

## FORSCHUNGSPROJEKT ProGram (2021-2023)

### Projektbeschreibung

<b>1 Projekttitel</b>	<b>ProGram</b> Professional Grammmaring in Austrian student teachers  <b>Leitung:</b> Univ.-Prof. Dr. Tom Rankin, JKU
<b>2a Projektleiter/in (PH intern)</b>	HS-Prof. Mag. Dr. Thomas Wagner, PHOÖ
<b>Welchem Fachbereich ist die/der Projektleiter/in zugeordnet?</b>	Sprachen-Literaturen-Kulturen
<b>2b Projektmitarbeiter/innen (PH intern)</b>	- Mag. Freya Aumaier, mitverwendet, Kollegium Aloisianum - Mag. Klaus Heissenberger, mitverwendet, Uni Wien
<b>2c Kooperationspartner/innen</b>	- Prof. Mag. Neil Stainthorpe, PHDL
<b>3a Forschungsschwerpunkt</b>	<input checked="" type="checkbox"/> Lehr- und Lernforschung <input type="checkbox"/> Schulentwicklung und Schulsystem <input type="checkbox"/> Lehrer/innenbildung und Professionalisierung <input type="checkbox"/> Lehrperson und Lehrer/innenbild <input type="checkbox"/> Inklusive Bildung/Pädagogik <input type="checkbox"/> Sonstiges:
<b>3b Sachgebiete</b>	6634 Kognitive Linguistik
siehe Anhang, Mehrfachantworten möglich	6624 Sprachunterrichtsforschung 6622 Psycholinguistik 6604 Angewandte Sprachwissenschaft 6605 Anglistik

#### 4 Abstract Deutsch

Sekundarstufenlehramtsstudierende im Unterrichtsfach Englisch müssen im Rahmen ihres Studiums sowohl die Eigenkompetenz (language proficiency) als auch ihr Metawissen über die Fremdsprache (pedagogical content knowledge) signifikant ausbauen. Am Ende des Lehramtsstudiums Englisch Cluster Mitte wird von ihnen erwartet, dass ihre Eigenkompetenz auf Niveau C2 liegt und dass sie gleichzeitig mit der Grammatik der englischen Sprache bestens vertraut sind und darüber verbal reflektieren und Auskunft geben können. Dies ist eine sehr spezifische Herausforderung und benötigt optimal darauf abgestimmte Studienarchitekturen und Lehrinhalte.

Aufbauen auf die poverty-of-stimulus-Theorie im Bereich von Lernaltersprachen untersucht das vorliegende Forschungsprojekt in diesem Zusammenhang den Aufbau von explizitem und implizitem grammatikalischen Wissen in Lehramtsstudierenden am Standort Linz. Dabei geht es einerseits darum, dass Lerner\*innen über instruierten und direkten Input hinaus und ohne direkt operationalisierbar linguistische Evidenzen Wissensbestände aufzubauen können, andererseits aber auch um die Parallelität im Aufbau expliziten und impliziten Wissens.

Ziel des vorliegenden Forschungsprojekts ist es daher, anhand der Grammatik der englischen Polaritätspartikel *any* zu zeigen, inwiefern deutschsprachige Lerner\*innen des Englischen auf B2+ Niveau explizites Grammatikwissen auf der Metaebene lernen und gleichzeitig implizites Sprachwissen erwerben, bzw. inwiefern sich die expliziten und impliziten Komponenten parallel entwickeln. Dazu

wird im vorliegenden Projekt eine Querschnittsstudie (Marsden, Whong, & Gil, 2018) repliziert. In der ursprünglichen Studie wurden implizites und explizites Wissen zur Partikel *any* bei arabischen Lerner\*innen auf drei verschiedenen Kompetenzstufen erhoben. Die Daten aus dieser Studie legen nahe, dass die arabischsprechenden Lerner\*innen ihr Grammatikwissen zu *any* einerseits aus instruierten Lernkontexten aufbauen, andererseits aber auch Feinheiten dieser Grammatik erworben haben, die weder instruiert noch aus sprachlichem Input ableitbar sind. Außerdem verläuft die Ausbildung der beiden Wissensbestände bei diesen Lerner\*innen dediziert nicht parallel.

Eine Replikation dieser Studie würde wichtige Rückschlüsse darüber zulassen, inwiefern Studierende anhand des derzeitigen sprachlichen und fachlichen Inputs aus den Modulen EB5 und EB6 im Curriculum Cluster Mitte sowohl ihre sprachliche Eigenkompetenz als auch, parallel dazu, ihr pedagogical content knowledge tatsächlich entwickeln. Dies wäre einerseits für die anstehende Curriculumsrevisionen zentral, andererseits aber auch sehr gut nutzbar in der Fortbildungen für Lehrpersonen aus der Sekundarstufe Englisch.

Schlüsselwörter:	<i>Explizites Wissen, implizites Wissen, Sprachbeherrschung, pedagogical content knowledge, poverty-of-stimulus, generative Grammatik, Curriculumsrevision</i>
------------------	--

#### 4 Abstract English

In the course of their study programme, secondary school teacher trainee students in Austria are supposed to build and develop both their personal language proficiency and their pedagogical content knowledge. When graduating, they are expected to reach C2 proficiency level and to have a profound knowledge of English grammar, enabling them to verbalise and reflect on this type of knowledge. This dual nature of language competence and pedagogical content knowledge is peculiar to such study programmes and thus necessitates a curriculum that is geared specifically towards such challenges. Developing pre-service teacher cognitions in useful ways has been shown to be potentially problematic, as degree programmes may not consistently be viewed as developing appropriate linguistic content knowledge (cf. Borg, 2006; Graus & Coppen, 2016, among others).

Theoretically grounded in the L2 poverty-of-stimulus approach, the present research project sets out to investigate the development of both explicit and implicit knowledge in Austrian teacher trainee students in Linz. In particular, the project will explore in how far learners build knowledge beyond instructed and observable linguistic evidence and in how far both explicit and implicit knowledge tend to develop in a parallel fashion.

Using the polarity item *any* as an under-researched test case, the present project, therefore, aims at exploring, in how far German-speaking B2+ learners learn explicit knowledge on a metalevel while simultaneously and parallelly acquiring and developing implicit knowledge. To this end, a cross-sectional study (Marsden, Whong, & Gil, 2018) will be replicated. The original study explored both explicit and implicit knowledge in Arabic-speaking learners of English on three different proficiency levels. Results from this study suggest that Arabic-speaking learners are able to derive, as expected, grammatical knowledge from instructed context, but quite surprisingly, also demonstrate detailed knowledge about aspects of this grammar which are neither instructed nor observable. Moreover, Arabic-speaking learners did not appear to build both types of knowledge parallelly.

A replication study would help understand, in how far the curricular modules EB5 and EB6 in fact support Cluster Mitte students in Linz in developing both their English L2 competence and their pedagogical content knowledge. On the one hand, such insights would be instrumental for the upcoming curricular revisions, on the other, they could be utilized for in-service teacher training courses.

Keywords (English):	<i>Explicit, implicit, language proficiency, pedagogical content knowledge, poverty-of-stimulus approach, generative grammar, curricular revision</i>
---------------------	---

## 5 Projektbeschreibung

### 5.1. Problemaufriss

Studierende des Lehramts Neu Englisch für Sekundarstufe müssen am Ende des Masters neben didaktischen Fähigkeiten das Kompetenzniveau C2 (CoE, 2001) und ein umfassendes Wissen über die englische Sprache und deren Grammatik vorweisen. Sowohl im Bereich des impliziten Sprachgebrauchs als auch im Bereich des expliziten Metawissens werden somit hohe Ansprüche gestellt, und eine professionelle Lehrer\*innenausbildung muss gewährleisten, dass diese Kompetenzstufen in der normalen Studienzeit sinnvoll erreicht werden können.

Grundsätzlich kann man davon ausgehen, dass grammatikalische Regelmäßigkeiten, die explizit in instruierten fremdsprachlichen Kontexten vermittelt bzw. aus sprachlichem Input induktiv erschlossen werden, auch größtenteils erworben werden (Hawkins & Franceschina, 2004; Montrul, Foote, & Perpiñan, 2008; White, Valenzuela, Kozłowska-Macgregor, & Leung, 2004). Gleichzeitig gibt es jedoch viele Hinweise darauf, dass erworbenes unbewusstes Grammatikwissen durch explizite Instruktion kaum beeinflussbar ist (Schwartz & Gubala-Ryzak, 1992); außerdem benötigt es zur induktiven Ableitung von Regelmäßigkeiten aus sprachlichem Input oft negative Evidenzen, welche aber in naturalistischen Kontexten natürlicherweise kaum vorkommen bzw. als Negativevidenz dort nicht sichtbar bzw. operationalisierbar sind. Diese Erkenntnisse führten von vielen Jahren zur sogenannten *no-interface* Position in der Debatte um das Verhältnis von erworbenem und erlerntem Sprachwissen (Schwartz, 1993). Gleichzeitig gibt es jedoch auch Evidenzen für eine *strong* bzw. *weak* Interface Positionen (Ellis, 2005), also ein variables aber systematisches Zusammenspiel beider Wissensbestände.

Der Aufbau von explizitem und implizitem sprachlichen Wissen ist somit seit Jahren Gegenstand intensiver Forschung (Andringa & Rebuschat, 2015). Wie entwickeln sich aber nun explizites und implizites Wissen über die Grammatik der englischen Sprache, wenn dieses Wissen nur zu einem Teil im instruierten Englischunterricht behandelt wird, ein Großteil davon jedoch weder in der Schule thematisiert wird noch aus naturalistischem Input induktiv ableitbar ist. Welche Rolle also spielt instruiertes fremdsprachliches Lernen in Bezug auf den systematischen Aufbau fremdsprachlichen expliziten als auch impliziten Wissens.

Neben psycholinguistischen und kognitiven Annäherungen an dieser Fragen widmet sich auch die generative linguistische Tradition diesem Aspekt des Fremdsprachenerwerbs (Rothman, 2008; Slabakova, 2002; Whong, Gil, & Marsden, 2013). Aus diesem Forschungsparadigma gibt es Studien, die nahelegen, dass selbst nicht beobachtbare Regelmäßigkeiten fremdsprachlich erworben werden können, besonders von fortgeschrittenen Lerner\*innen (Dekydtspotter & Hathorn, 2005; Kanno, 1998; Marsden, 2008, 2009), ein Phänomen, was auch als *L2 poverty-of-stimulus* bekannt geworden ist.

Aufbauend auf einer Studie aus dieser Tradition (Marsden, Whong, & Gil, 2018) und anhand eines konkreten Beispiels aus der englischen Grammatik (englische Polaritätspartikel *any*) soll in diesem Projekt der Aufbau des impliziten und expliziten Grammatikwissens bei Studierenden des Sekundarstufenlehramts Unterrichtsfach Englisch am Standort Linz untersucht werden. Die Polaritätspartikel *any* eignet sich deswegen gut für solche Untersuchungen, da diese Partikel einerseits in englischen Lehrbüchern der Unterstufe in Österreich (vgl. Gerngross et al., 2012) traditionell nur oberflächlich behandelt und vermittelt wird, andererseits die komplexen linguistischen Distributionsaspekte aus naturalistischem Sprachgebrauch nur schwer abgeleitet werden können (Chierchia, 2013; Giannakidou, 1998, 2001; Klima, 1964; Zwarts, 1995; Biber et al., 1999; Lin, 2015). Das heißt, ein substanzieller Teil des Wissens um die korrekte und idiomatische Verwendung von *any* kann von Lerner\*innen entweder nur induktiv aus sprachlichem Input abgeleitet werden oder ist schlichtweg nicht operationalisierbar. In Anlehnung an Originalstudie liegen folgende vier Leitfragen dem vorliegenden Forschungsprojekt zugrunde:

- (a) Haben Studierende Wissen zu nicht instruiert unterrichteten Regelmäßigkeiten, die im Sinne positiver Evidenz aus sprachlichem Input ableitbar wären?

- (b) Haben Studierende Wissen zu nicht instruiert unterrichteten Regelhaftigkeiten, die im Sinne positiver Evidenz nicht aus sprachlichem Input ableitbar wären?
- (c) Haben Studierende mehr Wissen über instruiert unterrichtete Regelhaftigkeiten im Vergleich zu nicht unterrichteten Regelhaftigkeiten?
- (d) Haben Studierende ihr explizites (bewusstes, verbalisierbares) und implizites (unbewusstes) Wissen über die korrekte und idiomatische Distribution der Polaritätspartikel *any* parallel entwickelt, oder gibt es substantielle Diskrepanzen?

Die Ergebnisse dieser Querschnittsstudie werden vor dem theoretischen Hintergrund der *poverty-of-stimulus* Debatte und dem theoretischen Modell des *Modular On-line Growth and Use of Language* (MOGUL, Sharwood Smith & Truscott, 2014a, 2014b; Truscott & Sharwood Smith, 2004) interpretiert.

## 5.2. Methoden

Die empirische Datenerhebung in diesem Projekt bedient sich eines paced / timed acceptability judgment tasks (AJT) sowie der gesteuerten Elizitation von sprachlichem Metawissen und ist als multifaktorielles, querschnittsartiges Kontrollgruppendesign angelegt. Einige Elemente des Forschungsdesigns von Marsden, Whong und Gil (2018) werden übernommen, andere werden neu angelegt, da es in der ursprünglichen Studie zu Unstimmigkeiten gekommen ist.

Erstens wollten die Autor\*innen eigentlich einen experimentellen Kontext untersuchen, in dem die Grammatik von *any* nicht instruiert unterrichtet wurde, als sprachlicher Input jedoch ableitbar gewesen wäre. Dieser Kontext wurde allerdings nie zur Erstellung und Auswertung von Testitems herangezogen. Zweitens scheint die Kreation der Testtokens nicht kohärent und auf die experimentellen Konditionen hin abgestimmt gewesen zu sein. Drittens ist das nicht ausbalancierte within-subject-Design der Originalstudie nicht unproblematisch, da alle Studierenden alle experimentellen Konditionen ausgesetzt waren (alle haben alle Sätze bewertet), was häufig carry-over bzw. Reihenfolgeeffekte mit sich bringt. Viertens waren die rating Kategorien in ihrer wörtlichen Ausformulierung unglücklich gewählt, da sie einen Fokus auf Korrektheit und Grammatikalität implizierten, wo eigentlich nach der Natürlichkeit der Konstruktion gefragt werden sollte. Fünftens wurden die unabhängigen Variablen eigentlich nie im Sinne eines multifaktoriellen Designs simultan, sondern eher sukzessive untersucht und berichtet. Sechstens legen aktuelle Studien nahe, dass auch rating Experimente mit Messwiederholungsdesign am besten mittels gemischter linearer Regressionsmodell anstatt traditioneller nicht-parametrischer Tests bzw. ANOVAs analysiert werden sollten (Myers, 2009; Bross, 2019a; 2019b). Und abschließend wurde die Erhebung der expliziten Kenntnisse zu *any* nicht wirklich mit den rating Ergebnissen in Beziehung gesetzt, ein Aspekt, der für die vorliegende Studie zentral sein wird. Daher sollen diese sechs Aspekte für die vorliegende Studie adaptiert werden.

Im Verlauf der ersten Phase des Forschungsprojekts wird entschieden, ob lediglich eine Gruppe von englischen L1-Sprecher\*innen als Kontrollgruppe in das Design mit aufgenommen und befragt wird, oder ob die explizite Vermittlung dieser Grammatik als between-groups Faktor hinzugezogen werden soll. In diesem Fall würden im WS 2021-2022 ca. 50% der Studierenden in der Lehrveranstaltung SeBEN05x02 „Advanced Grammar“ expliziten Unterricht zur Polaritätspartikel *any* bekommen, die andere Hälfte nicht; und sowohl das rating Experiment als auch die explizite Regelbeschreibung würden dann nach Gruppen untersucht.

### 5.2.1 Stichprobe

Daten werden voraussichtlich von einer Experimentalgruppe (Cluster Mitte Studierende Englisch aus 2-3 verschiedenen Kohorten,  $N = 100+$ ) sowie einer Kontrollgruppe (entweder englische Muttersprachler\*innen,  $N = 50+$  oder Kontrollgruppe ohne expliziten Unterricht) erhoben. Durch  $N > 100+$  und die Verwendung von Likert-Items sollte eine Trennschärfe / Power von ungefähr 80% gewährleistet sein (Sprouse & Almeida, 2012:21). Die Studierenden in der Experimentalgruppe haben alle das Kompetenzniveau B2+, welches durch ihren Schulabschluss und den Einstufungstest des Studienganges dokumentiert ist.

### 5.2.2 Experimentelle Konditionen und Stimuli

Es gibt insgesamt 4 positive oder idiomatische Verwendungskontexte, oder Satztypen, für *any*, quasi die ‚Grammatiken‘ für diese englische Polaritätspartikel. Außerdem gibt es im Sinne der poverty-of-stimulus Debatte zwei Arten der linguistischen Evidenz, nämlich dies sogenannte positive Evidenz (grammatikalische bzw. wohlgeformte Beispiele) sowie die negative Evidenz (ungrammatikalische bzw. nicht wohlgeformte Beispiele). Ersteres ist unmittelbar einleuchtend, da Lerner\*innen aus positiven Evidenzen Regelmäßigkeiten unmittelbar ableiten können. Negative Evidenzen sind interessant, weil sie dazu dienen könnten, Lerner\*innen auf die Begrenzung von Regelmäßigkeiten aufmerksam zu machen und durch induktive Regelableitungen Umfang und Limits der Regel zu verinnerlichen und Übergeneralisierungen zu vermeiden. Die vier Verwendungskontexte (1) bis (4) sowie die der Faktor der Wohlgeformtheit (a = wohlgeformt) und (b = nicht wohlgeformt) führen zu einem multifaktoriellen experimentalen Design und ergeben somit die folgenden acht experimentelle Konditionen.

	Grammatiken	Beispielsätze	Wohlgeformtheit	
			a	b
(1)	Fragesätze	<i>Do you want any cake?</i>	1a	1b
(2)	negierte Deklarativsätze	<i>Jenny doesn't want any cake.</i>	2a	2b
(3)	Satzkomplement negativer Verben	<i>Jenny denies that she ate any cake.</i>	3a	3b
(4)	Deklarativsätze negativer Adverbien	<i>Jenny hardly ate any cake.</i>	4a	4b

Dieses Design wird die Revision der Testitems aus Marsden, Whong und Gil (2018) vorgeben. Darüber hinaus gibt es noch einen weiteren Faktor, der nicht unmittelbar erhoben aber zur Interpretation der Daten herangezogen werden wird. Dieser Faktor beinhaltet die Möglichkeiten, als Lerner\*in mit den acht Konditionen in Berührung zu kommen, nämlich einerseits durch direkte Instruktion (I), wie im klassischen Englischunterricht (instruiert), oder andererseits durch positive Evidenz (II) im Sinne von beobachtbaren und operationalisierbaren linguistischen Evidenzen Sprachgebrauch (ableitbar). Dieser Faktoren hat also zwei Ausprägungen und lässt dadurch die folgenden 8 Deutungskontexte entstehen.

	Grammatiken	Beispielsätze	Erwerbskontexte	
			I instruiert	II ableitbar
(1)	(a) <i>any</i> in Fragesätzen	<i>Do you want any cake?</i>	+	+
	(b) <i>any</i> in affirmativen Deklarativsätzen	<i>Jenny wants any cake. *</i>	+/-	-
(2)	(a) <i>any</i> innerhalb negierter Sätze	<i>Jenny doesn't want any cake.</i>	+	+
	(b) <i>any</i> außerhalb negierter Sätze	<i>Anyone doesn't want any cake. *</i>	-	-
(3)	(a) <i>any</i> im Satzkomplement negativer Verben	<i>Jenny denies that she ate any cake.</i>	-	-
	(b) <i>any</i> im Satzkomplement positiver Verben	<i>Jenny thinks that she ate any cake. *</i>	-	-
(4)	(a) <i>any</i> mit negativen Adverbien	<i>Jenny hardly ate any cake.</i>	-	-
	(b) <i>any</i> mit positiven Adverbien	<i>Jenny probably ate any cake. *</i>	-	-

Die beiden Erwerbkontexte fließen entweder als hermeneutischer Interpretationshintergrund für die Daten ein oder werden explizit im Forschungsdesign durch Experimental- und Kontrollgruppe berücksichtigt. Der Argumentation von Marsden, Whong und Gil (2018) folgend wären die zwei Grammatiken (1a) und (2a) sowohl im instruierten Englischunterricht erlernbar als auch implizit aus sprachlichem Input ableitbar; (1b) stellt eine Mischform dar. Analysen von Lehrwerken und Korpusstudien zur Verwendung von *any* im Englischen legen demgegenüber nahe, dass die Grammatiken (2b) bis (4b) weder unterrichtet werden noch aus sprachlichem Input ableitbar sind. Obwohl diese Erwerbsszenarien theoretisch nicht besetzt sind, dienen sie theoretisch als potenzielle Erklärungs-szenarien.

Würde man nun für jede der acht experimentellen Konditionen zwei Testitems kreieren und durch ein Ablenkitem pro Kondition ergänzen, ergäben sich 24 Lexikalisierungstokens. Diese sind mit Beispielen in der folgenden Tabelle zusammengefasst:

Grammatik	Wohlgeformtheit	Instruierter Kontext.	Beobachtbarer Kontext	Nummer	Lexikalisierungstoken
Fragen	+	+	+	1.	Do you have any homework today?
	+	+	+	2.	Do you know anyone at that school?
	-	+	+	3.	I've heard any news about the campaign.
	-	+	+	4.	The policemen found anything at the crime scene.
Verneinung	+	+	+	5.	The teacher did not set any homework.
	+	+	+	6.	I don't know anyone at that school.
	-	-	-	7.	Anyone has not finished their homework.
	-	-	-	8.	Anyone did not follow the instructions.
Negatives Verb	+	-	-	9.	I'm sorry I said anything about your driving test.
	+	-	-	10.	I regret that I told anyone about our plans.
	-	-	-	11.	The journalist believed that he had taken any photos of the singer.
	-	-	-	12.	I suppose that you said anything to Sam last night.
Negatives Adverb	+	-	-	13.	James hardly ate anything at the party.
	+	-	-	14.	I seldom see anyone at the weekend.
	-	-	-	15.	Miss Jones probably heard anything at the staff meeting.
	-	-	-	16.	Bill probably wrote any notes during the lecture.
Ablenkitems	+			17.	I have never heard such beautiful music.
	+			18.	We will never see this place again.
	+			19.	Tourists often stay in this hotel.
	+			20.	Do you often play tennis in the summer?
	+			21.	I'm sorry that I missed your call.
	+			22.	Do you think that it will rain at the weekend?
	+			23.	I like watching movies.
	+			24.	Do you often buy train tickets online?

Bross (2019a; 2019b) und Myers (2009) folgend soll versucht werden, in der Endversion der Likert Items 4-6 Tokens für jede Kondition zu erstellen und diese im Sinne des Latin-Square-Designs so zu verteilen, dass jeder Teilnehmer / jede Teilnehmerin jede Kondition nur einmal präsentiert bekommt und sich die Testitems nicht wiederholen. Die Anzahl der neuen Testsätze wird abschließende proportional durch Ablenkitems ergänzt.

Für die empirischen Erhebungen gibt es zwei Messgrößen. Zunächst werden von jedem Studierenden *timed acceptability ratings* auf einer 4-stufigen Likert-Skala erhoben (*Ich bin mir sicher, dass das falsch ist, Ich glaube, dass das falsch ist, Ich glaube, dass das richtig ist, Ich bin mir sicher, dass das richtig ist*). Solche ratings haben sich als Erhebungsgröße für implizites sprachliches Wissen bewährt (Bowles, 2011; Ellis, 2005; Loewen, 2009). Im Gegensatz zur Originalstudie soll im vorliegenden Forschungsprojekt eine 4-stufige Likert-Skala ohne Ausweichoption zur Anwendung kommen, da die Möglichkeit, Unsicherheit beim Bewerten auszudrücken, bereits in der Likert-Skala eingebaut ist.

Danach folgen, wieder in Anlehnung an die Originalstudie, 2 Items zum expliziten Wissen über die Grammatik von *any* sowie 4 Items, welche die persönlichen Lernbiografien zu *any* bei jedem Studierenden in geschlossenen Fragen erheben.

### 5.2.3 Prozedur

Die Erhebungen sind für das Wintersemester 2021-2022 bzw. das Sommersemester 2022 geplant. Mittels mehrerer Onlinefragebögen in einer der gängigen Anwendungen (SoSciSurvey, WebExp; PsychoPy etc.) werden innerhalb dieses Semesters sowohl Kontextdaten als auch *paced acceptability ratings* sowie das Metawissen bei Studierenden und bei Muttersprachler\*innen erhoben. Aufgrund des Latin-Square-Designs werden vier Fragebogenvarianten erzeugt und dann für die statistische Analyse aggregiert.

### 5.2.4 Kodierung und Datenanalyse

Korrekte Antworten auf der 5-teiligen Likert-Skala werden numerisch wie folgt kodiert:

			<i>Ich bin mir sicher, dass das falsch ist</i>	<i>Ich glaube, dass das falsch ist</i>	<i>Ich glaube, dass das richtig ist</i>	<i>Ich bin mir sicher, dass das richtig ist</i>
	Grammatiken					
(5)	(a) Fragesätze G		-2	-1	1	2
	(b) Fragesätze U		2	1	-1	-2
(6)	(a) negierte Deklarativsätze G		-2	-1	1	2
	(b) negierte Deklarativsätze U		2	1	-1	-2
(7)	(a) negative Verben G		-2	-1	1	2
	(b) negative Verben U		2	1	-1	-2
(8)	(a) negative Adverbien G		-2	-1	1	2
	(b) negative Adverbien U		2	1	-1	-2

In einem ersten Analyseschritt werden die Antworten allesamt aufsummiert und als intervallskalierter Summenscore behandelt. Dieser Summenscore bildet die abhängige Variable und wird in einer mehrfaktoriellen ANOVA simultan auf *within-subject* Faktor Muttersprache sowie auf den *between-subject* Faktor Wohlgeformtheit hin modelliert (Haupt- und Interaktionseffekte), vor allem auch um die Vorgehensweise der Originalstudie nachzubilden. In einem zweiten Schritt wird dieser Summenscore aus den ratings mit dem Punktescore / der Kategorie aus der expliziten Regelerklärung in Beziehung gesetzt (mittels Korrelation und Regression). Die Regelerklärung wird dabei entweder in Schulnoten oder durch einen weiteren Summenscore operationalisiert, zum Beispiel aus Korrektheit, Genauigkeit und Verständlichkeit.

Im Gegensatz zur Auswertung in Marsden, Whong und Gil (2018) sollen die Daten aber nicht nur mittels *t*-Tests und ANOVAs, sondern zusätzlich mittels generalisierter linearer gemischter Regressionsmodelle analysiert werden (Bross, 2019; Cunnings, 2012; Gibson, Piantadosi & Fedorenko, 2011). Solche Modelle haben den Vorteil, dass sie nicht von Annahmen zur Verteilung abhängen,

orthogonale Designs ermöglichen und robust gegenüber random noise im Datensatz sind. Und mittels solcher Modelle kann man die höchstwahrscheinlich substanzielle Varianz, welche durch die Teilnehmer\*innen verursacht wird, als auch die Varianz, welche durch das repeated-measure Design und die verschiedenen Testitems verursacht werden, als sogenannte random effects modellieren und dadurch kontrollieren, sowohl hinsichtlich divergierender intercepts als auch slopes. Dies ist keinesfalls trivial, da divergierendes Antwortverhalten ein integraler Bestandteil solcher acceptability ratings ist und seit Jahrzehnten in der Literatur berichtet wird (Coward, 1997). Abschließend werden die Likert-Scores noch einer Itemanalyse unterzogen, um zu bestimmen, ob die verwendeten Testitems auch tatsächlich die acht Grammatiken hinreichend homogen abbilden.

Als Co-Variaten fließen eventuell noch das Geschlecht der Teilnehmer\*innen sowie die Englischnote des Schulzeugnisses bzw. der Punktescore des Einstufungstests der Teilnehmer\*innen mit in die Analysen ein.

### 5.3. Literatur

- Adringa, S., & Rebuschat, P. (2015). New directions in the study of implicit and explicit learning. *Studies in Second Language Acquisition*, 37, 185-196. doi:10.1017/S027226311500008X
- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., & Quirk, R. (1999). *Longman grammar of spoken and written English*. Harlow, UK: Pearson Education.
- Borg, S. (2006). *Teacher cognition and language education: Research and practice*. London: Continuum.
- Breiman L., Friedman J. H., Olshen R. A., & Stone, C. J. (1984). *Classification and regression trees*. Wadsworth.
- Bross, F. (2019a). *Acceptability ratings in linguistics: A practical guide to grammaticality judgments, data collection, and statistical analysis*. Version 1.02. Mimeo. Online: [www.fabianbross.de/acceptabilityratings.pdf](http://www.fabianbross.de/acceptabilityratings.pdf).
- Bross, F. (2019b). *Using mixed effect models to analyze acceptability rating data*. Version 1.0. Mimeo. Online: [www.fabianbross.de/mixedmodels.pdf](http://www.fabianbross.de/mixedmodels.pdf).
- Chierchia, G. (2013). *Logic in grammar: Polarity, free choice, and intervention*. Oxford, UK, and New York: Oxford University Press.
- Coward, W. (1997). *Experimental syntax. Applying objective methods to sentence judgments*. Thousand Oaks, London & New Delhi: Sage.
- Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research*, 28(3), 369–382.
- Dekydtspotter, L., & Hathorn, H. (2005). Quelque chose... de remarquable in English–French acquisition: Mandatory, informationally encapsulated computations in second language interpretation. *Second Language Research*, 21, 291–323.
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27, 305–352.
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27, 305-352.
- Giannakidou, A. (1998). *Polarity sensitivity as (non)veridical dependency*. Amsterdam, The Netherlands, and Philadelphia, PA: John Benjamins.
- Giannakidou, A. (2001). The meaning of free choice. *Linguistics and Philosophy*, 24, 659–735.
- Gibson, E. Piantadosi, S. & Fedorenko, K. (2011). Using Mechanical Turk to obtain and analyze English acceptability judgments. *Language and Linguistics Compass*, 5(8), 509–524.
- Graus, J., & Coppen, P. (2016). Student teacher beliefs on grammar instruction. *Language Teaching Research*, 20(5), 571-599.



- Hawkins, R., & Franceschina, F. (2004). Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish. In P. Prevost & J. Paradis (Eds.), *The acquisition of French in different contexts* (pp. 175–205). Amsterdam, The Netherlands, and Philadelphia, PA: John Benjamins.
- Hofmeister, P., Casasanto, L. S., & Sag, I. (2014). Processing effects in linguistic judgment data: (super-)additivity and reading span scores. *Language and Cognition*, 6(1), 111-145. Doi: <https://doi.org/10.1017/langcog.2013.7>
- Hothorn, T. Hornik, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*, 15(3), 651-674. 10.1198/106186006X133933
- Hulstijn, J. H. (2002). Towards a unified account of the representation, processing, and acquisition of L2 knowledge. *Second Language Research*, 18, 193-223
- Klima, E. S. (1964). Negation in English. In Jerry A. Fodor & Jerrold J. Katz (Eds.), *The structure of language, volume 3* (pp. 246–323). Englewood Cliffs, NJ: Prentice Hall.
- Leiner, D. J. (2019). *SoSci Survey* (Version 3.1.06) [Computer software]. Available at <https://www.sosicisurvey.de>
- Lin, J. (2015). *Acquiring negative polarity items*. Utrecht, The Netherlands: LOT.
- Marsden, H., Whong, M., & Gill, K. H. (2018). What's in the textbook and what's in the mind. Polarity item *any* in Learner English. *Studies in Second Language Acquisition*, 40, 91-118. doi:10.1017/S0272263117000018
- Montrul, S., Foote, R., & Perpiñan, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, 58, 503–553.
- Myers, J. (2009). The design and analysis of small-scale syntactic judgment experiments. *Lingua*, 19(3), 425-444.
- Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*, 51, 195-203. doi: 10.3758/s13428-018-01193-y
- Rothman, J. (2008). Aspect selection in adult L2 Spanish and the competing systems hypothesis when pedagogical and linguistic rules conflict. *Languages in Contrast*, 8, 74–106.
- Schwartz, B. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition*, 15, 147–163.
- Schwartz, B. D., & Gubala-Ryzak, M. (1992). Learnability and grammar reorganization in L2A: Against negative evidence causing the unlearning of verb movement. *Second Language Research*, 8, 1–38.
- Sharwood Smith, M., & Truscott, J. (2014b). *The multilingual mind: A modular processing perspective*. Cambridge, UK: Cambridge University Press.
- Slabakova, R. (2000). L1 transfer revisited: The L2 acquisition of telicity marking in English by Spanish and Bulgarian native speakers. *Linguistics*, 38, 739–770.
- Slabakova, R. (2002). The compounding parameter in Second Language Acquisition. *Studies in Second Language Acquisition*, 24, 507–540.
- Sprouse, J. & Almeida, D. (2012). *Power in acceptability judgment experiments and the reliability of data in syntax* (unpublished manuscript). University of California, Irvine & Michigan State University.
- Sprouse, J. & Almeida, D. (2017). Design sensitivity and statistical power in acceptability judgment experiments. *Glossa*, 2(1), 1-32. doi: <http://doi.org/10.5334/gjgl.236>
- Sprouse, J., & Almeida, D. (2017). Design sensitivity and statistical power in acceptability judgment experiments. *Glossa: A Journal of General Linguistics*, 2(1), 14. DOI: <http://doi.org/10.5334/gjgl.236>

- Truscott, J., & Sharwood Smith, M. S. (2004). Acquisition by processing: A modular perspective on language development. *Bilingualism Language and Cognition*, 7, 1–20.
- White, L., Valenzuela, E., Kozłowska-Macgregor, M., & Leung, Y. K. I. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics*, 25, 105–133.
- Whong, M., Gil, K.-H., & Marsden, H. (2013). *Universal Grammar and the second language classroom*. Dordrecht, The Netherlands: Springer.
- Zwarts, F. (1995). Nonveridical contexts. *Linguistic Analysis*, 25, 286–312.