

Adriana Hanulíková & Daniel Feldmeth-Müller
Universität Freiburg
Deutsches Seminar – Germanistische Linguistik

Abstract

This paper provides a detailed description of a project funded by an Assistant Professor grant from the Ministry of Science, Research and Arts in Baden-Württemberg. The aim of the project was to investigate individual differences in the comprehension and production of spoken language in youth population from diverse socioeconomic and language backgrounds. The main objective was to investigate the speech comprehension and production interface and to determine the relative contribution of participants' background as well as of language, auditory and cognitive skills to speech comprehension and production performance across various linguistic levels and tasks under noisy and thus more demanding listening conditions. We expected large individual variation across different tasks and participants that would allow for an analysis and a better understanding of the relative weight of diverse factors affecting language performance and language skills at different linguistic levels. To this end, a pilot and a main study were conducted. The main study comprised of 131 adolescents from three school types in Freiburg who completed a battery of tests assessing passive and active vocabulary knowledge, comprehension of utterances produced in accented and noise contexts, memory and inhibitory skills as well as a scene description task. In addition, a detailed questionnaire assessing the social and language background of the adolescents was included. This contribution provides a detailed description of each of these tests as well as the process of data collection. Finally, some preliminary descriptive results and an overview of ongoing studies and future research are presented.

Keywords: adolescents' language skills, individual variation, language production, speech comprehension

1 Einleitung

Dieser Beitrag beschreibt das im Rahmen des Juniorprofessuren-Programms des Ministeriums für Wissenschaft, Forschung und Kunst Baden-Württemberg gefördertes Projekt (No. 8025303102), in dem individuelle Unterschiede bezüglich kognitiver und sprachlicher Fähigkeiten Jugendlicher aus verschiedenen Schulen sowie sprachlichen und sozioökonomischen Hintergründen untersucht wurden. Konkret sollte einerseits der Zusammenhang zwischen Sprachproduktion und Sprachverstehen untersucht werden, andererseits das relative Gewicht unterschiedlicher Einflussfaktoren auf die sprachliche Leistung auf unterschiedlichen linguistischen Ebenen sowie unter erschwerten Hör- und Sprechbedingungen. Frau Hanulíková entwarf und leitete das Projekt, Herr Müller-Feldmeth hat als Post-Doc (2016-2018) u.a. die computerbasierte Umsetzung einiger Tests sowie die Organisation der Datenerhebung übernommen. Mehrere studentische und wissenschaftliche Hilfskräfte unterstützten uns bei der Datenerhebung an den Schulen in der Pilotphase (Anna Mormino, Katharina Truong), in der Hauptphase (Saskia Leymann, Lisa Hübner) und bei der Annotation der Produktionsdaten (Lisa Hübner, Dina Osime, Julia Haustein, Julia Weber).

Der Ausgangspunkt des Projekts war die Beobachtung, dass die aktuelle Forschung sich vor allem mit dem erfolgreichen Erwerb einer Zweit- oder Fremdsprache befasst und vergleichsweise weniger mit der Variation hinsichtlich der sprachlichen Fähigkeiten in einer Muttersprache¹ und der kognitiven Fähigkeiten in der Altersgruppe der Jugendlichen (z.B. Dabrowska, 2012; Darcy, Park, & Yang, 2015; Sparks & Ganschow, 1993; Sparks, Ganschow, & Patton, 1995; Sparks et al., 1997). Dabei variieren sprachliche Fähigkeiten in einer Muttersprache nicht nur stark, sie lassen auch den Erfolg beim Erwerb weiterer Sprachen vorhersagen. So wurde zum Beispiel gezeigt, dass schwächere Leistung in Lautdiskriminierung fremdsprachlicher Kontraste mit schwächerer Lautdiskriminierung in der Muttersprache zusammenhängt (Diaz et al., 2008). Es zeigte sich zudem, dass einige der schwächeren HörerInnen außergewöhnliche SprecherInnen waren (persönliches Gespräch, 2011), was eine Dissoziierung zwischen lautlicher Wahrnehmung und Sprechen nahelegt. Es ist deswegen von theoretischer

¹ Der Begriff „Muttersprache“ oder „MuttersprachlerIn“ ist nicht problemlos und wird z.B. bei Davies (2003) kritisch evaluiert.

Relevanz, die individuellen Unterschiede in unterschiedlichen sprachlichen Bereichen näher zu untersuchen. Auf welcher sprachlichen Ebene der Zusammenhang zwischen Perzeption und Verstehen einerseits und Produktion und SprecherIn andererseits besteht, bleibt weiterhin unklar. Eine systematische Untersuchung der unterschiedlichen sprachlichen Ebenen der linguistischen Analyse steht demnach aus. Vor allem die Bestimmung von sachdienlichen Faktoren bezüglich des Zusammenhangs zwischen Verstehen und Sprechen im Jugendalter und in sprachlichen Situationen, die dem Alltag nahekommen, ist in der Forschung unterrepräsentiert. In diesem Projekt sollte der Zusammenhang zwischen Sprachverarbeitung und Kognition teils explorativ und teils theoriegeleitet analysiert werden. Dafür wurden Freiburger SchülerInnen aus heterogenen sprachlichen und sozioökonomischen Hintergründen in einer Reihe von Experimenten getestet. Die erhobenen Daten wurden anhand von erweiterten statistischen Analysen auf unterschiedlichen sprachlichen Ebenen analysiert. Die Testung bestand aus einer Pilotstudie, die zur Optimierung der Experimente führte, und einer Hauptstudie, die hier im Detail beschrieben werden soll. Der vorliegende Beitrag dient der Vorstellung des Projekts, der Beschreibung der Stichprobe und der Methode sowie der Darstellung vorläufiger deskriptiver Teilergebnisse.

2 Beschreibung der Stichprobe und der Datenerhebung

2.1 Stichprobe

Insgesamt nahmen an der Hauptstudie 131 SchülerInnen der achten und neunten Klassen aus sechs Freiburger Schulen dreier Schularten (drei Werkreal-/Hauptschulen, eine Realschule, zwei Gymnasien) teil, sowie SchülerInnen unterschiedlicher Schularten aus einer Nachhilfschule. Von einer/m Versuchsteilnehmenden lag kein Fragebogen vor, weswegen die Daten aus weiterer Analyse ausgeschlossen wurden. Bei fünf anderen fehlten aufgrund von technischen Problemen einzelne Teiltests. Die finale Stichprobe bestand aus 130 Jugendlichen zwischen 12 und 17 Jahren (Durchschnittsalter 14.3 Jahre, SD .86, Median 14, 70 männlich). 43 SchülerInnen besuchten die achte und 87 die neunte Klasse. Die Verteilung der SchülerInnen auf den Schultyp ist der Tabelle 1 zu entnehmen. Für die Teilnahme haben Jugendliche einen Gutschein nach freier Wahl

erhalten (Buchhandlung, Kino oder Schwimmbad). Die Studie wurde von der Ethikkommission der Universität Freiburg genehmigt (202/16).

<i>Muttersprache/n-Gruppe</i>	<i>DreiSchularten</i>			<i>Total</i>
	Gymnasium	Haupt- /Werkrealschule	Realschule	
Einsprachige DaM	30 40.5 %	23 31.1 %	21 28.4 %	74 100 %
Mehrsprachige DaM	10 33.3 %	13 43.3 %	7 23.3 %	30 100 %
Mehrsprachige DaZ	2 7.7 %	17 65.4 %	7 26.9 %	26 100 %
Total	42 32.3 %	53 40.8 %	35 26.9 %	130 100 %

$$\chi^2=12.256 \cdot df=4 \cdot \text{Cramer's } V=0.217 \cdot p=0.016$$

Tabelle 1. Verteilung der SchülerInnen nach Muttersprache und Schulart (DaM = Deutsch als Muttersprache, DaZ = Deutsch als Zweitsprache).

Der sprachliche Hintergrund der SchülerInnen war heterogen und wurde anhand eines Fragebogens erfasst. SchülerInnen in der DaM-Gruppe (Deutsch als Muttersprache) gaben an, dass sie zu Hause Deutsch sprechen und Deutsch ihre einzige Muttersprache ist (mindestens ein Elternteil hatte ebenfalls Deutsch als Muttersprache). Einige wenige SchülerInnen in dieser Gruppe haben Grundkenntnisse einer anderen Sprache eines Elternteils angegeben. Bei einer genaueren Analyse der Fragebögen zeigte sich, dass bei vier der 78 SchülerInnen ausgiebige Kontakte zu einer weiteren Sprache im Elternhaus und Sprachkenntnisse in einer von einem Elternteil gesprochenen Sprache vorhanden waren, sodass eine Zuordnung zur mehrsprachigen Gruppe vorgenommen wurde. Die mehrsprachige DaM-Gruppe sprach zu Hause zusätzlich zu Deutsch noch eine weitere Sprache (dabei war entweder ein oder beide Elternteile fremdsprachig). Die DaZ-Gruppe (Deutsch als Zweitsprache) sprach zu Hause kein Deutsch, die deutsche Sprache wurde erst im Kindergarten oder beim Schuleintritt (außerhalb des Elternhauses) erworben. Der (nicht-deutsche) sprachliche Hintergrund der mehrsprachigen DaM-Gruppe sowie der DaZ-Gruppe war äußerst heterogen und umfasste beispielsweise die

Sprachen Albanisch, Arabisch, Chinesisch, Englisch, Französisch, Georgisch, Hindi, Kotokoli (Tem), Italienisch, Polnisch, Kurdisch, Romanes, Rumänisch, Russisch, Serbisch, Tamil, Twi.

So bestand die Stichprobe unter der Berücksichtigung der drei Sprachgruppen aus 74 SchülerInnen mit Deutsch als einziger Muttersprache (einsprachige DaM-Gruppe; 40 männlich), 30 SchülerInnen mit Deutsch sowie einer weiteren Muttersprache (mehrsprachige DaM-Gruppe; 15 männlich), und 26 SchülerInnen mit Deutsch als Zweitsprache (DaZ; 15 männlich).

2.2 Methoden und Design

Eine Experimentalsitzung bestand aus vier sprachlichen und zwei kognitiven Tests. Zusätzlich zu den Experimenten wurde ein ausführlicher Fragebogen erstellt, der wichtige Aspekte des sozioökonomischen Status als auch des sprachlichen Hintergrunds und der Persönlichkeitsmerkmale abdecken sollte. Dieser wurde vor der Experimentalsitzung ausgefüllt. Die folgenden Experimente und Tests wurden in dieser Reihenfolge (in Klammern die ungefähre Dauer) für eine Sitzung vorbereitet:

1. LexTALE (ca. 6 Minuten)
2. Stop signal Aufgabe (ca. 10 Minuten)
3. Bildbeschreibung mit Taboo Wörtern (ca. 10 Minuten)
4. Bildbenennung (ca. 8 Minuten)
5. Zahlenspanne rückwärts (ca. 8 Minuten)
6. Satzverstehen (ca. 25 Minuten)

LexTALE

Das passive Wortschatzwissen wurde anhand einer für unsere Zwecke angepassten Version des LexTALE Tests (Deutsche Version) erhoben (Lemhöfer & Broersma, 2012, www.lextale.com). Bei diesem Test handelt es sich um eine lexikalische Entscheidungsaufgabe, in der Teilnehmende Wörter hören und entscheiden müssen, ob es sich dabei jeweils um ein existierendes Wort des Deutschen handelt oder nicht. Ihre Antwort geben die Teilnehmenden durch die Wahl des Feldes ‚ja‘ oder ‚nein‘ auf dem Bildschirm. Es werden 40 existierende (*Wolle, staksig, monströs, Feist*) und 20 nicht existierende Wörter (*Termität, welstbar, plang, Mackel*) präsentiert. Der Grund für dieses

unausgewogene Verhältnis ist die kontrollierte Frequenz der 40 Wörter, von denen die Hälfte niedrige Gebrauchshäufigkeit aufweist. Die Wortlänge und Wortklassen variieren. Die nicht existierenden Wörter entsprechen den phonotaktischen und orthographischen Regeln des Deutschen und wurden durch das Vertauschen von Lauten oder Morphemen erstellt.

Um den Einfluss von unterschiedlichen Lesefähigkeiten zu vermeiden, wurden für die Zwecke dieses Projekts alle Items von einer standardnahen Sprecherin des Deutschen aufgenommen und den Jugendlichen auditiv über Lautsprecher (statt in schriftlicher Form) präsentiert. In der Praat Version des Tests (verfügbar unter <http://www.lextale.com>) wurden die schriftlichen Stimuli durch Audiodateien ersetzt. Nach der Instruktion, die auf dem Bildschirm präsentiert wurde, hörten die SchülerInnen jeweils ein Item und wählten ohne Zeitdruck jeweils eine Antwort aus. Der Test begann mit drei Übungsdurchgängen, die nicht in der Analyse berücksichtigt wurden. Die Reihenfolge der Items war für alle identisch. Für die Analyse wurde pro SchülerIn der Anteil korrekter Antworten anhand der Formel $(\text{Anzahl Wörter richtig}/40*100) + (\text{Anzahl Nonwörter richtig}/20*100) / 2$ berechnet (siehe Lemhöfer & Broersma, 2012).

Stop Signal Aufgabe

In der Stop-Signal Aufgabe werden anhand der Beurteilung geometrischer Formen inhibitorische Fähigkeiten getestet. Der Test wird mit Hilfe der Software STOP-IT (Verbruggen, Logan, & Stevens, 2008) durchgeführt. Während des Tests wird entweder ein Viereck oder ein Kreis auf dem Bildschirm gezeigt und die Aufgabe besteht darin, so schnell und korrekt wie möglich das Symbol durch eine entsprechende Tastaturtaste zu bestimmen (rote Taste für ein Viereck, blaue Taste für einen Kreis). Während der sogenannten Stop-Signal Durchgänge (25% aller Durchgänge) ertönt nach der Präsentation des Symbols ein Piepton (Stop-Signal, 750 Hz, 75 ms) und die TeilnehmerInnen sollen dann ihre motorische Reaktion (d.h. Tastendruck) unterdrücken. In der Instruktion wurde betont, dass man nicht abwarten sollte, ob ein Ton kommt, bevor man die Taste drückt. Man sollte stets so schnell wie möglich die richtige Taste drücken, außer wenn man einen Ton hört.

Jeder Durchgang beginnt mit einem Fixationszeichen (+) in der Mitte des Bildschirms (250 ms), das durch das jeweilige Symbol ersetzt wird. Die

TeilnehmerInnen haben 1250 ms Zeit für ihre Reaktion. Anschließend startet nach einem Intervall von 2000 ms der nächste Durchgang. Der Zeitpunkt der Präsentation des Stop Signals variiert (am Anfang immer 250 ms) je nach individuellen Antworten und Reaktionszeiten einer/s Teilnehmers/in, er verkürzt sich jeweils um 50 ms bei erfolgreicher Unterdrückung der Reaktion und verlängert sich um 50 ms bei erfolgloser Unterdrückung. Der Test beginnt mit einer Übungsphase (32 Durchgänge) und endet nach drei Blöcken mit je 64 Durchgängen. Für die Analyse wurden pro SchülerIn die Stop-Signal Reaktionszeit (SSRT) berücksichtigt. Diese wird durch die Subtraktion der durchschnittlichen Stop-Signal Verzögerung von der durchschnittlichen Reaktionszeit (für alle Durchgänge ohne Stop Signal) berechnet.

Bildbeschreibung mit Taboo Wörtern

In einer semispontanen Szenenbeschreibungsaufgabe untersuchten wir die Fähigkeiten im Bereich der Sprachproduktion und der verbalen Flexibilität. Wir nutzten das Taboo Paradigma und bauten dabei auf ein normiertes Projekt der Universität Aachen auf (Meffert et al., 2011), das wir für unsere Zwecke angepasst haben. Die Stimuli wurden uns vom AutorInnen-Team zur Verfügung gestellt. Dabei handelt es sich um schwarzweiße Zeichnungen komplexer Szenen aus dem täglichen Leben, auf welchen jeweils unterhalb und oberhalb der Szene zentrale Begriffe abgebildet werden, die für die Beschreibung nicht verwendet werden dürfen (sog. Taboo Wörter). Zum Beispiel sollten in einer Zirkusszene die Wörter „Clown“, „Elefant“, „Saxophone“, „Tiger“ und „Zirkus“ für die Beschreibung der Szene nicht verwendet werden. Es handelt sich dabei um Wörter, die in einer freien Beschreibungsaufgabe am häufigsten genannt wurden (siehe Meffert et al., 2011).

Wir haben drei Zeichnungen für die vorliegende Studie ausgewählt (Markt, Zirkus und Unfall), wobei nur die letzten zwei mit Taboo Wörtern präsentiert wurden. Pro Bild hatten die Jugendlichen jeweils zwei Minuten Zeit. Für den Übungsdurchgang wurde eine vierte Szene (Baustelle) gewählt, für die eine Minute zur Verfügung stand. SchülerInnen erhielten die Instruktion, dass es sich bei der Aufgabe um ein quasi-interaktives Spiel handelt, bei dem es jeweils zwei ähnliche Versionen einer Szene gibt. Das Bild sollte dabei so ausführlich wie möglich beschrieben werden, damit andere Mitspielende anhand der Beschreibung entscheiden können, um welche der beiden

Szenen es sich handelt. In der Übungsphase erhielten sie ein Beispiel von zwei Versionen einer Szene, in welchen die Unterschiede markiert wurden. Die Aufgabe war, in 30 Sekunden diese Unterschiede zu beschreiben. Nach der Übungsphase sahen die Jugendlichen ein neues Bild ohne Markierungen der Unterschiede. Sie wurden gebeten, die Szene so ausführlich wie möglich zu beschreiben und die 2 Minuten Zeit voll zu nutzen. Danach folgte jeweils eine Szene mit Taboo Wörtern und die Jugendlichen hatten wieder 2 Minuten Zeit pro Bild, sollten aber die Taboo Wörter nicht verwenden. Jeder Durchgang startete mit einer Instruktion, gefolgt von einem Fixationskreuz. Danach wurde die Szene für die Dauer von 2 Minuten eingeblendet. Das Experiment wurde mit DMDX (Forster & Forster, 2003, <http://www.u.arizona.edu/~kforster/dmdx/dmdx.htm>) programmiert. Für die Auswertung wurden alle Aufnahmen mit Praat annotiert.

Bildbenennung

Eine Bildbenennungsaufgabe wurde durchgeführt, um den aktiven Wortschatz und den lexikalischen Zugriff zu messen. SchülerInnen sahen auf dem Bildschirm 40 schwarz-weiße Bilder (Zeichnungen mit der Auflösung 300x300 Pixels). Die Zeichnungen wurden der online Datenbank *International Picture Naming Project* entnommen (Szekely et al. 2004, <https://crl.ucsd.edu/experiments/ipnp/1database.html>). Berücksichtigt wurden nur Bilder mit einer Rater-Übereinstimmung von 75% und höher (Variable *glex1* ≥ 0.75) und einer mittleren visuellen Komplexität (Variable *ovcjpg* 2+3 Quartil). Kontrolliert wurden die Gebrauchshäufigkeit (www.dlexdb.de) sowie die Größe der phonologischen Nachbarschaft (http://clearpond.northwestern.edu/clearpond_database.cgi). Aus den verbleibenden Items wurden 20 Samples zu je 40 Items erstellt. Jedes Sample enthielt 20 eher häufige Items und 20 eher seltene 20 Items (log. Typfrequenz in den oberen bzw. unteren 30% der verfügbaren Items). Aus den Samples wurde dasjenige ausgewählt, in dem sich die Items bezüglich weiterer Variablen am gleichmäßigsten verteilten.

Die Aufgabe bestand darin, das Bild so schnell und genau wie möglich zu benennen. Dabei erhielten die SchülerInnen die Instruktion, das Objekt ohne einen Artikel zu benennen (z.B. *Haus* statt *das/ein Haus*) und wenn möglich, Verzögerungen zu vermeiden. Das Experiment begann mit drei Übungsdurchgängen (mit den Abbildungen eines Kreuzes, Kreises, und Vierecks). Ein Durchgang begann mit einer kurzen Pause von 500 ms, darauf folgte ein Fixationskreuz (500 ms) und eine weitere kurze Pause (300

ms). Danach erschien das Bild zentriert auf dem Bildschirm. Über ein Mikrofon, das vor den SchülerInnen untergebracht war, wurden die Antworten auf dem Laptop aufgenommen. Das Experiment wurde mit DMDX programmiert und der *Voice Onset* (der Anfang der Objektbenennung) wurde mithilfe des *Voice Triggers* gemessen, der in der Software DMDX verfügbar ist.

Für die Datenanalyse wurden sowohl die Anzahl der richtig benannten Wörter als auch die Benennungszeiten berechnet. Die Benennungszeiten (*Voice Trigger*) wurden nach dem Experiment für jede/n ProbandIn anhand der CheckVocal Software (Protopapas, 2007) einzeln überprüft und, wo nötig, manuell angepasst. Die Antworten wurden von zwei RaterInnen annotiert.

Zahlenspanne rückwärts (Digit Span backwards)

Um das Arbeitsgedächtnis zu prüfen, wurde die Zahlenspanne rückwärts (Digit Span) mit der Software DMDX programmiert. Auf dem Bildschirm wurden im Abstand von einer Sekunde Abfolgen von Ziffern eingeblendet, welche im Anschluss in ein Feld am Computer in umgekehrter Abfolge eingetippt werden sollen. Die eingegebenen Zahlen erschienen beim Tippen nicht auf dem Bildschirm. Die Sequenzen bestanden aus zufällig generierten Abfolgen der Zahlen von 1-9 und wiederholten sich nie. Einige Anfangsabfolgen wie z.B. die Freiburger Postleitzahl und Vorwahl sowie die Beispielsequenzen aus der Instruktion wurden ausgeschlossen. Zudem wurden gleiche Ziffernfolgen oder Wiederholung von Ziffern (Bestimmung der maximalen Anzahl gleicher Ziffer: Sequenzlänge durch 3) innerhalb einer Sequenz vermieden. Es gab eine feste Randomisierung, die für alle ProbandInnen gleich war (nur bei technischen Problemen wurde auf alternative Randomisierungen zugegriffen). Für die Eingabe der Zifferabfolge gab es ein Time-out, das an die Längensequenz angepasst wurde (Berechnung: $4000 \text{ ms} + \text{Sequenzlänge} * 2000 \text{ ms}$). Bei drei Ziffern ergab sich somit ein Time-out von 10 Sekunden. Nach jeder Eingabe erhielten die Jugendlichen Feedback, ob die Ziffer korrekt waren und wie lang die nächste Zahlensequenz ausfallen wird. Nach einer richtigen Zifferwiedergabe erhöhte sich die Anzahl der Ziffer um eine Zahl. Bei einer inkorrekten Wiedergabe reduzierte sich die Anzahl der Ziffer im nächsten Durchgang um eine Zahl. Der Test startete mit einer Übungsphase und einer Sequenz von zwei Ziffern. Danach startete das eigentliche Experiment, das nach insgesamt drei

fehlerhaften Eingaben oder nach der Sequenz von 10 Ziffern endete. Das Experiment wurde für Sequenzen mit minimal zwei und maximal zehn Ziffern programmiert. Nach dem Experiment wurde die maximal erreichte Ziffer eingeblendet. Der maximale Digit Span Wert ergab sich aus der Zahl der längsten korrekt wiedergegebenen Abfolge.

Satzverstehen

Sprachverstehen unter erschwerten Hörbedingungen wurde anhand einer Satz wiederholungsaufgabe getestet. Die SchülerInnen hörten jeweils eine Äußerung und sahen ein Bild der angeblichen Sprecherin und ihre Aufgabe bestand darin, das Gehörte (ohne Zeitdruck) auf einen vorstrukturierten Antwortbogen aufzuschreiben. Im Anschluss an die Gesamtaufgabe haben SchülerInnen erneut jede Sprecherin gesehen und gehört und sollten unterschiedliche Bewertungen angeben (z.B. zur Verständlichkeit, Stärke des Akzents, soziale Klassenzugehörigkeit, kulturelle Identität, und zum ethnischen Hintergrund).

In der Studie wurden 36 kurze Äußerungen verwendet, die von einer deutschen Standardsprecherin, einer pfälzischen Dialektsprecherin und einer koreanischen Nicht-Muttersprachlerin der deutschen Sprache aufgezeichnet wurden. Die Aufzeichnungen wurden von Levy, Konieczny & Hanulíková (2019) zur Verfügung gestellt. Pro Experimentalliste gab es jeweils 12 Sätze pro Sprecherin. Die Sätze wurden kontrolliert und ausgeglichen bezüglich der Länge, der Wortklasse, der Wortfrequenz und der Semantik (für weitere Details siehe Levy et al., 2019). Um Erwartungseffekte zu reduzieren, waren alle Sätze semantisch bedeutungslos (z. B. *Ein Finger heiratet einen Kindergarten, Das Bild gibt dem Jungen einen Stuhl*).

Alle Sätze wurden für die Zwecke der vorliegenden Studie mit PRAAT (Boersma & Weenink 2012) bearbeitet und mit an die Sprachaufnahme angepasstem Hintergrundrauschen (*speech-shaped noise*) mit einem Signal-Rausch-Verhältnis von 0 dB versehen. Außerdem wurden pro Sprecherin durch die Anpassung der Grundfrequenz (F0) und der Formantendispersion jeweils zwei Stimmen erstellt (für eine ähnliche Vorgehensweise, siehe Hanulíková, 2018). Die Vorgehensweise war für eine theoriegeleitete Fragestellung zum Einfluss der SprecherInnenethnizität auf die Sprachverarbeitung nötig, und ermöglichte es, jeden Akzent mit jeweils zwei Bildern der angeblichen Sprecherin zu kombinieren. Die relevante Variable war dabei die

SprecherInnenethnizität (asiatisch oder weiß-europäisch). Somit konnte in einer Experimentalliste der koreanische Akzent in Stimme eins mit einem asiatischen Prime präsentiert werden und der koreanische Akzent in Stimme zwei mit einem weiß-europäischen Akzent. Bilder von SprecherInnen, die als visuelle Primes verwendet wurden, entnahmen wir der *Chicago Face Database* (Ma, Correll, & Wittenbrink, 2015). Ausgewählt wurden sechs Bilder von Frauen mit einem neutralen Gesichtsausdruck, die zur Hälfte der asiatischen und zur Hälfte der weiß-europäischen Ethnizität zugeordnet waren (siehe Abbildung 1). Die Frauen waren im Bezug auf Alter, Attraktivität und Vertrauenswürdigkeit vergleichbar (basierend auf Bewertungen, die in der Studie Ma et al., 2015 zur Verfügung stehen).

Für die Auswertung der Daten wurde der Anteil korrekt wiedergegebener Wörter berechnet. Abweichungen vom Zielwort wurden, wie in früheren Studien (z.B., Babel & Russell 2015; Levy et al., 2019), als korrekt eingestuft, wenn sie a) lediglich typographische Rechtschreibfehler beinhalteten (z.B., *Balon*, *Baloon*, *Ballong* für *Ballon*), b) Imitationen des Akzents darstellten (z.B., *Ent für Ente im Pfälzischen*), und c) morphosyntaktisch dem Kontext angepasst wurden (z.B., *Fahnen* für *Fahne*, *heirate* für *heiratet*).



Abbildung 1. Bilder, die als Primes in der Satzverstehensaufgabe verwendet wurden.

2.3 Datenerhebung

Die Datenerhebung bestand aus einer Pilotstudie und einer Hauptstudie. Die Pilotstudie diente der Optimierung der Hauptstudie. In den an einer Teilnahme interessierten Schulen wurden an die SchülerInnen zunächst Umschläge mit den Unterlagen (Informationsblatt, Fragebogen, Terminliste, Einverständniserklärung der Eltern) ausgeteilt. Alle SchülerInnen haben die Einverständniserklärung durch ihre Eltern (außer bei Volljährigen) unterschreiben lassen. Interessierte SchülerInnen gaben die ausgefüllten Unterlagen zurück und wurden dann zwecks Termin kontaktiert. Alle Tests wurden über

Laptops durchgeführt. Die Sprachaufnahmen wurden mittels eines Mikrofons (Monacor DM-500USB) direkt am Laptop gespeichert. Die auditiven Stimuli wurden über Lautsprecher (Inspire T 12) dargeboten. Alle Daten wurden anonymisiert gespeichert. Alle Jugendlichen wurden in einem frei verfügbaren Zimmer direkt in der Schule einzeln getestet. Sie erhielten Informationen zum Ablauf und erhielten vor jedem Test konkrete Instruktionen.

Pilotstudie

Vor der Durchführung der Hauptstudie wurde im Juni und Juli 2015 eine Pilotstudie an einer Gewerbeschule in Freiburg durchgeführt, die zur Optimierung des Designs und der Studie führte. Insgesamt nahmen 21 SchülerInnen im Alter von 16 bis 18 Jahren teil. Daten von zwei SchülerInnen konnten aufgrund technischer Probleme nicht berücksichtigt werden, sodass die finale Stichprobe 19 SchülerInnen (Durchschnittsalter 17.26; sieben Jungen) beinhaltete. Für die Teilnahme am Experiment erhielten alle eine Aufwandsentschädigung im Wert von 12 Euro. Das Experiment bestand aus den folgenden Aufgaben: Bildbenennung, Szenenbeschreibung, Grammatikalitätsurteile von Standard- und Dialektsätzen, LexTALE, Zahlenspanne rückwärts und Stop Signal. Die Auswertung der Daten ergab, dass die Aufgaben gut dafür geeignet waren, individuelle Unterschiede zu elizitieren. Eine Herausforderung bei der Zusammensetzung der Tests bestand darin, die Aufmerksamkeit der SchülerInnen durch langwierige Aufgaben nicht zu strapazieren. Da sich die Satzverarbeitungsaufgabe als zu lang erwiesen hatte, haben wir einen anderen Test zum Verstehen von akzentuierter Sprache mit einer zusätzlichen Fragestellung zur sozialen Kategorisierung der SprecherInnen gewählt. So wurde die Gesamtdauer auf ca. eine Stunde begrenzt, ohne an Aussagekraft der Teilerperimente zu verlieren. In diesem Zusammenhang war die Pilotphase äußerst konstruktiv.

Hauptstudie

Die Hauptdatenerhebung begann nach einem längeren administrativen Prozess im Juni 2016 und dauerte bis Ende 2017. Je nach Präferenz der Schulleitung und der Lehrkräfte wurden mit den an der Teilnahme interessierten SchülerInnen entweder direkt Termine außerhalb des Unterrichts vereinbart, oder den SchülerInnen wurde es ermöglicht, während der Unterrichtszeiten an der Studie teilzunehmen. Die Experimentalsitzungen fanden jeweils einzeln mit den SchülerInnen in einem Raum der Schule statt und

dauerten ca. 60 Minuten. Die Teilnehmenden erhielten als Aufwandsentschädigung einen Gutschein nach Wahl (Kino, Buchhandlung, Schwimmbad) im Wert von 10€.

3 Deskriptive Ergebnisse

Eine deskriptive Beschreibung der drei Sprachgruppen in Bezug auf ausgewählte Variablen aus dem Fragebogen und der Teiltests ist der Tabelle 2 zu entnehmen. Wie man sehen kann, verbrachten Jugendliche mit DaM deskriptiv etwas mehr Zeit mit Lesen als die anderen beiden Gruppen. Im Bezug auf den soziökonomischen Status und die passiven Wortschatzkenntnisse (LexTALE-Test) zeigen sich deskriptiv niedrigere Werte bei Jugendlichen mit DaZ in Vergleich zu den anderen beiden Gruppen.

	<i>Einsprachige mit DaM</i>			<i>Mehrsprachige mit DaM</i>			<i>Mehrsprachige mit DaZ</i>		
	<i>M</i>	<i>SD</i>	<i>R</i>	<i>M</i>	<i>SD</i>	<i>R</i>	<i>M</i>	<i>SD</i>	<i>R</i>
Alter	14.09	.81	12-16	14.52	.91	13-17	14.64	.81	13-16
Schulnote Deutsch	2.45	.65	1-4	2.77	.73	2-4	2.96	.54	2-4
Spaß am Deutsch	2.15	.90	1-4	2.03	.85	1-4	2.35	1.09	1-4
Lesen pro Woche	2.46	2.82	0-10	2.08	2.83	0-10	1.54	1.77	0-6
LexTALE	77.11	7.49	53-90	72.50	9.18	54-88	65.05	6.11	54-83
Sozioök. Status	.62	.17	.22-.96	.56	.17	.26-.92	.48	.21	.19-.81

Tabelle 2: Beschreibung der drei Sprachgruppen. Schulnote Deutsch (1-6, mit 1 als Bestnote), Spaß am Fach Deutsch (1 = ja; 2 = eher ja; 3 = eher nein; 4 = nein); Lesen pro Woche (0-10 Stunden pro Woche); M = Mittelwert; SD = Standardabweichung; R = Spannweite; DaM/DaZ = Deutsch als Muttersprache/Zweitsprache.

Vorläufige Ergebnisse im Hinblick auf den relevanten Forschungsstand zeigen einen positiven und schulunabhängigen Zusammenhang zwischen Verstehen und Sprechen, sowie zwischen Sprechen, Wortschatzkenntnissen und kognitiven Fähigkeiten. Wer also bei einer Szenenbeschreibung mehr erzählt, zeigt auch bessere Ergebnisse beim Verstehen (siehe Abbildung 2). Allerdings berücksichtigt diese Analyse lediglich die Gesamtanzahl der Substantive beim Erzählen. Dennoch konnten hier einige Mischfälle beobachten werden, also Personen, die viel und “gut” erzählen können aber beim Verstehen nicht gut abschneiden (der rechte untere Bereich der Abbildung 2), oder Personen, die zwar sehr gut beim Verstehen abschneiden, aber eine Szene nicht ausreichend beschreiben können (linker oberer Bereich der Abbildung 2). Um ein

differenzierteres Bild zu erhalten, stehen noch weiterführende Analysen bezüglich des Zusammenhangs zwischen Sprechen und Verstehen aus.

Bei den kognitiven Tests zeigen sich keine wesentlichen Gruppenunterschiede. Das Arbeitsgedächtnis war bei allen Gruppen vergleichbar gut, nur bei der kognitiven Kontrolle waren deskriptiv Bilinguale (DaM und DaZ) etwas besser als Monolinguale. Dieses Ergebnis ist aktuell, da in der Fachliteratur immer noch kontrovers diskutiert wird, unter welchen Umständen Bilingualismus zu kognitiven Vorteilen führen kann.

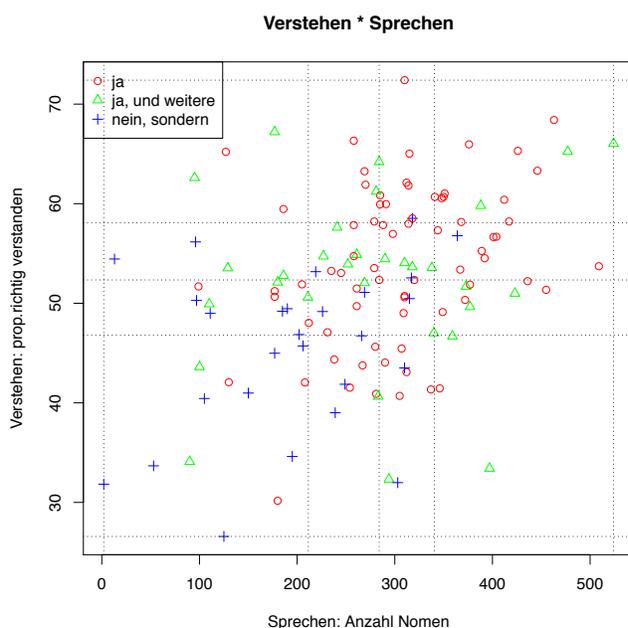


Abbildung 2. Der Zusammenhang zwischen Verstehen und Sprechen für die Gruppen DaM (ja), mehrsprachige DaM (ja, und weitere) und DaZ (nein, sondern).

Bei allen sprachlichen Tests zeigt sich, nicht überraschend, dass mit besseren Wortschatzkenntnissen bessere Ergebnisse erzielt werden. Auch der häufig beobachtete Einfluss von sozioökonomischen Faktoren (wie z.B. der Bildungshintergrund der Eltern) auf die Sprachentwicklung konnte repliziert werden: Bessere Wortschatzkenntnisse haben tendenziell diejenigen SchülerInnen, bei denen zu Hause mehr Bücher vorhanden sind (sprich in bildungsnahen Familien). Zwar schneiden im Durchschnitt Einsprachige deskriptiv besser ab als Mehrsprachige, und Mehrsprachige mit Deutsch als L1 besser als Mehrsprachige mit Deutsch als L2, allerdings gibt es innerhalb der Gruppen große Variation. So sind in unserem Sample nur insgesamt sieben Monolinguale, die bessere

Wortschatzkenntnisse aufweisen als die besten Bilingualen. Solide Wortschatzkenntnisse aber auch gutes Gedächtnis helfen beim Verstehen und Wiederholen von unbekanntem Akzent, was man in der Abbildung 3 sehen kann, in der signifikante positive Korrelationen zwischen Verstehen und Wortschatz (links) und Verstehen und Arbeitsgedächtnis (rechts) dargestellt sind.

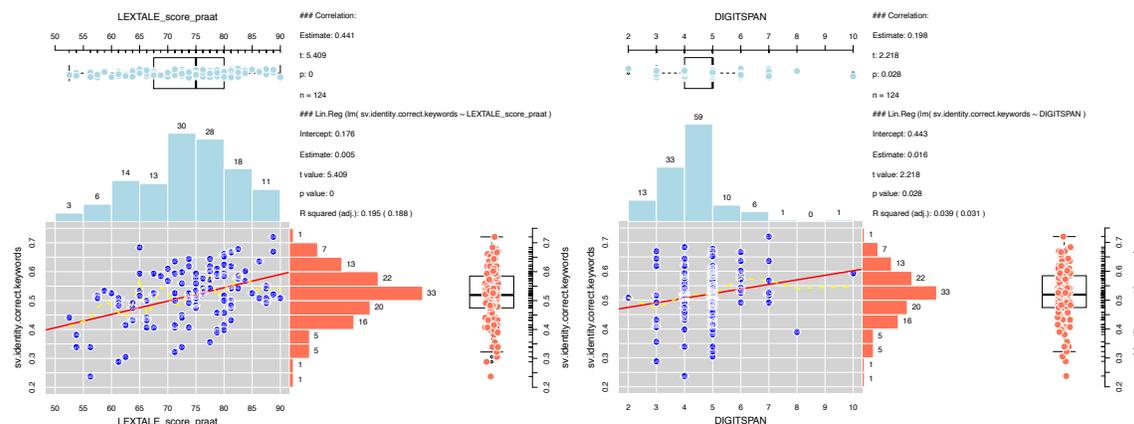


Abbildung 3. Der Zusammenhang zwischen Verstehen und Wortschatz (links) und Verstehen und Arbeitsgedächtnis (rechts).

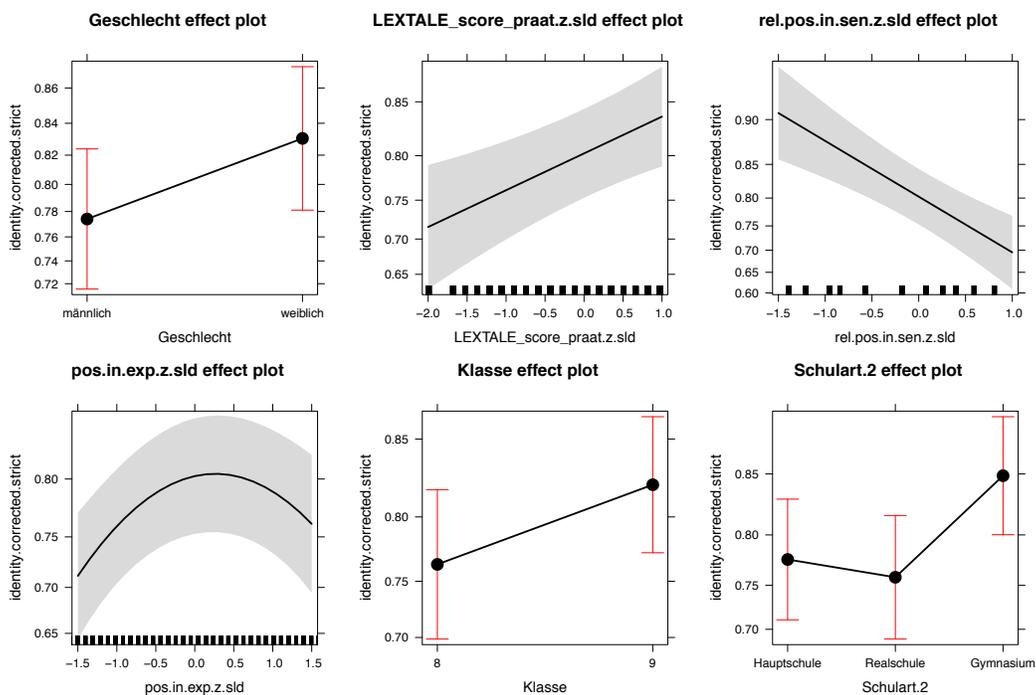


Abbildung 4. Effektplots für den Einfluss der folgenden Faktoren auf das Sprachverstehen: Geschlecht, passives Wortschatzwissen (LEXTALE), relative Position des Wortes im Satz (rel.pos.in.sen.z.sld), relative Position des Satzes im Experiment (pos.in.exp.z.sld), Klasse und Schulart. Die Fehlerbalken zeigen Konfidenzintervalle (95%).

Explorativ haben wir bei den Gesamtergebnissen auch immer wieder die Rolle weiterer Faktoren wie u.a. Schulart, Geschlecht oder Alter untersucht. Dies wird anhand der Abbildung 4 am Beispiel des Sprachverstehenstests illustriert. Man sieht im ersten Fenster, dass Mädchen im Verstehen besser als Jungen abschneiden. Das zweite Fenster rechts daneben zeigt, dass SchülerInnen mit besseren Wortschatzkenntnissen insgesamt auch besser verstehen. Daneben sieht man, dass man den Anfang eines Satzes besser versteht als das Ende des Satzes, was ein Nebeneffekt der semantisch unvorhersehbaren Sätze sein kann. In der zweiten Zeile sieht man im ersten Fenster, dass das Verstehen im Laufe des Experiments besser wird, zum Ende aber die Aufmerksamkeit und somit auch das Verstehen der Sätze wohl etwas abnimmt. Die nächsten zwei Fenster zeigen, dass Klasse 9 insgesamt besser abschneidet als Klasse 8, und dass GymnasiastInnen besser abschneiden als Haupt- und RealschülerInnen.

4 Ausblick

Die ersten Ergebnisse aus diesem Projekt wurden bereits bei zahlreichen nationalen aber auch internationalen Konferenzen in Großbritannien, Schweden und Canada vorgestellt. Aus den Teilergebnissen der Pilotstudie ging bereits 2019 eine Veröffentlichung zur Bewertung und Grammatikalität regionaler Syntax hervor (Hanulíková, 2019). Eine weitere Veröffentlichung zur Rechtschreibleistung von ein- und mehrsprachigen SchülerInnen im Vergleich erscheint demnächst als Buchkapitel (zusammen mit Caroline Ruppert). Eine Studie zum Sprachverstehen aus der Lebensspannenperspektive, die auf den Ergebnissen dieses Projekts sowie eines weiteren Folgeprojekts basiert, befindet sich unter Begutachtung. Drei weitere Aufsätze mit unterschiedlichen MitautorInnen zur Berechnung lexikalischer Diversität, zum Nomen-/Verb-Bias bei einsprachigen und mehrsprachigen SprecherInnen sowie zu Häsitationen und gefüllten und ungefüllten Pausen befinden sich in Vorbereitung.

Abgesehen von den Publikationen sind im Rahmen dieses Projekts zwölf theoriegeleitete Fragestellungen in Abschlussarbeiten bearbeitet worden und interessante explorative Fragestellungen in forschungsbasierten Seminaren beantwortet worden. So wurden zum Beispiel epistemische Marker ein- und mehrsprachiger SchülerInnen in den Szenenbeschreibungen analysiert, die Stellung der gefüllten Pausen innerhalb und

außerhalb von Äußerungen oder der Zusammenhang zwischen Ästhetik und Rechtschreibung untersucht.

Insgesamt lässt sich festhalten, dass das vorliegende Projekt primär der Grundlagenforschung galt. Es lassen sich daher nicht direkte Anwendungsaspekte für die Praxis ableiten. Korrelationen zwischen kognitiven und sprachlichen Fertigkeiten stellen zudem keine Kausalität dar. In Folgestudien könnte man stärker anwendungsbezogene Aspekte einbeziehen. Die große Heterogenität an deutschen Schulen lädt dazu ein, sich auch die Zusammensetzung der Schulen und Klassen genauer anzuschauen. Weiterhin aktuell ist nämlich die Frage, wie SchülerInnen aus benachteiligten Familien (unabhängig vom sprachlichen Hintergrund) effizient gefördert werden können. Diese Aufgabe kann weder den Eltern überlassen werden noch können LehrerInnen diese allein bewältigen. Sicherlich könnten mehr Lehrkräfte dabei helfen, gezielte Förderungsmaßnahmen anzubieten. Eine zusätzliche Möglichkeit wäre aber auch, Effekte der Lernumgebung und der Klassenzusammensetzung zu berücksichtigen und gezielt zu untersuchen. Konkret könnten quantitativ und qualitativ unterschiedliche Formen von Integrationsklassen untersucht werden. Zu erwarten ist, dass je höher der Anteil von sozial benachteiligten Kindern in einer Klasse, desto schlechter sind die Chancen bezüglich sprachlicher Entwicklung und schulischen Erfolgs im Vergleich zu besser gestellten Klassen. An den von uns getesteten Schulen und in unseren Ergebnissen haben wir beobachtet, dass die Gesamtergebnisse solcher Klassen bei sprachlichen Tests im unteren Bereich abschnitten. Die konkrete Frage wäre also, wie hoch der Anteil sein sollte, um deutliche Fortschritte bezüglich sprachlicher Fähigkeiten zu erwirken.

Literaturverzeichnis

- Babel, M., & Russell, J. (2015). Expectations and speech intelligibility. *The Journal of the Acoustical Society of America* 137(5), 2823–2833.
- Boersma, P., & Weenink, D. (2012). *Praat: Doing Phonetics by Computer* [Computer program]. Version 5.1.29, available from <http://www.praat.org/>.
- Dąbrowska, E. (2012). Different speakers, different grammars. Individual differences in native language attainment. *Linguistic Approaches to Bilingualism* 2/3, 219-253.
- Darcy, I., Park, H., & Yang, C. (2015). Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences* 40, 63-72.
- Davies, A. (2003). *The Native Speaker: Myth and Reality*. Clevedon: Multilingual Matters.
- Díaz, B., Baus, C., Escera, C., Costa, A., & Sebastián-Gallés, N. (2008). Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *PNAS* 105/42, 16083-1688.
- Forster, K.I., & Forster, J.C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers* 35, 116–124.
- Hanulíková, A. (2018). The effect of perceived ethnicity on text comprehension under clear and adverse listening conditions. *Linguistics Vanguard* 4/1. DOI: <https://doi.org/10.1515/lingvan-2017-0029>
- Hanulíková, A. (2019). Bewertung und Grammatikalität regionaler Syntax. Eine empirische Untersuchung zur Rolle der SprecherInnen und HörerInnen. *Linguistik Online*.
- Lemhöfer, K. & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods* 44, 325-343.
- Levy, H., Konieczny, L., Hanulíková, A. (2019). Processing of unfamiliar accents in monolingual and bilingual children: Effects of type and amount of accent experience. *Journal of Child Language* 46/2, 368-392.

- Ma, D.S., Correll, J., & Wittenbrink, B. (2015). The Chicago Face Database: A Free Stimulus Set of Faces and Norming Data. *Behavior Research Methods* 47(4), 1122-1135.
- Meffert, E., Tillmanns, E., Heim S., Jung, S., Huber, W., & Grande, M. (2011). Taboo: a novel paradigm to elicit aphasia-like trouble-indicating behaviour in normally speaking individuals. *Journal of Psycholinguistic Research* 40, 307-326.
- Protopapas, A. (2007). Check Vocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods* 39, 859–862.
- Sparks, R., & Ganschow, L. (1993). The effects of a multisensory structured language approach on the native language and foreign language aptitude skills of at-risk learners: A follow-up and replication study. *Annals of Dyslexia* 43, 194–216.
- Sparks, R., Ganschow, L., & Patton, J. (1995). Prediction of performance in first-year foreign language courses: Connections between native and foreign language learning. *Journal of Educational Psychology* 87, 638–655.
- Sparks, R., Ganschow, L., Patton, J., Artzer, M., Siebenhar, D., & Plageman, M. (1997). Prediction of foreign language proficiency. *Journal of Educational Psychology* 89, 549–561.
- Szekely, A., Jacobsen, T., D’Amico, S. et al. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language* 51/2, 247-250.
- Verbruggen, F., Loga, G. D., & Stevens, M. A. (2008). STOP-IT: Windows executable software for the stop-signal paradigm. *Behavior Research Methods* 40/2, 479-483.