- Gültekin-Karakoç, Nazan & Feldmeier, Alexis (2014), Analyse quantitativer Daten. In: Julia Settineri, Sevilen Demirkaya, Alexis Feldmeier, Nazan Gültekin-Karakoç & Claudia Riemer (Hrsg.), Empirische Forschungsmethoden für Deutsch als Fremd- und Zweitsprache: Eine Einführung. Stuttgart: UTB, 183–211.

- Han, ZhaoHong (2004), To be a native speaker means not to be a nonnative speaker. Review of Davies (2003). *Second Language Research*, 20: 2, 166–187.
- Hasler, Felix (2012), *Neuromythologie. Eine Streitschrift gegen die Deutungsmacht der Hirnforschung*. transcript.
- Henrici, Gert & Riemer, Claudia (2007), Zweitsprachenerwerbsforschung. In: Karl-Richard Bausch, Herbert Christ & Hans-Jürgen Krumm (Hrsg.), *Handbuch Fremdsprachenunterricht* (5. Aufl.). Tübingen [u. a.]: Francke: 38–43.
- Herder, Johann Gottfried (1795), *Sämtliche Werke*, Band XVII, hg. von Bernhard Suphan, Berlin.
- Hernig, Marcus (2010), Deutsch in China. In: Hans-Jürgen Krumm, Christian Fandrych, Britta Hufeisen & Claudia Riemer (Hrsg.), *Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch* (2. Aufl.). Berlin, New York: De Gruyter Mouton: 1637–1642.
- Herschensohn, Julia (2007), *Language Development and Age*. Cambridge: Cambridge University Press.
- Holliday, Adrian (2006), Native-Speakerism. *ELT-Journal*, 60: 4, 385–387.
- Holliday, Adrian (2008), Standards of English and politics of inclusion. *Language Teaching*, 41: 1, 119–130.
- Hoodgarzadeh, Mahzad (2010), Meine Muttersprache(n) bestimme ich selbst! Erste Ergebnisse einer Studie zum Verständnis des Muttersprachenbegriffs aus Sicht von Jugendlichen mit Einwanderungshintergrund. *Deutsch als Zweitsprache*, 3, 37–48.
- Hulstijn, Jan 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.
- Hulstijn, Jan H. (2012), The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition*, 15, 422–433.
- Hulstijn, Jan H. (2015), *Language proficiency in native and non-native speakers. Theory and research*. Amsterdam, Philadelphia: John Benjamins.
- Huneke, Hans-Werner & Steinig, Wolfgang (2000), *Deutsch als Fremdsprache. Eine Einführung*. Berlin: Erich Schmidt Verlag.

- Hyltenstam, Kenneth & Abrahamsson, Niclas (2012), High-Level L2 Acquisition, Learning, and Use. *Studies in Second Language Acquisition*, 34, 177–186.
- Ioup, Georgette L.; Boustagui, Elizabeth; El Tigi, Manal & Moselle, Martha (1994), Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition*, 16, 73–98.
- Johnson, Jacqueline S. (1992), Critical Period Effects in Second Language Acquisition: The Effect of Written versus Auditory Materials on the Assessment of Grammatical Competence. *Language Learning*, 42: 2, 217–248.
- Johnson, Jacqueline S. & Newport, Elissa L. (1989), Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60–99.
- Jun, Heesung Grace & Li, Jinrong (2010), Factors in raters' perceptions of comprehensibility and accentedness. In: John Levis & Kimberly LeVelle (Hrsg.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference*. Ames: Iowa State University, 53–66.
- Kellerman, Eric (1995), Age before beauty: Johnson and Newport revisited. In: Lynn Eubank, Larry Selinker & Michael Sharwood Smith (Hrsg.), *The current state of interlanguage: Studies in Honor of William E. Rutherford*. Amsterdam: Benjamins, 219–231.
- Keulers, Esther H. H.; Stiers, Peter & Jolles, Jelle (2011), Developmental changes between ages 13 and 21 years in the extent and magnitude of the BOLD response during decision making. *NeuroImage*, 54: 2, 1442–1454.
- Kim, Karl H. S.; Relkin, Norman R.; Lee, Kyoung-Min & Hirsch, Joy (1997), Distinct cortical areas associated with native and second languages. *Nature*, 388, 171–174.
- Klein, Wolfgang (1996), Language Acquisition at Different Ages. In: David Magnusson (Hrsg.), *Individual Development over the Lifespan: Biological and Psychosocial Perspectives*. Cambridge: Cambridge University Press, 88–108.
- Klein, Wolfgang (2001), Typen und Konzepte des Spracherwerbs. In: Gerhard Helbig, Lutz Götze, Gert Henrici & Hans-Jürgen Krumm (Hrsg.), *Deutsch als Fremdsprache*, 1. Halbband. Berlin: de Gruyter, 604–616.
- Klein-Braley, Christine (1997), C-Tests in the context of reduced redundancy testing. An appraisal. *Language Testing*, 14: 1, 47–84.

- Klein-Braley, Christine & Raatz, Ulrich (1984), A survey of research of the C-test. Language Testing and evaluation, Band 1, 134–146.
- Kleppin, Karin (1997), Fehler und Fehlerkorrektur. Berlin [u.a.]: Langenscheidt.
- Königs, Frank G. (2010), Zweitsprachenerwerb und Fremdsprachenlernen: Begriffe und Konzepte. In: Hans-Jürgen Krumm, Christian Fandrych, Britta Hufeisen & Claudia Riemer (Hrsg.), Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch (2. Aufl.). Berlin, New York: De Gruyter Mouton, 754–764.
- Lahti, Laura (eingereicht): „Ich fand die Verbkonjugation gar nicht so schlimm!“ Die mündliche Sprachkompetenz finnischer Deutschlernender unter dem Aspekt der grammatischen Korrektheit. Dissertation, Helsinki: Universität Helsinki.
- Larsen-Freeman, Diane & Long, Michael (1991), An Introduction to Second Language Acquisition Research. London: Longman.
- Larson-Hall, Jenifer (2010), A Guide to Doing Statistics in Second Language Research Using SPSS New York, NY [u. a.]: Routledge.
- Larson-Hall, Jenifer (2016), A Guide to Doing Statistics in Second Language Research Using SPSS and R (2 ed.). London, New York: Routledge.
- Lenneberg, Eric H. (1967), Biological Foundations of Language. New York:..
- Lenneberg, Eric H. (1972), Biologische Grundlagen der Sprache. Frankfurt am Main: Suhrkamp.
- Linck, Jared A.; Hughes, Meredith M.; Campbell, Susan G.; Silbert, Noah H.; Tare, Medha; Jackson, Scott R.; Smith, Benjamin K.; Bunting, Michael F.; Doughty, Catherine J. (2013), Hi-LAB: A New Measure of Aptitude for High-Level Language Proficiency. Language Learning, 63: 3, 530–566.
- Long, Michael (1990), Maturation Constraints on Language Development. Studies in Second Language Acquisition, 12: 3, 251–285.
- Long, Michael (2005), Problems with Supposed Counter-Evidence to the Critical Period Hypothesis. International Review of Applied Linguistics, 43, 287–317.
- MacIntyre, Peter D.; Mackinnon, Sean P. & Clément, Richard (2009), Toward the Development of a Scale to Assess Possible Selves as a Source of Language Learning Motivation. In: Zoltán Dörnyei & Ema Ushioda (Hrsg.), Motivation, Language Identity and the L2 Self. Bristol [u. a.]: Multilingual Matters, 193–214.
- MacWhinney, Brian (2006), Emergentism – Use often and with care. Applied Linguistics, 27: 4, 729–740.

- Marinova-Todd, Stefka H.; Marshall, D. Bradford & Snow, Catherine E. (2000), Three Misconceptions About Age and L2 Learning. *TESOL Quarterly*, 1.
- Mayberry, Rachel I. (1993), First-language acquisition after childhood differs from second-language acquisition. The case of American Sign Language. *Journal of Speech, Language, and Hearing Research*, 36: 6, 1258–1270.
- McDonald, Janet L. (2000), Grammaticality judgements in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics*, 21, 395–423.
- Meisel, Jürgen M. (2007), Mehrsprachigkeit in der frühen Kindheit: Zur Rolle des Alters bei Erwerbsbeginn. In: Tanja Anstatt (Hrsg.), *Mehrsprachigkeit bei Kindern und Erwachsenen: Erwerb, Formen, Förderung*. Tübingen: Narr, 93–113.
- Meisel, Jürgen M. (2011), *First and Second Language Acquisition. Parallels and Differences*. Cambridge: Cambridge University Press.
- Meißner, Franz-Joseph (2013), Mehrsprachigkeit in der Perspektive lebenslangen Lernens. In: Annette Berndt (Hrsg.), *Fremdsprachen in der Perspektive lebenslangen Lernens*. Frankfurt a. M. [u. a.]: Peter Lang, 75–88.
- Molnár, Heike (2010), Der Einfluss des Faktors Alter auf die Aussprachekompetenz in der L2. Ergebnisse einer Pilotstudie mit DaZ-Lernern. *Zeitschrift für Interkulturellen Fremdsprachunterricht*, 15: 1, 42–60.
- Moravcsik, Edith A. (2013), *Introducing language typology*. Cambridge: Cambridge University Press.
- Moyer, Alene (2004), *Age, Accent and Experience in Second Language Acquisition. An Integrated Approach to Critical Period Inquiry*. Clevedon: Multilingual Matters.
- Muñoz, Carmen (2006a), The Effects of Age on Foreign Language Learning: The BAF Project. In: Carmen Muñoz (Hrsg.), *Age and the Rate of Foreign Language Learning*. Clevedon: Multilingual Matters, 1–40.
- Muñoz, Carmen (Hrsg.) (2006b), *Age and the Rate of Foreign Language Learning*. Clevedon: Multilingual Matters.
- Muñoz, Carmen (2006c), The BAF Project: Research on the effects of age on foreign language acquisition. In: Cristián Abello-Contesse, Rubén Chacon-Beltrán, M. Dolores López-Jimenez & M. Mar Torreblanca-López (Hrsg.), *Age in L2 Acquisition and Teaching*. Frankfurt a. M. [u. a.]: Peter Lang, 81–92.
- Muñoz, Carmen & Singleton, David (2011), A critical review of age-related research on L2 ultimate attainment. *Language Teaching*, 44: 1, 1–35.

- Munro, Murray J. & Derwing, Tracey M. (1995), Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45: 1, 73–97.
- Munro, Murray J.; Derwing, Tracey M. & Morton, Susan L. (2006), The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, 28: 01, 111–131.
- Nitsch, Cordula (2007), Mehrsprachigkeit: Eine neurowissenschaftliche Perspektive. In: Tanja Anstatt (Hrsg.), *Mehrsprachigkeit bei Kindern und Erwachsenen: Erwerb, Formen, Förderung*. Tübingen: Narr, 47–68.
- Noels, Kimberly A. (2003), Learning Spanish as a second language: Learners' orientations and perceptions of their teachers' communication style. In: Zoltán Dörnyei (Hrsg.), *Attitude, orientations, and motivations in language learning*. Oxford: Blackwell, 97–136.
- Noels, Kimberly A.; Pelletier, Luc G.; Clément, Richard & Vallerand, Robert J. (2000), Why are you learning a second language? Motivational orientation and self-determination theory. *Language Learning*, 50: 1, 57–85.
- Norman, Jerry (1988), *Chinese*. New York [u. a.]: Cambridge University Press.
- Ortega, Francisco & Vidal, Fernando (2007), Mapping the cerebral subject in contemporary culture. *RECIIS. Electronic Journal of Communication, Information & Innovation in Health*, 1: 2, 255–259.
- Ortega, Lourdes (2009), *Understanding Second Language Acquisition*. London: Hodder Education.
- Ortega, Lourdes (2013), SLA for the 21st Century: Disciplinary Progress, Transdisciplinary Relevance, and the Bi/Multilingual Turn. *Language Learning*, 63, 1–24.
- Oyama, Susan (1976), A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, 5: 3, 261–283.
- Pagonis, Giulio (2009a), Der Altersfaktor in Theorie und Praxis. *Zeitschrift für Literaturwissenschaft und Linguistik*, 39: 153, 112–126.
- Pagonis, Giulio (2009b), Kritische Periode oder altersspezifischer Antrieb: Was erklärt den Altersfaktor im Zweitspracherwerb? Eine empirische Fallstudie zum ungesteuerten Zweitspracherwerb des Deutschen durch russische Lerner unterschiedlichen Alters. Frankfurt am Main [u. a.]: Lang.
- Paikeday, Thomas M. (2003), *The native speaker is dead!* Toronto, New York: Lexicography Inc.

- Paradis, Michel (2004), *A Neurolinguistic Theory of Bilingualism*. Amsterdam: Benjamins.
- Paradis, Michel (2009), *Declarative and procedural determinants of second languages*. Amsterdam, Philadelphia: John Benjamins.
- Patkowski, Mark (1980), The sensitive period for the acquisition of syntax in a second language. *Language Learning*, 30: 2, 449–472.
- Patkowski, Mark (2013), The critical period and parameter setting in five cases of delayed L1 acquisition. In: Leah Roberts, Anna Ewert, Miroslaw Pawlak & Magdalena Wrembel (Hrsg.), *EUROSLA Yearbook 2013*. Amsterdam, Philadelphia: John Benjamins, 1–21.
- Patton, Michael Quinn (2002), *Qualitative Evaluation and Research Methods* (3. Aufl.). London: Sage.
- Penfield, Wilder & Roberts, Lamar (1959), *Speech and brain-mechanisms*. Princeton: Princeton University Press.
- Perani, Daniela; Abutalebi, Jubin; Paulesu, Eraldo; Brambati, Simona; Scifo, Paola; Cappa, Stefano; Ferruccio, Fazio (2003), The role of age of acquisition and language usage in early, high-proficient bilinguals: an fMRI study during verbal fluency. *Human Brain Mapping*, 19: 3, 170–182.
- Petersen, Inger (2013), Entwicklung schriftlicher Argumentationskompetenz bei ein- und mehrsprachigen Oberstufenschüler/innen und Studierenden. In: Heike Brandl, Emre Arslan, Elke Langelahn & Claudia Riemer (Hrsg.), *Mehrsprachig in Wissenschaft und Gesellschaft. Mehrsprachigkeit, Bildungsbeteiligung und Potenziale von Studierenden mit Migrationshintergrund*. Bielefeld: BieColl (Bielefelder eCollection), 69–79.
- Pinker, Steven (1994), *The language instinct*. New York: W. Morrow and Co.
- Pinker, Steven (1996), *Der Sprachinstinkt. Wie der Geist die Sprache bildet*. München: Kindler.
- Pinker, Steven (2002), *The blank slate. The modern denial of human nature*. London: Lane.
- Pöppel, Ernst & Wagner, Beatrice (2010), *Je älter desto besser. Überraschende Erkenntnisse aus der Hirnforschung*. München: Gräfe und Unzer.
- Pulvermüller, Friedemann & Schumann, John H. (1994), Review Article: *Neurobiological Mechanisms of Language Acquisition*. *Language Learning*, 44: 4, 681–734.

- Riemer, Claudia (1997), Individuelle Unterschiede im Fremdsprachenerwerb. Eine Longitudinalstudie über die Wechselwirksamkeit ausgewählter Einflußfaktoren. Hohengehren: Schneider.
- Riemer, Claudia (2004), Zur Relevanz qualitativer Daten in der neueren L2-Motivationsforschung. In: Wolfgang Börner & Klaus Vogel (Hrsg.), *Emotion und Kognition im Fremdsprachenunterricht*. Tübingen: Narr, 35–65.
- Riemer, Claudia (2006), DaF-Lernende – alles Exoten? Motivationsforschung und Deutsch als Fremdsprache. In: Hans-Jürgen Krumm & Paul R. Portmann-Tselikas (Hrsg.), *Theorie und Praxis. Österreichische Beiträge zu Deutsch als Fremdsprache*. Innsbruck, Wien, Bozen: StudienVerlag, 43–58.
- Riemer, Claudia (2006), Der Faktor Motivation in der empirischen Fremdsprachenforschung. In: Almut Küppers & Jürgen Quetz (Hrsg.), *Motivation Revisited. Festschrift für Gert Solmecke*. Münster [u. a.]: LIT, 35–48.
- Riemer, Claudia (2010), Nativistische Ansätze. In: Hans-Jürgen Krumm, Christian Fandrych, Britta Hufeisen & Claudia Riemer (Hrsg.), *Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch (2. Aufl.)*. Berlin, New York: De Gruyter Mouton, 799–806.
- Robinson, Peter (2007), Aptitudes, abilities, contexts and practice. In: Robert DeKeyser (Hrsg.), *Practice in a Second Language: Perspectives from Applied linguistics and Cognitive Psychology*. Cambridge: Cambridge University Press, 256–286.
- Robinson, Peter (2013), Aptitude in Second Language Acquisition. In: Carol A. Chapelle (Hrsg.), *The Encyclopedia of Applied Linguistics*. Hoboken: Wiley-Blackwell, 1–5.
- Rost-Roth, Martina (2010), Affektive Variablen/Motivation. In: Hans-Jürgen Krumm, Christian Fandrych, Britta Hufeisen & Claudia Riemer (Hrsg.), *Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch (2. Aufl.)*. Berlin, New York: De Gruyter Mouton, 876–886.
- Ryan, Stephen (2009), Self and Identity in L2 Motivation in Japan: The Ideal L2 Self and Japanese Learners of English. In: Zoltán Dörnyei & Ema Ushioda (Hrsg.), *Motivation, Language Identity and the L2 Self*. Bristol [u. a.]: Multilingual Matters, 120–143.
- Schachter, Jacquelyn (1988), Second language acquisition and its relationship to Universal Grammar. *Applied Linguistics*, 9: 3, 219–235.

- Schlak, Torsten (2003), Gibt es eine kritische Periode des Spracherwerbs? Neue Erkenntnisse in der Erforschung des Faktors Alter beim Spracherwerb. Deutsch als Zweitsprache, 1, 18–23.
- Schlak, Torsten (2006), Der Frühbeginn Englisch aus psycholinguistischer Perspektive – Revisited. Fremdsprachen und Hochschule, 78, 7–25.
- Schlak, Torsten (2008), Fremdsprachenlernreignung: Tabuthema oder Forschungslücke? Zum Zusammenhang von Fremdsprachenlernreignung, Fremdsprachenlernen und Fremdsprachenvermittlung. Zeitschrift für Fremdsprachenforschung, 19, 3–30.
- Schlak, Torsten (2010), Spracherlernreignung. In: Wolfgang Hallet & Frank G. Königs (Hrsg.), Handbuch Fremdsprachendidaktik. Seelze, Stuttgart: Kallmeyer, Klett, 257–261.
- Schlak, Torsten; Zimmermann, Kerstin & Molnár, Heike (2010), Zur Entwicklung eines „C-Tests Wirtschaftsdeutsch“: Hinweise und Erfahrungen aus der Praxis. In: Annette Berndt & Karin Kleppin (Hrsg.), Sprachlehrforschung: Theorie und Empirie. Festschrift für Rüdiger Grotjahn. Frankfurt a. M.: Peter Lang, 13–22.
- Schlosser, Horst Dieter (2002), dtv-Atlas Deutsche Literatur (9. Aufl.). München: Deutscher Taschenbuch Verlag.
- Schön, Almut; Zimmermann, Kerstin & Johnson, Natalia (2012), Intrauniversitäre Kooperation – zur gemeinsamen Entwicklung eines C-Tests durch Sprachenzentrum und Sprachlehrforschung. Fremdsprachen und Hochschule, 86, 61–79.
- Schöningh, Ingo (2013), Was ist Mutter(s)Sprache? Zum Erwerb des Vietnamesischen von vietnamesisch-stämmigen Kindern in Deutschland. In: Yüksel Ekinci, Ludger Hoffmann, Kerstin Leimbrink & Lirim Selmani (Hrsg.), Migration, Mehrsprachigkeit, Bildung. Tübingen: Stauffenburg, 139–149.
- Segalowitz, Norman (2003), Automaticity and second languages. In: Catherine J. Doughty & Michael H. Long (Hrsg.), The handbook of second language acquisition. Oxford: Blackwell, 382–408.
- Segalowitz, Norman & Hulstijn, Jan H. (2005), Automaticity in bilingualism and second language learning. In: Judith F. Kroll & Annette M. B. DeGroot (Hrsg.), Handbook of bilingualism: Psycholinguistic approaches. Oxford: Oxford University Press, 371–388.
- Selinker, Larry (1972), Interlanguage. International Review of Applied Linguistics, 10, 209–231.

- Siebert-Ott, Gesa (2003), Muttersprachendidaktik – Zweitsprachendidaktik – Fremdsprachendidaktik. In: Ursula Bredel, Hartmut Günther, Jakob Ossner, Peter Klotz, Gesa Siebert-Ott (Hrsg.), *Didaktik der deutschen Sprache. Ein Handbuch. Band 1.* Paderborn, München [u. a.]: Ferdinand Schöningh: 30–41.
- Singleton, David (2005), The Critical Period Hypothesis: A Coat of Many Colours. *International Review of Applied Linguistics*, 43, 269–285.
- Singleton, David & Ryan, Lisa (2004), *Language Acquisition: The Age Factor* (2. Aufl.). Clevedon [u. a.]: Multilingual Matters.
- Skehan, Peter (1998), *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Slabakova, Roumyana (2013), Adult second language acquisition. A selective overview with a focus on the learner linguistic system. *Linguistic Approaches to Bilingualism*, 3(1), 48–72.
- Sobotta, Robert (2013), Deutschlernen mit älteren Menschen am Beispiel des Goethe-Instituts Dresden: „Deutsch 50 Plus“. In: Annette Berndt (Hrsg.), *Fremdsprachen in der Perspektive lebenslangen Lernens*. Frankfurt am Main [u. a.]: Peter Lang, 209–213.
- Sopata, Aldona (2004), *Universalgrammatik und Fremdsprachendidaktik*. Frankfurt a. M. [u. a.]: Peter Lang.
- Sow, Noah (2009), *Deutschland Schwarz Weiß. Der alltägliche Rassismus* (2. Aufl.). Goldmann.
- Steinmüller, Ulrich (2007), Teaching Foreign Languages in the 21st Century – Perspectives and Challenges. In: Constanze Göbel, Nikolaus Katsaounis, Chris Merkelbach, Judith Theuerkauf & Jianpei Yang (Hrsg.), *DaF-Didaktik aus internationaler Perspektive*. Frankfurt a. M.: IKO-Verlag, 15–27.
- Steinmüller, Ulrich (2010), Vielschichtig und leistungsfähig. Universitäre Bildung in der Volksrepublik China. *Forschung und Lehre*, 12, 890–892.
- Steinmüller, Ulrich (2013), Lernziele für Deutsch als Fremdsprache in China: Kulturkontrast oder interkulturelle Kompetenz? In: Ulrich Steinmüller & Fu Su (Hrsg.), *Chinesisch-deutsche Kulturbeziehungen*. Frankfurt a. M. [u. a.]: Peter Lang, 23–36.
- Stemmer, Brigitte (2010), A cognitive neuroscience perspective on learning and memory in aging. *Zeitschrift für Interkulturellen Fremdsprachunterricht*, 15: 1, 7–25.

- Stölting, Wilfried (1980), Die Zweisprachigkeit jugoslawischer Schüler in der Bundesrepublik Deutschland. Wiesbaden: Harrassowitz.
- Taguchi, Tatsuya; Magid, Michael & Papi, Mostafa (2009), The L2 Motivational Self System among Japanese, Chinese and Iranian Learners of English: A Comparative Study. In: Zoltán Dörnyei & Ema Ushioda (Hrsg.), *Motivation, Language Identity and the L2 Self*. Bristol [u. a.]: *Multilingual Matters*, 66–97.
- Tönjes, Andrea (unveröffentlichte Hausarbeit), Muttersprache als soziale Konstruktion. Historische Genese eines Begriffes und seine Implikationen in Forschung und Bildung. TU Berlin, Sommersemester 2012.
- Tomasello, Michael. (2008), *Origins of human communication*. Cambridge: MIT Press.
- Tomasello, Michael (2009), *Die Ursprünge der menschlichen Kommunikation*. Frankfurt am Main: Suhrkamp.
- Tran, Thuc-Lan (2009), *Der kommunikative Fremdsprachenunterricht und Lehr- und Lerntraditionen am Beispiel des Deutsch-als-Fremdsprach-Unterrichts in Vietnam*. Institut für Auslandsgermanistik / Deutsch als Fremd- und Zweitsprache, Friedrich-Schiller-Universität Jena, Dissertation.
- Ushioda, Ema (2001), *Language learning at university: Exploring the role of motivational thinking*. In: Zoltán Dörnyei & Richard W. Schmidt (Hrsg.), *Motivation and second language acquisition*. Honolulu: University of Hawaii Press, 99–124.
- Ushioda, Ema (2009), *A Person-in-Context Relational View of Emergent Motivation, Self and Identity*. In: Zoltán Dörnyei & Ema Ushioda (Hrsg.), *Motivation, Language Identity and the L2 Self*. Bristol [u. a.]: *Multilingual Matters*, 215–228.
- Vanhove, Jan (2013), *The Critical Period Hypothesis in Second Language Acquisition: A Statistical Critique and a Reanalysis*. *PLoS ONE*, 8(7), 1–15.
- Vermes, Geneviève (1998), *Schrifterwerb und Minorisierung als psychologisches Problem. Über Schwierigkeiten des Eintritts in die französische Schriftkultur für ethno- und soziolinguistische Minderheiten*. In: Ingrid Gogolin, Sabine Graap & Günther List (Hrsg.), *Über Mehrsprachigkeit*. Tübingen: Stauffenburg, 3–19.
- Vul, Edward; Harris, Christine; Winkielman, Piotr & Pashler, Harold (2009), *Puzzlingly High Correlations in fMRI Studies of Emotion, Personality and Social Cognition*. *Perspectives on Psychological Science*, 4: 3, 274–290.
- Vul, Edward & Pashler, Hal (2012), *Voodoo and circularity errors*. *NeuroImage*, 62, 945–948.

- Wartenburger, Isabella (2012), Mehrsprachigkeit. In: Barbara Höhle (Hrsg.), Psycholinguistik (2. Aufl.). Berlin: Akademie Verlag: 173–188.
- Wegmann, Johanna & Pomino, Jenny (2010), 60plus goes Web 2.0. Fremdsprachenlernen in einer medialgestützten Lernumgebung für Lerner im höheren Alter. Zeitschrift für Interkulturellen Fremdsprachenunterricht, 15: 1, 61–69.
- Weiner, Bernard (1992), Human motivation: metaphors, theories, and research. Newbury Park, Calif.: Sage.
- White, Lydia (1986), Implications of parametric variation for adult second language acquisition: an investigation of the pro-drop parameter. In: Vivian Cook (Hrsg.), Experimental Approaches to Second Language Acquisition. Oxford: Pergamon, 55–72.
- White, Lydia (2015), Linguistic Theory, Universal Grammar, and Second Language Acquisition. In: Bill VanPatten & Jessica Williams (Hrsg.), Theories in Second Language Acquisition. An Introduction (2. Aufl.). New York; London: Routledge, 34–53.
- Zappatore, Daniela (2006), Mehrere Sprachen – ein Gehirn. Einflussvariablen und Schlussfolgerungen für eine Didaktik der Mehrsprachigkeit. In: Werner Wiater (Hrsg.), Didaktik der Mehrsprachigkeit. Theoriegrundlagen und Praxismodelle. München: Ernst Vögel, 73–91.
- Zhao, Jin (2010), Kontrastive Analyse Mandarin – Deutsch. In: Hans-Jürgen Krumm, Christian Fandrych, Britta Hufeisen & Claudia Riemer (Hrsg.), Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch (2. Aufl.). Berlin, New York: De Gruyter Mouton: 627–634.
- Zippel, Wolfgang (2009), Semantik und Grammatik im Kopf: ein Forschungsüberblick. Unterschiede zwischen erwachsenen L1- und L2-Sprechern in der kortikalen Repräsentation semantischer und grammatikalischer Verarbeitungsprozesse und Konsequenzen für die Fremdsprachenvermittlung. Tübingen: Gunter Narr.

Verzeichnis der Internetquellen:

- Hulstijn, Jan (2014), BLC-HLC theory in a nutshell.
[\[http://www.uva.nl/binaries/content/documents/personalpages/h/u/j.h.hulstijn/en/downloads/downloads/assets%5B2%5D/asset?1395139039472\]](http://www.uva.nl/binaries/content/documents/personalpages/h/u/j.h.hulstijn/en/downloads/downloads/assets%5B2%5D/asset?1395139039472); letzter Zugriff am 05.09.2014]

- Krause, Ulrike-Marie & Stark, Robin (2010), Motivation. In: Rolf Arnold, Sigrid Nolda & Ekkehard Nüssli (Hrsg.), Online-Wörterbuch der Erwachsenenbildung (2. Aufl.). Stuttgart: UTB. [http://www.wb-erwachsenenbildung.de/online-woerterbuch/?title=Motivation&tx_buhutbedulexicon_main%5Bentry%5D=160&tx_buhutbedulexicon_main%5Baction%5D=show&tx_buhutbedulexicon_main%5Bcontroller%5D=Lexicon&cHash=7539863b6602c12a83a4c090d331320b; letzter Zugriff am 15.11.2014]
- Mai, Marina (2011), Mustermigranten werden weniger. Vietnamesen an Gymnasien. die tageszeitung, 02.11.2011. [<http://www.taz.de/!81094/>; letzter Zugriff am 14.02.2014]
- Meara, Paul M. (2005), Llama Language Aptitude Tests. Swansea: Lognostics. [http://www.lognostics.co.uk/tools/llama/llama_manual.pdf; letzter Zugriff am 14.02.2014]
- Rockmann, Ulrike, Rehkämper, Klaus & Leerhoff, Holger (2013), Bildung in Berlin und Brandenburg 2012. Ein indikatorengestützter Bericht zur Bildung im Lebenslauf. Amt für Statistik in Berlin-Brandenburg (AfS). [<https://www.statistik-berlin-brandenburg.de/home/pdf/RBB2012-VSB2012.pdf>; letzter Zugriff am 02.09.2014]
- Storz, Henning & Wilmes, Bernhard (2007), Grundlagendossier Migration. Die Bundeszentrale für politische Bildung. [<http://www.bpb.de/gesellschaft/migration/dossier-migration/56483/einbuengerung?p=all>; letzter Zugriff am 02.09.2014]
- URL 1: Wikipedia (Die freie Enzyklopädie) (2014), Gender-Gap (Linguistik). [[http://de.wikipedia.org/wiki/Gender_Gap_\(Linguistik\)](http://de.wikipedia.org/wiki/Gender_Gap_(Linguistik)); letzter Zugriff am 01.09.2014]
- URL 2: Science Magazine (2005), So much more to know... [<http://www.sciencemag.org/content/309/5731/78.2.full>; letzter Zugriff am 01.09.2014]
- URL 3: Council of Europe (2000), Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). [http://www.coe.int/t/dg4/linguistic/Cadre1_en.asp; letzter Zugriff am 01.09.2014]
- URL 4: Wikipedia (Die freie Enzyklopädie) (2014), Othering. [<http://de.wikipedia.org/wiki/Othering>; letzter Zugriff am 02.09.2014]

- URL 5: Statistisches Bundesamt – DESTATIS (2012), Bevölkerung und Erwerbstätigkeit. Bevölkerung mit Migrationshintergrund. Ergebnisse des Mikrozensus 2012.
[https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/MigrationIntegrations/Migrationshintergrund2010220127004.pdf;jsessionid=D33E0DE607B43B52831A7DD30E3D8488.cae3?__blob=publicationFile; letzter Zugriff am 02.09.2014]
- URL 6: Kommission der europäischen Gemeinschaften, Eine neue Rahmenstrategie für Mehrsprachigkeit. [<http://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52005DC0596&from=DE>; letzter Zugriff am 02.09.2014]
- URL 7: Netzwerk Deutsch (2010), Statistische Erhebungen 2010. Die deutsche Sprache in der Welt.
[https://www.goethe.de/resources/files/pdf37/Bro_Deutschlernerhebung_final2.pdf; letzter Zugriff am 25.06.2015]
- URL 8: Netzwerk Deutsch (2010), Statistische Erhebungen 2010. Die deutsche Sprache in der Welt. [<http://www.goethe.de/mmo/priv/5759818-STANDARD.pdf>; letzter Zugriff am 25.06.2015]
- URL 9: Deutscher Akademischer Austauschdienst (DAAD) und Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW) (2014), Wissenschaft weltweit: Herkunft ausländischer, Bildungsausländer-, Bildungsinländer-Studierender 2013 nach Erdteilen, Regionen, Herkunftsstaaten.
[<http://www.wissenschaftweltoffen.de/daten/1/2/1>; letzter Zugriff am 04.09.2014]
- URL 10: Deutscher Akademischer Austauschdienst (DAAD) und Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW) (2014), Wissenschaft weltweit: Mittelfristige Entwicklung – 1999 bis 2013 – der Zahl von Bildungsausländer-Studierenden der 2013 20 wichtigsten Herkunftsstaaten von Bildungsausländer-Studierenden.
[<http://www.wissenschaftweltoffen.de/daten/1/2/3>; letzter Zugriff am 04.09.2014]
- URL 11: TestDaF-Institut, onSET Deutsch (Online-Einstufungstest Deutsch als Fremdsprache). [<https://www.ondaf.de/ondaf/info/qualitaet.jsp>; letzter Zugriff am 21.09.2016]
- URL 12: ALTE (2005), Materials for the guidance of test item writers.
[http://www.alte.org/attachments/files/item_writer_guidelines.pdf; letzter Zugriff am 07.09.2014]
- URL 13: Wikipedia (Die freie Enzyklopädie) (2013), Akzent (Aussprache).

[[http://de.wikipedia.org/wiki/Akzent_\(Aussprache\)](http://de.wikipedia.org/wiki/Akzent_(Aussprache))]; letzter Zugriff am 04.09.2014]

URL 14: _lognostics, The LLAMA Language Aptitude Tests.

[<http://www.lognostics.co.uk/tools/llama/>]; letzter Zugriff am 13.02.2015]

URL 15: sourceforge.net (2014), Audacity. [<http://audacity.sourceforge.net>]; letzter Zugriff am 13.02.2015]

8. Anhang

Ausschreibungen Proband_innen-Suche	(Anhang S. 201-202)
Infoblatt für die Teilnehmer_innen	(Anhang S. 203)
Grammatiktest – Hinweise für die Teilnehmer_innen	(Anhang S. 204)
Grammatiktest mit Lösungsschablone	(Anhang S. 205)
S-C-Tests – Hinweise für die Testleiter_innen	(Anhang S. 206-207)
C-Tests – Hinweise für die Testleiter_innen	(Anhang S. 208-209)
S-C-Tests – Hinweise für die Teilnehmer_innen	(Anhang S. 210)
S-C-Tests	(Anhang S. 211-217)
C-Tests – Hinweise für die Teilnehmer_innen	(Anhang S. 218)
C-Tests	(Anhang S. 219-224)
Fragebogen zur Sprachlernbiografie (L2-Lerner_innen)	(Anhang S. 225-228)
Fragebogen zur Sprachlernbiografie (2L1)	(Anhang S. 229-232)
TESTATT-Skalen (Eigenschaftsprofil)	(Anhang S. 233-235)
Phonetik-Rating – Hinweise für die Bewerter_innen	(Anhang S. 236)
Fragebogen zu weiteren biografischen Hintergrundinformationen und zur Motivation	(Anhang S. 237-242)



Proband_innen gesucht!

Sprechen Sie **Mandarin** oder **Kantonesisch als Erstsprache** und haben Sie in erwachsenem Alter **Deutsch** gelernt?

Dann suchen wir Sie!

Im Rahmen eines Forschungsprojekts an der TU Berlin wird der Einfluss des Faktors Alter und der Ausgangssprache auf den Sprachlernerfolg untersucht.

Auf Wunsch informieren wir Sie nach der Auswertung des Tests darüber, wie Sie im Vergleich zu den anderen Proband_innen abgeschnitten haben.

Um an der Studie teilzunehmen, sollten Sie:

- Mandarin bzw. Kantonesisch als Erstsprache sprechen
- mit 16 Jahren oder älter angefangen haben, Deutsch zu lernen
- seit mindestens fünf Jahren in Deutschland oder einem deutschsprachigen Land leben
- über akademische Bildung verfügen bzw. studieren

Der Zeitaufwand beträgt insgesamt maximal 90 Minuten, für die Sie 10,- € als Aufwandsentschädigung erhalten. In dieser Zeit werden Sie einen Lückentext bearbeiten, einen Fragebogen beantworten, einen kurzen Text vorlesen und frei sprechen. Mit einzelnen Teilnehmer_innen wollen wir nach Auswertung der Sprachtests Interviews führen.

Interessiert?

Dann freuen wir uns auf eine E-Mail oder einen Anruf:

Simone Heine

simone.heine@campus.tu-berlin.de

Tel.: 030 / 314-73236

Mobil: 01573 / 8954740



Proband_innen gesucht!

Sind Sie zweisprachig aufgewachsen und haben schon als Kleinkind **Mandarin** oder **Kantonesisch** und **Deutsch** erworben?

Dann suchen wir Sie!

Im Rahmen eines Forschungsprojekts an der TU Berlin wird der Einfluss des Faktors Alter und der Ausgangssprache auf den Sprachlernerfolg untersucht.

Auf Wunsch informieren wir Sie nach der Auswertung des Tests darüber, wie Sie im Vergleich zu einsprachig aufgewachsenen Personen abgeschnitten haben.

Um an der Studie teilzunehmen, sollten Sie:

- mindestens 16 Jahre alt sein
- mit 0 bis 3 Jahren begonnen haben, Mandarin bzw. Kantonesisch zu lernen
- mit 0 bis 3 Jahren begonnen haben, Deutsch zu lernen
- über akademische Bildung verfügen, studieren oder ein Gymnasium besuchen

Der Zeitaufwand beträgt insgesamt maximal 90 Minuten, für die Sie 10,-€ als Aufwandsentschädigung erhalten. In dieser Zeit werden Sie einen Lückentext bearbeiten und in einem kurzen Interview Fragen zu Ihrer Sprachlernbiografie beantworten

Interessiert?

Dann freuen wir uns auf eine E-Mail oder einen Anruf:

Simone Heine

simone.heine@campus.tu-berlin.de

Tel.: 030 / 314-73236

Mobil: 01573 / 8954740

Liebe Teilnehmerinnen, liebe Teilnehmer!

Mit unserer Studie wollen wir dokumentieren, wie erfolgreich auch Erwachsene eine Fremdsprache erlernen können. Gleichzeitig interessieren wir uns dafür, welche Eigenschaften erfolgreiche Lerner aufweisen.

Wir möchten Sie bitten, sich bei allen Aufgaben bestmöglich anzustrengen.

Die Studie dient rein wissenschaftlichen Zwecken. Ihre Teilnahme ist selbstverständlich freiwillig, und Sie können die Mitarbeit jederzeit beenden. Um möglichst aussagekräftige Ergebnisse zu bekommen, würden wir uns aber freuen, wenn Sie möglichst lange an der Studie teilnehmen könnten.

Alle erhobenen Informationen bleiben selbstverständlich anonym! Anstelle Ihres Namens werden wir einen Teilnehmer-Code verwenden, den Sie selbst erstellen und der für die gesamte Studie gilt.

Der Teilnehmer-Code besteht aus:

- den ersten zwei Buchstaben des Vornamens Ihrer Mutter
- den ersten zwei Buchstaben des Vornamens Ihres Vaters
- Ihrem Geburtstag
- Ihrem Geburtsmonat

Hier sehen Sie ein Beispiel:

Vorname der Mutter	Vorname des Vaters	Geburtstag	Geburtsmonat
MARIA	PAUL	12.04.1970	12.04.1970
<input type="text" value="MA"/>	<input type="text" value="PA"/>	<input type="text" value="12"/>	<input type="text" value="04"/>

Bitte tragen Sie nun Ihren Teilnehmer-Code ein:

Vorname der Mutter	Vorname des Vaters	Ihr Geburtstag	Ihr Geburtsmonat
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

Wenn Sie Interesse daran haben, auch künftig an unserer Studie teilzunehmen, möchten wir Sie gegen Ende der heutigen Befragung um Ihre Email-Adresse und/oder Telefonnummer bitten, damit wir Sie gegebenenfalls in einer späteren Forschungsphase für weitere Fragen oder Untersuchungen kontaktieren können.

Wir danken Ihnen herzlich für Ihre Mitarbeit und Unterstützung!

Anhang

Grammatikaufgaben

Teilnehmer-Code:

Hinweise für die Teilnehmer

Tragen Sie bitte zuerst Ihren Teilnehmer-Code oben rechts ein! Dieser Code gilt auch für alle weiteren Aufgaben.

Bei den folgenden Aufgaben sollen Sie intuitiv mit Hilfe Ihres Sprachgefühls die Korrektheit grammatischer Strukturen beurteilen. Bitte beurteilen Sie daher jeden Satz möglichst schnell, sobald Sie ihn gelesen haben.

Bitte lesen Sie jeden Satz nur ein einziges Mal. Sie sollen Ihre Urteile nicht mehr überprüfen oder ändern. Sobald Sie mit der Bearbeitung aller Aufgaben fertig sind, drehen Sie Ihr Aufgabenblatt um.

Bitte kreuzen Sie an, ob Sie den jeweiligen Satz für „korrekt“ oder „nicht korrekt“ halten und ob Sie sich Ihrer Entscheidung „sicher“ oder „unsicher“ sind. Machen Sie pro Satz nur ein einziges Kreuz.

Beispiel:

	korrekt		nicht korrekt	
	sicher	unsicher	sicher	unsicher
Beispiel I: Ich verstehe nicht.	x			
Beispiel II: Ich verstehen nicht.			x	

Vielen Dank für Ihre Mitarbeit!

Grammatikaufgaben

	korrekt		nicht korrekt	
	sicher	unsicher	sicher	unsicher
Bsp. I: Ich verstehe nicht.	<input type="checkbox"/>			
Bsp. II: Ich verstehe nicht.	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	

1. Mein Freund ist ein starke Raucher.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
2. Nach der Schule komme nach Hause.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
3. Katja glaubt, dass sie klug ist.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Morgen wird es erklärt, dass er kommt.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
5. Alle meine Entchen sind hier.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Verstehen Deutsch gut.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
7. Auf der Autobahn wird es sehr schnell gefahren.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
8. Manfred fragt, ob er die Prüfung bestanden hat.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Katja glaubt, dass klug ist.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
10. Morgen wird regnen.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
11. Dass der Sommer angefangen hat, ist deutlich geworden.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. Wenn der Weg unbekannt ist, wird es gefragt.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
13. Er setzte sich zwischen seinen Schülern.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
14. Während des Krieges wird gelitten.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. Dass der Sommer angefangen hat, ist es deutlich geworden.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
16. Morgen wird erklärt, dass er kommt.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. Morgen wird es regnen.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. Wenn der Weg unbekannt ist, wird gefragt.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. Nach der Schule kommst du nach Hause.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. Er denkt immer an mich.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21. Während des Krieges wird es gelitten.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
22. Manfred fragt, ob die Prüfung bestanden hat.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
23. Morgen wird er die Arbeit beendet haben.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24. Um den Nachbarn nicht aufzuwecken, wird es geschlichen.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
25. Wir verstehen Deutsch gut.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26. Der Alkohol begann, seine Gesundheit zu untergraben.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27. Auf der Autobahn wird sehr schnell gefahren.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28. Dass Marie gewonnen hat, freut mich sehr.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29. Um den Nachbarn nicht aufzuwecken, wird geschlichen.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30. Der Student hat die Arbeit geschafft.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
31. Wir beauftragen ihn, die Post abgeholt zu haben.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
32. Dass Marie gewonnen hat, freut es mich sehr.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

C-Test

(Version 24.06.06)

Teil 1:

C-Test-Version mit Zeitlimitierung

Hinweise für die Testleiter

Es ist äußerst wichtig, dass sich die Testleiter genau an die Instruktionen und die angegebenen Zeiten halten!!!!!!!!!!!!!!!!!!!!

- Teilen Sie die Testbögen aus.
- Die TN werden aufgefordert, ihren Teilnehmer-Code auf die erste Seite des C-Test-Bogens zu schreiben. Kontrollieren durch Nachfragen!

Anweisung:

- **Bitte übertragen Sie Ihren Teilnehmer-Code in das Kästchen auf der ersten Seite des Tests.**
- **Haben Sie alle Ihren Teilnehmer-Code übertragen?**
- Weisen Sie die TN darauf hin, erst nach Aufforderung mit der Bearbeitung zu beginnen und erst nach weiterer Aufforderung umzublättern.

Anweisung:

- **Bitte beginnen Sie mit der Bearbeitung erst, wenn wir Sie dazu auffordern. Wir teilen Ihnen mit, wann Sie jeweils mit der Bearbeitung des nächsten Textes beginnen sollen.**
- Fordern Sie dann die TN auf, die Anweisungen zur Bearbeitung des C-Tests zu lesen. Die TN haben hierfür 1 1/2 Minuten Zeit.

Anweisung:

- **Bitte lesen Sie jetzt die Anweisungen zur Bearbeitung des C-Tests auf der ersten Seite des Tests. Sie haben dafür 1 1/2 Minuten Zeit.**

- Für jeden Text wird jeweils gemäß den unten aufgeführten Zeitvorgaben gesagt, wann die TN aufhören und mit der Bearbeitung des nächsten Textes beginnen sollen.

Anweisung:

- o **Bitte beginnen Sie jetzt mit dem ersten/nächsten Text.**

Zeitvorgaben für C-Test Version 30.4.06:

Text 1 "Krieg":	1:40 min
Text 2 "Linkshänder":	1:25 min
Text 3 "Astrologie":	1:20 min
Text 4 "Lenin-Denkmal":	1:55 min
Text 5 "Kunstleben":	1:05 min
Text 6 "Risiken der Technik":	1:25 min

- Nach der Bearbeitung des letzten Textes fordern Sie die TN auf, die beiden Fragen auf Seite 8 zu beantworten.

Anweisung:

- o **Bitte blättern Sie um und beantworten Sie die beiden Fragen auf der nächsten Seite!**

Nach Bearbeitung des zeitlimitierten C-Tests sammeln Sie die Testbögen ein.

Teil 2:
C-Test-Version ohne Zeitlimitierung

Hinweise für die Testleiter

Es ist äußerst wichtig, dass sich die Testleiter genau an die Instruktionen und die angegebenen Zeiten halten!!!!!!!!!!!!!!!!!!!!

- Teilen Sie die Testbögen für den nicht-zeitlimitierten C-Test aus.

Anweisung:

- **Wir teilen nun die Testbögen für den nicht-zeitlimitierten C-Test aus.**
- Fordern Sie die TN auf, den Teilnehmer-Code auf die 1. Seite des neu ausgegebenen Tests zu schreiben. Kontrollieren durch Nachfrage!

Anweisung:

- **Bitte übertragen Sie Ihren Teilnehmer-Code in das Kästchen auf der ersten Seite des neuen Tests.**
- **Haben Sie alle Ihren Teilnehmer-Code übertragen?**
- Fordern Sie dann die TN auf, die Anweisungen zur Bearbeitung des Tests zu lesen und mit der Bearbeitung der Texte zu beginnen.

Anweisung:

- **Bitte lesen Sie jetzt die Anweisungen zur Bearbeitung des C-Tests auf der ersten Seite des Tests und beginnen Sie mit der Bearbeitung der Texte.**

- Die Probanden haben pro Text 5 Minuten Zeit. Weisen Sie die TN nach dieser Zeit darauf hin, dass sie möglichst mit dem folgenden Text beginnen sollten.

Anweisung:

- **Sie sollten jetzt mit dem nächsten Text beginnen.**
- Am Ende sammeln Sie die Testbögen ein.

Anhang

C-Test

(Version 30.4.06 - zeitlimitiert)

Teilnehmer-Code:

Version mit Zeitlimitierung

Hinweise für die Teilnehmer

In den folgenden Texten fehlen bei einer Reihe von Wörtern am Wortende jeweils einige Buchstaben. Bitte ergänzen Sie möglichst sämtliche Lücken in sinnvoller Weise. Da Sie die Texte unter erheblichem Zeitdruck bearbeiten werden, werden viele von Ihnen nicht alle Lücken korrekt ergänzen können. Geben Sie sich trotzdem größte Mühe.

Arbeiten Sie bitte völlig unabhängig von den anderen Teilnehmern.

Es handelt sich um eine Untersuchung mit einer ausschließlich wissenschaftlichen Zielsetzung. Ihre Leistung wird nicht benotet!

Schreiben Sie bitte leserlich!

Sie haben pro Text zwischen 1 und 2 Minuten Zeit.

Die Testleiter teilen Ihnen mit, wann Sie jeweils mit der Bearbeitung des nächsten Textes beginnen sollen.

Bitte konzentrieren Sie sich immer nur auf die Bearbeitung eines einzigen Textes und blättern Sie nicht vor oder zurück.

Beispiel:

England ist eine alte Seefahrernation – kein Ort in Großbritannien ist weiter als 120 Kilometer vom Meer entfernt. Doch **die** _____ britischen Küstengewässer sind **tückisch** _____, selbst **für** _____ erfahrene Seeleute _____.

Vielen Dank für Ihre Mitarbeit!

1. Krieg

Dass der Krieg der Vater aller Dinge sei, sagt man heute nicht mehr. Wir erfa_____ jedoch täg_____, wie d_____ historischen Katast_____ als Kataly_____ ungeahnter u_____ ungewollter Veränd_____ im Gro_____ wie i_____ Kleinen wir_____. Diesen Wec_____ muss m_____ nicht n_____ als Ung_____, sondern ka_____ ihn au_____ als Cha_____, schließlich so_____ als Gew_____ sehen. I_____ diesem Si_____ ist sic_____ der Ti_____ der Lebenseri_____ Klaus von Bismarcks z_____ verstehen: "Aufbruch aus Pommern".

Bitte erst nach Aufforderung umblättern!!

2. Linkshänder

Das Schreibenlernen ist für viele Linkshänder eine besonders große Hürde. Unsere Sch_____ verläuft v_____ links na_____ rechts. D_____ bedeutet f_____ den Rechts_____ eine beq_____, ziehende Bewe_____. Für d_____ Linkshänder ab_____ wird e_____ eine schieb_____, die weit_____ schwieriger i_____. Diese Tec_____ muss sc_____ im Vorsch_____, zum Beis_____ beim Ma_____, geübt wer_____. Lob u_____ Geduld d_____ Eltern erlei_____ dem Links_____ das Schreib_____, außerdem gibt es spezielle Schreib- und Zeichengeräte sowie Scheren für Linkshänder, die helfen, die Anforderungen der Schule besser zu erfüllen.

Bitte erst nach Aufforderung umblättern!!

3. Astrologie

Die Astrologie begann ganz einfach – nämlich mit der Beobachtung des Himmels. Nachdem d_____ Mensch entd_____ hatte, d_____ aller Se_____ wie al_____ Unheil v_____ Himmel k_____, wandte e_____ seine Beobacht_____ intensiv d_____ himmlischen Gesc_____ zu u_____ begann, d_____ Zeichen d_____ Himmels z_____ erforschen. V_____ Himmel k_____ lebensspendende Wä_____, aber au_____ tödliche Hi_____. Der Men_____ vermutete hin_____ diesen Naturvo_____ himmlische, gött_____ Kräfte u_____ versuchte, beide zu erforschen, weshalb Astronomie und Astrologie also in den Anfängen gleich waren.

Bitte erst nach Aufforderung umblättern!!

4. Lenin-Denkmal

Das Lenin-Denkmal am Wiener Platz in Dresden, dem der Innenstadt zugekehrten Vorplatz des Hauptbahnhofs, wird jetzt abgebaut. Die Stadtver_____ haben ein_____ Mühen hin_____ sich, z_____ einem Ents_____ darüber z_____ kommen, w_____ aus d_____ Denkmal d_____ Gründers d_____ Sowjetunion wer_____ solle. Zunä_____ war e_____ einfacher Ab_____ geplant, d_____ den nachrevol_____ Stimmungen i_____ der Bevöl_____ entsprochen hä_____, zumal d_____ zu ei_____ umgestalteten Fußgän_____ erweiterte eins_____ Prachtstraße i_____ Oktober 1989 e_____ Brennpunkt der Revolution in Dresden war.

Bitte erst nach Aufforderung umblättern!!

5. Kunstleben

Jahrzehntlang war das Kunstleben in Russland streng organisiert und strukturiert: Alle wussten genau, welche Kunst erlaubt und welche verboten war. Jetzt sind die Kulturins zusammengebrochen und viele Künstler in den Westen gegangen. Daraus können man schließen, dass in diesem Bereich der Kunst ein Vakuum entsteht sei. Doch das Gegenteil ist der Fall. Bei einem kurzen Besuch in Moskau über von der unverminderten Vitalität der dortigen Kunstszene. Die sich ihrer Struktur in der letzten Zeit stark verändert hat, steht auf einem anderen Blatt.

Bitte erst nach Aufforderung umblättern!!

6. Risiken der Technik

Atomkraftwerke, Chemiefabriken oder die Gentechnik entziehen sich nicht nur der Kontrolle des Einzelnen; laufen sol_____ Anlagen Am_____, können d_____ Schäden f_____ die Allgem_____ tatsächlich s_____ unermesslich se_____, dass sel_____ Versicherungsunternehmen kapitu_____. Dieses Gefahrenp_____ verursacht Äng_____, die si_____ naturgemäß b_____ Störfällen verst_____, die i_____ Drehbuch d_____ Ingenieure u_____ Planer ni_____ vorgesehen wa_____ – zumal f_____ die Nai_____ unter d_____ Verbrauchern d_____ Strom a_____ der Stec_____ kommt und der Klebstoff aus der Tube.

Bitte erst nach Aufforderung umblättern!!

Beantworten Sie bitte noch folgende Fragen:

Haben Sie den Arbeitsanweisungen folgen können?

ja () nein ()

War es Ihnen möglich, Ihre Fähigkeiten im Test angemessen einzusetzen?

völlig () weitgehend () nur teilweise () überhaupt nicht ()

Vielen Dank für Ihre Mitarbeit!

C-Test

(Version 30.4.06- unlimitiert)

Teilnehmer-Code:

Version ohne Zeitlimitierung

Hinweise für die Teilnehmer

Sie bearbeiten jetzt dieselben Texte ein zweites Mal.

Diesmal sind für die Bearbeitung eines Textes mehrere Minuten vorgesehen. Dabei dürfen Sie jeden einzelnen Text mehrfach durchgehen. Damit Sie **alle** Texte bearbeiten können, geben wir Ihnen einen Hinweis, wann Sie jeweils mit dem nächsten Text beginnen sollten.

Geben Sie sich bitte wiederum größte Mühe und arbeiten Sie völlig unabhängig von den anderen Teilnehmern.

Schreiben Sie bitte leserlich!

Beispiel:

England ist eine alte Seefahrernation – kein Ort in Großbritannien ist weiter als 120 Kilometer vom Meer entfernt. Doch **die** _____ britischen Küstengewässer sind **tückisch** _____, selbst **für** _____ erfahrene Seeleute _____.

Vielen Dank für Ihre Mitarbeit!

1. Krieg

Dass der Krieg der Vater aller Dinge sei, sagt man heute nicht mehr. Wir erfa_____ jedoch täg_____, wie d_____ historischen Katast_____ als Kataly_____ ungeahnter u_____ ungewollter Veränd_____ im Gro_____ wie i_____ Kleinen wir_____. Diesen Wec_____ muss m_____ nicht n_____ als Ung_____, sondern ka_____ ihn au_____ als Cha_____, schließlich so_____ als Gew_____ sehen. I_____ diesem Si_____ ist sic_____ der Ti_____ der Lebenseri_____ Klaus von Bismarcks z_____ verstehen: "Aufbruch aus Pommern".

2. Linkshänder

Das Schreibenlernen ist für viele Linkshänder eine besonders große Hürde. Unsere Sch_____ verläuft v_____ links na_____ rechts. D_____ bedeutet f_____ den Rechts_____ eine beq_____, ziehende Bewe_____. Für d_____ Linkshänder ab_____ wird e_____ eine schieb_____, die weit_____ schwieriger i_____. Diese Tec_____ muss sc_____ im Vorsch_____, zum Beis_____ beim Ma_____, geübt wer_____. Lob u_____ Geduld d_____ Eltern erlei_____ dem Links_____ das Schreib_____, außerdem gibt es spezielle Schreib- und Zeichengeräte sowie Scheren für Linkshänder, die helfen, die Anforderungen der Schule besser zu erfüllen.

3. Astrologie

Die Astrologie begann ganz einfach – nämlich mit der Beobachtung des Himmels. Nachdem d_____ Mensch entd_____ hatte, d_____ aller Se_____ wie al_____ Unheil v_____ Himmel k_____, wandte e_____ seine Beobacht_____ intensiv d_____ himmlischen Gesc_____ zu u_____ begann, d_____ Zeichen d_____ Himmels z_____ erforschen. V_____ Himmel k_____ lebensspendende Wä_____, aber au_____ tödliche Hi_____. Der Men_____ vermutete hin_____ diesen Naturvo_____ himmlische, gött_____ Kräfte u_____ versuchte, beide zu erforschen, weshalb Astronomie und Astrologie also in den Anfängen gleich waren.

4. Lenin-Denkmal

Das Lenin-Denkmal am Wiener Platz in Dresden, dem der Innenstadt zugekehrten Vorplatz des Hauptbahnhofs, wird jetzt abgebaut. Die Stadtver_____ haben ein_____ Mühen hin_____ sich, z_____ einem Ents_____ darüber z_____ kommen, w_____ aus d_____ Denkmal d_____ Gründers d_____ Sowjetunion wer_____ solle. Zunä_____ war e_____ einfacher Ab_____ geplant, d_____ den nachrevol_____ Stimmungen i_____ der Bevöl_____ entsprochen hä_____, zumal d_____ zu ei_____ umgestalteten Fußgän_____ erweiterte eins_____ Prachtstraße i_____ Oktober 1989 e_____ Brennpunkt der Revolution in Dresden war.

5. Kunstleben

Jahrzehntlang war das Kunstleben in Russland streng organisiert und strukturiert: Alle wussten genau, welche Kunst erlaubt und welche verboten war. Jetzt sind die Kulturins zusammengebrochen und viele Künste in den Westen gegangen. Daraus können man schlussfolgern, dass in diesem Bereich der Kunst ein Vakuum entsteht sei. Doch das Gegenteil ist der Fall. Bei einem kurzen Besuch in Moskau überzeuge ich mich von der unverminderten Vitalität der dortigen Kunstszene. Die sich ihrer Struktur in der letzten Zeit stark verändert hat, steht auf einem anderen Blatt.

6. Risiken der Technik

Atomkraftwerke, Chemiefabriken oder die Gentechnik entziehen sich nicht nur der Kontrolle des Einzelnen; laufen sol_____ Anlagen Am_____, können d_____ Schäden f_____ die Allgem_____ tatsächlich s_____ unermesslich se_____, dass sel_____ Versicherungsunternehmen kapitu_____. Dieses Gefahrenp_____ verursacht Äng_____, die si_____ naturgemäß b_____ Störfällen verst_____, die i_____ Drehbuch d_____ Ingenieure u_____ Planer ni_____ vorgesehen wa_____ – zumal f_____ die Nai_____ unter d_____ Verbrauchern d_____ Strom a_____ der Stec_____ kommt und der Klebstoff aus der Tube.

**Fragebogen zum Sprachenprofil
(L2-Lerner_innen)**

Teilnehmer-Code:

1. Ihr Alter _____ Jahre

2. Ihr Geschlecht weiblich ()
männlich ()

3. In welchen Berufen arbeiten Sie bzw. haben Sie gearbeitet?

4. Welche Fächer studieren Sie bzw. haben Sie studiert?

5. Mit welcher Sprache sind Sie primär aufgewachsen?

a. _____

Falls Sie mehrsprachig aufgewachsen sind: mit welchen weiteren Sprachen?

b. _____

c. _____

6. Wo sind Sie aufgewachsen? Wenn in Deutschland, bitte Regionen und/oder Städte angeben.

Wenn Deutsch Ihre einzige Muttersprache ist, dann machen Sie bitte mit Frage 22 weiter.

Deutsch im Herkunftsland

7. Haben Sie Deutsch schon in Ihrem Herkunftsland gelernt?
(Bitte berücksichtigen Sie im Folgenden sowohl unterrichtliches als auch außerunterrichtliches Lernen!)

Ja () Nein () Wenn nein, bitte mit Frage 11 fortfahren!

Anhang

8. In welchem Alter haben Sie begonnen, Deutsch in Ihrem Herkunftsland zu lernen?
mit _____ Jahren
9. Wie lange haben Sie Deutsch in Ihrem Herkunftsland gelernt?
a. _____ Jahre und _____ Monate
b. Wie viele Stunden (à 60 Minuten) pro Woche durchschnittlich?
_____ Stunden pro Woche
10. Bitte beschreiben Sie stichpunktartig, wie Sie in Ihrem Heimatland vorwiegend Deutsch gelernt haben (z.B. Deutschunterricht an der Schule und/oder an der Universität, deutschsprachige Freunde, Tandem, Privatstunden, autodidaktisch etc.).

11. Spricht in Ihrer Familie jemand Deutsch als Muttersprache? Mutter ()
Vater ()
sonst jemand: _____

Deutsch in Deutschland bzw. in einem anderen deutschsprachigen Land

12. Wann haben Sie zum ersten Mal Deutschland oder ein anderes deutschsprachiges Land besucht?
mit _____ Jahren und _____ Monaten
13. Wann sind Sie nach Deutschland oder in ein anderes deutschsprachiges Land gezogen?
mit _____ Jahren und _____ Monaten
14. Wie lange leben Sie insgesamt schon in Deutschland oder in einem anderen deutschsprachigen Land?
_____ Jahre und _____ Monate

15. Haben Sie Deutschland/die deutschsprachigen Länder zwischendurch für mehr als 2 Monate verlassen?

a. Ja () Nein ()

b. Wenn ja, wie viele Monate insgesamt?

_____ Monate

16. Hatten Sie in Deutschland und/oder in den deutschsprachigen Ländern Deutschunterricht?

a. Ja () Nein ()

b. Wenn ja, wie lange?

_____ Jahre und _____ Monate

c. Wie viele Stunden (à 60 Minuten) pro Woche durchschnittlich?

_____ Stunden pro Woche

17. Wie häufig verwenden Sie die deutsche Sprache in Ihrem täglichen Leben auf einer Skala von 1 (sehr selten) bis 4 (sehr häufig)? Markieren Sie die entsprechende Zahl mit einem Kreis.

sehr selten	selten	häufig	sehr häufig
1	2	3	4

18. Wie häufig verwenden Sie Ihre Muttersprache in Ihrem täglichen Leben auf einer Skala von 1 (sehr selten) bis 4 (sehr häufig)? Markieren Sie die entsprechende Zahl mit einem Kreis.

sehr selten	selten	häufig	sehr häufig
1	2	3	4

19. Wie häufig verwenden Sie weitere Sprachen in Ihrem täglichen Leben auf einer Skala von 1 (sehr selten) bis 4 (sehr häufig)?

sehr selten	selten	häufig	sehr häufig
1	2	3	4

Anhang

20. Schätzen Sie bitte Ihre Deutschkenntnisse auf einer Skala von 1 (sehr niedrig, Anfänger) bis 6 (fast muttersprachlich) ein. Markieren Sie die entsprechende Zahl mit einem Kreis.

sehr niedrig	niedrig	mittel	hoch	sehr hoch	fast muttersprachlich
1	2	3	4	5	6

21. Wie wichtig ist es Ihnen, Deutsch auf einem sehr hohen Niveau zu sprechen? Bitte tragen Sie es auf einer Skala von 1 (völlig unwichtig) bis 4 (sehr wichtig) ein. Markieren Sie die entsprechende Zahl mit einem Kreis.

völlig unwichtig	unwichtig	wichtig	sehr wichtig
1	2	3	4

Weitere Fremdsprachenkenntnisse

22. Welche weiteren Fremdsprachen haben Sie gelernt?

Fremdsprache	Anfangsalter	Niveau von 1 (sehr niedrig, Anfänger) bis 6 (fast muttersprachlich)

23. Bitte nennen Sie uns noch Ihre E-Mail-Adresse und/oder Ihre Telefonnummer:

E-Mail:

Telefon:

(Diese Angaben sind selbstverständlich freiwillig. Wenn Sie Interesse haben, weiter an der Studie teilzunehmen, ist es für uns notwendig, Sie im Falle von Fragen oder weiteren Untersuchungen kontaktieren zu können.)

Vielen Dank für Ihre Mitarbeit!

Fragebogen zum Sprachenprofil (2L1)

1. Ihr Alter _____ Jahre

2. Ihr Geschlecht weiblich ()
männlich ()

3. In welchen Berufen arbeiten Sie bzw. haben Sie gearbeitet?

4. Welche Fächer studieren Sie bzw. haben Sie studiert?

5. Mit welchen Sprachen sind Sie aufgewachsen?

6. Wo sind Sie aufgewachsen? (bitte Länder, Regionen und/oder Städte angeben sowie Veränderungen wie Umzüge, längere Aufenthalte und jeweils den ungefähren Zeitpunkt bzw. die Zeitspanne dafür angeben)

7. Wie alt waren Sie beim ersten Kontakt mit der deutschen Sprache?

8. Mit wem haben Sie Deutsch gesprochen? War/en das ein/alle Muttersprachler?

9. Wie viel Zeit haben Sie an einem durchschnittlichen Tag mit dieser Person/diesen Personen verbracht? (wenn es einschneidende Veränderungen gab, bitte auch notieren)

11. In welchen Sprachen sind Sie alphabetisiert worden?

Anhang

12. Wie würden Sie das Verhältnis der Sprachen, mit denen sie aufgewachsen sind, beschreiben?

Sprechen: relativ ausgewogen
 Deutsch war dominant
 andere Sprache(n) waren dominant

Schreiben: relativ ausgewogen
 Deutsch war dominant
 andere Sprache(n) waren dominant

13. Wie lange leben Sie insgesamt schon in Deutschland oder in einem anderen deutschsprachigen Land?

_____ Jahre und _____ Monate

14. Wie häufig verwenden Sie derzeit die deutsche Sprache in Ihrem täglichen Leben auf einer Skala von 1 (sehr selten) bis 4 (sehr häufig)? Markieren Sie die entsprechende Zahl mit einem Kreis.

sehr selten	selten	häufig	sehr häufig
1	2	3	4

15. Wie häufig verwenden derzeit Sie die andere(n) Sprache(n), mit der (denen) Sie aufgewachsen sind, in Ihrem täglichen Leben auf einer Skala von 1 (sehr selten) bis 4 (sehr häufig)? Markieren Sie die entsprechende Zahl mit einem Kreis.

Sprache 1:

sehr selten	selten	häufig	sehr häufig
1	2	3	4

Sprache2:

sehr selten	selten	häufig	sehr häufig
1	2	3	4

Sprache3:

sehr selten	selten	häufig	sehr häufig
1	2	3	4

16. Wie häufig verwenden Sie derzeit Ihre Fremdsprachen in Ihrem täglichen Leben auf einer Skala von 1 (sehr selten) bis 4 (sehr häufig)?

Sprache 1:				
sehr selten	selten	häufig	sehr häufig	
1	2	3	4	
Sprache2:				
sehr selten	selten	häufig	sehr häufig	
1	2	3	4	
Sprache3:				
sehr selten	selten	häufig	sehr häufig	
1	2	3	4	

17. Wie häufig nutzen Sie deutschsprachige Medien?

	nie	sehr selten	selten	häufig	sehr häufig
TV	0	1	2	3	4
Radio	0	1	2	3	4
Deutsche Musik	0	1	2	3	4
Zeitungen	0	1	2	3	4
Zeitschriften	0	1	2	3	4
Bücher	0	1	2	3	4

18. Welche Sprachen werden in Ihrer Wohnung gesprochen?

- Deutsch
- die andere(n) Sprache(n), mit der (denen) ich aufgewachsen bin
- andere Sprachen
- Ich wohne allein.

Anhang

19. Welche Sprachen werden da, wo Sie derzeit arbeiten/studieren, gesprochen?

- Deutsch
- die andere(n) Sprache(n), mit der (denen) ich aufgewachsen bin
- andere Sprachen

20. Schätzen Sie bitte Ihre Deutschkenntnisse auf einer Skala von 1 (sehr niedrig, Anfänger) bis 6 (fast muttersprachlich) ein. Markieren Sie die entsprechende Zahl mit einem Kreis.

sehr niedrig	niedrig	mittel	hoch	sehr hoch	fast muttersprachlich
1	2	3	4	5	6

21. Wie wichtig ist es Ihnen, Deutsch auf einem sehr hohen Niveau zu sprechen? Bitte tragen Sie es auf einer Skala von 1 (völlig unwichtig) bis 4 (sehr wichtig) ein. Markieren Sie die entsprechende Zahl mit einem Kreis.

völlig unwichtig	unwichtig	wichtig	sehr wichtig
1	2	3	4

22. Welche weiteren Fremdsprachen haben Sie gelernt?

Fremdsprache	Niveau von 1 (sehr niedrig, Anfänger) bis 6 (fast muttersprachlich)

23. Wenn sie Interesse an weiteren Untersuchungen haben, nennen Sie uns bitte noch Ihre E-Mail-Adresse und/oder Ihre Telefonnummer:

E-Mail:

Telefon:

(Diese Angaben sind selbstverständlich freiwillig. Wenn Sie Interesse haben, weiter an der Studie teilzunehmen, ist es für uns notwendig, Sie im Falle von Fragen oder weiteren Untersuchungen kontaktieren zu können.)

Vielen Dank für Ihre Mitarbeit!

Eigenschaftsprofil

Uns interessiert, wie Sie sich im Fremdsprachenunterricht zumeist fühlen bzw. gefühlt haben.

Bitte **kreuzen** Sie an, inwieweit die folgenden Aussagen auf Sie persönlich zutreffen?

trifft sicher nicht zu trifft sicher zu

1 - 2 - 3 - 4

x

- | | |
|---|---------------|
| 1. Ich zittere, wenn ich weiß, dass ich im Fremdsprachenunterricht gleich aufgerufen werde. | 1 - 2 - 3 - 4 |
| 2. Ich denke dauernd daran, dass die anderen Teilnehmer besser in Fremdsprachen sind als ich. | 1 - 2 - 3 - 4 |
| 3. Ich gerate in Panik, wenn ich ohne Vorbereitung im Fremdsprachenunterricht reden muss. | 1 - 2 - 3 - 4 |
| 4. Es ist mir unangenehm, wenn ich an einem Gespräch im Fremdsprachenunterricht teilnehmen soll. | 1 - 2 - 3 - 4 |
| 5. Auch wenn ich mich auf den Fremdsprachenunterricht gut vorbereitet habe, habe ich Angst. | 1 - 2 - 3 - 4 |
| 6. Ich habe Herzklopfen, wenn ich im Fremdsprachenunterricht aufgerufen werde. | 1 - 2 - 3 - 4 |
| 7. Ich werde nervös und ganz durcheinander, wenn ich im Fremdsprachenunterricht spreche. | 1 - 2 - 3 - 4 |
| 8. Es ist mir unangenehm, wenn ich vor den anderen Mitlernern die Fremdsprache sprechen muss. | 1 - 2 - 3 - 4 |
| 9. Im Fremdsprachenunterricht fühle ich mich angespannter und nervöser als in anderen Fächern. | 1 - 2 - 3 - 4 |
| 10. Ich habe immer das Gefühl, dass die anderen in meinem Kurs die Fremdsprache besser sprechen als ich. | 1 - 2 - 3 - 4 |
| 11. Ich werde nervös, wenn Lehrende mir Fragen stellen, die ich nicht schon vorher vorbereitet habe. | 1 - 2 - 3 - 4 |

Anhang

Wir möchten Ihre Selbsteinschätzung zu folgenden Eigenschaften erfahren.

Bitte **kreuzen** Sie an, inwieweit die folgenden Aussagen auf Sie persönlich zutreffen!

trifft sicher nicht zu

trifft sicher zu

1 - 2 - 3 - 4

x

- | | |
|---|---------------|
| 1. Ich bin abenteuerlustig. | 1 - 2 - 3 - 4 |
| 2. Ich bin schnell entschlossen. | 1 - 2 - 3 - 4 |
| 3. Ich bin leicht provozierbar. | 1 - 2 - 3 - 4 |
| 4. Ich bin spontan. | 1 - 2 - 3 - 4 |
| 5. Ich bin gern mit anderen Menschen zusammen. | 1 - 2 - 3 - 4 |
| 6. Ich gehe viel aus. | 1 - 2 - 3 - 4 |
| 7. Ich schreie zurück. | 1 - 2 - 3 - 4 |
| 8. Ich bin amüsierfähig. | 1 - 2 - 3 - 4 |
| 9. Ich bin sehr lebhaft. | 1 - 2 - 3 - 4 |
| 10. Ich bin in Gegenwart anderer kommunikativ. ... | 1 - 2 - 3 - 4 |
| 11. Ich handle gern schnell. | 1 - 2 - 3 - 4 |
| 12. Ich nutze jede Gesprächsgelegenheit. | 1 - 2 - 3 - 4 |
| 13. Ich brauche persönliche Kontakte. | 1 - 2 - 3 - 4 |
| 14. Ich bin selbstbewusst. | 1 - 2 - 3 - 4 |
| 15. Ich kann gut aus mir herausgehen. | 1 - 2 - 3 - 4 |
| 16. Ich kann Partys in Schwung bringen. | 1 - 2 - 3 - 4 |

Warum haben Sie Deutsch gelernt? bzw. Warum lernen Sie Deutsch?
 Bitte **kreuzen** Sie an, inwieweit die folgenden Aussagen auf Sie persönlich zutreffen !

trifft sicher nicht zu
trifft sicher zu

1 - 2 - 3 - 4
x

Ich habe Deutsch gelernt ... bzw. Ich lerne Deutsch ...

- | | |
|---|---------------|
| 1. ... für meinen zukünftigen Beruf. | 1 - 2 - 3 - 4 |
| 2. ... um in deutschsprachige Länder reisen zu können. | 1 - 2 - 3 - 4 |
| 3. ... weil ich die Deutschen mag. | 1 - 2 - 3 - 4 |
| 4. ... um die deutsche Lebensweise und das Land, in dem diese Sprache
gesprochen wird, besser zu verstehen. | 1 - 2 - 3 - 4 |
| 5. ... weil es mir Spaß macht. | 1 - 2 - 3 - 4 |
| 6. ... um ein höheres Einkommen zu erzielen. | 1 - 2 - 3 - 4 |
| 7. ... weil ich diese Sprache mag. | 1 - 2 - 3 - 4 |
| 8. ... um in meinem Leben eine größere Anzahl deutschsprachiger Menschen
kennen zu lernen. | 1 - 2 - 3 - 4 |
| 9. ... weil es für mein Studium wichtig ist. | 1 - 2 - 3 - 4 |
| 10. ... weil ich gerne Sprachen lerne. | 1 - 2 - 3 - 4 |
| 11. ... um eine höhere Bildung zu erlangen. | 1 - 2 - 3 - 4 |
| 12. ... weil mir Deutschland gefällt. | 1 - 2 - 3 - 4 |
| 13. ... weil ich im Studium erfolgreich sein möchte. | 1 - 2 - 3 - 4 |
| 14. ... weil mich die deutsche Kultur und Sprache interessieren. | 1 - 2 - 3 - 4 |
| 15. ... weil ich im Beruf erfolgreich sein möchte. | 1 - 2 - 3 - 4 |

Anhang

Liebe Bewerberin, lieber Bewerber

mit unserer Studie wollen wir dokumentieren, wie erfolgreich auch Erwachsene eine Fremdsprache erlernen können. Gleichzeitig interessieren wir uns dafür, welche Eigenschaften erfolgreiche Lernende aufweisen. Neben der Überprüfung des grammatischen, lexikalischen und morpho-syntaktischen Wissens (Grammatik, Wortschatz, Satzbau und Wortendungen) der Testpersonen, wollen wir auch eine Einschätzung der Phonetik, das heißt der Aussprache der Testpersonen. Deswegen möchten wir Sie bitten die folgenden ca. 30 sekundigen Ausschnitte aus den Interviews, die mit den Testpersonen geführt wurden, zu bewerten. Die Sprechproben sind von Personen, die

- als Erwachsene Deutsch gelernt haben oder
- zweisprachig aufgewachsen (zwei Erstsprachen, eine davon ist Deutsch) sind oder
- mit Deutsch als Erstsprache aufgewachsen sind (Kontrollgruppe)

Der Focus dieser Bewertung liegt auf der Phonetik bezogen auf Akzent und Verständlichkeit, weswegen wir Sie bitten, sich darauf zu konzentrieren. Versuchen Sie, sich bei Ihrem Urteil nicht von eventuellen Grammatik- oder Wortschatz-Fehlern leiten zu lassen. Bitte hören Sie sich die Audio-Dateien jeweils nur einmal an und bewerten Sie nach Ihrem ersten Eindruck. Hören Sie sich die Dateien nur in Ausnahmefällen zweimal an, bspw. wenn Sie sehr unsicher sind, wie Sie urteilen sollen oder beim Hören gestört wurden.

Akzent bezeichnet die – meist unbewusste und ungewollte – Übertragung von Aussprachegewohnheiten der Erst- bzw. Muttersprache oder der vorrangig gebrauchten Sprache auf eine später erlernte (Fremdsprache) oder weniger gebrauchte Sprache. Die Abweichungen können die Lautformung, Intonation, Betonungsmuster und Satzrhythmus betreffen. Auch die Aussprachegewohnheiten eines Dialekts können von den Sprecher_innen auf die Hochsprache übertragen werden.

Verständlichkeit bezeichnet hier die phonetische Verständlichkeit in dem Sinne, ob die Äußerungen klar zu verstehen sind (s. Bewertungsskala) und nicht die inhaltliche Verständlichkeit in dem Sinne, ob das Gesagte einen Sinn ergibt. Inhaltlich sind die Sätze der Sprechproben teilweise aus dem Zusammenhang gerissen, da sie zusammengeschnitten sind.

Die Bewertung erfolgt in drei Schritten:

- Beurteilen Sie zuerst die Sprechprobe mit dieser fünf-stufigen Skala:
 - 1 = sehr starker Akzent
 - 2 = starker Akzent
 - 3 = weniger starker Akzent
 - 4 = leichter Akzent
 - 5 = kein Akzent (akzentfrei)
- Kreuzen Sie an, ob Sie den Akzent als fremdsprachlich oder dialektbedingt wahrnehmen:
Der Akzent klingt für mich fremdsprachlich oder dialektbedingt
- Beurteilen Sie die Verständlichkeit der Sprechprobe mit dieser fünf-stufigen Skala:
 - 1 = Die Person ist nicht zu verstehen
 - 2 = Die Person ist schwer zu verstehen, ich musste mich beim Hören sehr anstrengen
 - 3 = Die Person ist stellenweise schwer zu verstehen, ich musst mich beim Hören teilweise anstrengen¹
 - 4 = Die Person ist zu verstehen, ich musste mich beim Hören aber etwas anstrengen
 - 5 = Die Person spricht klar und deutlich und ist gut zu verstehen
- In dem Feld „Begründen Sie Ihre Meinung“, haben Sie die Gelegenheit eine kurze Begründung Ihrer Entscheidung zu schreiben, wenn Sie das möchten.

Wir möchten Sie bitten, sich bei allen Aufgaben bestmöglich zu konzentrieren.

Wir danken Ihnen herzlich für Ihre Mitarbeit und Unterstützung!

¹ Hier ist gemeint, dass z.B. einzelne Wörter nicht zu verstehen sind

Hintergrundinformationen

1. a) Haben Sie in Deutschland studiert? ja nein
b) Falls ja, wie viele Semester? ____
2. Fühlen Sie sich „locker“, wenn Sie mit Deutschen zusammen sind?
a) beruflich ja nein
b) privat ja nein
3. Haben bzw. hatten Sie einen deutschen (oder deutschsprachigen) Partner oder Partnerin?
 ja nein

Wenn ja, sprechen bzw. sprachen Sie überwiegend Deutsch miteinander?
 ja nein
4. Haben bzw. hatten Sie deutsche (oder deutschsprachige) Freunde, Freundinnen und Bekannte mit denen Sie regelmäßig Zeit verbringen bzw. verbrachten?
 ja nein

Wenn ja, sprechen bzw. sprachen Sie überwiegend Deutsch miteinander?
 ja nein
5. Haben Sie vor, noch mehr Sprachen zu lernen?
 ja nein
6. Was war/ist Ihr Ziel beim Deutschlernen?
 Deutsch wie eine Muttersprachlerin/ein Muttersprachler zu sprechen
 Sehr gut Deutsch zu sprechen
 Gut Deutsch zu sprechen
 Ausreichend Deutsch zu sprechen
7. Ich kann gut Sprachen/Dialekte/Akzente nachmachen.
 ja nein
8. Wenn mich Leute aus meinem deutschsprachigen Umfeld korrigieren, dann
 freue ich mich im Allgemeinen darüber.
 ist mir das im Allgemeinen egal.
 ärgert mich das im Allgemeinen.
9. Ich habe Leute aus meinem deutschsprachigen Umfeld oft darum gebeten, mich zu korrigieren
 ja nein
10. Wenn mich Leute korrigieren, dann
 übernehme ich die Korrektur und versuche sie mir zu merken.
 merke ich mir die Korrektur vielleicht, vielleicht aber auch nicht.
 merke ich mir die Korrektur nicht.

Anhang

11. Welcher der Bereiche (Sprechen, Hören, Lesen, Schreiben, Grammatik, Wortschatz, Aussprache) war /ist Ihnen besonders wichtig beim Deutschlernen? (Mehrfachantworten möglich)

a) **Sprechen**

- alltägliche Kommunikation
- berufliche Kommunikation
- akademische Kommunikation

b) **Hören**

- alltägliche Kommunikation
- berufliche Kommunikation
- akademische Kommunikation

Was genau hören Sie?

- Radio
- Fernsehen (hören und sehen)
- Filme (hören und sehen)
- Hörbücher
- Vorträge
- anderes, z. B. _____

c) **Lesen**

- alltägliche Kommunikation
- berufliche Kommunikation
- akademische Kommunikation

Was genau lesen Sie?

- Zeitung
- Internet (Foren, Chats etc. wenn Sie häufig Nachrichten im Internet lesen, kreuzen Sie bitte „Zeitung“ an)
- E-Mails
- Briefe
- Formulare
- Verträge
- Bücher
 - Sachliteratur
 - Fachliteratur
 - Belletristik (Romane etc.)
 - Lyrik
- anderes, z. B. _____

d) **Schreiben**

- alltägliche Kommunikation
- berufliche Kommunikation
- akademische Kommunikation

Was genau schreiben Sie?

- E-Mails
- Briefe

- wissenschaftliche Texte
- Notizen
- anderes, z. B. _____

e) **Grammatik**

- Möglichst fehlerfrei sprechen und schreiben war/ist mein Ziel
- Hauptsache, ich verstehe und werde verstanden

f) **Wortschatz/Vokabeln**

- einen ausdifferenzierten Wortschatz zu haben, war/ist mein Ziel
- einen großen Wortschatz zu haben, war/ist mein Ziel
- für mich ist der Wortschatz nicht so wichtig, zur Not kann ich ein fehlendes Wort auf Englisch (oder einer anderen Sprache) verwenden und/oder umschreiben

g) **Aussprache**

- Aussprache war/ist mir nur insofern wichtig, um verstanden zu werden
- eine gute Aussprache war/ist mir wichtig, um gut verstanden zu werden
- eine sehr gute Aussprache war/ist mir besonders wichtig, mein Ziel war/ist es, dass ich akzentfrei Deutsch spreche

12. Wie häufig nutzen Sie deutschsprachige Medien?

	sehr selten	selten	häufig	sehr häufig
TV	1	2	3	4
Radio	1	2	3	4
Deutsche Musik	1	2	3	4
Zeitungen	1	2	3	4
Zeitschriften	1	2	3	4
Bücher	1	2	3	4
Internet	1	2	3	4

13. Welche Sprachen werden in Ihrer Wohnung gesprochen? (Mehrfachantworten möglich)

- Deutsch
- die Sprache, mit der ich aufgewachsen bin
- andere Sprachen
- Ich wohne allein.

Anhang

14. Welche Sprachen werden da gesprochen, wo Sie derzeit arbeiten/studieren?
(Mehrfachantworten möglich)

- Deutsch
- die andere(n) Sprache(n), mit der (denen) ich aufgewachsen bin
- andere Sprachen

15. Welche Sprachen werden da gesprochen, wo Sie Ihre Freizeit verbringen?
(Mehrfachantworten möglich)

- Deutsch
- die Sprache, mit der ich aufgewachsen bin
- andere Sprachen

Warum haben Sie Deutsch gelernt? Bzw. Warum lernen Sie Deutsch

Bitte kreuzen Sie an, inwiefern die folgenden Aussagen auf Sie persönlich zutreffen.
Dafür steht Ihnen eine Skala von 1 - 4 zur Verfügung:

- 1 = trifft sicher nicht zu
- 2 = trifft eher nicht zu
- 3 = trifft eher zu
- 4 = trifft sicher zu

16.	Als ich damals Deutsch-Unterricht (in Deutschland) hatte, habe ich auch außerhalb des Klassenzimmers versucht, viel Deutsch zu sprechen.	1	2	3	4
17.	Als ich damals Deutsch-Unterricht hatte, habe ich immer die Hausaufgaben gemacht.	1	2	3	4
18.	Mein Ziel war es, so gut Deutsch zu beherrschen wie eine Person, die mit Deutsch aufgewachsen ist.	1	2	3	4
19.	Ich habe Deutsch gelernt, weil enge Freundinnen und/oder Freunde von mir dachten, dass das wichtig ist.	1	2	3	4
20.	Deutsch lernen war wichtig für mich, weil ich dachte, dass es eines Tages nützlich wäre, um einen guten Job zu bekommen oder viel Geld zu verdienen.	1	2	3	4
21.	Als ich damals Deutsch-Unterricht hatte und über meine Zukunft nachdachte, wusste ich, dass Deutschkenntnisse wichtig sind.	1	2	3	4
22.	Ich habe hart gearbeitet, um Deutsch zu lernen.	1	2	3	4
23.	Gefällt es Ihnen, Deutsch zu sprechen?	1	2	3	4
24.	Ich kann ehrlich sagen, dass ich mein Bestes gegeben habe, um Deutsch zu lernen.	1	2	3	4

25.	Deutsch lernen war wichtig für mich, um die Anerkennung meiner Familie zu gewinnen.	1	2	3	4
26.	Als ich damals Deutsch-Unterricht hatte, war ich mir sicher, dass ich eine Fremdsprache lernen kann.	1	2	3	4
27.	Ich glaube, dass ich der Typ bin, der sich schnell ängstlich und unwohl fühlt, wenn er mit jemandem in der Fremdsprache sprechen muss.	1	2	3	4
28.	War/Ist es Ihnen wichtig, deutsche Literatur zu lesen?	1	2	3	4
29.	Ich musste Deutsch lernen, weil ich sonst meine Eltern enttäuscht hätte.	1	2	3	4
30.	Es war mir extrem wichtig, Deutsch zu lernen.	1	2	3	4
31.	Deutsch lernen war wichtig für mich, weil ich dachte, dass mit einem hohen Sprachniveau viel Geld zu verdienen ist.	1	2	3	4
32.	Als ich damals Deutsch-Unterricht hatte, war das Deutsch-Lernen einer der wichtigsten Aspekte meines Lebens.	1	2	3	4
33.	Ich sehe oft deutsches Fernsehen, um meine Sprachkenntnisse zu verbessern.	1	2	3	4
34.	Deutsch lernen war mir wichtig, weil ich ohne gute Noten in Deutsch als schwache Studentin bzw. schwacher Student eingeschätzt worden wäre.	1	2	3	4
35.	Als ich damals Deutsch-Unterricht hatte, stellte ich mir oft vor, dass ich jemand bin, der Deutsch fließend verwenden kann.	1	2	3	4
36.	Wenn ich ein deutsches Lied im Radio höre, höre ich oft sorgfältig zu und versuche, jedes Wort zu verstehen.	1	2	3	4
37.	Deutsch lernen war wichtig für mich, weil ich glaubte, dadurch eines Tages einen guten Job zu bekommen.	1	2	3	4
38.	Deutsch lernen war wichtig, weil mein Umfeld das von mir erwartet hat.	1	2	3	4
39.	Ich lernte Deutsch, weil ich sonst meinen Universitäts-Abschluss nicht bekommen hätte.	1	2	3	4
40.	Als ich damals Deutsch-Unterricht hatte, fühlte ich immer, dass meine Klassenkameradinnen und/oder Klassenkameraden besser waren als ich.	1	2	3	4
41.	Als ich damals Deutsch-Unterricht hatte, habe ich versucht, auch durch deutsche Musik, deutsches Fernsehen etc. Deutsch zu lernen.	1	2	3	4

Anhang

42.	Deutsch lernen war wichtig für mich, um die Anerkennung meiner Freundinnen bzw. Freunde zu gewinnen	1	2	3	4
43.	Wenn es einen Deutschkurs gäbe, der mich auf meinem Niveau noch verbessern könnte, würde ich daran teilnehmen.	1	2	3	4
44.	Fühlen Sie sich mit Ihren deutschen Sprachkenntnissen respektiert?	1	2	3	4
45.	Deutsch lernen war wichtig für mich, um die Anerkennung meiner Lehrerinnen und/oder Lehrer zu gewinnen.	1	2	3	4
46.	Als ich damals Deutsch-Unterricht hatte, konnte ich mir vorstellen, in der Zukunft Deutsch zu verwenden.	1	2	3	4
47.	Deutsch lernen war wichtig für mich, um die Anerkennung meiner Chefin und/oder meines Chefs zu gewinnen.	1	2	3	4
48.	Eine Fremdsprache zu lernen, ist eine schwierige Aufgabe für mich.	1	2	3	4
49.	Ich lernte Deutsch, weil ich keine schlechten Noten bekommen wollte.	1	2	3	4
50.	Als ich damals Deutsch-Unterricht hatte, wusste ich, dass ich Deutsch für meine Zukunft brauchen würde.	1	2	3	4
51.	Als ich damals Deutsch-Unterricht hatte, befürchtete ich, dass die anderen Schülerinnen und/oder Schüler über mich lachen könnten, wenn ich Deutsch sprach.	1	2	3	4
52.	Als ich damals Deutsch-Unterricht hatte, konnte ich mir vorstellen, es mit deutschen Freunden zu sprechen.	1	2	3	4
53.	Meine Eltern glauben, dass Fremdsprachenkenntnisse der Schlüssel zum Erfolg sind.	1	2	3	4
54.	Ich bin ein Typ, der sich sehr anstrengt, um Deutsch zu lernen.	1	2	3	4
55.	War/Ist es Ihnen wichtig, deutsche Filme zu verstehen?	1	2	3	4

Die Zielstellung der vorliegenden Studie ist die Identifizierung der Faktoren, die individuelle Leistungsunterschiede beim L2-Lernen von Späterwerber/-innen (Erwerbsbeginn ≥ 16 Jahre) beeinflussen. Dabei wird die Sprachkompetenz von Späterwerber/-innen (Erwerbsbeginn ≥ 16 Jahre) gemessen und mit der von Personen mit doppeltem Erstspracherwerb (Erwerbsbeginn = 0-3 Jahre) und monolingualen Erstsprachler/-innen verglichen. Auf der Basis dieser Leistungsmessung werden die Profile erfolgreicher Späterwerber/-innen unter Berücksichtigung von kognitiven, affektiv-motivationalen und sozio-kulturellen Einflussgrößen spezifiziert. Die Ergebnisse der Frühbilingualen variieren stärker als die der Erstsprachler/-innen. Dies hat zur Folge, dass es mehr Späterwerber/-innen gibt, die innerhalb des Leistungsbereichs der Frühbilingualen abscheiden. Des Weiteren wird ein deutlicher Alterseffekt festgestellt. Nach dem Faktor Alter weisen die Faktoren Motivation der Lerner/-innen, Deutsch auf hohem Niveau zu beherrschen, und die Aufenthaltsdauer im Land der Zielsprache den größten Zusammenhang auf.

...everybody has a view of the role of age in language learning, and, curiously enough, most of these views are wrong." (Dörnyei 2009: 233)

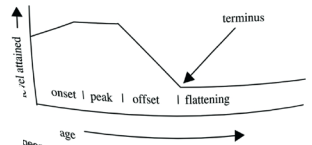
...stands it is like the mythical hydra, whose heads are so many that the capacity to produce new heads is rendered it impossible to cut them off." (Singler 2009: 280)

1. Forschungsstand Faktor

...suggero che la Babelica sia divenuta Curia e che il suo destino è quello di un'isola di lingua.

Eine kritische Periode für den Zweitspracherwerb und das Fremdsprachenlernen?

„Daher lässt sich ferner nachweisen, dass zu der Zeit, zu der der primäre Spracherwerb aufhört, das Gehirn seinen Reifezustand erreicht hat und die zerebrale Lateralisation unumkehrbar festgelegt ist. [...] Als Arbeitshypothese nehmen wir deshalb an, dass die allgemeinen nichtspezifischen Stufen der Hirnreifung Vorbedingungen und begrenzende Faktoren für die Sprachentwicklung darstellen.“ (Lenneberg 1972: 208)



geometrische Eigenschaften einer CP (Abb. aus van Bostel/Bongaerts/Coppen 2005: 356)

...Zusammenhang mit dem Einfluss des Alters auf den L2-Lernerfolg werden

- Neurolinguistische Reifung
- Motivation und Erwerbsanliegen
- Sprachlernleistung
- Kognitive Entwicklung
- Lernstil - Lerntyp
- expliziter vs. impliziter Erwerb
- Quantität und Qualität des

„Die erreichte Sprachkompetenz lässt während der kritischen Periode biologisch bedingt linear nach, es kommt zu einem „cut-off point“ in der statistischen Auswertung am Endpunkt der kritischen Periode, der Späterwerb nach Abschluss der kritischen Periode ist primär von sehr unterschiedlichen Umweltbedingungen abhängig, und die durchschnittlich erreichte Sprachkompetenz lässt dadurch diskontinuierlich nach.“ (Schlack 2003: 18)

ISBN 978-3-7376-0272-3

9 783737 602723 >