

Disfluencies und Reparaturen bei Muttersprachlern und Lernern – eine kontrastive Analyse

Disfluencies and repairs in native and learner
speech – a contrastive analysis

Masterarbeit

zur Erlangung des akademischen Grades
Master of Arts (M. A.)
im Fach Linguistik

Humboldt-Universität zu Berlin
Philosophische Fakultät II
Institut für deutsche Sprache und Linguistik

Eingereicht von: Malte Belz

Erstgutachterin: Prof. Dr. Anke Lüdeling
Zweitgutachterin: Prof. Dr. Christine Mooshammer

Berlin, den 13.11.2013

tacendo loqui videbantur

Inhaltsverzeichnis

0	Einführung	1
1	Theoretischer Überblick	3
1.1	Disfluencies	3
1.2	Reparaturen	8
1.2.1	Schema	16
1.2.2	Subreparaturen	19
1.3	Statistische Hypothesen	21
2	Korpusbasierte Analyse	25
2.1	Methodik und Korpusdesign	25
2.2	Annotation	27
2.2.1	Vorhandene Annotation	28
2.2.2	Hinzugefügte Annotation	29
2.3	Verarbeitungspipeline	31
3	Auswertung	33
3.1	Ergebnisse	33
3.1.1	Interregna bei Muttersprachlern und Lernern	33
3.1.2	Gefüllte Pausen im Interregnum	36
3.1.3	Rolle in der Map Task	36
3.1.4	Subreparaturen	37
3.1.5	Limitationen und Mixed Models	40
3.2	Diskussion	43
3.2.1	Implikationen der Ergebnisse für die Erwartungen	44
3.2.2	Allgemeine Diskussion	51
3.3	Zusammenfassung und Ausblick	54
	Literatur	56
A	Graphen und Tabellen	62
B	Annotationsrichtlinien	73

Tabellenverzeichnis

1.1	Terminologische Variation bei der Beschreibung von Reparaturen	18
1.2	Vereinfachtes Subreparaturklassifikationsschema	20
2.1	Verdopplung der Annotationsebenen für verschachtelte Reparaturen	30
3.1	Übersicht der Datengrundlage sowie der Frequenzen für Reparaturen mit und ohne Interregnum	33
3.2	Lineares Modell für L1/L2 für gefüllte Pausen im Interregnum .	36
3.3	Ergebnis eines <i>Linear-Mixed-Effects</i> -Modells für die Wahrscheinlichkeit eines Interregnums	43
3.4	Korrelationen zwischen Tokenlänge, Silbenlänge und Äußerungsdauer in L1	43
3.5	Korrelationen zwischen Tokenlänge, Silbenlänge und Äußerungsdauer in L2	44
A.1	Frequenzen für muttersprachliche Reparaturen mit und ohne Interregnum	66
A.2	Frequenzen für lernersprachliche Reparaturen mit und ohne Interregnum	66
A.3	Frequenzen für mutter- und lernersprachliche Reparaturen mit gefüllter Pause im Interregnum	66
A.4	Frequenzen der L1- und L2-Reparaturen aufgeschlüsselt nach Interregnum und Rolle	67
A.5	Frequenzen der Subreparaturbeziehungen in Interaktion mit Wortarten bei L1- und L2-Reparaturen mit und ohne Interregnum	72

Abbildungsverzeichnis

2.1	Map-Task-Karten eines Instructors und eines Instructees	26
3.1	Interregnahäufigkeit bei L1-Reparaturen	34
3.2	Interregnahäufigkeit bei L2-Reparaturen	35
3.3	Rollenverteilung der Reparaturen für Fähigkeitsfaktor und Interregnumsausprägung	38
3.4	Verteilung der Subreparaturen für L1 und L2	39
3.5	Interaktion der Subreparaturen mit Wortartengruppe je Fähigkeitsniveau und Interregnumsausprägung	41
A.1	Interregnahäufigkeit bei L1- und L2-Reparaturen im Vergleich	62
A.2	Korrelationen zwischen den Tokenlängen von Reparandum und Reparans bei L1- und L2-Reparaturen	63
A.3	Korrelationen zwischen der Silbenanzahl von Reparandum und Reparans bei L1- und L2-Reparaturen	64
A.4	Korrelationen zwischen Äußerungsdauer von Reparandum und Reparans bei L1- und L2-Reparaturen	65
A.5	Rollenverteilung der Reparaturen für L1 und L2	67
A.6	Rollenverteilung der Reparaturen für Fähigkeitsfaktor und Interregnumsausprägung je Sprecher	68
A.7	Interaktion der Subreparaturen mit Wortart je Fähigkeitsniveau	69
A.8	Interaktion der Subreparaturen mit Wortart je Fähigkeitsniveau für einzelne Sprecher	70
A.9	Interaktion der Subreparaturen mit Wortart je Fähigkeitsniveau für einzelne Sprecher	71

Abkürzungsverzeichnis

ADJA	Attributives Adjektiv
ADJD	Prädikatives Adjektiv
ADV	Adverb
BeMaTaC	Berlin Map Task Corpus
DIP	Delayed Interruption for Planning Hypothesis
GER	Gemeinsamer Europäischer Referenzrahmen für Sprachen
H0	Nullhypothese
H1	Alternativhypothese
i	Einfügung als Subreparaturkategorie
IR	Interregnum
L1	Muttersprache
L2	Fremdsprache
MIR	Mit Interregnum
NE	Eigenname
NN	Nomen
OIR	Ohne Interregnum
r	Wiederholung als Subreparaturkategorie
RD	Reparandum
RS	Reparans
rs	Zusammengefasste Kategorien <i>r</i> und <i>s</i>
s	Substitution als Subreparaturkategorie
VMFIN	Finites Modalverb
VVFIN	Finites Vollverb
VVIMP	Imperativ
VVINFIN	Infinites Vollverb

0 Einführung

Unterscheiden sich fortgeschrittene Deutschlerner noch von Muttersprachlern des Deutschen, was ihre Reparatur- und Disfluency-Produktion betrifft? Dies ist die zentrale Fragestellung dieser Arbeit. Im Mittelteil von Reparaturen (bspw. *rechte äh linke Seite*) lassen sich oft Disfluencies wie *äh/ähm* oder stille Pausen finden. Reparaturen können aber auch direkt ohne Verzögerung ausgeführt werden. Ins Zentrum der Untersuchung rückt daher vor allem der interne Aufbau von Reparaturen: Unterscheiden sich Muttersprachler und Lerner in der Frequenz ihrer Reparaturen mit und ohne Disfluencies? Welche Teile der Reparatur werden wiederholt oder ersetzt, was wird neu eingefügt? Äußern Lerner ähnlich häufig dieselben Muster wie Muttersprachler? Welchen Einfluss haben Wiederholungen, Ersetzungen und Neueinfügungen auf die Existenz von Disfluencies innerhalb von Reparaturen? Welche Rolle spielen dabei Wortarten der offenen und geschlossenen Klassen?

Disfluencies wie ungefüllte oder gefüllte Pausen und auch Reparaturen sind typisch für gesprochene Sprache, sowohl für Äußerungen in der Muttersprache (L1) als auch für Äußerungen in der Ziel- bzw. Fremdsprache (L2). Durch eine kontrastive Analyse der Disfluencies und Reparaturen von Lernern und Muttersprachlern können sich Hinweise auf die für Lerner schwierigen Muster ergeben. Zudem lässt sich vergleichen, ob die Umstände und Bedingungen, welche zu Disfluencies führen, in beiden Gruppen ähnlich oder verschieden sind. Nur bei Lernern beobachtbare Reparaturmuster lassen beispielsweise Rückschlüsse auf die unterschiedlichen psycholinguistischen Prozesse innerhalb des L2-Sprachproduktionssystems zu, die noch von der L1-Sprachproduktion abweichen.

In einem quantitativ-qualitativen Ansatz wird mit korpuslinguistischen Methoden die Art der Annäherung der spontansprachlichen Merkmale der L2-Sprecher an die spontansprachlichen Merkmale der L1-Sprecher in einer gleichbleibenden Umgebung untersucht. Von spontansprachlicher muttersprachlicher Kompetenz würde man bei Lernern dann ausgehen, wenn die untersuchten Merkmale oder Muster von denen der Zielsprache nicht mehr signifikant abweichen. Um die oben genannten Fragestellungen zu beantworten, benötigt man möglichst in einem natürlichen Rahmen geäußerte und komparierbare spontane Sprachdaten. Die Auswertung erfolgt daher mit schon vorhandenen spontansprachlichen Daten aus dem Berlin Map Task Corpus, in denen jeweils

ein aktiver Dialogpartner einem passiven Dialogpartner einen Weg beschreibt. Aufgrund dieses Erhebungsdesigns wird auch untersucht, ob der eingenommenen Rolle im Gespräch ein signifikanter Einfluss auf die Disfluencies innerhalb von Reparaturen zugeschrieben werden kann.

Die Arbeit gliedert sich in drei Hauptteile. Zu Beginn wird ein Überblick über den relevanten wissenschaftlichen Diskurs für Disfluencies und Reparaturen gegeben, mithilfe dessen zu Ende des ersten Teils statistische Hypothesen über die Frequenzen und Muster in den Daten aufgestellt werden. Im zweiten Teil wird die korpuslinguistische Methodik zur Aufbereitung der Daten erläutert, bevor im dritten Teil die Auswertung vorgestellt wird. Dort werden zunächst die Ergebnisse der deskriptiven und analytischen Statistik präsentiert, bevor deren Implikationen für die Hypothesen diskutiert werden. Abschließend folgt eine allgemeinere Diskussion und eine Zusammenfassung der gewonnenen Erkenntnisse.

1 Theoretischer Überblick

1.1 Disfluencies

Sprachliche Produktion unterliegt unter anderem zeitlicher Linearität, Flüchtigkeit und konstantem Planungsdruck. Diese Bedingungen führen zu einem normalen Phänomen der gesprochenen Sprache: Disfluencies.¹

Speech disfluencies are generally defined as phenomena that interrupt the flow of speech and do not add propositional content to an utterance. This includes long pauses, repeated words or phrases, restarted sentences, and the fillers *uh* and *um* (Fox Tree 1995: 709).

Unter dem Terminus ‚Disfluencies‘ wird also eine ganze Reihe von Phänomenen gesprochener Sprache subsumiert, wie bspw. gefüllte Pausen (*äh*, *ähm*), ungefüllte Pausen, Wort- und Silbenabbrüche, Reparaturen mit beispielsweise Wiederholungen, Einfügungen, Substitutionen und Löschungen, aber auch selbstadressierte Fragen und Prolongationen (vgl. u. a. Eklund 2004; Hartsuiker & Notebaert 2010; Shriberg 1994). Reparaturen (im Sinne von Selbstreparaturen, vgl. Schegloff et al. (1977)) gliedern sich weiter auf in drei Teile – Reparaturandum (RD), Interregnum (IR) und Reparans (RS) –, wobei jeder Teil einer Reparatur wiederum Disfluencies enthalten kann. Allerdings ist dies oft auf das Interregnum beschränkt. Phänomene wie Wiederholungen, Einfügungen und Substitutionen werden als Spezifikation des Reparans betrachtet.

Beispiel 1 aus dem Berlin Map Task Corpus (BeMaTaC², vgl. Abschnitt 2.1), anhand dessen später auch die Forschungsfrage mit ihren Auspezifizierungen (vgl. Abschnitt 1.3) überprüft wird, soll diese reparaturspezifischen Begriffe kurz illustrieren, bevor das Reparaturschema genauer in Abschnitt 1.2.1 ausgeführt wird. Das Reparaturandum enthält den linken Teil, das Reparans den

¹Disfluencies existieren auch in der geschriebenen Sprache, dort sind sie aber nicht mehr nachvollziehbar. Sofern man Aphasien oder andere pathologischen Sprachstörungen wie bspw. Stottern beschreiben möchte, wird in der Literatur die Schreibweise mit ‚y‘ gewählt: *Dysfluencies*.

²BeMaTaC ist ein Korpus gesprochener Sprache und besteht aus Dialogen von Muttersprachlern und Lernern. Jeweils zwei Dialogpartner erklären sich dabei einen Weg durch eine Karte.

rechten Teil einer Reparatur. Im Interregnum können beispielsweise ungefüllte und gefüllte Pausen stehen. Dieses Kapitel betrachtet Disfluencies auch vor diesem Hintergrund.

	Reparandum	Interregnum	Reparans		
(1)	<i>nach links</i>	0,5s	<i>waagerecht</i>	<i>nach</i>	<i>links</i>
			Einfügung	Wiederholung	Wiederholung

Im Allgemeinen erreicht der Disfluency-Anteil in natürlicher Kommunikation ca. 5–10 % (Shriberg 2001; vgl. auch Bortfeld et al. 2001; Clark 1996; Levelt 1989; Shriberg 1994). Der oft zitierte Näherungswert für eine unterspezifizierte spontane Kommunikationsart liegt bei etwa sechs Disfluencies pro 100 Wörtern (Eklund 2004; Fox Tree 1995). Damit stellen Disfluencies ein hochfrequentes Phänomen dar.

Clark (1996) geht davon aus, dass es eine ‚ideal delivery‘ gibt – eine ‚ideale Auslieferung‘ bzw. Produktion einer Äußerung, die dann zum Ausdruck käme, wenn die performanztechnischen Umstände nicht wären. Zu diesen zählen das Entscheiden, was gesagt wird, das Formulieren und das Artikulieren; diese Prozesse stören die ideale Produktion (Clark 1996: 254). Die Abweichung von einer ‚idealen Auslieferung‘ wird oft als Kern einer Disfluency-Definition erachtet, ungeachtet der Tatsache, dass gesprochene Sprache zu einem großen Teil eben gerade nicht ideal geäußert wird, sondern vielen kommunikativen und kognitiven Zwängen unterliegt. Diesen Gedanken bringt folgendes Zitat auf den Punkt:

Why does one speak? It is certainly not to provide a demo of syntactically well-formed utterances. Rather, one speaks in order to communicate with one another effectively and accurately whatever one wishes. (O’Connell & Kowal 2005: 557)

Eine vollständige ‚Verflüssigung‘ gesprochener Sprache scheint realiter nicht erreichbar und weder natürlich noch (mit Rücksicht auf den Perzipienten) erstrebenswert zu sein. Gegen die ideale Auslieferung beziehungsweise eine konstante Flüssigkeit gesprochener Sprache sprechen mindestens vier Gründe. Schon die pulmonale Bedingtheit für die Sprachproduktion erschwert dieses Vorhaben: Sprecher müssen atmen. Auch wird die Verstehenskapazität von Hörern beschränkt durch die Dichte der Rede pro Zeiteinheit. Das Wechseln des Rederechts in Dialogen (*Turn-Taking*) gestattet keine Kontinuität. Viertens können Disfluencies auch gezielt eingesetzt werden, um bestimmte rhetorische Effekte bei Hörern zu erzielen (O’Connell & Kowal 2005: 557).

Einiges spricht also dafür, dass Disfluencies ein natürlicher Bestandteil von gesprochener Sprache sind. Sie kommen cross-linguistisch vor (vgl. beispielsweise Eklund 2004 für Schwedisch, Silber-Varod & Maruyama 2013 für Hebräisch und Japanisch und Deme & Markó 2013 für Ungarisch) und auch in monologischen Sequenzen wie bspw. Narrativa (Tavakoli 2010). Aufgrund dieses ubiquitären Vorkommens kann man schließen, dass auch Lerner einer Fremdsprache Disfluencies produzieren und sich ab einem gewissen Niveau hinsichtlich ihrer Disfluency-Produktion auch muttersprachlich verhalten sollten. Obwohl in der Literatur meines Wissens keine umfassenden Niveaustudien diesbezüglich erfasst sind, kann solch eine hohe Kompetenz wohl erst ab Niveau C1–C2 des GER³ erwartet werden. Dabei sollte allerdings nicht vergessen werden, dass Sprecher eine starke Neigung besitzen, ihr Pausenprofil aus der Erstsprache in die Zweitsprache zu übernehmen (Raupach 1980: 270), was prinzipiell für alle Disfluencies gelten sollte. Dies für Reparaturen als eine im Allgemeinen unter die Typologie der Disfluencies subsumierte Gattung zu überprüfen, emergiert dabei als interessantes Desiderat. Dass dabei der Sprachstand der Lerner auf Vergleichbarkeit und zur besseren Einordnung der Ergebnisse kontrolliert werden muss, versteht sich von selbst (vgl. Watanabe & Rose 2012: 482).

Neben den somatischen Auslösern für Disfluencies gibt es unzählige weitere bedingende Faktoren und Funktionen von Disfluencies (vgl. Eklund 2004; Corley & Stewart 2008). Hierunter fallen insbesondere zwei Faktoren, die eine bedeutende Rolle für die Produktion von Disfluencies spielen: die *Lexical-Retrieval*-Hypothese und die *Floor-Holding*-Hypothese.

Verzögerungen im Sprachfluss, wie sie Pausen darstellen, korrelieren sehr wahrscheinlich mit der Unsicherheit darüber, was man als nächstes sagen möchte (Corley & Stewart 2008: 6). Gefüllte und ungefüllte Häsitationspausen entstehen besonders häufig vor dem Abruf von Inhaltswörtern aus dem mentalen Lexikon (Hartsuiker & Notebaert 2010; Maclay & Osgood 1959; Tannenbaum et al. 1965). Eine dergestalt großes mentales Lexikon sorgt demnach für mehr kognitiven Aufwand und lässt die Sprache weniger flüssig werden, was bedeutet, dass mehr Disfluencies produziert werden sollten. Indizien hierfür lassen sich aus der Studie von Schachter et al. (1991) zu Disfluencies in verschiedenen wissenschaftlichen Disziplinen herauslesen.

Schachter et al. (1991) zeigen in ihrer Untersuchung, dass Dozenten an der Columbia-Universität in den Natur-, Sozial- und Geisteswissenschaften sich signifikant in der Anzahl der produzierten gefüllten Pausen unterscheiden. Während individuelle Effekte ausgeschlossen werden konnten, ergab eine stichprobenweise Auszählung der Wörter pro Disfluency, dass „[t]he greater the num-

³Gemeinsamer Europäischer Referenzrahmen für Sprachen (<http://www.europaeischer-referenzrahmen.de/>).

ber of different words, the greater the number of *uhs* and *ahs*“ (Schachter et al. 1991: 366). Der Effekt scheint also von der Synonymdichte der einzelnen Disziplinen abzuhängen. Für eine kontrastive Analyse mag es demnach sinnvoll erscheinen, das Register als potenziell konfundierenden Faktor konstant zu halten. Diese Beschränkung wiederum motiviert das Heranziehen zweier eben daraufhin ausgelegter gleich konzipierter Korpora (vgl. Abschnitt 2.1 auf Seite 25).

In einem mit Wortarten versehenen Korpus ließe sich, wie später in Abschnitt 3.1.4 auch demonstriert wird, mittels einer der Einfachheit halber eher groben Einteilung von Wortklassen überprüfen, ob beispielsweise das Vorhandensein eines Inhaltswortes überzufällig häufig eine Disfluency im Antezedens entstehen lässt. Ein typisches Muster, welches hierfür infrage kommt, sind Reparaturen, anhand welcher man „Lernersprache auf [...] Strategien der Selbstkorrektur sowie Neu- und Umformulierungen“ hin genauer untersuchen könnte (Götz 2007: 82). Als Antezedentia würde man dann die Interregna beziehungsweise deren Inhalt werten, welche vor dem Reparans auftreten.

Obwohl in letzter Zeit intensiv über Disfluencies geforscht wurde, ist eine Disambiguierung der *Floor-Holding*-Hypothese nicht gelungen. Unter dieser versteht man die oftmals ambige Eigenschaft einer gefüllten Pause, sowohl eine Häsitiation als Planungssignal ausdrücken zu können als auch mittels einer intendierten Vokalisation von Stille konversationsanalytisch das Rederecht (*Turn*) beizubehalten zu versuchen (Maclay & Osgood 1959). Die *Floor-Holding*-Hypothese als ein häufiges Argument für die Erklärung von Disfluencies wird für meine Untersuchung demnach ausgeblendet⁴, kann für konversationsanalytische oder pragmatische Studien mit diesem Korpus aber durchaus von Interesse sein.

Es ist schon länger bekannt, dass Disfluencies bewusst gesteuert und unterdrückt werden können, was Siegel et al. (1969) erstmals experimentell nachwies. In der Debatte um die mehrdeutige Funktion von Pausen zwischen strategischem Einsatz und dem Anzeigen kognitiver Planungsschwierigkeiten hat Fischer (1999) vorgeschlagen, Pausen als ikonisch zu betrachten. „Pausen haben somit eine Bedeutung, die mit ‚I am thinking‘ umschrieben werden könnte“ (ebd.: 253). Solch ein pragmatisch-holistischer Ansatz stellt für diese Studie einen guten Kompromiss dar. Disfluencies können also unter Umständen aktiv in die Kommunikation eingebunden werden. Meistens entstehen sie jedoch ohne allzu großes Bewusstsein in bestimmten Kontexten.

⁴Dies bedeutet, dass solche Disfluencies auch nicht in die Analyse der Daten mit eingehen dürfen, was durch dementsprechende Annotationen (Abschnitt 2.2) und Annotationsrichtlinien (vgl. Anhang B) gewährleistet wird.

Dass Disfluencies so vielfältig von Sprechern verwendet werden können, motiviert nochmals eine genauere Analyse im kontrastiven Vergleich von Muttersprachlern und Lernern. Solch ein cross-linguistischer kontrastiver Ansatz mit L1- und L2-Korpora kann beispielsweise zu einem besseren Verständnis der Probleme während des Lernens und Sprechens von L2 beitragen (Dechert 1980: 272). Die Identifikation von Gemeinsamkeiten und Unterschieden zwischen L1- und L2-Disfluencies kann zu einem besseren Verständnis der Sprachproduktion im Allgemeinen und der L2-Sprachproduktion im Besonderen führen (vgl. Rieger 2003a: 41). Dieses Forschungsfeld ist – insbesondere für das Deutsche – bisher noch wenig ausgeleuchtet worden. Interessant ist insbesondere, inwiefern sich fortgeschrittene Lerner hinsichtlich der Verwendung ihrer Disfluencies noch von Muttersprachlern unterscheiden.

[T]here is little comparative research evidence to demonstrate in what respect(s) L2 learner's fluency is different from that of native speakers (Tavakoli 2010: 72).

Tavakoli untersucht, ob es Unterschiede bei monologischen Bildnacherzählungen in den Pausenmustern von englischen Muttersprachlern und Lernern des Englischen auf B2-Niveau gibt. Sie vergleicht dazu die Anzahl und Stelle ihrer Pausen sowie ihre Gesamtfrequenzen.⁵ Die Lerner und Muttersprachler des Englischen unterscheiden sich in Tavakolis Studie nicht signifikant in der Anzahl oder Länge der ungefüllten Pausen, sondern vor allem bezüglich ihrer Position. So pausieren die Lerner nur in Mittsatzpositionen statistisch häufiger als Muttersprachler Tavakoli (ebd.: 77). Wie reliabel diese Erkenntnis ist, mag dahingestellt sein, denn es existieren keine nachvollziehbaren Richtlinien, wie genau eine ‚Mittsatzposition‘ in ihrer Studie definiert wird. Gerade in der gesprochenen Sprache können klassische Satzdefinitionen in die Irre führen. Tavakoli (ebd.: 75f.) beobachtet auch, dass Lerner „oft“ (sic – der statistische Zusammenhang bleibt unklar) pausieren, bevor sie ein Wort wiederholen oder wenn sie ein Wort oder einen Ausdruck substituieren. Ihre (nicht statistisch nachgeprüfte) Beobachtung bezieht sich also auf das Phänomen einer reparaturinternen Pause im Interregnum, auf die im Reparans eine Wiederholung oder Substitution folgt. Dieses Phänomen könnte mutatis mutandis (B2-Niveau bei Tavakoli vs. C1-Niveau in BeMaTaC) auch in meinen Daten auftreten, wobei zu diskutieren bliebe, welcher Veränderung im Versuchsdesign ein gefundenes Ergebnis zuzuschreiben wäre bzw. wie reliabel dies angesichts der beschriebenen Unzulänglichkeiten sein kann.

⁵Dabei verwendet sie einen Cut-Off von 400 ms für stille Pausen, was im Lichte neuerer Forschung kritisiert werden kann. So empfehlen De Jong & Bosker (vgl. 2013) nach einer umfassenden Pausenuntersuchung, den oft verwendeten Cut-Off von 250 ms für eine Häsitationspause auch weiterhin zu verwenden.

Sprecher einer Zweitsprache unterscheiden sich von Muttersprachlern hinsichtlich ihrer Sprachproduktion in Grammatik, Wortschatz und Automatisierung. Zu Beginn des Zweitspracherwerbs mag dies trivial erscheinen. Betrachtet man jedoch den fortschreitenden Kenntniserwerb der Zweitsprache in allen linguistischen Domänen, so verwundert es, dass selbst fast bilinguale Sprecher von Muttersprachlern als Nicht-Muttersprachler identifiziert werden können, obwohl Syntax und Wortschatz auf C2-Niveau des GER angelangt sind. Dies kann unter anderem mithilfe von Disfluencies erklärt werden und wurde beispielsweise beschrieben von Eklund (2004: 146) („disfluencies will reveal a foreign origin“) und Götz (2007: 70) („*foreign-soundingness* von gesprochener Lernaltersprache“).

Diese Untersuchung beleuchtet nur einen kleinen Teil der Disfluency-Produktion näher. Jedoch lassen sich aus dieser Detailnähe durchaus Schlüsse für die Ausprägung des L2-Sprachproduktionssystems ziehen. Es wurde gezeigt, dass fortgeschrittene deutsche Lerner des Englischen sich unterscheiden: 80 % benutzen signifikant mehr gefüllte Pausen als der Durchschnitt der Muttersprachler (Brand & Götz 2013: 124). Für das Deutsche sollte für die Lernaltersprache daher auch ein erhöhtes Disfluency-Vorkommen bei Lernern zu erwarten sein, was in den Hypothesen 1.1–1.3 überprüft werden wird (Abschnitt 1.3 auf Seite 21). In ihrer Analyse der zeitlichen Variablen der Sprechflüssigkeit wie bspw. Sprechrate in Interaktion mit den lernaltersprachlichen Fehlern konnten Brand & Götz keine klaren Trends erkennen. Sie wünschen sich aber eine multivariate Analyse der Sprachproduktion, um auf diese Weise explorativ mögliche Korrelationen sichtbar zu machen (ebd.: 133f.). Auf diese Arbeit lässt sich dieser Gedanke zwar nicht im gleichen Sinne übertragen – Reparaturen sind nicht immer Fehler und werden auch nicht so annotiert (vgl. Abschnitt 2.2) –, jedoch motiviert dies die Extraktion und den Korrelationstest einiger zeitlichen Messungen der untersuchten Reparaturen, wie Silbenanzahl, Äußerungsdauer und Tokenanzahl.

1.2 Reparaturen

[R]epairs differ from the other types of disfluencies, and in a way they merit—and require—a study of their own (Eklund 2004: 266).

Reparaturen rückten über die linguistische Konversationsanalyse in den Fokus der Forschung. Schegloff et al. (1977) fassen den Begriff einer Äußerungskorrektur erstmals weiter und führen den Terminus der (fremd- oder selbstinitiierten) Selbstreparatur ein, welcher der (fremd- oder selbstinitiierten) Fremd-

reparatur⁶ gegenübergestellt wird. Sie genügen damit dem Anspruch, verschiedene Subtypen von Selbstreparaturen in die Erklärung des Phänomens mit einzubeziehen, die über die reine Korrektur von Fehlern hinausgehen, wie das Tip-of-the-tongue-Phänomen, Häitationen und explizite Editierungen. Psycholinguistisch lassen sich aus Produktionsfehlern wie den Selbstreparaturen Schlüsse für die Sprachproduktion ziehen, da sie Evidenz für nicht direkt beobachtbare Prozesse liefern. Historisch wurden dazu vor allem Versprecher herangezogen (Fromkin 1971). Der Begriff ‚Reparatur‘ wird in dieser Arbeit auf die selbstinitiierten Selbstreparaturen beschränkt, weil diese Kategorie sowohl frequenter als die der Fremdreparatur ist (Colman & Healey 2011) als auch Einsichten in die selbstgesteuerten Strukturen genau eines Sprechers gibt.

Es ist nicht immer leicht, Reparaturen von anderen Phänomenen der gesprochenen Sprache zu unterscheiden. So wird in Beispiel 2 keine Reparatur, sondern ein Neustart geäußert (*denn du triffst ja wieder...*). Beispiel 3 zeigt eine mehrmalige Wiederholung, die dabei auch grammatisch verändert wird. Dass hier eine Disfluency bzw. eine Art von Reparatur vorliegt, ist intuitiv klar – sie genau zu fassen, ist dennoch schwierig. Es handelt sich um drei Wiederholungen der Präposition und des Artikels, wobei nach der ersten modifizierenden Reparatur (*der* zu *dem*) die modifizierte Phrase nochmals mit dem Nomen wiederholt wird. Verschachtelte Reparaturen werde ich aber auch als Reparaturen, was bedeutet, dass man dieses Phänomen in einer Annotation berücksichtigen muss, um die Frequenzen nicht zu verfälschen.

- (2) ziehst du einen Kreis [...] um den Arzt 0,5⁷ herum 0,8 bis du dann denn du triffst ja wieder automatisch auf deine hergekommene Linie (L1:2011-12-14-B)⁸
- (3) gehst du [...] also aus der aus dem aus dem Fünfundreißiggradwinkel sozusagen ne grade Linie (L1:2011-12-14-B)

Ob Wiederholungen (Repetitionen) eine Reparatur darstellen, ist umstritten und ist abhängig von Semantik und Intonation der Wiederholung. Wenn weder der Inhalt des Reparans vom Reparandum abweicht, noch F0 besonders betont wird oder ein Kontrastfokus erkennbar ist (vgl. Brennan & Shober (2001: 293)), kann kaum von einer Reparatur die Rede sein. Somit ist Bsp. 4 auf der nächsten Seite keine Reparatur, Bsp. 5 auf der nächsten Seite aufgrund des hörbaren Kontrastakzents hingegen schon. Bei Shriberg (1994: 10) ist jedoch

⁶Eine Fremdreparatur wird nicht vom Sprecher selbst, sondern von jemand anderem getätigt.

⁷Stille Pause in Sekunden.

⁸Die Zitation der Korpusdaten erfolgt nach dem Schema ⟨L1 bzw. L2⟩⟨:⟩⟨Gesprächs-Identifikationsnummer⟩ in der von mir erweiterten Version von BeMaTaC.

eindeutig von einer exakten Wiederholung als Reparatur die Rede. Engelhardt et al. (2010) ziehen aus ihren psycholinguistischen Experimenten den Schluss, Wiederholungen und gefüllte Pausen von Reparaturen zu unterscheiden. In jedem Fall ist das Phänomen weiterhin als Disfluency zu klassifizieren (vgl. auch Clark & Wasow 1998: 202; Maclay & Osgood 1959). Lerner sprachlich sind die Umstände einer exakten Wiederholung von Interesse: Laut Tavakoli (2010) dient die Wiederholung dem Gewinn zusätzlicher Planungszeit, wenn konzeptuell oder grammatisch schwierigere Passagen überwunden werden müssen.

Auch Rieger betont, dass nicht jede Wiederholung zur Klasse der Reparaturen gehört. Sprecher können Wiederholungen eines Tokens auch als Emphase zur Intensivierung desselben oder eines anderen Wortes verwenden (Bsp. 6–8). Diese Beispiele sind zwar im Rahmen der *Floor-Holding*-Hypothese als Wiederholungen analysierbar (Rieger 2003b: 51), wurden aber für meine Untersuchung ausgeschlossen.

- (4) und jetzt gehst du schräg bis du am Ziel bist 0,2 also schräg nach rechts unten (L1:2012-11-08-B)
- (5) geradeaus [RD nach links] [IR 0,6] [RS aso_{i1} eine_{i2} gerade_{i3} Linie_{i4} nach_{r1} links_{r2}]⁹ (L1:2012-11-08-B)
- (6) und links [...] links [...] links [...] und dann [...] rechts (L2:2013-05-02-A)
- (7) jajaja kein Problem kein Problem okay (L2:2013-04-19-A)
- (8) nach Westen nach Westen nach Westen (L1:2012-11-08-A)

Unabhängig von der auslösenden Motivation der Reparaturen und der linguistischen Analyse werden oft drei Klassen mithilfe der Beziehung zwischen Reparandum und Reparans unterschieden: „fresh starts, modification repairs, and abridged repairs“ (Heeman & Allen 1999: 3; vgl. auch Shriberg 1994: 10). So zeigen Neustarts keine oder nur eine geringe Korrelation zwischen Reparandum und Reparans. Modifikationsreparaturen tendieren hingegen zu einer größeren Wortkorrespondenz zwischen Reparandum und Reparans, während verkürzte Reparaturen kein Reparandum besitzen (Heeman & Allen 1999: 4). Neustarts zeigen typischerweise eine geringe Reparandum-Reparans-Wortkorrelation, was sich in der Einfügung von besonders viel neuem Material zeigt. In dieser Arbeit werden nur die *modification repairs* näher untersucht und mit dem Begriff ‚Reparatur‘ bezeichnet.

⁹Die Reparaturteile werden für RD, IR und RS in eckigen Klammern zusammengefasst. Jedes Token im RS trägt als Subskript seine Annotation.

Reparaturen sind alltägliche Bestandteile der gesprochenen Sprache. Oft werden sie jedoch nicht aktiv wahrgenommen. Sie entstehen durch kognitive Kontrollmechanismen der Sprachplanung, die auf verschiedenen Stufen der Sprachproduktion eine Überprüfung erlauben. Diese Kontrollmechanismen nennt man *Monitoring*. Wenn während der Sprachproduktion noch andere komplexe Aufgaben erledigt werden, so ist es wahrscheinlicher, dass die Sprachproduktion fehlerbehafteter wird oder ganz abbricht. Häufig kommt es bei (semi-)komplexen Aufgaben wie Wegbeschreibungen zu einer erhöhten Korrektur- bzw. Reparaturrate. Sprecher überprüfen ihre geplante und geäußerte Sprache konstant mithilfe des Monitoring-Prozesses. Von Levelt (1983, 1989) stammt die sogenannte *Perceptual Loop Theory of Monitoring*. Diese postuliert eine innere Monitor-Schleife zur Kontrolle der internen Sprache und eine äußere Monitor-Schleife zur Kontrolle der geäußerten Rede. Daher können Reparaturen nach Levelt (1983) auch overt und covert auftreten. Kovertre Reparaturen werden durch den internen Monitoring-Loop ausgelöst und vor der Artikulation repariert, overtre Reparaturen hingegen erst nach der Artikulation. Kovertre Reparaturen sind extrem schwierig experimentell nachzuweisen, weswegen sich die vorliegende Arbeit auf overtre Reparaturen konzentriert.

Im Rahmen der *Perceptual Loop Theory* werden Reparaturen dadurch ausgelöst, dass das Monitoring eine falsche oder unpassende Äußerung entdeckt. Indem dieser Fehler vom Sprecher *on-line*, d. h. während der Sprachproduktion, korrigiert wird, entsteht im Fluss der gesprochenen Sprache eine Reparatur. Die kognitiven Ressourcen zur Sprachproduktion sind begrenzt. Je mehr Ressourcen man für die Sprachplanung einsetzt, desto weniger stehen für das Monitoring zur Verfügung (Postma 2000).

Die Untersuchung von Reparaturen bei Lernern ist von Bedeutung für Theorien der Zweitsprachproduktion und des Zweitspracherwerbs. Um diese Theorien weiter verbessern zu können, ist es „essential to have more firmly theory-based, large-scale investigations which focus on self-repair by L2 speakers at different levels of proficiency [...]“ (Hest et al. 1997: 112). In dieser Studie werde ich mich auf fortgeschrittene Lerner beschränken und diese mit muttersprachlichen Daten vergleichen. Während Lerner im Anfangsstadium des Fremdspracherwerbs Selbstreparaturen anders als Muttersprachler verwenden, verwenden fortgeschrittene Lerner eher Strategien, wie sie auch Muttersprachler anwenden (Rieger 2003a: 41). Weiterhin wird angenommen, dass sich zusammen mit einem höheren Kompetenzniveau in der Zielsprache auch das metalinguistische Bewusstsein erhöht und der Aufmerksamkeitsfokus sich weg von lexikalischer, grammatischer und phonologischer Performanz hin zur Diskursebene verschiebt (ebd.: 44). Rieger analysiert Aufnahmen mündlicher Sprachtests von zehn Lernern des Deutschen der Stufe „intermediate“ qualitativ. Nach einem

Kompetenz-Grading der Probanden und der Auswertung der Aufnahmen produzieren Probanden der unteren Stufen „unfilled pauses, quasi-lexical fillers [bspw. *uhm* – M. B.], and very few repetitions of lexical items“, während demgegenüber Probanden mit dem besten Grading „German lexical fillers [bspw. *ja/also* – M. B.], repetitions of lexical items [...] and quasi-lexical fillers“ verwenden (Rieger 2003a: 44). Dies spricht für eine Annäherung der „intermediate“-Sprecher an ein muttersprachliches Niveau.

Im Allgemeinen sind die Sprachproduktionsprozesse und ihre einzelnen Module bei Lernern weniger automatisiert als bei Muttersprachlern (Kormos 1999; Wiese 1984). Dies betrifft aber nicht alle Module (in der Terminologie des modularen Sprachproduktionsmodells von Levelt 1989) im gleichen Maße, denn während beispielsweise die *Message*¹⁰ ebenso schnell wie in der Erstsprache produziert werden kann, beanspruchen die morphosyntaktischen Formulierungen und der Lemmazugriff hohe kognitive Ressourcen, welche für andere Tätigkeiten nicht mehr zur Verfügung stehen. So erwartet man bei Lernern mittleren Niveaus, welche mehr Ressourcen für die linguistische Genauigkeit ihrer Aussage verbrauchen, dass weniger Aufmerksamkeit und Arbeitsspeicherkapazität für das Monitoring ihrer Aussage zur Verfügung steht. Damit einhergehend erwartet man mehr Fehler und weniger Selbstreparaturen als etwa bei fortgeschrittenen Lernern, die die linguistische Formulierung in der Zweitsprache schon besser automatisieren und mehr Kapazitäten für ihr Monitoring zur Verfügung haben. Declerck & Kormos (2012) testen diese Hypothesen mit Lernern auf mittlerem und fortgeschrittenem Niveau (B1 und C1 des GER¹¹) in einem Dual-Task-Experiment¹². Die fortgeschrittenen Lerner produzieren in jeder getesteten Kategorie signifikant weniger Fehler als die B1-Lerner. Der Sprachproduktionsmechanismus der B1-Lerner arbeitet demnach weniger effizient als der der C1-Lerner. Die Anzahl der korrigierten Fehler jedoch ist für die C1-Lerner in den Kategorien Lexik, Syntax und Phonologie signifikant höher als bei den B1-Lernern. Dies lässt schließen, dass die Monitoringprozesse mit höherer Sprachkompetenz in der Zweitsprache effizienter funktionieren. Desweiteren indiziert ein Vergleich der Interregna (in der Terminologie von Declerck & Kormos „cut-off to repair times“), dass sich die C1-Lerner auch schneller korrigieren, was mit der höheren Automatisierung in der Zweitsprache erklärt werden kann (Declerck & Kormos 2012: 794).

Mit fortschreitender Kompetenz ergibt sich bei Lernern eine Annäherung an muttersprachliche Verwendungsweisen der Reparaturen. Dies wird auch durch

¹⁰Die *Message* wird vor dem Modul *Formulator* konzipiert, das Modul *Formulator* wird vor dem Modul *Articulator* ausgeführt (Levelt 1989)

¹¹Getestet mit dem Oxford Quick Placement Test.

¹²Aufgabe der Probanden war es, den Weg eines sich durch ein visuell präsentiertes Netzwerk von Abbildern verschiedener Objekte bewegendem roten Punktes zu beschreiben.

die Aufmerksamkeit bedingt, die für das Monitoring verfügbar ist. Probanden, die das zweite Mal dieselbe Aufgabe in der Zweitsprache durchführen, tendieren dazu, sich öfter zu korrigieren. Dies deutet darauf hin, dass bei wiederholter Ausübung einer Aufgabe mehr Aufmerksamkeit für das Monitoring zur Verfügung steht (Bygate 1996, zit. n. Kormos 1999). Durch die Wiederholung setzt also eine zunehmende Automatisierung ein. Da Automatisierung als Abwesenheit von bewusster Kontrolle definiert wird (Segalowitz & Hulstijn 2005), reduziert sich durch diesen Vorgang auch die Belastung des Arbeitsspeichers durch das Monitoring.

Strategien zur Behebung von vom Sprecher nichtintendierten Äußerungen können erklären, wieso bestimmte Typen von Reparaturen auftreten. So ist bekannt, dass Sprecher eine Unterbrechung des Redeflusses verzögern, bis sie die Reparantia für größere und kleinere Reparanda fertig geplant haben. Mit Blackmer & Mitton (1991: 192) formuliert: „it seemed that speech was interrupted when the repair was ready rather than when the problem was detected“. Mit dieser Strategie lässt sich die sofortige Null-Millisekunden-Reparatur erklären (Seyfeddinipur et al. 2008: 840 ff.), was einer Reparatur ohne Interregnum gleichkommt. Levelts lang angenommene Theorie zum Abbruch bei Reparaturen – die *Main Interruption Rule*, nach der der Sprachfluss sofort gestoppt wird, sobald die Möglichkeit einer Reparatur entdeckt wird (Levelt 1983: 56) – wurde schon von Blackmer & Mitton (1991: 192) infrage gestellt. Seyfeddinipur et al. können mittels der *Delayed Interruption for Planning Hypothesis* (DIP) auch gleich die sogenannten *Editing Terms* bzw. Editierungsaudrücke erklären (*also, ich meine, nein quatsch*). Diese fungieren als Hilfsmittel zur Aufrechterhaltung der Fluency, wenn Sprecher nicht schnell genug reparieren können. Editierung und Disfluencies stellen also eine Verzögerungsstrategie dar, deren Ergebnis eine Reparatur mit Interregnum ist. Solche Hilfsmittel könnten allerdings auch Prolongationen schon im Reparandum darstellen, was in dieser Untersuchung aber ausgeblendet wird.¹³ Die DIP baut auf L1-Daten auf, daher können empirische Untersuchungen Indizien für diese Strategien bei Lernern liefern.

Um eine Reparatur erfolgreich durchzuführen, muss das Monitoring ein Reparandum identifizieren und die Sprachproduktion ein Reparans planen und speichern. Sind genügend Ressourcen im Arbeitsspeicher frei, so geschieht dies problemlos, kann mit der DIP erklärt werden und man beobachtet sehr gerin-

¹³Interessant wäre beispielsweise, ob in Reparaturen ohne IR im letzten Token des RD eine Prolongation auftaucht, welches eine Planungsphase anzeigen könnte. Das Interregnum würde dann sozusagen in die Prolongations-Disfluency des RD inkorporiert. Leider ergab die Suche in BeMaTaC mithilfe der Prolongationsannotation von Myriam Klapi nur vier annotierte Prolongationen für L1 in solch einer RD-Position, so dass der Frage in dieser Arbeit nicht weiter nachgegangen wird.

ge gefüllte und ungefüllte Zeitspannen zwischen Reparandum und Reparans. Für Lerner jedoch muss man eine höhere Belastung des Arbeitsspeichers annehmen. Dies kann dazu führen, dass Reparanda nicht erkannt werden oder Reparantia langsamer (oder später) als in L1 geplant werden, was zu längeren gefüllten oder ungefüllten Pausen (oder deren Kombination) im Interregnum führt als bei Muttersprachlern. Diese direkte Folge einer höheren kognitiven Auslastung und einer geringeren Automatisierung lassen die Reparaturen von Lernern wiederum weniger muttersprachlich klingen. Obige quantitative Vorhersagen dürfen jedoch nicht ohne die qualitative Überprüfung des Reparaturauslösers betrachtet werden. Sprecher unterbrechen nämlich ihre Rede bei offensichtlichen Fehlern sofort (auch innerhalb eines Wortes), während sie hingegen die Unterbrechung des Redeflusses verzögern und das aktuelle Wort zu Ende äußern, wenn selbiges nur unangemessen ist, obwohl es grammatisch korrekt sein mag (Seyfeddinipur et al. 2008: 837), wie Bsp. 9 illustriert. Dies muss aber nicht immer der Fall sein (s. Bsp. 10). Tatsächlich scheint es also nicht leicht vorherzusagen, aus welchen Gründen eine Äußerung abgebrochen und repariert wird. Daher wird diese (auch annotatorisch aufwändige) Betrachtung für die weitere Vorgehensweise ignoriert.

(9) [RD nach links] [IR 0,5] [RS waagrecht_{i1} nach_{r1} links_{r2}] (L1:2012-01-19-A)

(10) [RD du z/¹⁴] [IR 0,6] [RS du_{r1} zeichnest_{s1}] (L2:2013-04-18-D)

Bei Lernern besteht zwar ein Zielkonflikt zwischen Fluency und Korrektheit, das Monitoring fordert in L2 aufgrund der geringeren Automatisierung aber bewusste Aufmerksamkeit (Declerck & Kormos 2012). Aufgrund der beschränkten Arbeitsspeicherressourcen sind auch die Reparaturplanungen im Sinne der DIP eingeschränkt. Dies impliziert, dass Lerner weniger 0-ms-Reparaturen produzieren als Muttersprachler und in Folge – interpretiert man eine 0-ms-Reparatur als fehlendes Interregnum – in ihren Reparaturen mehr Interregna verwenden sollten als Muttersprachler (siehe Hypothese 1.1 in Kap. 1.3).

Da Disfluencies ein typisches Phänomen von Spontansprache sind und spontane Sprache durch eine geeignete Methode eliziert werden muss, ohne dass die Probanden zu sehr dem *observer's paradox* erliegen (d. h. ihre Natürlichkeit möglicherweise verfälscht wird, weil sie sich beobachtet fühlen), wird zur Elizitierung oft eine Map-Task-Aufgabe herangezogen. So auch für BeMaTaC (eine genaue Erklärung der Map Task findet sich in Abschnitt 2.1 auf Seite 25). In diesem Kontext entstehen oft Situationen, in denen der Sprachplan geändert werden muss, beispielsweise durch zusätzliche Informationen, um die Äußerung informativer zu machen. Dies wiederum kann in Reparaturen resultieren

¹⁴Mit Schrägstrich wurde in BeMaTaC ein Abbruch annotiert.

(„self-corrections“) (Hartsuiker & Notebaert 2010: 169). Aufgrund der häufigen Verwendungsweise dieser Methode ist die Asymmetrie der beiden Sprecherrollen während eines Gesprächs als beeinflussender Faktor für die Art der Äußerungen evident geworden. Die Rolle des Instruierenden beinhaltet nicht nur mehr Äußerungsmaterial, sondern wird auch reflektiert in der unterschiedlichen Verteilung der verwendeten Reparaturtypen (Colman & Healey 2011: 1566). Zudem finden Colman & Healey signifikant mehr direkte Korrekturen als fremd-initiierte Fremd-Reparaturen. Kein Effekt wird für die Bekanntheit der Sprecher gefunden auf den Gebrauch von Reparaturen gefunden. Reparaturen sind für Colman & Healey nicht nur „noise in the signal“ sind, sondern sie erzeugen systematische Muster.

Weil die Instruierenden nicht nur besser informiert sind, sondern diese Information auch konzeptualisieren und in einen Sprachplan umsetzen müssen, weisen sie im Vergleich zu den Instruierten eine signifikant höhere Disfluency-Rate auf (Branigan et al. 1999: 389). In derselben Studie untersuchten Branigan et al. auch die Art einer Reparatur (jedoch als Ganzes und nicht auf Wortebene) mit „repeat, delete, insert, substitute“ (ebd.: 387), ohne jedoch eine genau Richtlinie für diese Kategorien anzugeben. In den Kategorien Repetition, Insertion und Substitution waren die Instruierenden signifikant häufiger disfluent als die Instruierten.

Auch Bortfeld et al. (2001) und Belz & Klapi (2013) zeigen, dass bei Map-Task-Quasi-Experimenten, wie sie für das Zustandekommen der hier verwendeten spontansprachlichen Daten gebraucht wurden, signifikante Unterschiede in der Verwendung von Disfluencies in den unterschiedlichen Rollenzuweisungen (Instructor/Instructee) innerhalb eines Map-Task-Gesprächs bestehen.

In der Studie von Branigan et al. gibt es zwar keinen signifikanten Effekt des Faktors **Blickkontakt**¹⁵ in der Ausprägung *ja/nein* während einer Map Task; allerdings finden sie eine signifikante Interaktion zwischen **Blickkontakt** und **Geschlecht** des Sprechers. Auch werden Repetitionen signifikant häufiger in der Ausprägung ohne Blickkontakt gemessen (Branigan et al. 1999: 390). In den hier untersuchten Daten hatten die Sprecher keinen Blickkontakt, daher sollte eine insgesamt höhere Repetitionshäufigkeit erwartet werden; aufgrund einer fehlenden Kontrollstudie kann die genaue Zahl aber nicht quantifiziert werden.

Die oben beschriebene Diskussion um die kognitiv-häsitationelle (*Lexical Retrieval*) vs. konversationsintendierte (*Floor Holding*) Unterscheidung der Disfluency-Bedingtheit zeigt – neben vielen weiteren Theorien zur Produktion

¹⁵Wenn Faktoren bzw. Variablen gemeint sind, werden sie zur besseren Unterscheidung in *äquidistanter Schrift* gesetzt. Faktorausprägungen bzw. Varianten von Variablen werden *hervorgehoben*.

von Disfluencies – die vielfältigen Interpretationsmöglichkeiten der einzelnen Konkretisierungen von Disfluencies. Gerade hierin liegt für eine quantitativ-qualitative Analyse eine große Herausforderung, denn jede Studie birgt die Gefahr, eigene Klassifikationen aus diesen Beobachtungen emergieren zu lassen. Um solche interpretativen Fehler zu minimieren, wird der Fokus hier auf die Art der sprachplanerischen Disfluency als Reparatur gelegt. Natürlich gilt auch für das weiter unten auszuarbeitende Schema, dessen linker und rechter Rand als *fuzzy* interpretiert werden kann – die spätere Analyse kann diese Marginalia jedoch gut ignorieren¹⁶, da die für diese Studien relevanten Disfluencies innerhalb einer Reparatur liegen: die Interregna, welche in der Mitte der Reparatur zu finden sind (oder eben nicht). Dies gilt auch für die später hinzukommenden Annotationen der semantischen Beziehungen des eingefügten Teils in Bezug zum reparierten Teil der Reparatur: die Art der Beziehung ist dank der Annotationsrichtlinien im Wesentlichen klar nachvollziehbar.

1.2.1 Schema

Ein weit verbreitetes Schema für Reparaturen ist ihre Dreiteilung inklusive eines Abbruchmoments. Es wurde in Abschnitt 1.1 schon angerissen und wird nun genauer expliziert. Levelt (1983: 44) nennt den ersten Teil der Reparatur die „original utterance [Kursivierung hier u. i. F. gelöscht – M. B.]“, welche das Reparandum („das zu Reparierende“) enthält. Bei Levelt spannt sich die original utterance von der letzten Satzgrenze vor dem Reparandum bis zum Abbruchmoment („moment of interruption“). Das Reparandum definiert nur genau die Einheit innerhalb der original utterance, die später repariert wird. Sprecher können sich innerhalb oder nach Äußerung des Reparandum unterbrechen, was Levelt „delay of interruption“ nennt. Nach dem Abbruchmoment folgt die Editierungsphase, welche Disfluencies enthalten kann. Schließlich folgt die eigentliche Reparatur („repair“).

Der Terminus ‚Reparatur‘ ist im Deutschen nicht unproblematisch, da er sowohl die Gesamthandlung einer sprachlichen Korrektur beschreibt, als auch den dritten Teil innerhalb des Reparaturschemas. Daher werde ich für die eigentliche Korrekturspanne der Äußerung den Terminus *Reparans* übernehmen (vgl. auch Eklund 2004: 164). Die terminologische Problematik veranschaulicht auch Tabelle 1.1 auf Seite 18. Im Englischen besteht hier eine terminologische Unsicherheit, besonders auch durch die Verwendung von *repair* (als *Reparans*) als Teil von *repair* (als Gesamtphänomen). Durch den Terminus *Reparatur*

¹⁶Die rechten Ränder der *Reparantia* sind allerdings u. U. recht dehnbar, was sich möglicherweise auf eine Frequenzanalyse auswirkt. Hier sei daher nochmals auf die Priorität rigider Annotationsrichtlinien hingewiesen.

für das Phänomen und *Reparans* für das Reparierende kann diese Ambiguität behoben werden.

Das Modell von Shriberg (1994) vereinfacht die Betrachtungsweise, indem die einzelnen Teile über die Zeitspanne der Aktionen definiert werden. Shriberg bezieht sich auf Levelts Schema, nimmt aber einige Änderungen vor. Sie nennt den Abbruchmoment „interruption point“ und die Editierungsphase, die durch Disfluencies oder Diskursmarker gefüllt werden kann, Interregnum. Das Interregnum muss nicht mit Sprache gefüllt werden, sondern kann auch eine stille Pause beinhalten.

[Interregnum] specif[ies] the temporal region from the end of the reparandum to the onset of the repair even if this region contains no editing term (Shriberg 1994: 8).

Das von mir angewandte Schema für Reparaturen folgt schließlich Shriberg und enthält drei Bestandteile ohne den Abbruchpunkt. Im Reparandum (RD) steht der Ausdruck, welchen der Sprecher korrigieren möchte. Im Interregnum (IR) befinden sich Disfluencies, die eine Reparatur indizieren, wie ungefüllte Pausen (Stille), gefüllte Pausen (*äh, ähm*), Diskursmarkierer (*sozusagen, beziehungsweise*) und explizite Editierungen (*ich meine*). Das Interregnum ist fakultativ, da Reparaturen auch direkt vom Reparandum ohne Planungszeit ins Reparans übergehen können. Das Reparans (RS) enthält den neuen, reparierten Ausdruck. Gefüllte und ungefüllte Pausen zwischen Reparandum und Reparans stehen nach dieser Vorgehensweise im Interregnum (Bsp. 11) und können somit unterschieden werden von anderen Pausen im Gespräch (Bsp. 12).

(11) und [RD dem] [IR 0,2 äh] [RS der_{s1} Sanduhr_{i1}] (L1:2011-12-14-B)

(12) asso ja genau äh warte äh 0,7 von dir also (L1:2012-01-19-A)

Die Definition ist intuitiv sicherlich verständlich, allerdings ist methodisch prima facie noch unklar, wo nun genau das Reparandum beginnt, welches im Reparans manipuliert wird. Dies kann nur über die Struktur des Reparans und deren Abgleich mit dem Reparandum erkannt werden und wird in den Annotationsrichtlinien festgelegt.

[H]uman listeners cannot reliably predict upcoming disfluencies in the region of the onset of the reparandum (Nakatani & Hirschberg 1994: 1607).

Tabelle 1.1: *Terminologische Variation bei der Beschreibung von Reparaturen.*

In dieser Arbeit:	Reparandum	Abbruchmoment	Interregnum	Reparans
In der Literatur:				
Levelt (1989)	reparandum	moment of interruption	interregnum	repair
Blackmer & Mitton (1991)	–	cuttoff	cut-off to repair	–
Shriberg (1994)	reparandum	interruption point	interregnum	repair
Nakatani & Hirschberg (ebd.)	reparandum interval	interruption site	disfluency interval	repair interval
Fox & Jasperson (1995)	repaired segment	–	–	repairing segment
Heeman & Allen (1997, 1999)	reparandum	interruption point	editing terms	alteration
Lickley & Bard (1998)	reparandum	interruption point	editing term or pause	repair
Core (1999)	reparandum	interruption point	editing term	alteration
Heeman & Allen (1999)	reparandum	interruption point	editing term	alteration
Brennan & Shober (2001)	reparandum	interruption site	edit interval	repair interval
Rieger (2003b)	repairable	repair initiation	repair initiation	repairing segment
Eklund (2004)	Reparandum	suspension point	–	Reparans
Besser & Alexandersson (2007)	reparandum	–	interregnum	reparans
Fox et al. (2012)	repair initiation	–	–	repair proper

1.2.2 Subreparaturen

Eine Reparatur im eigentlichen Wortsinne ist jedoch typischerweise eine Modifikation des zuvor Gesagten. Diese Relation wird selten in eine Untersuchung einbezogen, dabei ist sie ein integraler struktureller Bestandteil von Reparaturen. Es scheint, als ob Reparaturen dazu tendieren, Wortkorrespondenz zwischen Reparandum und Reparans zu erhalten, dennoch aber oft eine Mischform im Reparans produzieren. Dort kommt es zu exakten Wiederholungen (Repetitionen), zu weniger exakten Wiederholungen (Substitutionen) und zu Ergänzungen (Insertionen). Diese Kategorien verwenden auch Hartsuiker & Notebaert (2010: 171), wobei sie auch Löschungen hinzunehmen. Im Gegensatz zu Hartsuiker & Notebaert betrachte ich Repetitionen nicht als separate Kategorie, sondern als Subkategorie von Reparaturen. Explizit auf die semantische Beziehung zwischen Reparandum und Reparans fokussiert postulieren aber schon Fox & Jasperson (1995) für das Englische eine Reparatursyntax, die diese Klassen genauer spezifizieren. Im Rahmen ihrer Untersuchung finden Fox & Jasperson sieben mögliche Reparaturklassen, welche das Reparandum (Fox & Jasperson: „repaired segment“) und das Reparans (Fox & Jasperson: „repairing segment“) in Beziehung zueinander setzen. Für meine Subreparaturklassifikation orientiere ich mich an Fox & Jasperson, vereinfache aber die von ihnen verwendeten Definitionen, die in drei Subtypen kondensieren: Typ *r* für Repetitionen, Typ *s* für Substitutionen und Typ *i* für Insertionen. Mit diesen drei Typen lassen sich Fox & Jaspersons sieben Klassen darstellen (s. Tabelle 1.2 auf der nächsten Seite). Mehr noch, die drei per Token¹⁷ definierten und pro Reparans durchnummerierten Typen erlauben durch ihre Kombination praktisch jede beliebige Reparaturrealität der Sprecher abzubilden und sind so deutlich flexibler in der Anwendung. Die Beispiele 13–18 werden den Typen in Tabelle 1.2 zugeordnet.

- (13) äh der Mitte [RD des] [IR 0,1] [RS des_{r1}] äh Nagelbildes (L1:2011-12-14-A)
- (14) gibts nämlich [RD ein] [IR 0,3] [RS keinen_{s1}] rechten Winkel (L1:2012-01-19-A)
- (15) [RD in ne] [RS in_{r1} ne_{r2}] nörd/ (L2:2013-04-18-D)
- (16) [RD die m/] [RS die_{r1} Mitte_{s1}] von der Oberfläche (L2:2013-04-18-D)
- (17) [RD unter dem] [RS unter_{r1} dem_{r2} Bild_{i1}] (L2:2013-04-18-D)

¹⁷Token in Bezug auf das verwendete Korpus bezieht sich auf eine Tokenisierung auf Wortebene, die jedoch, da sie sich auf die Dipl-Ebene bezieht, durchaus auch Ausdrücke wie *haste* beinhalten kann.

Tabelle 1.2: Vereinfachtes Subreparaturklassifikationsschema.

Fox & Jaspersen: 90	Vereinfachte Typologie	Beispiel
A) recycle word	r (bspw. r1) ^a	Bsp. 13 auf der vorherigen Seite
B) replace word	s (bspw. s1)	Bsp. 14 auf der vorherigen Seite
C) recycle prior phrase including word	r (bspw. r1 r2)	Bsp. 15 auf der vorherigen Seite
D) recycle prior phrase, replace word	rs (bspw. r1 s1)	Bsp. 16 auf der vorherigen Seite
E) recycle prior phrase, add new elements	ri (bspw. r1 r2 s1)	Bsp. 17 auf der vorherigen Seite
F) change syntactic framework	ir (bspw. i1 i2 r1 r2)	Bsp. 18
G) abandon structure, start new structure	i (bspw. i1 i2)	

^a Zählung innerhalb des Reparans. Wird in Abschnitt 2.2.2 auf Seite 29 nochmals erläutert.

(18) [RD bis zu de/] [RS also_{i1} gerade_{i2} bis_{r1} zu_{r2} dem_{s1}] (L2:2013-04-18-D)

Die Struktur von Insertions-Reparaturen des Englischen untersuchen Wilkinson & Weatherall (2011) vor allem qualitativ. Drei Viertel ihrer untersuchten Reparaturen enthalten einen Abbruch im Reparandum (ebd.: 68) und die Hälfte des insertierten Materials besteht aus Adjektiven, gefolgt von Adverbien (ein Viertel) und anderen, nicht weiter beschriebenen Wortarten. Leider beschreiben sie nicht, welchen Anteil die Insertions-Reparatur an anderen Reparaturarten (Repetition etc.) ausmacht. Jedoch scheint die Kategorie der offenen Wortklassen endemisch für Insertionen zu sein. So kommen sie zu dem Schluss, dass die häufigsten Verwendungsweisen von Insertions-Reparaturen das Spezifizieren zur Disambiguierung (ebd.: 80) und das Intensivieren der Aussage (ebd.: 82) sind. Wilkinson & Weatherall verwenden den Begriff der Insertion im strikten Sinne:

Insertion repair does not substitute one referent for another, or one formulation for another: Instead it retains and builds on the original formulation [...] (Wilkinson & Weatherall 2011: 87).

In dieser Arbeit wird der Begriff ‚Insertion‘ für Material stehen, das sowohl im Sinne einer klassischen Insertionsreparatur vorkommt, als auch außerhalb des Insertionsrahmens verwendet wird, um das Reparans weiter zu spezifizieren. Da die klassische Insertionsreparatur wie in Bsp. 19 eher selten vorkommt

(dennoch aber nachgerade als Lehrbuchreparatur fungiert), kann durch die so geschehene Definitionserweiterung (s. Bsp. 20) ein weiterer Bereich der quantitativen Analyse abgedeckt werden.

- (19) [RD nach links] [RS waagerecht_{i1} nach_{r1} links_{r2}] (L1:2012-01-19-A)
- (20) [RD dann gehs de jetzt] [IR ähm 0,5 (?)¹⁸] [RS jetzt_{r1} machs_{i1} de_{r2}
wieder_{i3} ne_{i4} Waagerechte_{i5}] (L1:2012-11-02-B)

Die von mir gewählte Klassifikation beschränkt sich also vornehmlich auf die Beziehung des ersetzenden Materials zum ersetzten Material. Dies hat zum einen den Vorteil, dass die nicht unerheblich aufwändige Annotation leichter anzuwenden ist. Damit werden hoffentlich weniger Zweifelsfälle auftreten als bei einer linguistisch-inhaltlichen Annotation, die Gesamtqualität der händischen Annotation wird damit gesteigert. Eine linguistische Analyse ist dennoch nicht ausgeschlossen, denn dank der Annotation im Stand-Off-Format kann leicht eine Verbindung mit den annotierten Wortarten hergestellt werden.

1.3 Statistische Hypothesen

Die in den ersten beiden Kapiteln beschriebene Literatur über die mutter- und lernersprachliche Produktion und Verwendung von Disfluencies und Reparaturen weist – vor allem auch für die deutsche Sprache – noch einige Lücken auf. Aus den beschriebenen und genauer beleuchteten Phänomenen stelle ich im Folgenden Erwartungen an die Daten auf, die daraufhin mit deskriptiver und analytischer Statistik ausgewertet werden. Da analytische Statistik optimalerweise nicht post hoc Fragen an die Daten stellen darf, beschreibe ich die Erwartungen an die Daten in sorgfältig formulierten Null- (H0) und Alternativhypothesen (H1) in ihren Ausprägungen so genau wie möglich. Die Nullhypothesen sind dabei gerichtet, um mit der relativ kleinen empirischen Grundlage ein Höchstmaß an Aussagekraft zu erlangen.

Die Aufmerksamkeit der Untersuchung gilt der Explorierung von mutter- und lernersprachlichen Disfluencies, Reparaturen und ihren Interaktionen. Um die Interaktion von Disfluencies in und mit Reparaturen genauer zu untersuchen, muss man die Art und Häufigkeit ihrer Vorkommen und eventuellen Abhängigkeiten innerhalb der Reparaturen betrachten, ergo das Interregnum.

Für Muttersprachler und Lerner erwarte ich jeweils mehr Sprecher, die mit Gebrauch einer Disfluency, also eines Interregnums, reparieren, als Sprecher, die ohne Disfluency reparieren.

¹⁸Steht laut BeMaTaC-Richtlinien für vom Transkribenden nicht verstandene Äußerungen.

Hypothese 1.1: Interregnum-Häufigkeit bei Muttersprachlern.

H0: $P(SIR) \leq 0,5$

SIR: Für einen Sprecher gilt $MIR > OIR$.

MIR: Häufigkeit der Reparaturen mit Interregna.

OIR: Häufigkeit der Reparaturen ohne Interregna.

Sprecher mit $OIR = MIR$ werden ausgeschlossen.

In Worten: Die Wahrscheinlichkeit, dass ein L1-Sprecher mehr Reparaturen mit IR als ohne IR äußert, ist höchstens so groß wie die Wahrscheinlichkeit, dass ein L1-Sprecher mehr Reparaturen ohne IR als mit IR äußert.

H1: $P(SIR) > 0,5$

In Worten: Die Wahrscheinlichkeit, dass ein L1-Sprecher mehr Reparaturen mit IR als ohne IR äußert, ist größer als die Wahrscheinlichkeit, dass ein L1-Sprecher mehr Reparaturen ohne IR als mit IR äußert.

Hypothese 1.2: Interregnum-Häufigkeit bei Lernern (vgl. Hyp. 1.1).

Hypothese 1.2 entspricht Hypothese 1.1, aber für die Gruppe der Lerner.

Für den Vergleich von Lernern und Muttersprachlern sollten Lerner nach Declerck & Kormos (2012) jedoch mehr Interregna verwenden als Muttersprachler. Muttersprache und Zielsprache werden im Folgenden als Ausprägungen des Faktors **Fähigkeitsniveau**¹⁹ zusammengefasst, *MIR* und *OIR* als Ausprägungen des Faktors **Interregnum**.

Hypothese 1.3: Muttersprachler und Lerner im Vergleich der Interregnum-Häufigkeit.

H0: Muttersprachler und fortgeschrittene Lerner unterscheiden sich nicht in der Verteilung ihrer verwendeten Interregna dergestalt, dass Lerner höchstens so viele MIR-Reparaturen produzieren wie Muttersprachler. Die Verteilung der Variablenausprägungen der Variable **Reparatur** variiert nicht in Abhängigkeit von der Variable **Fähigkeitsniveau**.

H1: Muttersprachler und fortgeschrittene Lerner unterscheiden sich in der Verteilung ihrer verwendeten Interregna dergestalt, dass Lerner mehr MIR-Reparaturen produzieren als Muttersprachler. Die Verteilung der Variablenausprägungen der Variable **Reparatur** variiert in Abhängigkeit von der Variable **Fähigkeitsniveau**.

¹⁹Begriff übernommen von Wiese (1983).

Nach der Ermittlung der Proportionshäufigkeit der verwendeten Interregna stellt sich die Frage nach deren interner Struktur. Nach Belz & Klapi (2013: 10) verwenden fortgeschrittene Lerner zwar doppelt so häufig gefüllte Pausen wie Muttersprachler. Unklar ist jedoch, ob dieses Muster auch innerhalb von Reparaturen gilt, was sich innerhalb des Interregnums überprüfen lässt. Dieses kann wiederum Disfluencies aller Art enthalten. In Hypothese 2 betrachte ich jedoch nur gefüllte Pausen innerhalb von Interregna.

Hypothese 2: Gefüllte Pausen im Interregnum.

H0: $P(FP|L2) \leq P(FP|L1)$

FP: Das IR enthält eine gefüllte Pause (*äh/ähm*).

In Worten: Fortgeschrittene Lerner verwenden höchstens gleich häufig gefüllte Pausen im IR wie Muttersprachler.

H1 $P(FP|L2) > P(FP|L1)$

FP: Das IR enthält eine gefüllte Pause (*äh/ähm*).

In Worten: Fortgeschrittene Lerner verwenden häufiger gefüllte Pausen im IR als Muttersprachler.

Nach der Ermittlung der Proportionshäufigkeit von Interregna sowie von Interregna mit gefüllten Pausen stellt sich die Frage nach deren Abhängigkeit im Zusammenspiel mit dem Reparandum und besonders dem Reparans, da Sprachplanung immer nach vorne gerichtet ist. Hier wird zunächst die Besonderheit des genutzten Korpus berücksichtigt. Bortfeld et al. (2001) und Belz & Klapi (2013) weisen darauf hin, dass die in der Map Task eingenommene Rolle Auswirkungen auf Disfluencies hat. Als *Instructor* äußert man signifikant häufiger Disfluencies als als *Instructee*. Innerhalb des Faktors **Fähigkeitsniveau** wird dies als Alternativhypothese aufgestellt. Für das Vorkommen eines Interregnums deutet die im Theorieteil beschriebene Literatur an, dass es beispielsweise einen Zusammenhang zwischen Inhaltswörtern und dem Vorkommen einer Disfluency gibt (Hartsuiker & Notebaert 2010), dass aber auch die Relation der reparierten Segmente zueinander (Tavakoli 2010) Disfluencies beeinflussen können.

Hypothese 3: Anzahl der Reparaturen in der Gesprächsrolle in der Map Task.

H0: Die Verteilung der Reparaturen zwischen *L1* und *L2* bezüglich ihrer Rolle ist gleich.

H1: Die Verteilung der Reparaturen zwischen *L1* und *L2* bezüglich ihrer Rolle ist verschieden.

Im Anschluss an Hypothese 3 erfolgt eine explorative Auswertung, um möglichen

cherweise vorhandene Unterschiede zwischen Muttersprachlern und Lernern in ihrer Gesprächsrolle und der Anzahl der geäußerten Interregna in genau dieser Ausprägung genau beschreiben zu können. Analysiert wird also die Kombination zwischen den Faktoren **Rolle**, **Fähigkeitsfaktor** und **Interregnum**.

Möglicherweise können die im Rahmen dieser Arbeit erhobenen Daten Aufschluss über die Distinktion von *MIR*- vs. *OIR*-Reparaturen geben. Die Beziehung des Reparans zum Reparandum, wie es in den Subreparaturebenen annotiert wird, sowie die damit verbundenen Wortarten der Subreparaturtoken können aufzeigen, ob möglicherweise eine wortartenspezifische und/oder relationspezifische Komplexität bei der Sprachplanung in L1 und L2 auf dem Niveau C1–C2 unterschieden werden kann. In Hypothese 4 werden daher a priori keine Vorhersagen getroffen, sondern ich werte den Einfluss der in der Subreparaturebene annotierten inhaltlichen Beziehung zwischen RD und RS auf das Vorhandensein eines IR explorativ a posteriori aus. Betrachtet werden die Faktoren **Fähigkeitsniveau**, **Interregnum** und **Subreparaturen** sowie die zugeordneten **Wortarten**.

Hypothese 4: Einfluss der Subreparaturen.

Welchen Einfluss haben die Subreparaturen in Verbindung mit Wortarten darauf, ob Lerner oder Muttersprachler ein Interregnum vor dem Reparans äußern? Die Auswertung erfolgt explorativ.

2 Korpusbasierte Analyse

2.1 Methodik und Korpusdesign

Methodisch erfolgt die Untersuchung mittels eines kontrastiven (Lüdeling 2007: 34), quantitativ-qualitativen und korpusbasierten (Lemnitzer & Zinsmeister 2010: 37), explorativen Ansatzes im Gebiet spontansprachlichen Sprechens. Ein kontrastiver Vergleich ermöglicht Erkenntnisse über die Abweichung der Lerner Sprache gegenüber der Zielsprache hinsichtlich der Verwendung von Pausen und gefüllten Pausen, über Unterschiede in der Länge und Art der Reparaturen, sowie über die Interaktion von gefüllten Pausen und Reparaturen (oder Reparaturarten). Klassischerweise ermittelt man in Texten die „durchschnittliche Satz- oder Wortlänge“ und ob „Lerner [...] eine Klasse von Wörtern oder eine Konstruktion im Vergleich zu häufig [...] oder zu selten [...] anwenden“ (Lüdeling 2007: 34).

Für die Untersuchung wird das Berlin Map Task Corpus (BeMaTaC) (Giesel et al. 2013; Sauer & Lüdeling 2013) herangezogen. Das muttersprachliche Subkorpus besteht in der für diese Arbeit verfügbaren Version aus 12 aufgabenorientierten Dialogen (66 Minuten, 11.192 Tokens) mit insgesamt 16 Probanden (Version 2013-01). Acht Probanden führen jeweils zwei Dialoge miteinander (= 8 Dialoge), in denen sie die Rolle des Instructors wechseln; von acht weiteren Probanden wird mangels Aufbereitung jeweils nur ein einzelnes Gespräch ohne Wechsel untersucht (= 4 Dialoge).

Im neuen, lernersprachlichen Subkorpus BeMaTaC-L2 sind fünf¹ Gespräche zwischen Deutschlernern mittleren Niveaus enthalten, deren Muttersprache Englisch ist (77 Minuten, 21.330 Tokens). Der Sprachstand der Lerner wurde anhand eines C-Tests (Eckes 2010) ermittelt. Alle Sprecher erreichten mindestens das Niveau C1.

Die Erhebung der Daten erfolgte im Rahmen der Q-Tutorien² der BeMaTaC-Gruppe mittels einer Map-Task-Aufgabe, wie sie bei Anderson et al. (1991)

¹Zur Entstehungszeit dieser Arbeit konnte nur auf fünf transkribierte Dialoge im L2-Korpus zugegriffen werden.

²Ein Q-Tutorium der Humboldt-Universität zu Berlin ist eine studentisch initiierte, konzipierte und durchgeführte Veranstaltungsform, in der die Studierenden ein selbst gewähltes Thema forschungsorientiert bearbeiten. <http://bolognalab.hu-berlin.de/projekte-des-bologna.labs/q-programm/q-tutorien>, besucht am 10.11.2013.

für das HCRC³-Korpus erstmals beschrieben wurde. Zwei Sprecher sehen eine Karte mit Wegpunkten vor sich. In der einen Karte ist ein Weg durch diese Punkte eingezeichnet, in der anderen nicht. Der Sprecher mit dem eingezeichneten Weg (= Instructor) gibt Anweisungen, die der andere Sprecher (= Instructee) umsetzen muss, um den Weg mit einem Stift vom Startpunkt zum Zielpunkt nachzuzeichnen. Die so erhobenen Daten enthalten quasi-spontan elizitierte Sprache. Die Sprecher haben keinen Blickkontakt und sehen auch nicht die Karte ihres Gegenübers, müssen die Aufgabe also sprachlich lösen, ohne nonverbale Gestik oder Mimik mit einzubeziehen. Der Instructee kooperiert durch Feedback oder Rückfragen. Dies wird durch eine Abweichung eines Wegpunktes je Karte sogar explizit eliziert. Abbildung 2.1 zeigt das eine von zwei Kartensets mit Abweichung des Wegpunktes am linken mittleren Rand der Karte (Nägel vs. Schornsteinfeger). Das aufgabenorientierte Erhebungsdesign hat zur Folge, dass der Instructor wesentlich mehr Tokens äußert als der Instructee.

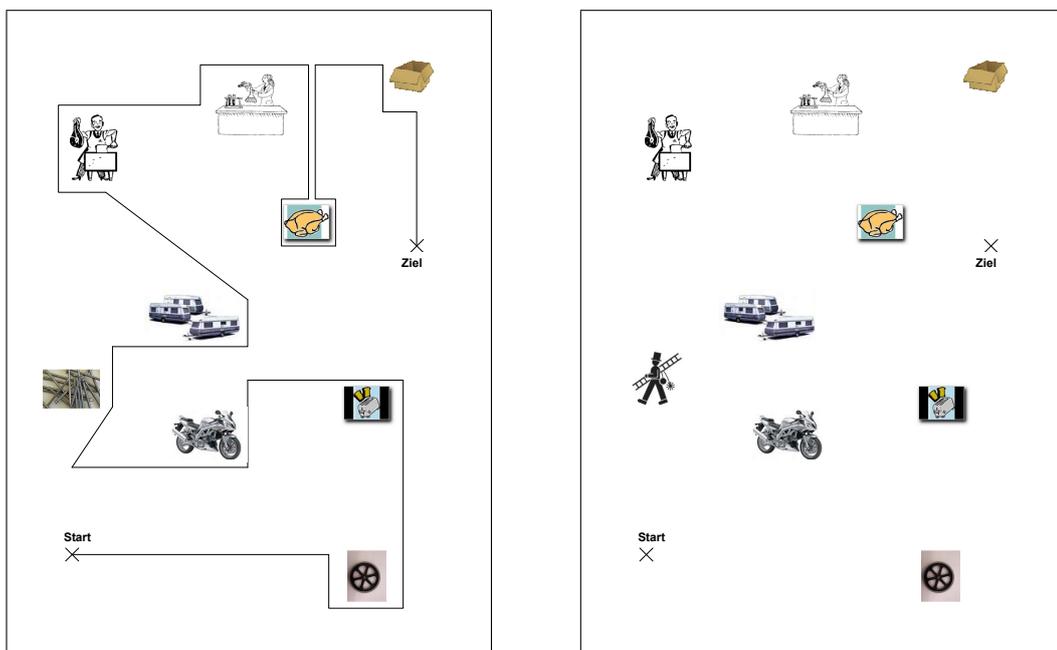


Abbildung 2.1: Typische Karte eines Instructors mit eingezeichnetem Weg (links) und eines Instructees ohne eingezeichneten Weg (rechts) in einem Map-Task-Quasi-Experiment.

Mit der Map-Task-Methode wird für beide Faktorausprägungen der Sprache (Muttersprache vs. Zielsprache bzw. L1 vs. L2) sichergestellt, dass die Be-

³Human Communication Research Centre, Edinburgh.

dingungen der Sprachelizitation exakt gleich sind. So können konfundierende Effekte wie Registereffekte (vgl. Schachter et al. 1991) und rhetorische Effekte (vgl. Siegel et al. 1969) ausgeschlossen oder minimiert werden.

Da die Gespräche unter kontrollierten Bedingungen in einer unnatürlichen Umgebung (Phonetiklabor) mit expliziten Anweisungen verbunden waren, muss man immer auch von einem *Observer's Paradox* ausgehen, welches die Daten verzerren und so unter Umständen falsche Rückschlüsse aus der Stichprobe auf die Grundgesamtheit nach sich ziehen kann. Hier lässt sich wenigstens feststellen, dass sich die Probanden besonders zu Beginn und Ende der Gespräche ihrer Umgebung bewusst waren, durch die Wegbeschreibungsaufgabe kognitiv aber immerhin so hinreichend beansprucht wurden, dass man von originär elizierten Interaktionsdaten ausgehen können sollte.

2.2 Annotation

Zur Erhebung quantitativer Vergleichsdaten müssen die verwendeten Daten transkribiert und mit Annotationen zu den Phänomenen versehen werden, damit sie ausgezählt werden können. Dies ist nicht trivial, da jede Kategorisierung mit einer Interpretation der Daten einhergeht (Lüdeling 2007: 29). Daher sind im Anhang B die Annotationsrichtlinien mit Beispielen aufgeführt. Da die Daten von mir selbst nach meinen Richtlinien annotiert wurden, habe ich kein umfangreiches Inter-Annotator-Agreement getestet. Jedoch wurde ein Gespräch fremdannotiert, um die Eindeutigkeit und Nachvollziehbarkeit der Richtlinien zu testen und im Anschluss zu gewährleisten.⁴

Die Gespräche sind mit Praat (Boersma 2001) transkribiert worden.⁵ Trotz einiger in der Literatur beschriebenen Tools und Parser zum automatischen Auffinden und Markieren von Reparaturen in Texten (Core 1999; Heeman & Allen 1999; Nakatani & Hirschberg 1994) ist keines der beschriebenen Werkzeuge online verfügbar gemacht worden.⁶ Da es zudem nicht unwahrscheinlich ist, dass die entwickelten Tools in der deutschen Sprache nur eingeschränkt funktionieren würden, wurde das vorliegende kleine Korpus händisch annotiert.

Durch die händische Transkription kann nicht angenommen werden, dass die Transkription ein vollständiges Abbild des Schallsignals ergibt. Dies ist

⁴Hierfür herzlichen Dank an Carolin Odebrecht.

⁵Die L2-Daten wurden teilweise von Myriam Klapi und mir sowie von Studierenden aus dem Q-Tutorium transkribiert. Die L1-Daten lagen schon fertig transkribiert vor.

⁶DialogueView (Yang et al. 2008), ein Tool für die automatische Annotation von (u. a.) Reparaturen, ist auch nach intensiver Suche nicht auffindbar und wird nur über tote Links referiert (<http://www.cslu.ogi.edu/DialogueView/>, besucht am 05.11.2013).

bedingt durch die perzeptiv-kategorielle Automatisierung der Transkribenden. Unsere kognitive Verarbeitung blendet sehr schnelle und kurze Reparaturen automatisch aus.

Die Annotation geschah über von mir entwickelte Annotationsrichtlinien⁷ zu Reparaturen (s. Anhang B) mithilfe des EXMARaLDA-Tools (Schmidt & Wörner 2009) in einem Multilayer-Stand-Off-Format. Sowohl die muttersprachlichen als auch die lernersprachlichen Daten wurden komplett von mir annotiert. Durch das parallele Anhören der Gespräche können auch prosodische Hinweise auf eine Reparatur miteinbezogen werden (Nakatani & Hirschberg 1994; vgl. auch Wichmann 2008: 196), wo die Transkription keine Rückschlüsse auf die Prosodie zulässt. Das Korpus ist mithilfe des TreeTaggers (Schmid 1994) und des STTS-Tagsets⁸ auch nach Wortarten annotiert worden.⁹

2.2.1 Vorhandene Annotation

Schon im Korpus 2013-01 vorhanden sind die Annotationsebenen *dipl* (die Transkription), *norm* (die Normalisierung), *lemma* (die Lemma-Ebene), *pos* (Wortarten, vgl. Abschnitt 2.3), *utt* (eine Utterance-Ebene) und *extra* (extralinguistische Ereignisse), jeweils für den Instructor und den Instructee. Diese Ebenen wurden auch für die L2-Daten erhoben.

Transkriptionsmängel wurden aus Zeit- und Konsistenzgründen nicht behoben. Aufgrund der universellen Transkription, die nicht spezifisch auf ein Phänomen konzentriert ist, können für bestimmte Phänomene leichte Abweichungen im Verständnis und in der Transkription und Annotation des Signals zu erwarten sein. Stille Pausen sind beispielsweise in der L1-Version noch manuell annotiert worden, in der L2-Version hingegen automatisch mit vereinzelt manuellen Korrekturen, so dass die Qualität dort besser sein dürfte. Als untere Schwelle (Cut-Off) für eine Pause wurde im L1-Korpus 100 ms festgelegt¹⁰. Dieser Wert ist eine semi-arbiträre Festlegung – Pausen unterhalb dieses Wertes sind oft artikulatorisch bedingt. So kann Atmen beispielsweise als Pau-

⁷Ursprünglich für die Erhebung im Rahmen des Q-Tutoriums „Berlin Map Task Corpus – Korpusdesign und gesprochene Sprache“ vorgesehen, konnte die Annotation in diesem Rahmen jedoch nur für zwei halbe Gespräche realisiert werden, was sich erst im Laufe dieser Arbeit herausstellte. Daher habe ich aus Konsistenzgründen auch diese Daten selbst annotiert. Die im Tutorium annotierten Ebenen wurden trotzdem übernommen und mit dem Suffix *_tut* kenntlich gemacht.

⁸Stuttgart-Tübingen-Tagset

(<http://www.ims.uni-stuttgart.de/forschung/ressourcen/lexika/TagSets/stts-table.html>, besucht am 08.11.2013).

⁹Hierfür besonderen Dank an Simon Sauer.

¹⁰Nach den BeMaTaC-Transkriptionsrichtlinien (<https://u.hu-berlin.de/bematac>, besucht am 11.11.2013).

se interpretiert werden, ist aber durch Annotation in der extralinguistischen Ebene laut den Transkriptionsrichtlinien von BeMaTaC gekennzeichnet. Auch Plosivverschlüsse sind problematisch. Für die Auswertung ist hier aber gerade die häsitationelle, kognitiv bedingte Pauseneinfügung von Interesse und nicht die durch den motorisch-artikulatorischen Apparat bedingte. Dies ist jedoch nicht immer eindeutig zu unterscheiden. Es erscheint dennoch wünschenswerter, eine möglichst exhaustive Annotation zu schaffen, die dann durch manuelle Einschränkung in der späteren statistischen Auswertung auch auf Effekte hinsichtlich dieses Schwellenwertes getestet werden kann.

2.2.2 Hinzugefügte Annotation

Für jedes Gespräch werden die Reparaturen nach den Annotationsrichtlinien im Anhang B auf den Ebenen *instructor_repair* bzw. *instructee_repair* und *instructor_subrep* bzw. *instructee_subrep* annotiert. Obwohl in den Annotationsrichtlinien noch die Annotationsart ‚Deletion‘ (<d>) für das Reparaturandum ausgewiesen ist, hat sich diese Kategorie als nicht zielführend herausgestellt und wurde nicht weiter verfolgt. Die *_repair*-Annotationen sind Spannenannotationen mit Bezug auf die jeweilige diplomatische Transkription (*instructor_dipl* und *instructee_dipl*), da das Reparaturschema unter jedem seiner Teile RD, IR und RS im Prinzip beliebig viele Token zulassen kann. Ein Reparaturandum wird durch den Tag <rd>, ein Interregnum durch den Tag <ir> und ein Reparans durch den Tag <rs> gekennzeichnet.

Die *_subrep*-Annotationen als Tokenannotationen in Bezug zur jeweiligen diplomatischen Transkription *_dipl* und zur jeweiligen Reparaturannotation *_repair* werden nur im Bereich der Reparans-Spanne <rs> annotiert, eben weil sie die Relation jedes Tokens im Reparans zum Reparaturandum deutlich machen.

Das typische Reparaturmuster folgt dem oben definierten Schema RD–IR–RS bzw. RD–RS. Aufgrund dieses starren Korsetts konnten fehlerhafte Annotationen mit dem Korpussuchtool ANNIS (Zeldes et al. 2009) schnell gefunden und aus der Analyse ausgeschlossen werden. Beispiele für die Anwendung der Annotationen finden sich im Abschnitt B auf Seite 75. Die qualitative Auswertung eines zufällig ausgewählten und von einer Zweitannotatorin nur anhand der Richtlinien annotierten Gesprächs zeugt davon, dass das zwei- bzw. dreiteilige Schema sowie die Subreparaturenkategorisierung gut nachvollzogen werden konnten¹¹.

Das zwei- bzw. dreiteilige Schema deckt auch einen wichtig Punkt einer empirischen Auswertung ab. Da für diese Arbeit schließlich Interregna als Einheit,

¹¹Im Rahmen dieser Arbeit konnte kein quantitatives Maß zur Qualitätssicherung bei der Korpusannotation verwendet werden.

in der sich Disfluencies manifestieren, ins Zentrum gerückt sind, wird auch die Analyse daraufhin ausgerichtet sein. Für eine vollständigen Auswertung darf jedoch nicht nur dann ein Interregnum annotiert werden, wenn es tatsächlich vorkommt, sondern muss auch dann annotiert werden, wenn es nicht vorkommt. Nichtexistierende Interregna werden mittels meines erarbeiteten theoretischen und annotatorischen Schemas mit RD – RS bzw. ⟨rd⟩ ⟨rs⟩ abgedeckt. Damit wird das Vorgehen für diese Variable dem Prinzip der *Accountability* gerecht – „reports of the occurrences of a variant must be accompanied by reports of all non-occurrences“ (Labov 2004: 7).

Neues Material im Reparans wird in meiner Annotation als Einfügung mit ⟨i1⟩, ⟨i2⟩ etc. getaggt. So können potenzielle Neustarts leicht gefunden und bei Bedarf näher untersucht werden. Nach einer ersten Einsicht in die Daten scheinen tatsächliche Neustarts jedoch seltener zu sein als Einfügungen¹² (vgl. Fox & Jaspersen 1995: 90, Typ E).

Die Subreparaturen erhalten eine Zählziffer, die für jedes Reparans feststellt, wie oft die annotierte Relation *r*, *s* oder *i* auftritt. So wird sowohl gezählt, wie häufig eine der drei Klassen in einem Reparans aufgetreten ist, als auch ein spezifisches Muster für ein einzelnes Reparans vergeben.

Wiederholungen, wie sie mit der *Floor-Holding*-Hypothese erklärt werden können, werden nicht annotiert (vgl. Bsp. 6–8 auf Seite 10). Um dem Phänomen der verschachtelten Reparatur wie in Bsp. 3 auf Seite 9 Rechnung zu tragen, ist eine Duplizierung der Reparatur-Ebene methodisch notwendig (vgl. Tabelle 2.1).

Tabelle 2.1: *Beispiel für die Verdopplung der Annotationsebenen `_repair` und `_subrep` für verschachtelte Reparaturen im Sinne des in Abschnitt 1.2.1 auf Seite 16 beschriebenen Schemas. Die `instructor_dipl`-Ebene enthält die Transkription.*

instructor_dipl	<i>aus der</i>	<i>aus dem</i>	<i>aus dem Fünfundreißiggradwinkel</i>		
instructor_repair	rd	rs			
instructor_subrep		r1	s1		
instructor_repair2		rd		rs	
instructor_subrep2			r1	r2	i1

Nach Abschluss dieser Arbeit werden die hier erarbeiteten Annotationsebenen mit einer CC-BY-Lizenz¹³ versehen und BeMaTaC zur Verfügung gestellt. Die neue Version mit den L2-Daten wird dann 2013-02 heißen.

¹²Für zukünftige Arbeiten lässt sich jedoch leicht eine neue Annotationsebene mit dieser Information in das Korpus einfügen.

¹³Creative Commons: Namensnennung (<http://creativecommons.org/licenses/by/2.0/de/>, besucht am 09.11.2013).

2.3 Verarbeitungspipeline

Die mit zusätzlichen Annotationen ergänzten Daten habe ich aus dem EXMARaLDA-Format mit SaltNPepper (Zipser & Romary 2010) inklusive der dazugehörigen Timeline in das relANNIS-Format gebracht und in ANNIS exploriert. Im Anschluss wurden die Treffer mit dem ANNIS-*GridExporter* exportiert und mit zwei Perl-Skripten¹⁴ und einem von mir erstellten *R*-Skript (R Core Team 2012) in eine auswertbare Tabelle konvertiert, wobei jede Reparatur-Instanz mit ihren relevanten Metadaten verknüpft ist.

Für Reparaturen mit Interregnum wurden über ANNIS Suchanfragen folgender Art gestellt und mit dem *GridExporter* gespeichert:

```
instructor_repair="rs" & instructor_repair="ir" &
instructor_repair="rd" & #1.#2 & #2.#3 &
meta::source="2011-12-14-A.TextGrid"
```

Für Reparaturen ohne Interregnum wurden über ANNIS Suchanfragen folgender Art gestellt und mit dem *GridExporter* gespeichert:

```
instructor_repair="rs" & instructor_repair="rd" & #1.#2 &
meta::source="2011-12-14-A.TextGrid"
```

Diese Suchanfragen wurden für jedes vorhandene Gespräch für alle Instructor- und Instructee-Ebenen variiert, inklusive der duplizierten Ebenen *_repair2* und *_subrep2*. Die exportierten Treffen enthielten außerdem weitere zur Auswertung benötigte Informationen, wie die Wortartenannotation *_pos* und die Timeline. Ein Beispiel für die Exportoptionen des *GridExporters* ist folgendes:

```
Left Context: 0
Right Context: 0
Annotation Keys: instructor_dipl, instructor_lemma, instructor_pos,
instructor_repair, instructor_subrep, time
```

Nach Generierung der Daten aus dem Korpus wurde manuell eine Ebene mit der Silbenanzahl für Reparanda und Reparantia hinzugefügt. Zur Auszählung wurde das hierarchische Silbenmodell – bestehend aus Onset, Koda und Reim – zugrunde gelegt (Vater 1992). Silbische Nasale wie in *n bisschen*, *so n* oder *mitm* wurden ebenso wie Abbrüche (*h/*, *b/* etc.) als eine Silbe gewertet.

¹⁴<http://strawberryperl.com>, besucht am 11.11.2013. Hierfür vielen Dank an Amir Zeldes.

Mithilfe von R wurde die jeweils vorhandene Tokenanzahl in RS und RD gezählt und gespeichert, sodass diese später in die Auswertung miteingehen kann.

Heuristisch muss jedes Reparandum und Reparans aus mindestens einem Token bestehen. Nach einer Konsistenzprüfung in R auf der Silben- und Subreparaturannotation wurde der L1-Datensatz um vier¹⁵ fehlerhafte Instanzen, der L2-Datensatz um drei¹⁶ fehlerhafte Instanzen reduziert, welche keinerlei Token enthielten.

¹⁵Davon drei auf *_repair* und eine auf *_repair2*.

¹⁶Davon zwei auf *_repair* und eine auf *_repair2*.

3 Auswertung

3.1 Ergebnisse

Die Darstellung der Ergebnisse erfolgt hauptsächlich anhand der Variablen *Fähigkeitsniveau* mit den Ausprägungen *L1/L2*, der Variablen *Rolle* mit den Ausprägungen *Instructor/Instructee*, der Variablen *Subreparatur* mit den Ausprägungen *i/r/s* und der Variablen *IR* mit den Ausprägungen *MIR¹/OIR²*.

Am Anfang der Auswertung erfolgt ein grobgerasteter, der Einfachheit halber übergeneralisierter Blick auf die Daten, indem die Gruppen *L1* und *L2* so behandelt werden, als ob sie homogen erscheinen. Dies erlaubt zunächst die Ausblendung der geringen Sprechermenge sowie ein Vergleichen der für die Auswertung relevanten Verteilungen mit einfachen χ^2 -Distributions-Test. In Abschnitt 3.1.5 erfolgt eine feinere, fast holistische Betrachtung mit dem individuellen Einfluss der Sprecher in einem Gemischten Modell. Die beiden möglichen Ansätze werden schließlich in Abschnitt 3.2.1 gegenübergestellt und diskutiert.

Tabelle 3.1: Übersicht der Datengrundlage aus dem Korpus *BeMaTaC* sowie der Frequenzen für Reparaturen mit und ohne Interregnum (*IR*). Die Zahlen in Klammern zeigen die relative Frequenz aller Reparaturen pro Zeile.

	OIR	MIR	Gesamt	Token	% Rep.	Dauer	Gespräche	Pbdn.
L1	141 (0,59)	98 (0,41)	239	11.192	0,02	66 min	12	16
L2	148 (0,58)	109 (0,42)	257	21.330	0,01	77 min	5	6

3.1.1 Interregna bei Muttersprachlern und Lernern

Tabelle 3.1 zeigt, dass die 16 Muttersprachler insgesamt weniger Reparaturen äußerten als die fünf Lerner (239 in L1 vs. 257 in L2). Das Verhältnis von Reparaturen mit *IR* zu Reparaturen ohne *IR* ist sehr ähnlich. Obwohl nur fünf Lerner 16 Muttersprachlern gegenübergestellt werden, ist die absolute

¹Mit Interregnum.

²Ohne Interregnum.

Reparaturfrequenz der Lerner höher. Vor einer voreiligen Analyse muss jedoch bedacht werden, dass die Lerner wesentlich längere Gespräche führten als die Muttersprachler und die Datengrundlage der Lerner fast doppelt so viele Token umfasst wie die der Muttersprachler.

Hypothese 1.1 stellt Vorhersagen für die IR-Häufigkeit bei Muttersprachlern auf. Die Nullhypothese besagt, dass die Wahrscheinlichkeit, dass ein L1-Sprecher mehr Reparaturen mit IR als ohne IR äußert, höchstens so groß ist, wie die Wahrscheinlichkeit, dass ein L1-Sprecher mehr Reparaturen ohne IR als mit IR äußert.

Insgesamt äußerten 12 von 16 Muttersprachlern mehr Reparaturen ohne IR als mit IR. Abbildung 3.1 verdeutlicht die Verteilung, wobei die Breite der Säulen die Anzahl der Instanzen eines Sprechers visualisiert; die exakten Aufschlüsselungen je Proband befinden sich in Tabelle A.1 auf Seite 66. Die Überprüfung der Hypothese 1.1 erfolgte mittels eines gerichteten Binomialtests. Die zu überprüfende gerichtete Nullhypothese (Ohne IR \leq Mit IR) ergibt vier Erfolge bei 14 Versuchen.³ Der Test erbrachte ein eindeutig nicht signifikantes Ergebnis ($p_{gerichtet} = 0,97$). Damit kann die Nullhypothese nicht abgelehnt werden.

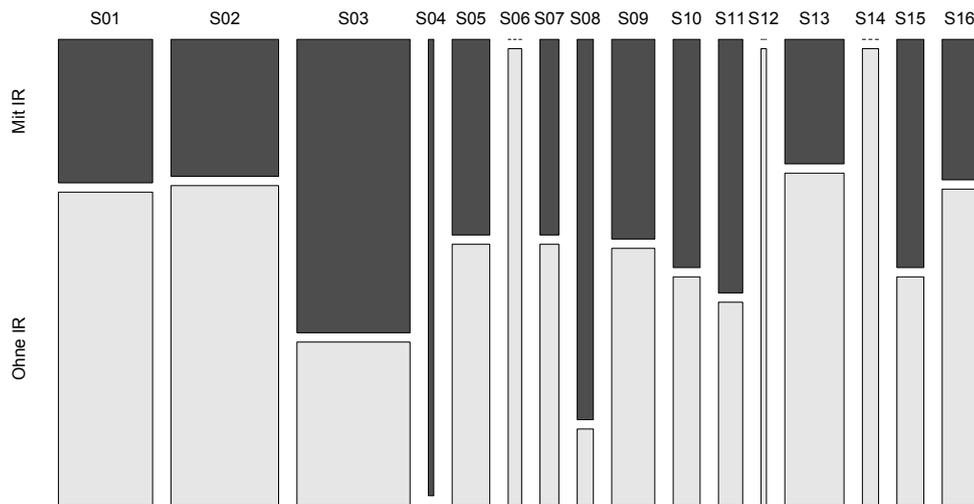


Abbildung 3.1: Übersicht über die Verteilung von Reparaturen mit und ohne Interregnum (IR) der einzelnen muttersprachlichen Sprecher (S01–S16) in BeMaTaC. Die Blöcke stellen die individuelle Gesamtzahl der Reparaturen eines Sprechers mit und ohne IR dar. Je breiter ein Block, desto mehr Reparaturen äußert dieser Sprecher im Verhältnis zu den anderen.

³Von den 16 Probanden wurden zwei ausgeschlossen, da sie jeweils gleich viel Reparaturen mit und ohne Interregnum äußerten ($OIR = IR$).

In Hypothese 1.2 wird die gerade betrachtete Frage noch einmal für die Lerner gestellt. Der Nullhypothese entspricht die Formulierung, dass die Wahrscheinlichkeit, dass ein L2-Sprecher mehr Reparaturen mit IR als ohne IR äußert, höchstens so groß ist wie die Wahrscheinlichkeit, dass ein L1-Sprecher mehr Reparaturen ohne IR als mit IR äußert.

Die sechs Lerner äußerten 257 Reparaturen, davon 148 ohne Interregnum. Zwei Lerner äußerten mehr Reparaturen mit Interregnum. Die Nullhypothese 1.2 kann nach einem gerichteten Binomialtest bei zwei Erfolgen und 6 Versuchen jedoch wiederum nicht abgelehnt werden ($p_{gerichtet} = 0,89$). Abbildung 3.2 verdeutlicht die Verteilung. Die aufgeschlüsselten Werte je Proband befinden sich in Tabelle A.2 auf Seite 66.

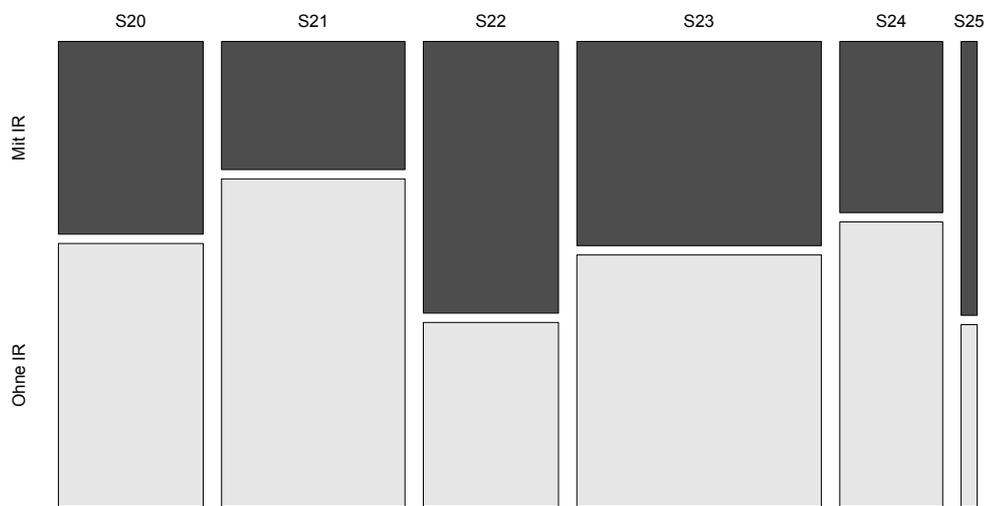


Abbildung 3.2: Übersicht über die Verteilung von Reparaturen mit und ohne Interregnum (IR) der einzelnen fortgeschrittenen lernersprachlichen Sprecher (S20–S25) in BeMaTaC. Die Blöcke stellen die individuelle Gesamtzahl der Reparaturen eines Sprechers mit und ohne IR dar. Je breiter ein Block, desto mehr Reparaturen äußert dieser Sprecher im Verhältnis zu den anderen.

In Hypothese 1.3 wird schließlich untersucht, ob sich Muttersprachler und Lerner signifikant unterschiedlich verhalten, was die Variation der Reparaturen angeht. Die Nullhypothese geht davon aus, dass sich Muttersprachler und Lerner nicht verschieden verhalten: Muttersprachler und fortgeschrittene Lerner unterscheiden sich nicht in der Häufigkeit ihrer verwendeten Interregna dergestalt, dass Lerner höchstens so viele MIR-Reparaturen produzieren wie Muttersprachler. Die Nullhypothese kann hier jedoch nicht abgelehnt werden, da die Frequenzverteilung (wie in Abbildung A.1 auf Seite 62 dargestellt) nicht signifikant von der erwarteten Verteilung abweicht ($\chi^2=0,1$; $df = 1$; $p = 0,75$).

3.1.2 Gefüllte Pausen im Interregnum

Zur Überprüfung der Hypothese 2 (inwiefern unterscheiden sich die Lerner und Muttersprachler bezüglich der Einfügung von gefüllten Pausen in ein Interregnum) wurden alle in den Interregna vorkommenden gefüllten Pausen der Form *äh* und *ähm* gezählt. Für L2 wurden außerdem die davon abweichenden Formen *eh*, *ahm* und *hh* mit aufgenommen. Dies ist möglich, da die Hypothese keine Bedeutungsunterschiede der äußeren Formen postuliert, und somit noch nicht muttersprachlich angepasste Formen in die Analyse miteinbezogen werden können. Tabelle A.3 auf Seite 66 enthält die Ergebnisse je Sprecher für L1 und L2. Das Bild ist sehr homogen: Muttersprachler äußerten insgesamt 28, Lerner 32 Formen einer gefüllten Pause. Die statistische Überprüfung mittels eines generalisierten linearen Modells mit der Anzahl der gefüllten Pausen als abhängiger Variable und dem Fähigkeitsfaktor als Prädiktor ergibt keinen signifikanten Unterschied ($p = 0,82$), s. auch Tabelle 3.2.

Tabelle 3.2: *Ergebnis des linearen Modells für L1 und L2 für gefüllte Pausen im Interregnum als abhängige Variable und Fähigkeitsfaktor als Prädiktor.*

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-3,24	0,19	-16,83	0,00
LL2	0,06	0,26	0,23	0,82

3.1.3 Rolle in der Map Task

Hypothese 3 untersucht die Auswirkung der in der Wegbeschreibung eingenommenen Rollen der Sprecher auf die Äußerung eines Interregnums innerhalb einer Reparatur. Diese Rollen führen zu einer unterschiedlichen Reparaturanzahl in den beiden Gruppen. Die Präferenz ist gemäß einem χ^2 -Unterschiedstest signifikant ($\chi^2 = 19,2$; $df = 1$; $p < 0,001$). Muttersprachler reparierten als Instructor signifikant häufiger als Lerner (s. auch Abbildung A.5 auf Seite 67). Ein genaueres Bild ergibt sich jedoch aus der Analyse der Variable *Interregnum* in der Bedingung *MIR/OIR* (Abbildung 3.3 auf Seite 38). Aus der Abbildung wird graphisch ersichtlich, was aus den absoluten Häufigkeiten (in Tabelle A.4 auf Seite 67) inferiert werden kann: der Unterschied zwischen L1 und L2 besteht auch fort in der Bedingung mit zuvor geäußertem Interregnum (Quadrant *L1/MIR* vs. *L2/MIR*; $\chi^2 = 14,8$; $df = 1$; $p < 0,001$). Als signifikant unterschiedlich erweist sich auch der Vergleich von L1 und L2 in der Bedingung ohne zuvor geäußertes Interregnum (Quadrant *L1/OIR* vs. *L2/OIR*, $\chi^2 = 5,7$; $df = 1$; $p = 0,01$).

Innerhalb des **Fähigkeitsniveaus** verwenden die Muttersprachler in der Instructor- und in der Instructee-Bedingung signifikant häufiger Reparaturen ohne Interregnum ($\chi^2 = 6,6; df = 1; p = 0,01$). Bei den Lernern hingegen trägt der Faktor **Rolle** nicht signifikant dazu bei, ob eine *MIR*- oder *OIR*-Reparatur geäußert wird ($\chi^2 = 0,32; df = 1; p = 0,56$). Abbildung 3.3 verdeutlicht auch die ähnliche (weil nicht signifikante) Verhaltensweise von Muttersprachlern in der Bedingung *OIR* und Lernern in der Bedingung *MIR* bezüglich der Rollenverteilung (vgl. *L1/OIR* vs. *L2/MIR*; $\chi^2 = 2,3; df = 1; p = 0,12$).

3.1.4 Subreparaturen

In Hypothese 4 wird nach dem Einfluss der Subreparaturtypen auf die Verwendung eines Interregnums innerhalb einer Reparatur gefragt. Abbildung 3.4 auf Seite 39 zeigt die Verteilung der Subreparaturen für L1 und L2. Die Verteilungen unterscheiden sich signifikant je Fähigkeitsfaktor ($\chi^2 = 45,8; df = 2; p < 0,001$). Jedoch verhalten sich die Kategorien *s* und *r* einander nicht signifikant unähnlich ($\chi^2 = 0,8; df = 1; p = 0,35$). Neben der semantischen Motivierung (in der Diskussion, Abschnitt 3.2.1 auf Seite 47) spricht dies zusätzlich nicht gegen ein Zusammenfassen der Kategorien *r* und *s* in eine neue Kategorie *rs*, so dass hier nur noch zwei Variablenausprägungen kontrastieren, nämlich *rs* und *i*, welche sich wiederum signifikant unterscheiden ($\chi^2 = 43,5; df = 1; p < 0,001$).

Die mit dem STTS-Tagset annotierten Wortarten wurde aufgrund der oben beschriebenen Literatur reduziert auf einen Faktor mit zwei Ausprägungen: Inhalts- (*C*) und Funktionswort (*F*).⁴ Abbildung A.7 auf Seite 69 zeigt die unterschiedliche Verteilung der Interaktionen zwischen den semantischen Beziehungen der Reparaturteile *i* und *rs* und den verwendeten Wortartenhauptkategorien *C* und *F*. Diese ist signifikant unterschiedlich ($\chi^2 = 40,6; df = 3; p < 0,001$). Muttersprachler fügen neue Informationen im Reparans in 52% der Fälle als Inhaltswort ein, während Lerner dies nur in 33% ihrer Reparaturen tun (s. auch Tabelle A.5 auf Seite 72). Auch dieser Unterschied ist signifikant ($\chi^2 = 13,9; df = 1; p < 0,001$).

Ein weitaus interessanteres Bild ergibt sich, wenn man die Interaktionen nicht nur für die Bedingung **Fähigkeitsniveau**, sondern auch für die Bedingung **Interregnum ja/nein** betrachtet (s. Abbildung 3.5 auf Seite 41). Die Muster *L1/OIR* und *L1/MIR* verhalten sich signifikant verschieden ($\chi^2 =$

⁴„Inhalt“ als Merkmal offener Wortklassen wird dabei den Tags NN, NE, ADJA, ADJD, ADV, VVFIN, VMFIN, VVIMP und VVINFIN zugesprochen – die restlichen Wortarten fallen in die Residualkategorie Funktionswort. Auxilia wurden vorerst nicht miteinbezogen. Partizipien und Modalverben im Infinitiv kommen in keiner der annotierten Reparantia vor.

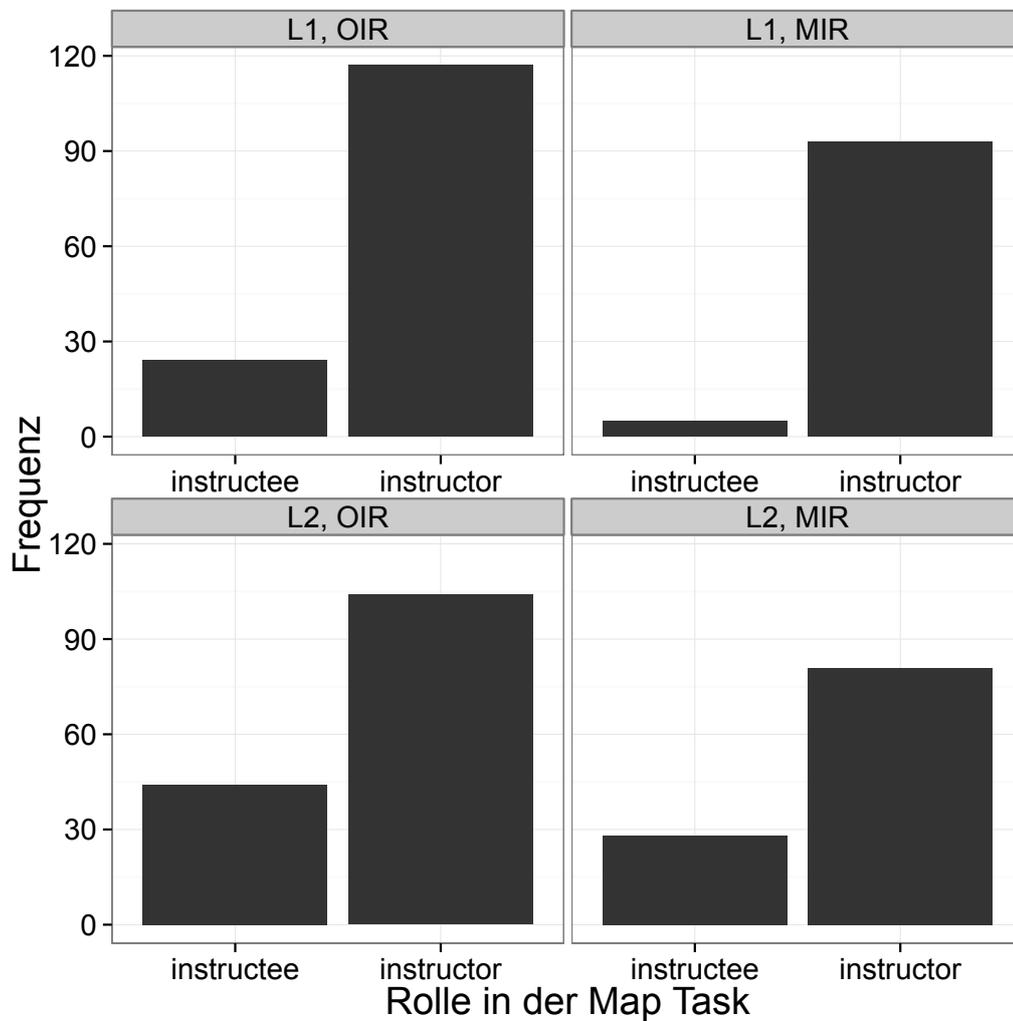


Abbildung 3.3: Rollenverteilung der Reparaturen für Muttersprachler (L1) und fortgeschrittene Lerner (L2) für Reparaturen mit Interregnum (MIR) und ohne Interregnum (OIR) in BeMaTaC. Instructor/Instructee entsprechen den in der Map Task eingenommenen Rollen.

	L1	L2
i	250 (0.41)	88 (0.21)
r	229 (0.37)	197 (0.47)
s	132 (0.22)	133 (0.32)

Abbildung 3.4: Verteilung der Subreparaturen für Muttersprachler (L1) und fortgeschrittene Lerner (L2) in BeMaTaC. Die Subreparaturkategorien relatieren die Tokens im Reparans zum Reparandum, Kategorien sind r – Repetition, s – Substitution und i – Insertion.

24,2; $df = 3$; $p < 0,001$). In der *OIR*-Bedingung tendieren L1-Sprecher dazu, Inhaltswörter eher in der Kategorie Wiederholung/Substitution zu äußern (87 vs. 68), während sie in der *MIR*-Bedingung Inhaltswörter eher mit der Kategorie einer Insertion verwenden (79 vs. 51). Dieser Unterschied ist signifikant ($\chi^2 = 7,4$; $df = 1$; $p < 0,01$). Dieser Unterschied besteht für Lerner nicht ($L2/OIR/i.C$ 22 vs. $L2/OIR/rs.C$ 62 und $L2/MIR/i.C$ 37 vs. $L2/MIR/rs.C$ 56; $\chi^2 = 3,0$; $df = 1$, $p = 0,07$). Das Gesamtmuster $L2/OIR$ vs. $L2/MIR$ verhält sich dennoch signifikant verschieden ($\chi^2 = 10,4$; $df = 3$; $p < 0,02$).

Lerner tendieren dazu, wenn man nur die Kategorie *rs* betrachtet, in *OIR/rs.F&rs.C* mehr Repetitionen/Substitutionen (173) zu äußern als in *MIR/rs.F&rs.C* (136). Dieser Unterschied ist signifikant ($\chi^2 = 4,4$; $df = 1$; $p < 0,04$). Aus Tabelle A.5 auf Seite 72 geht hervor, dass Lerner mit IR 136 Mal in *rs* reparieren, aber 173 Mal ohne IR, also häufiger ohne IR als mit IR.

Jedoch scheint es, als ob L2-Sprecher in der *MIR*-Bedingung ein Reparaturverhalten zeigen, das dem von L1 in der *OIR*-Bedingung sehr ähnelt ($\chi^2 = 0,24$; $df = 1$; $p = 0,6$). Der χ^2 -Test zeigt hier, dass das Verhältnis zwischen *i.C* und *rs.C* in der $L1/OIR$ -Bedingung (2. Quadrant des Graphen in Abbildung 3.5) dem zwischen *i.C* und *rs.C* in der $L2/MIR$ -Bedingung (4. Quadrant des Graphen) ähnlich ist. Dies betrifft freilich nicht nur die Verhältnisse für die *C*-Kategorie: die beiden Quadranten $L1/OIR$ und $L2/MIR$ sind sich in allen Ausprägungen sehr ähnlich ($\chi^2 = 0,9$; $df = 3$; $p = 0,8$).

3.1.5 Limitationen und Mixed Models

Die bisher durchgeführte Analyse vernachlässigt, dass die individuellen Sprechereigenschaften, wie die Sprechgeschwindigkeit oder die Reparaturfrequenz, nicht ausreichend einfließen. Sprecher können trotz eines gemeinsamen Fähigkeitsniveau stark unterschiedliche Sprechweisen aufweisen. Dies muss berücksichtigt werden, wenn man Aussagen über die Grundgesamtheit einer Sprechergemeinschaft treffen möchte. Haben wenige Sprecher einen großen Anteil an den erzeugten Daten, so besteht die Gefahr, dass die Analyse nicht – mit dem üblichen Bedacht – auf eine hypothetisch vorhandene Grundgesamtheit generalisiert werden kann, sondern vielmehr fälschlicherweise die Eigenschaften der einzelnen Sprecher projiziert.

Daher werden nun explizit die Verhaltensweisen der einzelnen Sprecher betrachtet, weil die von einem Probanden geäußerten Reparaturen womöglich nicht unabhängig voneinander sind (eben weil Sprecher zu bestimmten idiosynkratischen Sprechweisen tendieren). Für Hypothese 1.1 bestätigt ein χ^2 -Test der L1-Daten (Tabelle A.1 auf Seite 66), dass die Sprecher sich signifikant verschieden verhalten ($\chi^2 = 32,9$; $df = 15$; $p < 0,01$).⁵ Auch in Hypothese 1.2 erwartet man idiosynkratische Verhaltensweisen der Lerner. Ein χ^2 -Test der Verteilung der L2-Daten ergibt kein signifikantes Ergebnis ($\chi^2 = 10,9$; $df = 5$; $p = 0,052$).⁶ Der p -Wert ist jedoch zu klein, als dass man wirklich davon überzeugt sein kann, dass die Sprecher sich gleich verhalten. Alles weist darauf hin, dass die L2-Sprecher sich – wie die L1-Sprecher – unterschiedlich verhalten.

Man muss also zumindest davon ausgehen, dass die Muttersprachler sich in ihren Reparatureigenschaften unterschiedlich verhalten, und dass man mit mehr Lernern möglicherweise verlässlichere Nachweise für eine unterschiedliche Verhaltensweise erhalten wird. Dies indiziert eine weitere Analyse für die individuellen Einflüsse der einzelnen Sprecher. Die Auswirkungen der einzelnen Probanden und ihrer Gewichtung für meine Ergebnisse werden exemplarisch im Anhang nochmals veranschaulicht, indem die in Abschnitt 3.1 vorgestellten Graphen der Hypothese 3 und 4 mit dem jeweiligen Anteil der einzelnen Sprecher dargestellt werden: für Hypothese 3 in der Abbildung A.6 auf Seite 68 und für Hypothese 4 in der Abbildung A.8 auf Seite 70 und der Abbildung A.9 auf Seite 71.

⁵Da die χ^2 -Näherung aufgrund einiger Null-Datenpunkte der Instructors/Instructees infrage gestellt werden kann, wird das Ergebnis abgesichert durch 2000 Monte-Carlo-Simulationen nach Hope (1968): $\chi^2 = 32,9$; $df = \text{NA}$; $p < 0,01$.

⁶Auch nach 2000 Monte-Carlo-Simulationen nicht ($\chi^2 = 10,9$; $df = \text{NA}$; $p = 0,05$).

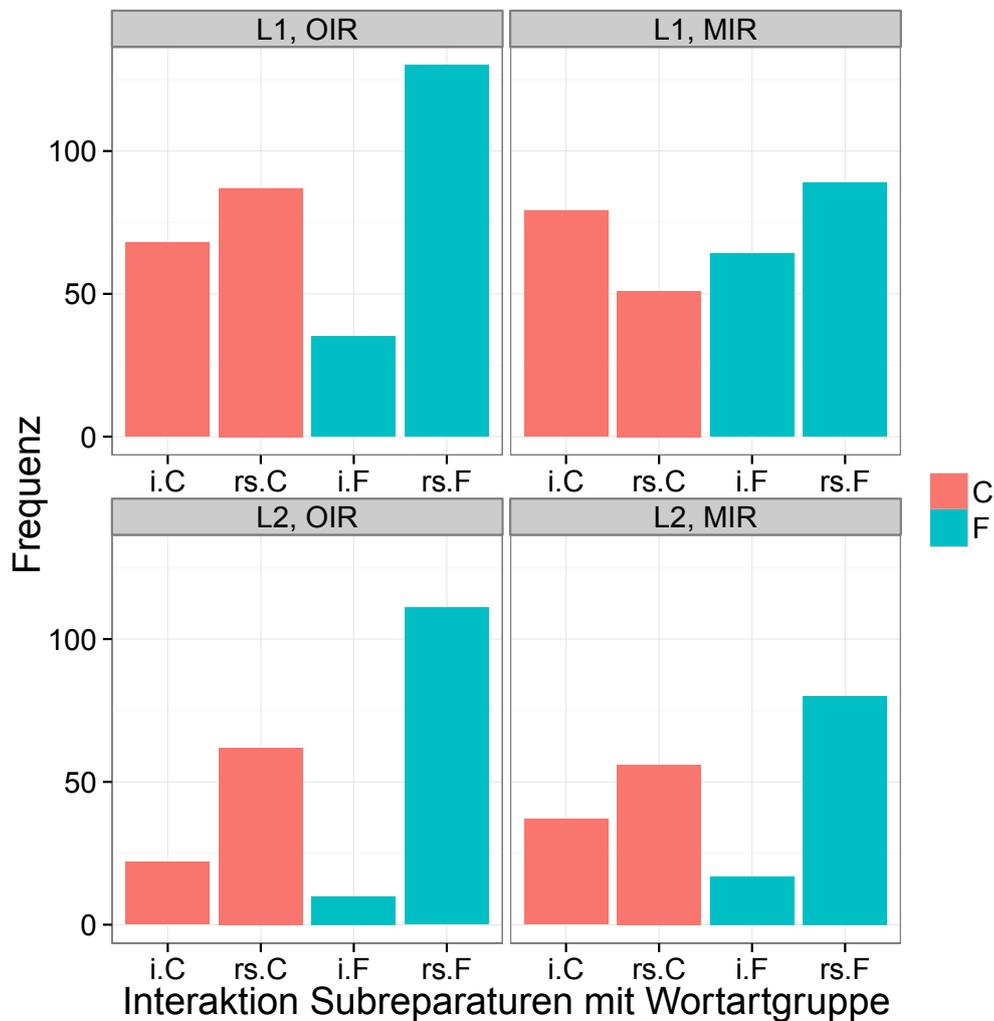


Abbildung 3.5: Interaktion der Subreparaturen *i* (Insertion) und *rs* (Repetitionen/Substitutionen) mit Inhaltswortarten (*C*) und Funktionswortarten (*F*). Die Balken zeigen die Frequenzen für jede Interaktion in der Bedingung *L1* vs. *L2* und *OIR* vs. *MIR* aller Reparaturen in *BeMaTaC*.

Um meine Analyse durch Modelle abzusichern, die auf Kosten höherer Komplexität die Varianz zwischen den Sprechern besser berücksichtigt, verwende ich das *lme4*-Paket (Bates et al. 2012) in *R*, um ein Gemischtes Modell (bzw. *Linear-Mixed-Effects*-Modell oder *LME*-Modell) zu rechnen. Während der χ^2 -Test immer nur zwei Verteilungen miteinander vergleichen kann, bieten *LME*-Modelle den Vorteil, dass sie mehrere Faktoren gleichzeitig auswerten können. Als feste Effekte (*fixed effects*) gehen die Faktoren ein, von denen man annimmt, dass sie reproduzierbare Prädiktoren für das untersuchte Phänomen sind. Zufällige Effekte (*random effects*) können zusätzlich einfließen, so dass hier die unabhängigen Eigenschaften eines jeden Sprechers für sich genommen und nach der relativen Größe dieses Sprechers in den Daten berücksichtigt werden kann.

Mit dem Modell untersuche ich hier alle Faktoren, die möglicherweise einen Einfluss auf das Vorkommen eines Interregnums haben. In das Modell gehen also als fixe Effekte die Prädiktoren **Fähigkeitsfaktor**, **Rolle**, **Subreparatur** und **Wortart** mit ein. Als zufällige Effekte werden die einzelnen Sprecher gewertet. Zu Beginn der Auswertung steht ein volles Modell mit dreifachen Interaktionen der fixen Effekte, welches schrittweise anhand einer *backward selection* auf Grundlage des AIC (*Akaike Information Criterion*, vgl. Akaike 1974) reduziert wurde. Dies ist notwendig, um genau die Anzahl bedingender Faktoren zu erhalten, welche im Sinne von Ockhams Rasiermesser – *entia non sunt multiplicanda praeter necessitatem*⁷ – die Daten am besten beschreiben. Die *random slopes* jedes Prädiktors, d. i. die individuelle Auswirkung jedes Sprechers auf einen der Prädiktoren, wurden sukzessive mit vergleichenden Log-Likelihood-Tests getestet.

Am Ende der Auswahlprozesse steht ein Modell, welches als *random slope* den Faktor **Rolle** in Abhängigkeit eines einzelnen Sprechers miteinbezieht. Als fixe Effekte bilden tatsächlich nur zwei Faktoren die Daten am besten nach: **Rolle** und **Subreparatur**. Tabelle 3.3 auf der nächsten Seite zeigt die erwarteten Werte für eine Reparatur mit Interregnum. Das Modell zeigt, welche der untersuchten Faktoren in welcher Ausprägung denn nun einen signifikanten positiven oder negativen Anteil daran haben, ob ein Interregnum geäußert wird oder nicht. Als Prädiktor mit überzufälligem Einfluss auf die Variantenausprägung *MIR* des Faktors **Interregnum** hat sich nur der Faktor **Subreparatur** in der Ausprägung *rs* herausgestellt. Der Faktor **Rolle** erreicht das Signifikanzniveau nicht.

Für die Messungen der zeitlichen Dauer der Reparaturteile wurden drei Maße erhoben und miteinander verglichen. Die Äußerungsdauer wurde automa-

⁷Es ist unklar, ob Ockham dies je so äußerte. Der Sinnspruch bringt Ockhams Heuristik jedoch gut auf den Punkt.

Tabelle 3.3: Ergebnis der Schätzungen der Prädiktoren für die Wahrscheinlichkeit eines Interregnums in einem schrittweise reduzierten LME-Modell mit den fixen Effekten *Rolle* und *Subreparaturen* sowie dem zufälligen Effekt der Rolle in Abhängigkeit eines Sprechers in *BeMaTaC*.

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0,59	0,63	-0,94	0,35
roleinstructor ^a	1,09	0,62	1,75	0,08
value.rsrs ^b	-0,74	0,15	-4,78	0,00 ***

Signifikanzen: $p < 0,001$ ***

^a Einfluss der *Rolle* in der Ausprägung *Instructor*.

^b Einfluss der *Subreparaturen* in der Ausprägung *rs*.

tisch mithilfe der Timeline ermittelt. Die Tokenlänge gibt Aufschluss über die Anzahl der geäußerten Tokens und wurde automatisch mit R berechnet. Schließlich wurde für jede Reparatur auch die Silbenlänge je Einheit ermittelt (vgl. Abschnitt 2.3). Jeder der RS-Maße korreliert in einem mittleren bis hohen Bereich positiv mit seinem RD-Pendant. Bei den muttersprachlichen Reparaturen ist die Korrelation insgesamt etwas höher ausgeprägt (s. Tabelle 3.4) als bei den lernersprachlichen Reparaturen (Tabelle 3.5 auf der nächsten Seite). Sowohl für L1 als auch für L2 scheinen Reparantia dazu zu tendieren, länger als ihr jeweiliges Reparaturandum zu sein. Anschaulich wird dies dargestellt in den Abbildungen A.2 bis A.4 auf den Seiten 63–65.

Tabelle 3.4: Korrelationen zwischen Tokenlänge, Silbenlänge und Äußerungsdauer zwischen Reparaturandum (RD) und Reparans (RS) innerhalb der Reparaturen von Muttersprachlern in *BeMaTaC*.

	RS-Tokens	RS-Silben	RS-Zeit
RD-Tokens	0,69	0,62	0,55
RD-Silben	0,59	0,61	0,54
RD-Zeit	0,47	0,52	0,54

3.2 Diskussion

Die ursprünglich geplante Subreparatur-Annotation für Löschungen im Reparaturandum ($\langle\langle d\# \rangle\rangle$) wurde nicht in die Analyse miteinbezogen, da erstens kaum Instanzen annotiert wurden und zweitens die Aussagekraft für eine Analyse zu gering erschien. Die Nummerierung der Subreparaturen nach der in einem

Tabelle 3.5: Korrelationen zwischen Tokenlänge, Silbenlänge und Äußerungsdauer zwischen *Reparandum (RD)* und *Reparans (RS)* innerhalb der Reparaturen von fortgeschrittenen Lernern in *BeMaTaC*.

	RS-Tokens	RS-Silben	RS-Zeit
RD-Tokens	0,57	0,49	0,38
RD-Silben	0,51	0,62	0,43
RD-Zeit	0,44	0,38	0,38

Reparans vorkommenden Tokenstelle (bspw. ⟨r1⟩ ⟨i1⟩ ⟨r2⟩) ist zwar im Korpus annotiert, hat sich für eine Auswertung im Rahmen der hier darstellbaren Möglichkeiten jedoch als zu mächtig herausgestellt. Für die Ergebnisse wurde daher die Ziffer an der zweiten Stelle dieser Annotationen ignoriert.

3.2.1 Implikationen der Ergebnisse für die Erwartungen

Interregna bei Muttersprachlern und Lernern

Die Nullhypothese 1.1 (die Wahrscheinlichkeit, dass ein L1-Sprecher mehr Reparaturen mit IR als ohne IR äußert, ist höchstens so häufig wie die Wahrscheinlichkeit, dass ein L1-Sprecher mehr Reparaturen ohne IR als mit IR äußert bzw. $P(SIR) \leq 0,5$) kann nicht abgelehnt werden.⁸ Es ist somit wahrscheinlicher, dass ein L1-Sprecher in meiner Untersuchung weniger Reparaturen mit IR äußert als Reparaturen ohne IR ($MIR < OIR$). Anders formuliert: Muttersprachler reparieren im *BeMaTaC*-Korpus häufiger direkt, ohne eine Häsitationsphase als Verzögerung zu nutzen. Da die Erwartungen an die Verhaltensweisen der Daten heuristisch als Alternativhypothese aufgestellt werden, entspricht die Tatsache, dass die Nullhypothese eben gerade nicht abgelehnt werden kann, nicht meinen Erwartungen. Eine plausible Erklärung für dieses Verhalten ist möglicherweise die hohe Automatisierung in der Sprachproduktionsplanung von Muttersprachlern, die auf diese Weise auch Reparaturen sehr schnell planen können.

Die Nullhypothese 1.2 (die Wahrscheinlichkeit, dass ein L2-Sprecher mehr Reparaturen mit IR als ohne IR äußert, ist höchstens so häufig wie die Wahrscheinlichkeit, dass ein L2-Sprecher mehr Reparaturen ohne IR als mit IR äußert bzw. $P(SIR) \leq 0,5$) kann nicht abgelehnt werden.⁹ Es ist somit wahrscheinlicher, dass ein Lerner in meiner Untersuchung weniger Reparaturen mit IR äußert als Reparaturen ohne IR ($MIR < OIR$). Dies wiederum bedeutet,

⁸Vgl. Abschnitt 3.1.1 auf Seite 33, Abbildung 3.1 auf Seite 34 und Tabelle A.1 auf Seite 66.

⁹Vgl. Abschnitt 3.1.1 auf Seite 33, Abbildung 3.2 auf Seite 35 und Tabelle A.2 auf Seite 66.

dass Lerner, wie oben schon bei den Muttersprachlern nachgewiesen wurde, auch häufiger direkt reparieren, ohne ein Interregnum einzusetzen. Dies entspricht nicht der in der Alternativhypothese formulierten Erwartung, weswegen die Nullhypothese auch nicht abgelehnt werden kann. Lerner verhalten sich hier also ähnlich wie die Muttersprachler, was nicht erwartet wurde, da von einem geringeren Automatisierungsgrad bei der L2-Sprachplanung auszugehen ist. Eine Erklärung dieses Ergebnisses könnte zum einen die Stichprobengröße von nur sechs lernersprachlichen Sprechern sein, die möglicherweise erweitert werden muss, um die Nachhaltigkeit der Tests zu prüfen. Hierfür müsste das BeMaTaC-Korpus mit weiteren L2-Gesprächen ergänzt werden. Für diese Erklärung spricht insbesondere, dass der χ^2 -Distributions-Test gegen idiosynkratische Verhaltensweisen der Lerner spricht (vgl. Abschnitt 3.1.5), was – wie man bei den Muttersprachlern gesehen hat – jedoch nicht der Realität einer Sprechergrundgesamtheit entsprechen kann. Die Muttersprachler verhalten sich im gleichen Test tatsächlich signifikant unterschiedlich.

Falls jedoch mit einem erweiterten Korpus, also einer größeren Stichprobe, und den dann ausgeräumten statistischen Vorbehalten die Nullhypothese 1.2 wiederum nicht abgelehnt werden kann, so würde dies implizieren, dass die Sprachkenntnisse der Lerner (die ja auf mindestens C1 getestet wurden) tatsächlich schon so weit fortgeschritten sind, dass keine bedeutenden Unterschiede in dem (oberflächlichen) Äußerungsverhalten auftreten. Dies ließe eine ebenfalls höhere, fast schon muttersprachliche Automatisierung der Sprachproduktion annehmen. Falls jedoch die Ähnlichkeit der L2-Sprecher untereinander in einer größeren Sprecherstichprobe fortbestünde, so kann dies dahingehend interpretiert werden, dass die Fernwirkung desselben Herkunftssprachsystems (hier Englisch) der Lerner in salienter Weise persistiert.

Die Nullhypothese 1.3 (Muttersprachler und fortgeschrittene Lerner unterscheiden sich nicht in der Häufigkeit ihrer verwendeten Interregna) kann wiederum nicht abgelehnt werden.¹⁰ L1- und L2-Sprecher in dieser Untersuchung unterscheiden sich folglich nicht signifikant in der Häufigkeit ihrer verwendeten Interregna, im Gegenteil: beide Gruppen verhalten sich fast identisch (vgl. Abbildung A.1 auf Seite 62). In jedem Falle sollte dieses Ergebnis auf seine Replizierbarkeit mit einer größeren L2-Stichprobe überprüft werden. Sollte sich die größere Stichprobe konsistent zu diesem Ergebnis verhalten, so impliziert dies, dass die L2-Sprecher auf einem muttersprachlichen Niveau mit der ihnen hier gestellten Wegbeschreibungsaufgabe umgehen können. Zumindest für ihre mit dem Schema *Reparandum – Interregnum – Reparans* gemessenen Verhaltensweisen kann dann kein Unterschied zu Muttersprachlern mehr festgestellt werden. Dies widerspricht den Vorhersagen von Declerck & Kormos (2012),

¹⁰Vgl. Abschnitt 3.1.1 auf Seite 33 und Abbildung A.1 auf Seite 62

welche interessanterweise in ihrer Studie ebenfalls C1-Lerner untersuchten.

Gefüllte Pausen im Interregnum

Die Nullhypothese 2 (fortgeschrittene Lerner verwenden höchstens gleich häufig gefüllte Pausen im IR wie Muttersprachler) kann nicht abgelehnt werden.¹¹ Demnach verwenden Lerner nicht häufiger gefüllte Pausen innerhalb eines Interregnums, wie es in der Alternativhypothese 2 als Erwartungshaltung formuliert wurde. Ihr Disfluency-Verhalten scheint also bei einer genaueren (exemplarischen) Betrachtung der Disfluencies zwischen Reparaturandum und Reparans nicht von dem muttersprachlichen Frequenzniveau abzuweichen.

Dass die Frequenzen der gefüllten Pausen im Interregnum sich nicht unterscheiden, ist unerwartet. Mehrere Schlüsse drängen sich daher auf: Zum einen ist es möglich, dass das experimentelle Design nicht stimulierend genug war. Eine Kontrollstudie in der englischen Ausgangssprache der Sprecher könnte hierzu Aufschluss geben. Weiterhin wäre denkbar, dass sich das festgestellte Mindestniveau C1 bzgl. der Häsitationspausen nicht mehr signifikant von dem der Muttersprachler unterscheidet, Lerner gefüllte Pausen vor Reparantia also zumindest frequenzbasiert in ähnlicher Weise wie Muttersprachler verwenden. Gestützt würde diese Annahme von den ebenso ähnlichen Ergebnissen der oben getesteten Hypothesen 1.1–1.3.

Rolle

Die Nullhypothese 3 (die Verteilung der Reparaturen zwischen *L1* und *L2* bezüglich ihrer Rolle ist gleich) muss nach einem χ^2 -Unterschiedstest abgelehnt werden.¹² Dies zieht ein Eintreten der Alternativhypothese nach sich, die besagt, dass die Verteilung der Reparaturen zwischen *L1* und *L2* bezüglich ihrer Rolle verschieden sei. Impliziert wird hiermit, dass sich die in der Map-Task eingenommene Rolle für Muttersprachler und Lerner also verschieden auf ihr Reparaturverhalten auswirkt. Selbst wenn sich Lerner in einer Rolle befinden (nämlich als Instructee), in der sie per se weniger Sprache planen und produzieren müssen, verhalten sie sich doch anders als Muttersprachler, denn sie reparieren in 28 % aller Reparaturen als Instructee, verglichen mit 12 % der L1-Reparaturen.

Die explorierende Post-hoc-Analyse der Faktoren **Interregnum**, **Rolle** und **Fähigkeitsniveau** (Abbildung 3.3 auf Seite 38) hat ein detailliertes Bild zu Tage gefördert. Die Ergebnisse deuten an, dass die Unterschiede im Reparaturverhalten bezüglich der Rollenverteilung in den Kombinationen *L1/OIR* vs.

¹¹Vgl. Abschnitt 3.1.2 auf Seite 36, Tabelle 3.2 auf Seite 36 und Tabelle A.3 auf Seite 66.

¹²Vgl. Abschnitt 3.1.3 auf Seite 36, Abbildung A.5 auf Seite 67 Tabelle A.4 auf Seite 67.

L2/OIR (d. i. die Variante ohne Interregnum) und *L1/MIR* vs. *L2/MIR* (d. i. die Variante mit Interregnum) robust bleiben. Es scheinen also im Gegensatz zu den Hypothesen 1 und 2 doch Unterschiede im Verhalten der beiden Fähigkeitsgruppen ausgemacht werden zu können.

Der Vergleich innerhalb der beiden Fähigkeitsgruppen zeigt, dass Muttersprachler sowohl in der Instructor- als auch in der Instructee-Bedingung signifikant häufiger mit Interregnum reparieren als ohne Interregnum. In der Lernergruppe wirkt sich der Faktor *Rolle* hingegen nicht systematisch aus, favorisiert also in keiner Ausprägung von *Rolle* eine Variante von *IR*. Die Disfluencies innerhalb von Reparaturen sind somit bei Lernern bzgl. der Rollen auch weniger systematisch verteilt.

Interessanterweise scheinen sich – trotz der statistischen Unterschiede zwischen L1 und L2 – die Muttersprachler und Lerner nicht ganz unähnlich zu sein, jedoch hinsichtlich eines kaum vorhersehbaren Unterschiedes. So gibt es statistisch keinen Unterschied zwischen der Reparaturverteilung der Muttersprachler bzgl. ihrer Rolle in der Bedingung *OIR* und der Lerner bzgl. ihrer Rolle in der Bedingung *MIR*. Möglicherweise ist dies ein Indiz für die der Verhaltensweise in der Zielsprache angepassten Strukturen der Lerner, die zwar als Muster abrufbar sind, hingegen zeitlich noch nicht in der muttersprachlichen Geschwindigkeit berechnet oder geplant werden können. Dieser Schluss wird für die Subreparaturen in Hypothese 4 nochmals aufgegriffen.

Subreparaturen

Hypothese 4 beziehungsweise die Frage nach dem Einfluss der Subreparaturen auf die Verwendung eines Interregnums innerhalb einer Reparatur wurde in Abschnitt 3.1.4 explorativ ausgewertet. Dabei wurde zunächst festgestellt, dass die Verteilung der untersuchten Subreparaturen Einfügung (*i*), Ersetzung (*s*) und Wiederholung (*r*) in L1 und L2 signifikant verschieden verteilt sind (Abbildung 3.4 auf Seite 39). In einer genaueren Berechnung erwies sich der Unterschied zwischen Wiederholung und Substitution aber als nicht signifikant, weswegen diese beiden Kategorien in einer Kategorie *rs* vereinigt wurden. Dieser Schritt erleichterte nicht nur die weitere Analyse, er ist auch qualitativ gut zu motivieren. Während bei einer Wiederholung laut meiner Definition in den Annotationsrichtlinien das reparierte Material exakt dem ersetzten entspricht, so bleiben bei Substitutionen zwar nicht alle, aber doch auch einige Merkmale des ersetzten Materials erhalten – so enthält das betreffende Token im Reparans bspw. einen Teil des ersetzten Materials (Bsp. 21) oder andere Merkmale wie dieselbe Wortart des ersetzten Tokens (Artikelwort in Bsp. 22 auf der nächsten Seite, möglicherweise Verb in Bsp. 23 auf der nächsten Seite).

(21) [RD re/] [IR 0,2] [RS rechten_{s1}] (L2:2013-04-18-C)

(22) [RD an der] [RS an_{r1} seiner_{s1} rechten_{i1} Seite_{i2}] (L1:2011-12-14-A)

(23) genau du [RD b/] [RS gehst_{s1}] (L1:2011-12-14-A)

Für eine detailliertere Analyse wurden die getaggtten Wortarten in zwei Klassen eingeteilt, nämlich Inhalts- und Funktionswörter (C und F), die in etwa mit der Beschreibung von offenen und geschlossenen Klassen korrespondieren. Aus Abschnitt 1.1 geht hervor, dass Disfluencies besonders vor einem Wort der C -Klasse häufiger vorkommen, weswegen sich die Betrachtung auf diese Gruppe konzentrierte.

Aus dem Graph der Interaktion von Subreparaturen mit Wortarten (Tabelle A.5 und Abbildung A.7 auf Seite 69) erschließt sich zunächst bildlich, dass sich die Lerner doch in der Art ihrer verwendeten Beziehungen in Verbindung mit dem geäußerten wortartenspezifischen Material von den Muttersprachlern unterscheiden. Diese Präferenzen sind auch statistisch haltbar. Lerner verwenden signifikant häufiger Wiederholungen/Substitutionen (rs) als Insertionen in der C -Klasse und in der F -Klasse, während Muttersprachler sich in der C -Klasse ausgeglichen verhalten. Aus Hypothese 1 aber ist ersichtlich, dass sich die beiden Gruppen in ihrer Gesamtzahl von Interregna nicht unterscheiden. Folglich äußern Lerner zwar mehr Wiederholungen/Substitutionen, sie äußern aber auch ähnlich viele Disfluencies wie Muttersprachler. Dies ist unerwartet, wenn man davon ausgeht, dass Wiederholungen/Substitutionen einfacher zu äußern sind, weil sie keine neuen Informationen mit neuen Merkmalen darstellen.

Der Graph mit den detaillierten Interaktionen von Wortartenklasse und Subreparaturenklasse in Verbindung mit der Information über den Fähigkeitsfaktor und den IR-Gebrauch ist noch aufschlussreicher (Abbildung 3.5 auf Seite 41). Muttersprachler zeigen bei Einfügungen von neuem Material als Inhaltswort ($i.C$) und bei Wiederholungen/Substitutionen von nicht gänzlich neuem Material (geteilte Merkmale) als Inhaltswort ($rs.C$) ähnliche Verhältnisse bei OIR -Reparaturen wie die Lerner bei MIR -Reparaturen. Dies impliziert, dass die Sprachverarbeitung von L2 trotz des fortgeschrittenen Fähigkeitsniveaus immer noch mehr Zeit in Anspruch nimmt, um die gleichen Muster zu berechnen, die auch Muttersprachler verwenden.

Tavakoli (2010) beobachtete, dass Lerner oft zu pausieren schienen, bevor sie ein Wort wiederholten oder neu einfügten, also in MIR -Reparaturen. Tatsächlich äußern Lernern häufiger rs als Insertionen; dies wäre ein Grund dafür, dass Tavakoli mehr Pausen vor Wiederholungen und Substitutionen beobachtet hat, weil diese einfach per se häufiger geäußert werden. Die Lerner in meinen Daten unterscheiden sich in $L2/OIR$ und $L2/MIR$ aber tatsächlich nicht dergestalt, wie es Tavakoli beobachtet hat, sondern genau andersherum: Die Lerner tendieren dazu, signifikant häufiger eben gerade keine Disfluencies vor

Wiederholungen/Substitutionen einzusetzen. Der Beobachtung, dass sie „oft“ pausieren, lässt sich also entgegenhalten, dass sie dies zwar 136 Mal tun (ein IR einfügen). Signifikant häufiger jedoch fügen sie keine Disfluency vor einer Wiederholung ein (vgl. Abschnitt 3.1.4 auf Seite 39).

Obwohl ‚nur‘ zwei Variablen (**Fähigkeitsfaktor** und **Interregnum**) mit zwei weiteren Variablen interagieren (**Wortart** und **Subreparaturen**) ist die Konkretisierung solcher Muster in Beispielen nicht leicht. Beispiel 24 und 25 könnten eine Realisierung dieser Muster darstellen.

(24) [RD an der] [RS an_{r1} seiner_{s1} rechten_{i1} Seite_{i2}] (L1:2011-12-14-A)

(25) [RD bis zum] [IR 0,3 ähm 1,1][RS bis_{r1} zum_{r1} Ecke_{i2}] (L2:2013-04-19-A)

Hierbei ist natürlich der Kontext nicht unwichtig. Jedoch kann man davon ausgehen, dass aufgrund der gleichbleibenden Taskschwierigkeit die kontextuellen und sprachinhärenten Faktoren bei L1 und L2 ähnlich gesetzt sind (Ahmadian et al. 2012). Dass Muttersprachler signifikant häufiger eine Einfügung im Reparans mit einem Inhaltswort verbinden als Lerner, ist möglicherweise mit der größeren Kenntnis der Muttersprachler an Inhaltswörtern zu erklären. In jedem Falle sollte die Untersuchung mit einer größeren Datengrundlage wiederholt werden, damit reliable Aussagen über eine potenzielle Grundgesamtheit der L2-Sprechsprache des Deutschen gewonnen werden können.

Für das Muster *i.C* und *rs.C* verhalten sich Lerner mit Interregnum nicht signifikant verschieden von Muttersprachlern ohne Interregnum. Diese Analyse birgt jedoch einige Imponderabilitäten. Dabei ist weniger die zwar linguistisch motivierte, aber in ihrer Kategorisierung immer problematische Dichotomisierung von Inhaltswörtern vs. Funktionswörtern gemeint, als vielmehr die Frequenzen der Subreparaturen, die in ihren Interaktionen von *i* und *rs* mit *C* und *F* gegenübergestellt werden. Die Annotation der Subreparaturen erfolgte jeweils innerhalb eines Reparans und tokenbasiert. Die *i*- und *rs*-Tags dann jedoch ohne ihre direkte Umgebung miteinander zu vergleichen, vernachlässigt eben diesen Einfluss des Kontextes, der sowohl linguistisch als auch probabilistisch determiniert sein kann, d. h. die Kategorien der aufeinanderfolgenden Token der Wortarten als auch der Subreparaturen sind möglicherweise für ein Reparans nicht unabhängig voneinander. Um diesen Vorbehalt abzuschwächen, wurden die Daten im nächsten Schritt mit einem *Linear-Mixed-Effects*-Modell auf Konsistenz geprüft.

Limitationen und Mixed Models

Die für Hypothese 1, 3 und 4 durchgeführten χ^2 -Distributions-Test vergleichen Verteilungen miteinander. Dabei wird über die Anteile der einzelnen Sprecher

hinweggeneralisiert. In Abschnitt 3.1.5 auf Seite 40 wurde jedoch gezeigt, dass es gute Gründe gibt, die Sprecher nicht als homogen zu betrachten: Ihre individuellen Verteilungen für die Äußerungen von *OIR*- und *MIR*-Reparaturen unterscheiden sich signifikant. Aus diesem Grund ist es notwendig geworden, ein *LME*-Modell mit Sprechern als randomalen Effekten zu rechnen. Die Ergebnisse des Modells können nun mit denen der Distributionstests verglichen werden und so Aussagen über die Validität der Ergebnisse treffen.

Die Vorhersagen aus dem *LME*-Modell sprechen den Faktoren **Fähigkeitsniveau**, **Rolle** und **Wortart** keinen Einfluss auf das Vorhandensein eines Interregnums zu. Ob ein Sprecher der Ausprägung *L1* oder *L2* in der Kategorie *Instructor* oder *Instructee* zugeordnet ist, scheint seine Interregnum-Häufigkeit demnach ebenso wenig zu beeinflussen wie die Wortart seiner geäußerten Wörter im Reparaturansatz. Der einzig verbliebene signifikante Effekt des Modells ist der Faktor **Subreparatur** in der Ausprägung *rs* mit einem negativen Schätzer bzw. Logit (*Estimate* in der Tabelle 3.3 auf Seite 43) von -0,74. Dies bedeutet, dass das Vorkommen einer Wiederholung oder Substitution die Wahrscheinlichkeit verringert, dass ein Interregnum geäußert wird.

Dass *L2* kein signifikanter Einfluss auf die Wahrscheinlichkeit, ein Interregnum zu äußern, ausgesprochen wird, entspricht der in Hypothese 1.3 aufgestellten und mit dem dort durchgeführten Test nicht abzulehnenden Nullhypothese. *L1*- und *L2*-Sprecher unterscheiden sich folglich nicht in der Häufigkeit ihrer Interregna.

Nach Hypothese 3 hätte es aber eine signifikante Interaktion zwischen **Fähigkeitsniveau** und **Rolle** geben müssen. Jedoch ist im Modell weder das eine noch das andere signifikant. Dass der Unterschied zwischen *L1* und *L2* und zwischen *Instructor* und *Instructee* in diesem Modell keine Rolle zu spielen scheint, ist auch ein Indiz dafür, dass zu wenig Sprecher für eine Wirkung im Modell vorhanden waren und dass die gefundenen Ergebnisse aus Hypothese 3 unbedingt mit mehr *L2*-Sprechern überprüft werden sollten. Der nicht vorhandene Einfluss des Fähigkeitsniveaus hätte noch mit der muttersprachlichen Kompetenz der Lerner erklärt werden können. Aber dass selbst die Gesprächsrolle keinen signifikanten Einfluss ausübt, weicht erheblich von den in der Literatur gefundenen Ergebnissen ab (vgl. bspw. Bortfeld et al. (2001)) und spricht eher für die zu geringe Menge an Lernerdaten.

Im Vergleich zu den χ^2 -Tests aus Hypothese 4 zeigt das Modell nun, welche Ausprägung einen Einfluss auf das Vorhandensein eines Interregnum hat: nämlich *rs*. Allerdings wird nichts über die Vergleiche dreier Faktoren ausgesagt, wie bspw. in Abschnitt 3.1.4, da die dreifachen Interaktionen alle keine Signifikanzen ergaben.

Trotz der unerwarteten Einflussnahme nur eines Faktors (nämlich Subreparatur) auf das Vorhandensein von Interregna lässt sich mindestens feststellen, dass die interne Relation zwischen Reparans und Reparandum eine signifikante Auswirkung auf die Phase vor dem Äußern des Reparans haben kann. Dies ist zumindest ein Indiz dafür, dass neben vielen weiteren möglichen Faktoren die Relation des geäußerten Materials zueinander eine bedeutende Rolle spielt. Das Ergebnis ist auch nachvollziehbar: Die Haupteigenschaft von Wiederholungen und Substitutionen ist ja gerade, dass sie so oder in Teilen schon einmal geäußert worden sind. Ihre Merkmale sind daher kognitiv salient und können womöglich schneller in Planungsprozesse eingebunden werden, so dass auch das Interregnum als mentaler ‚Puffer‘ weniger wahrscheinlich benötigt wird.

3.2.2 Allgemeine Diskussion

Die Frage, inwiefern sich Lerner in ihrem Disfluency- und Reparaturgebrauch noch von Muttersprachlern unterscheiden, wurde schließlich sowohl mit deskriptiver als auch mit analytischer Statistik beleuchtet. Mit letzterer konnte zudem auch explorativ vorgegangen werden. Die durchgeführten Analysen zeigen ein leicht ambivalentes Bild, welches vornehmlich auf die kleine Datengrundlage zurückzuführen sein mag. Für eine detaillierte Analyse der vorliegenden Daten wurde für die Auswertungen der Hypothesen 1–4 simuliert, dass sich die Sprecher homogen verhalten und daher mit Distributionstests miteinander verglichen werden können. Der Vorteil einer solchen Annäherung an die Daten ist, dass unterschiedliche Muster zwischen zwei Gruppen sehr detailliert untersucht werden können. Damit versteht sich diese Arbeit auch als Wegbereiter für Analysen mit einer größeren Datengrundlage, welche ohne diese Vorbehalte noch verlässlichere Aussagen über die Muster deutscher L1- und L2-Sprecher treffen kann.

Die am leichtesten feststellbare Disfluency, eine gefüllte Pause wie bspw. *äh* oder *ähm*, die als Ansatz für eine Vertiefung der Disfluency-Strukturen innerhalb des Interregnums herangezogen wurde (Hypothese 2), ergab keinen Unterschied zwischen Muttersprachlern und Lernern. Weitere Analyseschritte für Pausenmuster innerhalb des Interregnums bleiben offen – Effekte sind aber nicht ausgeschlossen. In jedem Fall sollte eine detaillierte Überlegung bezüglich einer Auswertung stiller Pausen getroffen werden, deren Anfang und Ende auch wieder der Interpretation unterliegt und vielfach diskutiert wurde (vgl. De Jong & Bosker (2013)). Die vorliegende Arbeit hat sich daher nach einer kurzen Betrachtung des Interregnums auch nur auf die Faktoren konzentriert, die überhaupt erst bedingen, dass ein solches geäußert wird.

So scheint der für eine Map Task vorhergesagte Einfluss der Rolle auf die

Äußerungen im Gespräch, wie sie von Bortfeld et al. (2001) und Belz & Klapi (2013) gefunden wurden, bestätigt zu werden. Zumindest für deutsche C1-C2-Lerner haben sich die Vorhersagen aus den Schlussfolgerungen von Tavakoli (2010) hingegen nicht bewahrheitet: Lerner äußerten vor Wiederholungen/Substitutionen signifikant weniger häufig Interregna, i. e. Disfluencies. Auch die Implikationen von Declerck & Kormos (2012), dass aufgrund geringerer Automatisierung weniger Aufmerksamkeit für das Monitoring zur Verfügung steht, konnte nicht nachvollzogen werden, da hierfür mehr *MIR*-Reparaturen zu erwarten gewesen wären.

Dass sich die fortgeschrittenen Lerner hinsichtlich ihrer Disfluency-Produktion innerhalb von Reparaturen für die binären Vorkommen von Interregna selbst oder von gefüllten Pausen innerhalb von Interregna (vgl. Hypothese 1–2) nicht von den Muttersprachlern unterscheiden, war zunächst unerwartet. Ließe man die geringe L2-Sprechermenge außer Acht, so hätte man die Ergebnisse der Tests auch als Bestätigung des muttersprachlichen Sprechniveaus der Lerner auffassen können. Unabhängig von ihrer sonstigen phonetisch transferierten perzipierbaren *foreign-soundingness* müssten zumindest die Unterschiede gegenüber Muttersprachlern auf dem Gebiet der Disfluencies innerhalb von Reparaturen abgelehnt werden. Ein weiterer Aspekt, der die Ähnlichkeit der Reparaturverwendungen aufzeigt, ist die jeweils positive Korrelation der Reparantia mit den Reparanda, was die Silbenanzahl, die Tokenanzahl und die Äußerungszeit in L1 und L2 betrifft. Verfälschen könnte das Ergebnis allerdings eine potenzielle Ähnlichkeit deutscher und englischer Disfluency-Profile, wie sie bei Raupach (1980) allgemein für Erst- und Zweitsprache postuliert wurde. Dann könnte man nicht mehr unterscheiden, ob die Ähnlichkeit aus der Herkunftssprache transferiert wurde oder ob sich die Sprachkompetenz der Zielsprache angepasst hat.

Nichtsdestoweniger hat die fortgeschrittene Analyse unter Hinzuziehung weiterer Faktoren, wie der Rolle (Hypothese 3) oder der Interaktion von Subreparaturrelationen und Wortarten (Hypothese 4), eine Menge unterschiedlicher, aber deswegen nicht weniger interessanter ähnlicher Strukturen in der Reparaturverwendung von Muttersprachlern und Lernern aufgezeigt. Wie Lerner mit sprachlichen Reparaturen umgehen und inwiefern sie sich hier von Muttersprachlern unterscheiden, wurde für einige ausgewählte Phänomene herausgearbeitet. Lerner auf C1-C2-Niveau unterscheiden sich auf den ersten (oberflächlichen) Blick nicht mehr von Muttersprachlern, was zunächst konsistenterweise für ihr Niveau spricht. Gesprächsrolle und Subreparaturbeziehungen in Verbindung mit Wortarten zeigen jedoch weiterhin Muster, die nicht mit denen der Muttersprachler übereinstimmen. Mitunter am besten ließ sich der Unterschied zwischen C1-C2 und dem zu erreichenden Muttersprachlerniveau

aber durch die Muster herausarbeiten, welche zwar schon mit der muttersprachlichen Vergleichsgruppe übereinstimmen, aber gegenüber jener, welche die Muster direkt ohne Interregnum produzieren, zusätzliche Planungszeit und damit Disfluencies benötigten, um diese zu produzieren. Diese Muster finden sich in der Interaktion von *Instructor* mit Interregna (Hypothese 3) und in der Interaktion der Substitutionen/Wiederholungen und Einfügungen mit den offenen und geschlossenen Wortklassen.

Die Ähnlichkeiten der Lerner und Muttersprachler bei einer Auswertung der Daten mit einem Gemischten Modell zeigen sich darin, dass beide Gruppen dazu tendieren, eher keine Disfluencies vor Reparantia mit wiederholten oder substituierten Funktionswörtern einzusetzen. Reparaturen ohne Interregna müssen schon fertig geplant sein, damit sie sofort und ohne Zeitverzögerung geäußert werden können (Seyfeddinipur et al. 2008). Dass gerade für die Kombination einer geschlossenen Wortklasse (*F*) als Wiederholung oder Substitution in einem Reparans kein signifikanter Unterschied für ein Einfügen von Interregna zwischen L1 oder L2 zu finden ist, könnte die Vermutung stützen, dass solche Relationen möglicherweise kognitiv leichter zu verarbeiten sind als Relationen der Kategorie Insertion oder der Kategorie der offenen Wortklassen. Dies würde auch die *Lexical-Retrieval*-Hypothese stützen.

Dass sich L1 und L2 gerade in Hypothese 4 aber doch in vielen Fällen noch deutlich unterscheiden, zeigt nochmals die Komplexität des Fremdspracherwerbs. Dank der Subreparaturannotation konnte die Art der Verhältnisse zwischen Reparandum und Reparans erstmals so für das Deutsche empirisch dokumentiert und ausgewertet und mit den verwendeten Wortarten verknüpft werden. Die Schwierigkeit der statistischen Auswertung besteht auch darin, dass jedes Wort in einen bestimmten – mehr oder weniger wahrscheinlichen – Kontext eingebettet ist. Dieser Kontext konnte bei dieser Auswertung jedoch nicht berücksichtigt werden; die Analyse musste sich mit einer rein frequenzbasierten Auswertung zufrieden geben. Trotz dieser Vorbehalte lässt sich also – mit gebotener Vorsicht – feststellen, dass zwar weiterhin abweichende Strukturen bei L1 und L2 existieren. Jedoch scheinen die C1-C2-Lerner in gewisser Weise bestimmte Muster der Zielsprache (nämlich für die Reparaturumgebung) angenommen zu haben, während die sprechsprachliche Verarbeitung hingegen weiterhin auf zusätzliche Verarbeitungszeit angewiesen ist. Dies wiederum könnte auf weniger automatisierte psycholinguistische Prozesse – wie Zugriff auf das mentale Lexikon oder allgemeine Produktionsprozesse – zurückzuführen sein.

Letztlich darf man die Ergebnisse auch nicht singulär als alleinigen Anspruch für eine – wie auch immer geartete – Grundgesamtheit von Lernern generalisieren, was Rissanen (2008: 65) als „God’s truth fallacy“ formuliert. Die hier

gewonnenen Ergebnisse sollten demnach nur so verstanden werden, dass sich C1-C2-Deutschler unter bestimmten Umständen in den von mir untersuchten disfluency-beeinflussenden Faktoren in gesprochener Sprache ähnlich und/oder verschieden zu deutschen Muttersprachlern verhalten – sie mögen nicht völlig repräsentativ für ihre Grundgesamtheit sein.

3.3 Zusammenfassung und Ausblick

In dieser Studie wurde anhand eines Korpus mutter- und lernersprachlicher Daten eine quantitativ-qualitative Untersuchung von Disfluencies innerhalb von Reparaturen durchgeführt. Die Ergebnisse und ihre Auswertung geben Aufschluss darüber, ob Interregna von bestimmten Faktoren abhängen oder nicht. Solche empirischen Studien wurden meines Wissens bisher für das Deutsche in Verbindung mit der Kontrastierung von L1/L2 noch nicht angestrengt. Die gewonnenen Erkenntnisse konnten und können daher sowohl Hypothesen zur deutschen Lernersprache bestätigen beziehungsweise ablehnen als auch neue Ansatzpunkte für die Betrachtung von Disfluencies im Umfeld gesprochener Lernersprache, Kontext und psycholinguistischen Prozessen aufzeigen sowie Anreize für weiterführende Untersuchungen setzen.

Die Arbeit hat gezeigt, dass Muttersprachler und Lerner sich hinsichtlich der Häufigkeit von Interregna innerhalb ihrer Reparaturen anscheinend nicht mehr unterscheiden. Dieses Ergebnis gilt für beide durchgeführten Analysen, sowohl ohne die Berücksichtigung individueller Sprechereigenschaften (Abschnitt 3.1.1 auf Seite 33) als auch mit deren Berücksichtigung (Abschnitt 3.1.5 auf Seite 40). Wenn Lerner ähnliche Strukturen wie Muttersprachler im Reparans verwenden, so scheint es, dass sie dennoch ein Interregnum vorschalten müssen, während Muttersprachler sofort reparieren können. In dem homogenen Modell in Abschnitt 3.1.4 auf Seite 37 scheint dies sowohl für Inhalts- als auch für Funktionswörter in ihrer jeweiligen Funktion als Einfügung und Wiederholung/Substitution zu gelten. Das zusätzlich gerechnete Gemischte Modell kommt zu einem Ergebnis, welches die Ähnlichkeit von L1 und L2 indirekt bestätigt, eben weil es keine Signifikanz für diesen Faktor feststellen kann (Tabelle 3.3 auf Seite 43). Jedoch zeigt es zusätzlich, dass vor Funktionswörtern, welche als Wiederholung oder Substitution geäußert werden, die Wahrscheinlichkeit für ein Interregnum geringer wird – unabhängig von L1 und L2.

Kritisch zu betrachten ist die ungenügende qualitative Untersuchung des Interregnums, was (außer in Hypothese 2) jeweils als ein binäres Disfluency-Konglomerat betrachtet wird (entweder es existiert – oder nicht). In anschließenden Studien könnte die interne Gliederung dieses Reparaturteils genauer untersucht werden. Auch kein eindeutiger Schluss konnte für die Unterschei-

dung der Sprecherrolle in der Wegbeschreibung gezogen werden. Eine größere Datenmenge wäre hierfür von Vorteil.

Vertieft werden muss die Analyse der Wortartenmuster mit den Subreparaturrelationen. Wünschenswert wäre hierbei eine neuerliche Auswertung mit einem Modell, welches die Position der Tokens innerhalb des Reparans mit einbezieht. Auch eine Verknüpfung mit der linguistischen Einheit ‚Äußerung‘ (*Utterance*) birgt Potenzial für die Ergänzung und Verifizierung von in der Literatur zu findenden Thesen (bspw. „Mittsatzposition“, vgl. Tavakoli 2010). Mit den vorliegenden Daten ließe sich dies gut überprüfen, da Äußerungseinheiten in BeMaTaC in einer Utterance-Ebene annotiert wurden. Der Vergleich mit Lernern einer anderen Herkunftssprache kann Aufschluss darüber geben, ob die Analyse möglicherweise durch einen Transfer der Pausen- bzw. Reparaturprofile aus der Herkunftssprache korrumpiert wurde.

Schließlich steht eine Ergänzung der vorliegenden Daten mit weiteren L2-Gesprächen im Raum, um der relativ kleinen Probandenmenge und damit möglichen idiosynkratischen Mustern noch weniger Vorbehalte einräumen zu müssen. Ein Vergleich mit Lernern des darunterliegenden Niveaus B2 ließe Schlüsse über die Steigerung der Lernerkompetenz für das betrachtete Phänomen zu. Dieser Vergleich wäre auch fruchtbar für quantitative Überprüfungen qualitativer Ergebnisse, bspw. für die Wiederholungsfrequenz lexikalischer Entitäten im Übergang von B2/C1 für die B2-Studie von Rieger (2003a). Eine weitgreifende Analyse könnte zudem Erkenntnisse herbeiführen, die auch aus einer informationsstrukturellen Perspektive desideriert sein mögen. Als Stichworte seien hier ‚bekannte Information‘ (man denke an Wiederholungen und Substitutionen) und ‚neue Information‘ (wie bspw. Einfügungen) genannt. Nicht zuletzt bleibt auch die Frage offen, inwieweit sich C1-C2-Lerner überhaupt dem muttersprachlichen Disfluency-Niveau bezüglich ihrer Frequenzen und Muster anpassen können.

Literatur

- Ahmadian, Mohammad J.; Abdolrezapour, Parisa & Ketabi, Saeed (2012): Task difficulty and self-repair behavior in second language oral production. In: *International Journal of Applied Linguistics* 22(3). 310–330.
- Akaike, Hirotugu (1974): A New Look at the Statistical Model Identification. In: *IEEE Transactions on Automatic Control* AC-19(6). 716–723.
- Anderson, Anne H.; Bader, Miles; Bard, Ellen G.; Boyle, Elizabeth; Doherty, Gwyneth; Garrod, Simon; Isard, Stephen; Kowtko, Jacqueline; McAllister, Jan; Miller, Jim; Sotillo, Catherine; Thompson, Henry S. & Weinert, Regina (1991): The HCRC Map Task Corpus. In: *Language and Speech* 34(4). 351–366.
- Bates, Douglas; Maechler, Martin & Bolker, Ben (2012): *lme4: Linear mixed-effects models using Eigen and Eigenfaces*. URL: <http://cran.r-project.org/web/packages/lme4/lme4.pdf> (besucht am 28.09.2013).
- Belz, Malte & Klapi, Myriam (2013): Pauses following Fillers in L1 and L2 German Map Task Dialogues. In: Eklund, Robert (Hrsg.): *Proceedings of DiSS 2013*. 9–12.
- Besser, Jana & Alexandersson, Jan (2007): A comprehensive disfluency model for multi-party interaction. In: Keizer, Simon; Bunt, Harry & Paek, Tim (Hrsg.): *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue*. Antwerp, Belgium. 182–189.
- Blackmer, Elizabeth R. & Mitton, Janet L. (1991): Theories of monitoring and the timing of repairs in spontaneous speech. In: *Cognition* 39(3). 173–194.
- Boersma, Paul (2001): Praat, a system for doing phonetics by computer. In: *Glott International* 5(9). 341–345.
- Bortfeld, Heather; Leon, Silvia D.; Bloom, Jonathan E.; Schober, Michael F. & Brennan, Susan E. (2001): Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender. In: *Language and Speech* 44(2). 123–147.
- Brand, Christiane & Götz, Sandra (2013): Fluency versus accuracy in advanced spoken learner language: A multi-method approach. In: Gilquin, Gaëtanelle & Cock, Sylvie de (Hrsg.): *Errors and disfluencies in spoken corpora*. (= Benjamins Current Topics, Bd. 52). Amsterdam, Philadelphia: John Benjamins. 117–137.

- Branigan, Holly; Lickley, Robin & McKelvie, David (1999): Non-linguistic influences on rates of disfluency in spontaneous speech. In: *Proceedings of the 14th International Conference of Phonetic Sciences*. 387–390.
- Brennan, S. & Shober, Michael F. (2001): How Listeners Compensate for Disfluencies in Spontaneous Speech. In: *Journal of Memory and Language* 44(2). 274–296.
- Bygate, Martin (1996): Effects of task repetition: Appraising the developing language of learners. In: Willis, Jane & Willis, Dave (Hrsg.): *Challenge and change in language teaching*. Oxford: Heinemann. 136–146.
- Clark, Herbert H. (1996): *Using language*. Cambridge: Cambridge Univ. Press.
- Clark, Herbert H. & Wasow, Thomas (1998): Repeating Words in Spontaneous Speech. In: *Cognitive Psychology* 37. 201–242.
- Colman, Marcus & Healey, Patrick G. T. (2011): The distribution of repair in dialogue. In: Carlson, Laura; Hoelscher, Christoph & Shipley, Thomas F. (Hrsg.): *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society. 1563–1568.
- Core, Mark G. (1999): Dialog Parsing: from Speech Repairs to Speech Acts. Diss. Rochester, New York.
- Corley, Martin & Stewart, Oliver W. (2008): Hesitation Disfluencies in Spontaneous Speech: The Meaning of um. In: *Language and Linguistics Compass* 2(4). 589–602.
- De Jong, Nivia H. & Bosker, Hans Rutger (2013): Choosing a threshold for silent pauses to measure second language fluency. In: Eklund, Robert (Hrsg.): *Proceedings of DiSS 2013*. 17–20.
- Dechert, Hans W. (1980): Pauses and intonation as indicators of verbal planning in second-language speech productions: Two examples from a case study. In: Dechert, Hans W. & Raupach, Manfred (Hrsg.): *Temporal variables in speech*. The Hague: Mouton. 271–285.
- Declerck, Mathieu & Kormos, Judit (2012): The effect of dual task demands and proficiency on second language speech production. In: *Bilingualism: Language and Cognition* 15(4). 782–796.
- Deme, Andrea & Markó, Alexandra (2013): Lengthenings and filled pauses in Hungarian adults' and children's speech. In: Eklund, Robert (Hrsg.): *Proceedings of DiSS 2013*. 21–24.
- Eckes, Thomas (2010): Der Online-Einstufungstest Deutsch als Fremdsprache (onDaF): Theoretische Grundlagen, Konstruktion und Validierung. In: Grotjahn, Rüdiger (Hrsg.): *Der C-Test: Beiträge aus der aktuellen Forschung*. (= Language testing and evaluation, Bd. 18). Frankfurt u. a.: Lang. 125–192.

- Eklund, Robert (2004): Disfluency in Swedish human-human and human-machine travel booking dialogues. Diss. Linköping, Sweden: Linköpings Universitet.
- Engelhardt, Paul E.; Corley, Martin; Nigg, Joel T. & Ferreira, Fernanda (2010): The role of inhibition in the production of disfluencies. In: *Memory & Cognition* 38(5). 617–628.
- Fischer, Kerstin (1999): Die Ikonizität der Pause: Zwischen kognitiver Last und kommunikativer Funktion. In: Wachsmuth, Ipke & Jung, Bernhard (Hrsg.): *Proceedings der KogWis99*. Gesellschaft für Kognitionswissenschaft. 250–255.
- Fox Tree, Jean E. (1995): The Effects of False Starts and Repetitions on the Processing of Subsequent Words in Spontaneous Speech. In: *Journal of Memory and Language* 34(6). 709–738.
- Fox, Barbara A.; Benjamin, Trevor & Mazeland, Harrie (2012): Conversation Analysis and Repair Organization: Overview. In: Chapelle, Carol A. (Hrsg.): *The Encyclopedia of Applied Linguistics*. Oxford: Blackwell Publishing Ltd.
- Fox, Barbara A. & Jasperson, Robert (1995): A Syntactic Exploration of Repair in English Conversation. In: Davis, Philip W. (Hrsg.): *Alternative linguistics*. (= Amsterdam studies in the theory and history of linguistic science. Series IV, Current issues in linguistic theory, Bd. 102). Amsterdam, Philadelphia: J. Benjamins. 77–134.
- Fromkin, Victoria A. (1971): The non-anomalous nature of anomalous utterances. In: *Language*. 27–52.
- Giesel, Linda; Klapi, Myriam; Krüger, Daisy; Nunberger, Isabelle; Rasskazova, Oxana & Sauer, Simon (2013): *A deeply annotated multimodal map-task corpus of spoken learner and native German*. Poster bei der DGfS-CL in Potsdam. URL: <http://u.hu-berlin.de/bematac> (besucht am 07. 11. 2013).
- Götz, Sandra (2007): Performanzphänomene in gesprochenem Lernerenglisch: Eine korpusbasierte Pilotstudie. In: *Zeitschrift für Fremdsprachenforschung* 18(1). 67–84.
- Hartsuiker, Robert J. & Notebaert, Lies (2010): Lexical Access Problems Lead to Disfluencies in Speech. In: *Experimental Psychology (formerly Zeitschrift für Experimentelle Psychologie)* 57(3). 169–177.
- Heeman, Peter A. & Allen, James F. (1997): Intonational Boundaries, Speech Repairs and Discourse Markers: Modeling Spoken Dialog. In: *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics*.

- Heeman, Peter A. & Allen, James F. (1999): Speech Repairs, Intonational Phrases and Discourse Markers: Modeling Speakers' Utterances in Spoken Dialog. In: *Computational Linguistics* 25(4). 527–571.
- Hest, Erna van; Poulisse, Nanda & Bongaerts, Theo (1997): Self-Repair in L1 and L2 Production: an Overview. In: *ITL International Journal of Applied Linguistics* (117-18). 85–115.
- Hope, Adery C. A. (1968): A simplified Monte Carlo significance test procedure. In: *Journal of the Royal Statistical Society. Series B (Methodological)*. 582–598.
- Kormos, Judit (1999): Monitoring and Self-Repair in L2. In: *Language Learning* 49(2). 303–342.
- Labov, William (2004): Quantitative Analysis of Linguistic Variation. In: Ammon, Ulrich (Hrsg.): *Sociolinguistics*. (= Handbücher zur Sprach- und Kommunikationswissenschaft, Bd. 1). Berlin & New York: W. de Gruyter. 6–21.
- Lemnitzer, Lothar & Zinsmeister, Heike (2010): *Korpuslinguistik: Eine Einführung*. 2. Aufl. (= Narr-Studienbücher). Tübingen: Narr.
- Levelt, Willem J. M. (1983): Monitoring and self-repair in speech. In: *Cognition* 14(1). 41–104.
- Levelt, Willem J. M. (1989): *Speaking: From intention to articulation*. Cambridge, Mass. [u.a.]: MIT Press.
- Lickley, Robin & Bard, Ellen G. (1998): When Can Listeners Detect Disfluency in Spontaneous Speech? In: *Language and Speech* 41(2). 203–226.
- Lüdeling, Anke (2007): Das Zusammenspiel von qualitativen und quantitativen Methoden in der Korpuslinguistik. In: Kallmeyer, Werner & Zifonun, Gisela (Hrsg.): *Sprachkorpora - Datenmengen und Erkenntnisfortschritt*. Berlin/New York: Mouton de Gruyter. 28–48.
- Maclay, Howard & Osgood, Charles E. (1959): Hesitation Phenomena in Spontaneous English Speech. In: *Word* 5. 19–44.
- Nakatani, Christine H. & Hirschberg, Julia (1994): A corpus-based study of repair cues in spontaneous speech. In: *Journal of the Acoustical Society of America* 95(3). 1603–1616.
- O'Connell, Daniel C. & Kowal, Sabine (2005): Uh and Um Revisited: Are They Interjections for Signaling Delay? In: *Journal of Psycholinguistic Research* 34(6). 555–576.
- Postma, Albert (2000): Detection of errors during speech production: a review of speech monitoring models. In: *Cognition* 77(2). 97–132.
- R Core Team (2012): *R: A language and environment for statistical computing*. Wien.

- Raupach, Manfred (1980): Temporal variables in first and second language speech production. In: Dechert, Hans W. & Raupach, Manfred (Hrsg.): *Temporal variables in speech*. The Hague: Mouton. 263–270.
- Rieger, Caroline L. (2003a): Disfluencies and hesitation strategies in oral L2 tests. In: Eklund, Robert (Hrsg.): *Proceedings of DiSS'03, Disfluency in Spontaneous Speech Workshop*. (= Gothenburg Papers in Theoretical Linguistics, Bd. 90). 41–44.
- Rieger, Caroline L. (2003b): Repetitions as self-repair strategies in English and German conversations. In: *Journal of Pragmatics* 35(1). 47–69.
- Rissanen, Matti (2008): Corpus Linguistics and Historical Linguistics. In: Lüdeling, Anke & Kytö, Merja (Hrsg.): *Corpus Linguistics*. (= Handbücher zur Sprach- und Kommunikationswissenschaft, Bd. 29.1). Berlin: Walter de Gruyter. 53–68.
- Sauer, Simon & Lüdeling, Anke (2013): BeMaTaC: A Flexible Multilayer Spoken Dialogue Corpus for Contrastive SLA Analyses. In: *ICAME 34*. 46–47.
- Schachter, Stanley; Christenfeld, Nicholas; Ravina, Bernard & Bilous, Frances (1991): Speech disfluency and the structure of knowledge. In: *Journal of Personality and Social Psychology* 60(3). 362–367.
- Schegloff, Emanuel A.; Jefferson, Gail & Sacks, Harvey (1977): The preference for self-correction in the organization of repair in conversation. In: *Language* 53(2). 361–382.
- Schmid, Helmut (1994): Probabilistic Part-of-Speech Tagging Using Decision Trees. In: *Proceedings of International Conference on New Methods in Language Processing*.
- Schmidt, Thomas & Wörner, Kai (2009): EXMARaLDA - Creating, analysing and sharing spoken language corpora for pragmatic research. In: *Pragmatics* 19(4). 565–582.
- Segalowitz, Norman & Hulstijn, Jan (2005): Automaticity in Bilingualism and Second Language Learning. In: Kroll, Judith F. & Groot, Annette M. B. de (Hrsg.): *Handbook of bilingualism*. Oxford [u.a.]: Oxford Univ. Press. 371–388.
- Seyfeddinipur, Mandana; Kita, Sotaro & Indefrey, Peter (2008): How speakers interrupt themselves in managing problems in speaking: Evidence from self-repairs. In: *Cognition* 108(3). 837–842.
- Shriberg, Elizabeth E. (1994): Preliminaries to a Theory of Speech Disfluencies. Diss. Berkeley: University of California.
- Shriberg, Elizabeth E. (2001): To ‘errrr’ is human: ecology and acoustics of speech disfluencies. In: *Journal of the International Phonetic Association* 31(1). 153–169.

- Siegel, Gerald M.; Lenske, Joanne & Broen, Patricia (1969): Suppression of normal speech disfluencies through response cost. In: *Journal of Applied Behavior Analysis* 2(4). 265–276.
- Silber-Varod, Vered & Maruyama, Takehiko (2013): The linguistic role of hesitation disfluencies: evidence from Hebrew and Japanese. In: Eklund, Robert (Hrsg.): *Proceedings of DiSS 2013*. 67–70.
- Tannenbaum, Percy H.; Williams, Frederick & Hillier, Carolyn S. (1965): Word Predictability in the Environments of Hesitations. In: *Journal of verbal learning and verbal behavior* 4(2). 134–140.
- Tavakoli, Parvaneh (2010): Pausing patterns: differences between L2 learners and native speakers. In: *ELT Journal* 65(1). 71–79.
- Vater, Heinz (1992): Zum Silben-Nukleus im Deutschen. In: Eisenberg, Peter; Ramers, Karl Heinz & Vater, Heinz (Hrsg.): *Silbenphonologie des Deutschen*. Tübingen: Narr. 100–133.
- Watanabe, Michiko & Rose, Ralph L. (2012): Pausology and Hesitation Phenomena in Second Language Acquisition. In: Robinson, Peter (Hrsg.): *The Routledge encyclopedia of second language acquisition*. London: Routledge. 480–483.
- Wichmann, Anne (2008): Speech corpora and spoken corpora. In: Lüdeling, Anke & Kytö, Merja (Hrsg.): *Corpus Linguistics*. (= Handbücher zur Sprach- und Kommunikationswissenschaft, Bd. 29.1). Berlin: Walter de Gruyter. 187–207.
- Wiese, Richard (1983): *Psycholinguistische Aspekte der Sprachproduktion: Sprechverhalten und Verbalisierungsprozesse*. (= Papiere zur Textlinguistik, Bd. 44). Hamburg: Buske.
- Wiese, Richard (1984): Language Production in Foreign and Native Languages: Same or Different? In: Dechert, Hans W.; Möhle, Dorothea & Raupach, Manfred (Hrsg.): *Second Language Productions*. (= Tübinger Beiträge zur Linguistik, Bd. 240). Narr. 11–25.
- Wilkinson, Sue & Weatherall, Ann (2011): Insertion Repair. In: *Research on Language and Social Interaction* 44(1). 65–91.
- Yang, Fan; Heeman, Peter A.; Hollingshead, Kristy & Strayer, Susan E. (2008): DialogueView: annotating dialogues in multiple views with abstraction. In: *Natural Language Engineering* 14(1).
- Zeldes, Amir; Ritz, Julia; Lüdeling, Anke & Chiarcos, Christian (2009): AN-NIS: A Search Tool for Multi-Layer Annotated Corpora. In: *Proceedings of Corpus Linguistics*.
- Zipser, Florian & Romary, Laurent (2010): A model oriented approach to the mapping of annotation formats using standards. In: *Proceedings of the Workshop on Language Resource and Language Technology Standards*.

A Graphen und Tabellen

	L1	L2
Mit IR	98 (0.41)	109 (0.42)
Ohne IR	141 (0.59)	148 (0.58)

Abbildung A.1: Übersicht über die Verteilung von Reparaturen mit und ohne Interregnum (IR) der Muttersprachler (L1) und fortgeschrittenen Lerner (L2) in BeMaTaC. Die Blöcke stellen die individuelle Gesamtzahl der Reparaturen aller L1- oder L2-Sprecher mit und ohne IR dar.

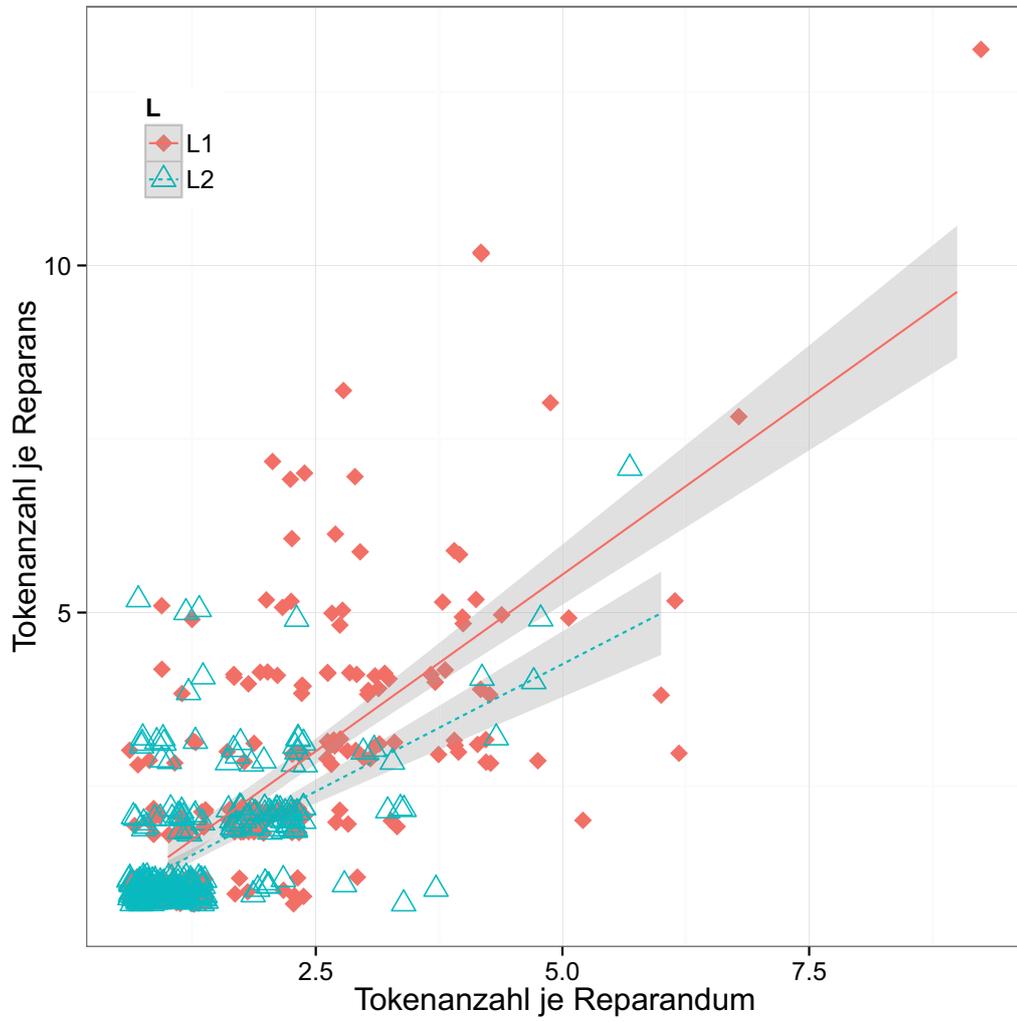


Abbildung A.2: Korrelationen zwischen den Tokenlängen von Reparaturandum (RD) und Reparans (RS) innerhalb der Reparaturen von Muttersprachlern (L1) und fortgeschrittenen Lernern (L2) in BeMaTaC.

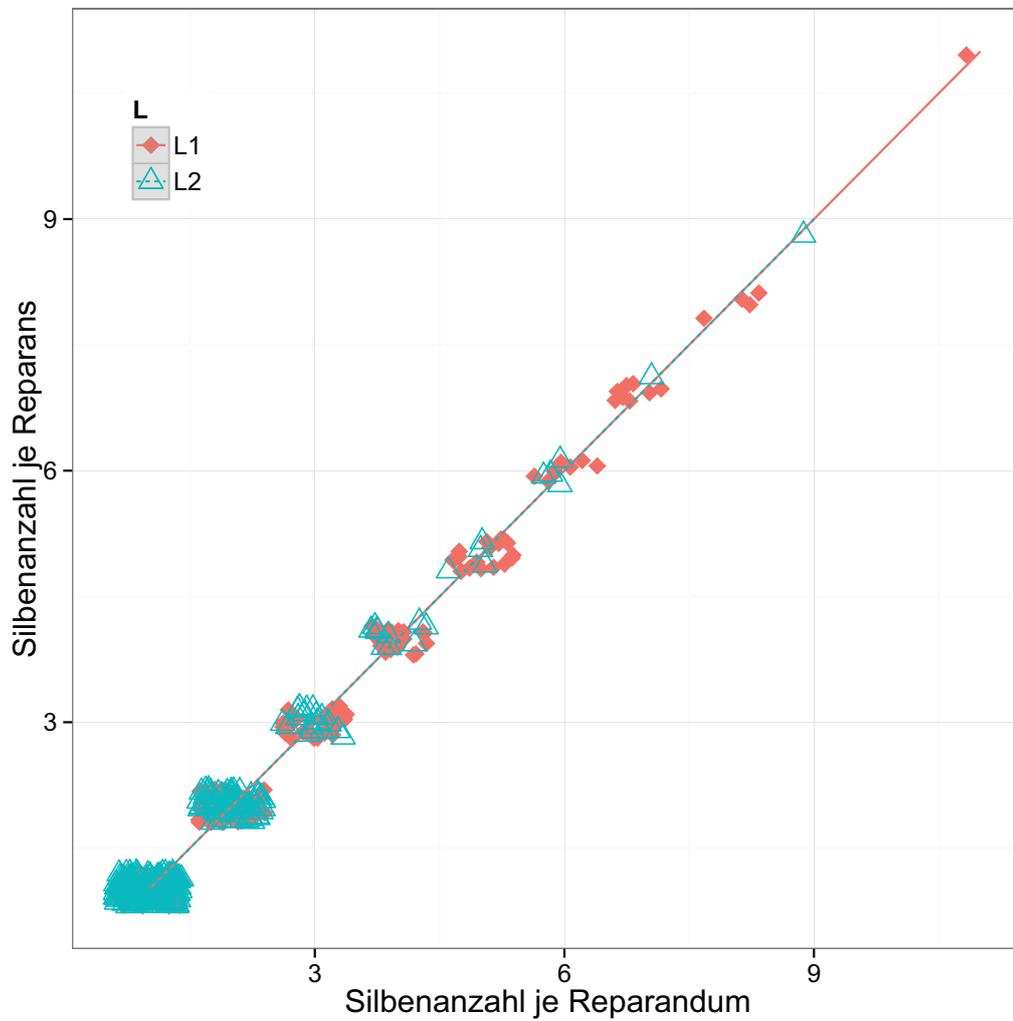


Abbildung A.3: Korrelationen zwischen der Silbenanzahl von Reparaturandum (RD) und Reparans (RS) innerhalb der Reparaturen von Muttersprachlern (L1) und fortgeschrittenen Lernern (L2) in BeMaTaC.

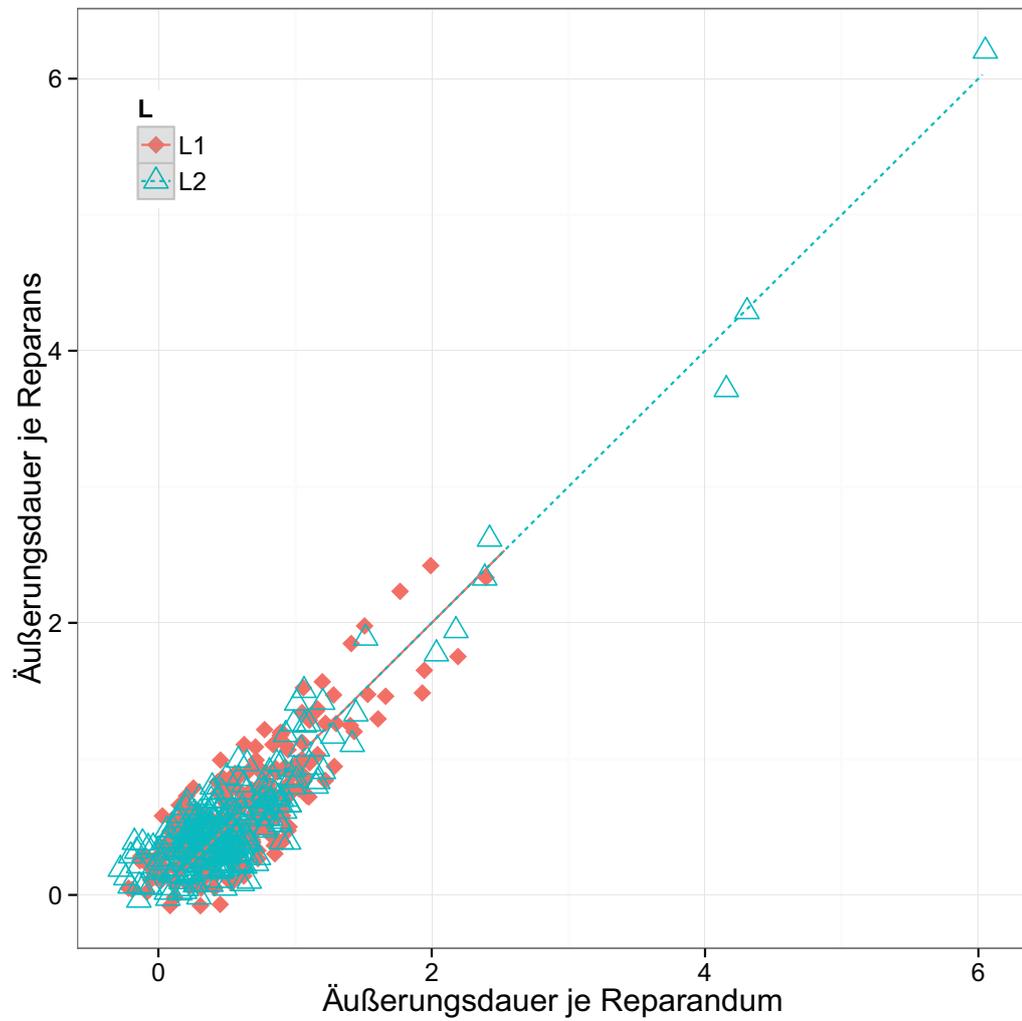


Abbildung A.4: Korrelationen zwischen der Äußerungsdauer von Reparatur (RD) und Reparans (RS) innerhalb der Reparaturen von Muttersprachlern (L1) und fortgeschrittenen Lernern (L2) in BeMaTaC.

Tabelle A.1: Frequenzen für die von Muttersprachlern geäußerten Reparaturen in BeMaTaC mit und ohne Interregnum (IR). S01–S16 bezeichnen die einzelnen Probanden.

	S01	S02	S03	S04	S05	S06	S07	S08	S09	S10	S11	S12	S13	S14	S15	S16	Total
Mit IR	11	12	27	2	6	0	3	5	7	5	5	0	6	0	5	4	98
Ohne IR	24	28	15	0	8	5	4	1	9	5	4	2	16	6	5	9	141
Total	35	40	42	2	14	5	7	6	16	10	9	2	22	6	10	13	239

Tabelle A.2: Frequenzen für die von fortgeschrittenen Lernern geäußerten Reparaturen in BeMaTaC mit und ohne Interregnum (IR). S20–S25 bezeichnen die einzelnen Probanden.

	S20	S21	S22	S23	S24	S25	Total
Mit IR	19	16	25	34	12	3	109
Ohne IR	26	41	17	42	20	2	148
Total	45	57	42	76	32	5	257

Tabelle A.3: Frequenzen für die von fortgeschrittenen Lernern und Muttersprachlern geäußerten Reparaturen in BeMaTaC, welche ein Interregnum mit einer gefüllten Pause enthalten. S01–S16 bezeichnet die einzelnen L1-Probanden, S20–S24 die L2-Probanden.

	S01	S02	S03	S05	S07	S11	S13	S15	S16	S20	S21	S22	S23	S24	
FP	5	3	11	1	1	1	1	2	3	1	3	6	11	7	5

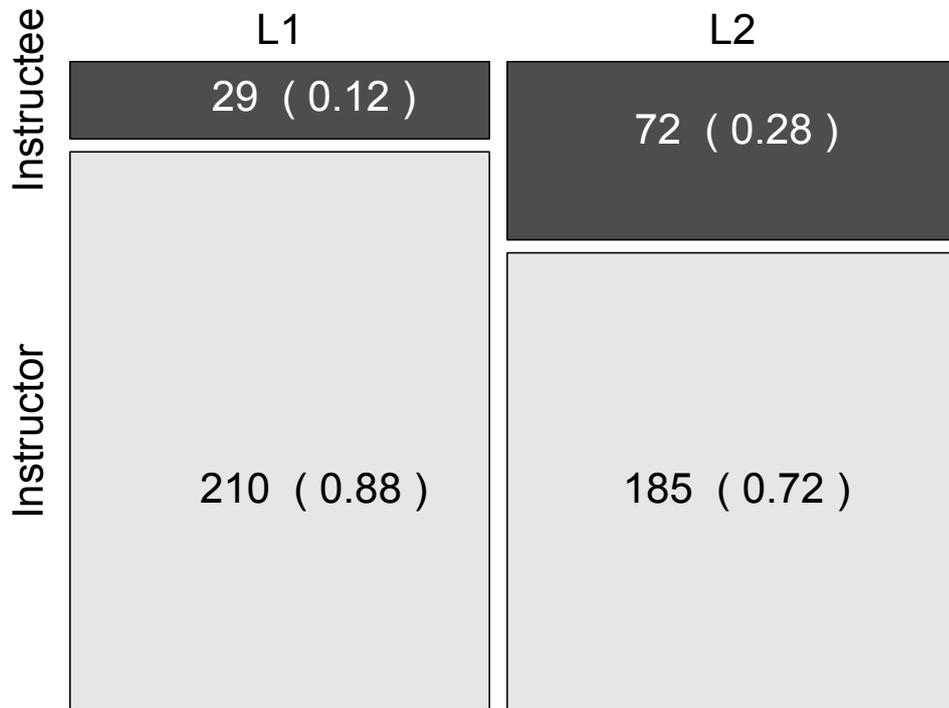


Abbildung A.5: Rollenverteilung der Reparaturen für Muttersprachler (L1) und fortgeschrittene Lerner (L2) in BeMaTaC. Instructor/Instructee entsprechen den in der Map-Task-Aufgabe eingenommenen Rollen, die relativen Frequenzen je L-Gruppe sind in Klammern angegeben.

Tabelle A.4: Frequenzen der Reparaturen in der Rolle Instructor vs. Instructee bei Muttersprachlern und Lernern in BeMaTaC für die Bedingungen Mit IR/Ohne IR. Die relativen Frequenzen sind in Klammern jeweils für die Zeilen C und F je Fähigkeitsfaktor angegeben.

	L1		L2	
	Instructor	Instructee	Instructor	Instructee
Gesamt	210 (0,87)	29 (0,13)	185(0,71)	72 (0,29)
Ohne IR	117	24	104	44
Mit IR	93	5	81	28

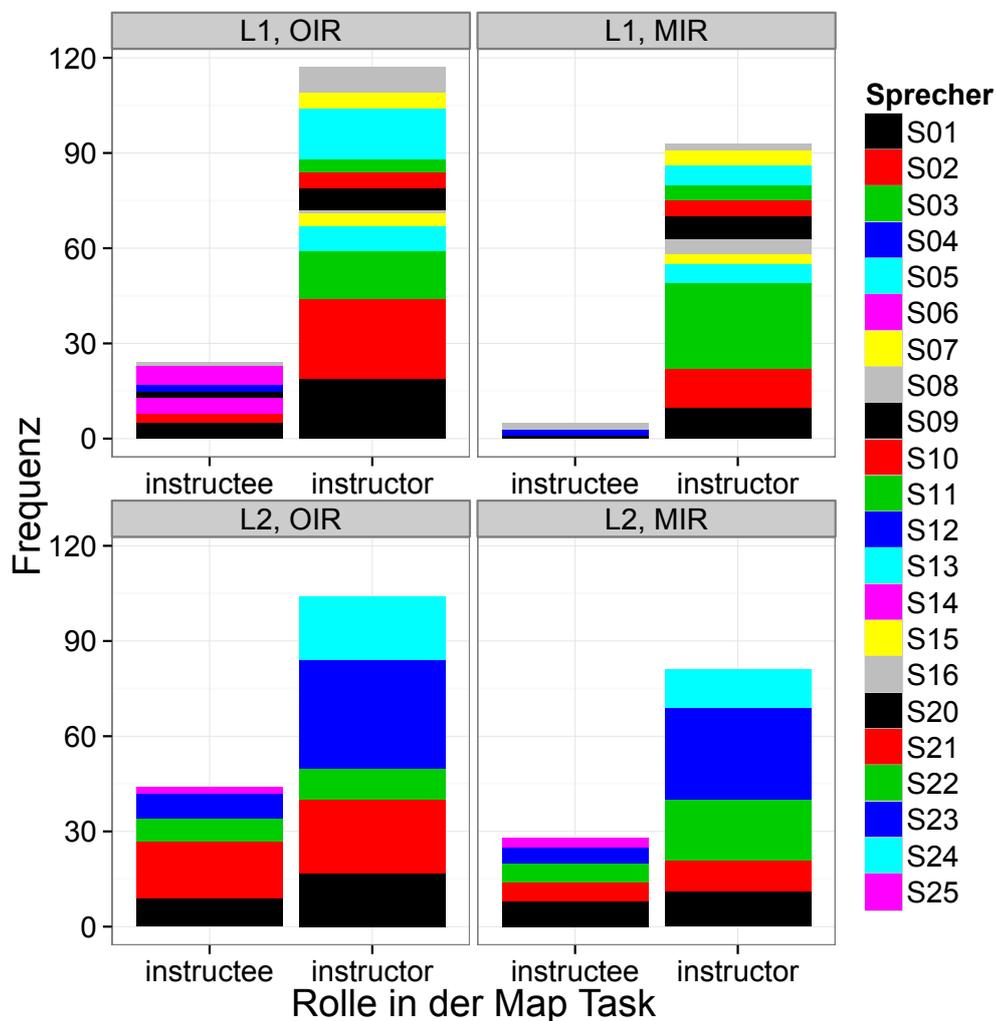


Abbildung A.6: Rollenverteilung der Reparaturen für Muttersprachler (L1) und fortgeschrittene Lerner (L2) für Reparaturen mit Interregnum (MIR) und ohne Interregnum (OIR) je Sprecher in BeMaTaC. Instructor/Instructee entsprechen den in der Map Task eingenommenen Rollen. Die Legende zeigt die Farben für die einzelnen Sprecher, welche sich systematisch wiederholen. S01–S16 sind Muttersprachler, S20–S25 Lerner.

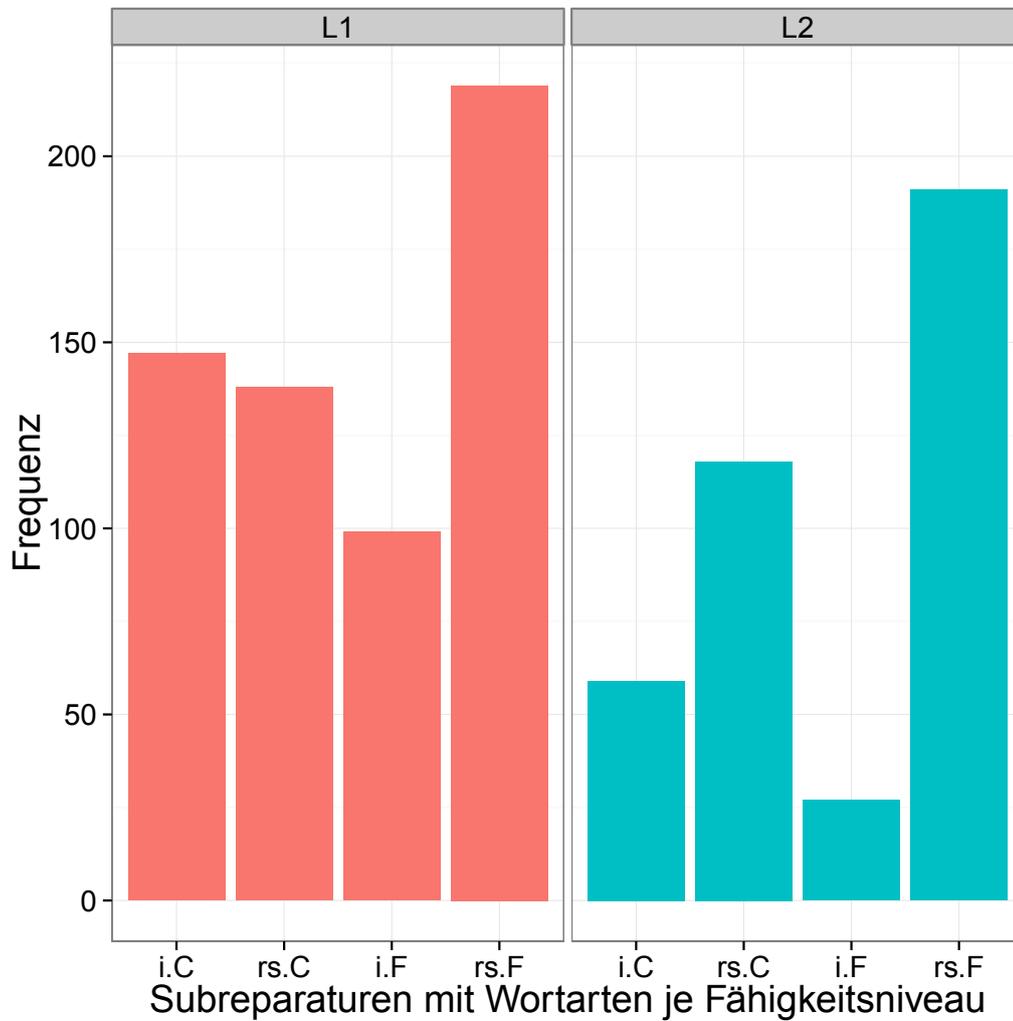


Abbildung A.7: *Interaktion der Subreparaturen i (Insertion) und rs (Repetitionen/Substitutionen) mit Inhaltswortarten (C) und Funktions- bzw. Residualwortarten (F). Die Balken zeigen die Frequenzen für jede Interaktion in L1 (rot) und L2 (blau) aller Reparaturen in BeMaTaC.*

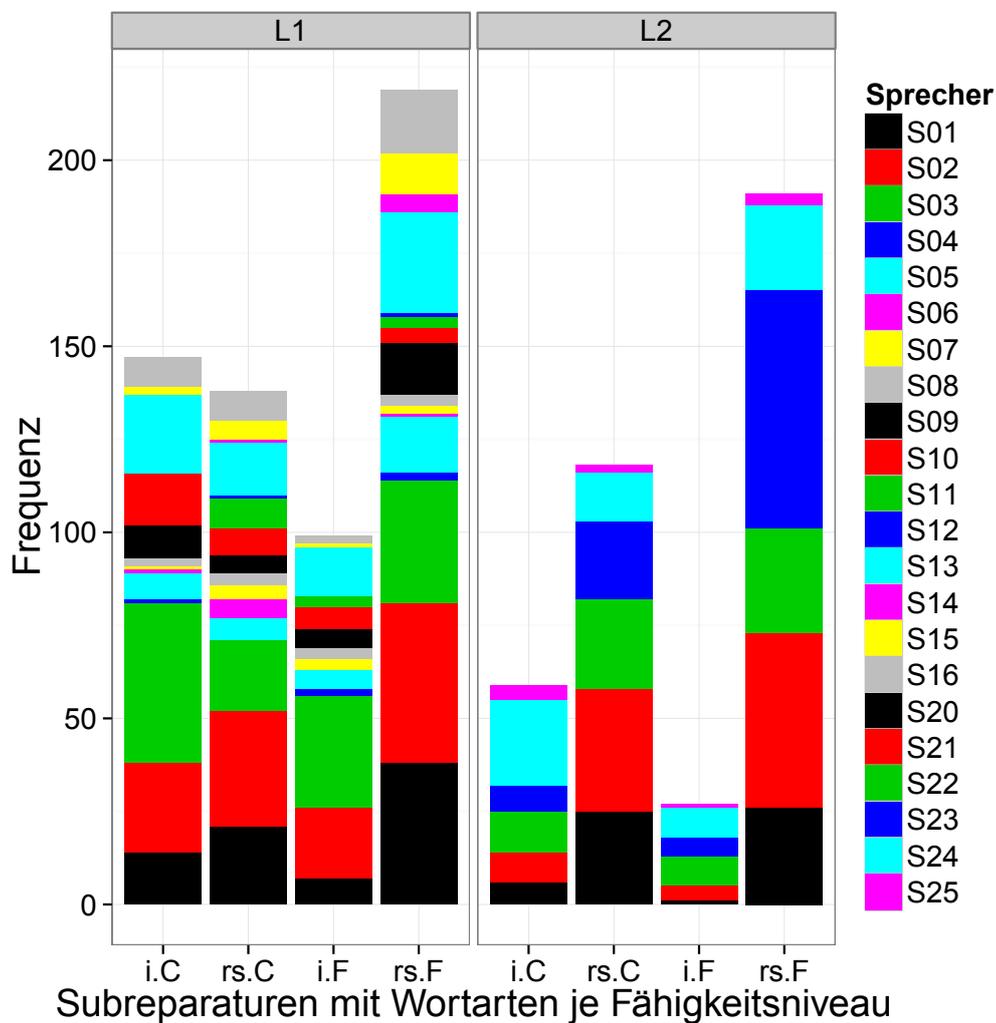


Abbildung A.8: Interaktion der Subreparaturen *i* (Insertion) und *rs* (Repetitionen/Substitutionen) mit Inhaltswortarten (*C*) und Funktions- bzw. Residualwortarten (*F*). Die Säulen zeigen die Frequenzen für jede Interaktion in *L1* und *L2* aller Reparaturen in *BeMaTaC* je Sprecher. Die Legende zeigt die Farben für die einzelnen Sprecher, welche sich systematisch wiederholen. *S01–S16* sind Muttersprachler, *S20–S25* Lerner.

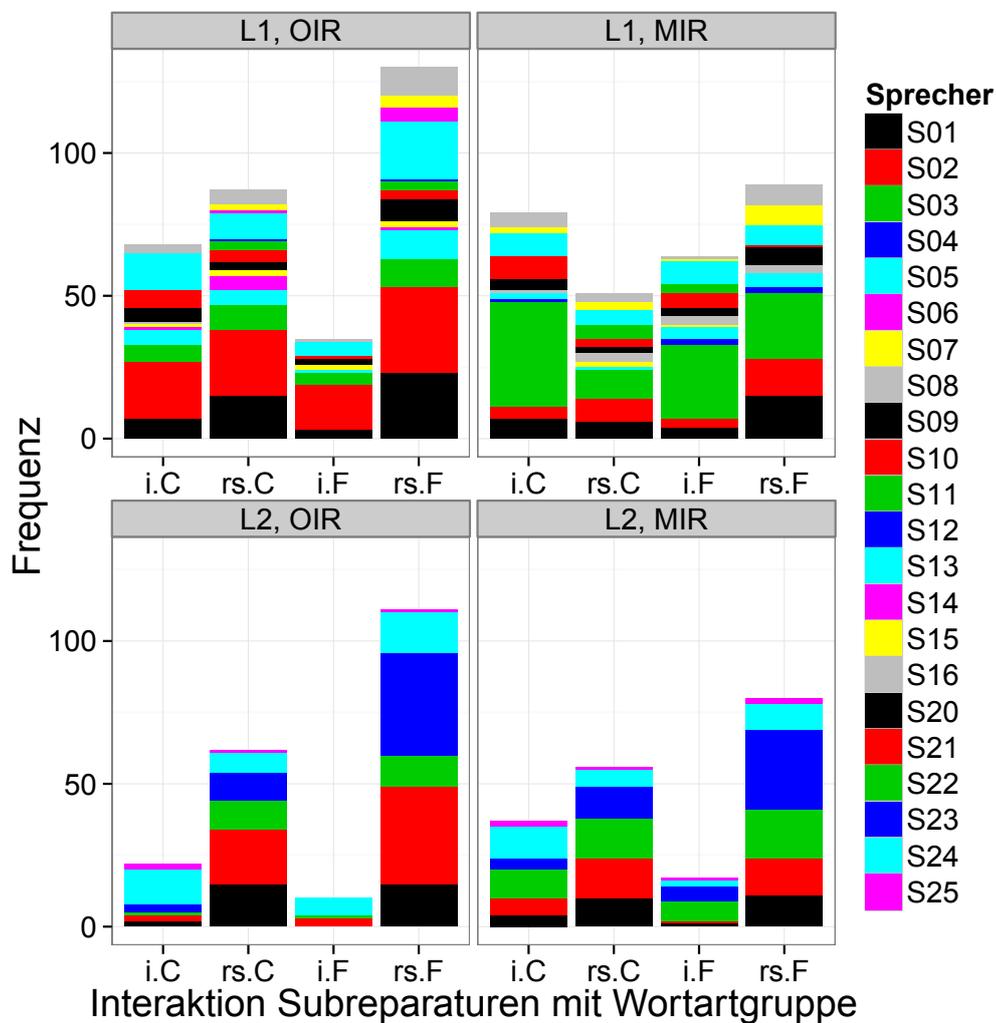


Abbildung A.9: Interaktion der Subreparaturen *i* (Insertion) und *rs* (Repetitionen/Substitutionen) mit Inhaltswortarten (C) und Funktions- bzw. Residualwortarten (F). Die Säulen zeigen die Frequenzen für jede Interaktion in L1 und L2 aller Reparaturen in BeMaTaC. Die Legende zeigt die Farben für die einzelnen Sprecher, welche sich systematisch wiederholen. S01–S16 sind Muttersprachler, S20–S25 Lerner.

Tabelle A.5: *Frequenzen der Interaktionen zwischen den Subreparaturbeziehungen i (Insertion) und rs (Repetition/Substitution) und den reduzierten Wortartenkategorien C (Inhaltswörter) und F (Funktionswörter) bei Muttersprachlern und Lernern in BeMaTaC für die Bedingungen Mit IR/Ohne IR. Die relativen Frequenzen sind in Klammern jeweils für die Zeilen C und F je Fähigkeitsfaktor angegeben.*

		L1		L2	
		i	rs	i	rs
Gesamt	C	147 (0,52)	138 (0,48)	59 (0,33)	118 (0,67)
	F	99 (0,31)	219 (0,69)	27 (0,12)	191 (0,88)
Ohne IR	C	68	87	22	62
	F	35	130	10	111
Mit IR	C	79	51	37	56
	F	64	89	17	80

B Annotationsrichtlinien

Reparaturen	Beschreibung	Tags
instructor_repair instructee_repair	Ebene für die Annotation von Reparaturandum, Interregnum, Reparans. Bezugsebene für die Segmentierung ist ⟨instructor_dipl⟩ bzw. ⟨instructee_dipl⟩.	rd, ir, rs
Reparandum	Der Teil der Aussage, der vom Sprecher später repariert wird. Meist folgt danach ein Abbruch, eine ungefüllte Pause oder eine gefüllte Pause. Wenn rd annotiert wird, muss auch rs annotiert werden. Spannenannotation.	rd
Interregnum (fakultativ!)	Interregnum ist das, was zwischen Reparaturandum und Reparans steht, meist Stille (vgl. Ebene ⟨break⟩), gefüllte Pausen (vgl. Ebene ⟨df⟩), explizite Editierungen (bspw. <i>oder</i> , <i>beziehungsweise</i> , <i>Quatsch</i> , <i>ich meine</i> , Bsp. 5). Mehrere vorhandenen Phänomene bilden auch nur ein Interregnum (Bsp. 1, 3, 5). Wird ir annotiert, muss unmittelbar davor ein rd und unmittelbar danach ein rs stehen (Bsp. 3). Spannenannotation.	ir
Reparans	Die Reparatur der Aussage. Ihr geht entweder ein Reparaturandum voraus oder ein Interregnum. Die Reparatur hört dort auf, wo die reparierte Sinneinheit (bspw. Silbe, Wort, Phrase) zu Ende ist oder eine neue Utterance beginnt. Dies ist v. a. bei Insertions nicht immer eindeutig (s. Insertion). Ist kein Interregnum vorhanden, beginnt rs unmittelbar nach rd (Bsp. 5). Spannenannotation.	rs
Subreparaturen		
instructor_subrepair instructee_subrepair	Ebene für die Annotation von konkreten Reparaturphänomenen. Bezugsebene für die Segmentierung ist ⟨instructor_dipl⟩ bzw. ⟨instructee_dipl⟩.	d⟨n⟩, i⟨n⟩, r⟨n⟩, s⟨n⟩
Deletion (Löschung)	Nur im Reparaturandum zu annotieren. Damit wird getaggt, was im späteren Reparans nicht mehr vorhanden ist und nicht substituiert wurde (Bsp. 8). Tokenannotation. n steht für eine Zahl von 1 bis n, also d1, d2 etc.	d1, d2, ...

Insertion (Einfügung)	Nur im Reparans zu annotieren. Damit wird getaggt, was im Reparandum noch nicht vorhanden war, sondern neu hinzukam. Im Zweifelsfall muss man untersuchen, ob die fragliche Insertion ohne die Reparatur auch geäußert worden wäre. Falls ja, ist sie nicht mehr Teil des Reparans und wird nicht annotiert. Tokenannotation. n steht für eine Zahl von 1 bis n, also i1, i2 etc.	i1, i2, ...
Substitution (Ersetzung)	Nur im Reparans zu annotieren. Damit wird getaggt, was im Reparans in Bezug auf das Reparandum ersetzt worden ist. Sobald das neue Token nicht mehr exakt das alte Token darstellt, ist es eine Substitution (Bsp. 1). Tokenannotation. n steht für eine Zahl von 1 bis n, also s1, s2 etc.	s1, s2, ...
Repetition (Wiederholung)	Nur im Reparans zu annotieren. Damit wird getaggt, was im Reparans in Bezug auf das Reparandum gleich blieb. Wiederholung von identischen Tokens oder Wiederholung mit Verkürzung des Tokens (Bsp. 1). Tokenannotation. n steht für eine Zahl von 1 bis n, also r1, r2 etc. (Bsp. 7).	r1, r2, ...
Residuen „Weiß-nicht-Tag“	Wird vergeben, wenn man sich zwischen den Subreparaturtags nicht entscheiden kann. Tokenannotation. Bitte sparsam und gewissenhaft vergeben!	x

Anmerkung: Jede Ebene `_repair` bzw. `_subrep` kann zu `_repair2` bzw. `_subrep2` verdoppelt werden, um verschachtelten Reparaturen besser Rechnung zu tragen.

Beispiele

1. Beispiel

instructor_dipl	dann	schlägst	ähm		schä/	schlägst	du	sozusagen
break				1,5				
instructor_repair		rd		ir		rs		
instructor_subrep					x	r1	i1	

Das Reparans besteht hier aus der Auflösung der Enklise *schlägst* und der Fehlaussprache *schä/* während dieses Prozesses. Der Diskursmarkierer *sozusagen* zählt nicht mit ins Reparans, er hätte auch zusammen mit der Enklise realisiert werden können.

2. Beispiel

instructor_dipl	des	Kai/	des	Kaiserbrötchens
instructor_repair		rd		rs
instructor_subrep			r1	s1

3. Beispiel

instructor_dipl	also	du	hast	ähm					du	bist	am	Start	
instructee_dipl						ich	bin	am	Start				
break					0,2								
instructor_repair			rd			ir				rs			
instructor_subrep										r1	s1	i1	i2

Das Reparans besteht aus der vom Instructee übernommenen Feedbackphrase *du bist am Start*. Im Reparans gelten *am* und *Start* als Insertionen, da diese Aussage vom Instructee übernommen wurde und nicht zwingend geäußert hätte werden müssen.

4. Beispiel

instructor_dipl		du		gehst		sozusagen		einen		ähm		in		einem		Fünfundvierziggradwinkel
instructor_repair								rd		ir				rs		
instructor_subrep												i1		s1		

Hier endet das Reparans bei *einem*, da der Kasus repariert wurde (*einen*/Akk -> *in einem*/PP).

5. Beispiel

instructor_dipl		und		machst		rechts		oder		beziehungsweise		gehst		rechts		herum
instructor_repair						rd				ir				rs		
instructor_subrep												s1		r1		i1

Die VP *machst rechts* kann auch alleine stehen. *herum* ist nicht zwingend schon mental geplant worden, daher zählt es als Insertion mit ins Reparans.

6. Beispiel

instructor_dipl		und		gehst		äh		grade/		grad/		in		gerader		Linie		nach		rechts
instructor_repair								rd						rs						
instructor_subrep												i1		s1		i2				

grade/ und *grad/* werden schnell hintereinander abgebrochen und bilden ein gemeinsames Reparandum. Sie werden durch eine Adverbialphrase ersetzt, daher reicht das Reparans von *in* bis *Linie*.

7. Beispiel

instructor_dipl		einen		runden		Kreis		um		den		Zah/		um		den		Arzt		herum
instructor_repair								rd						rs						
instructor_subrep														r1		r2		s1		

Hier ist die Reparatur eigentlich auf das Nomen begrenzt, doch der Sprecher ersetzt die ganze PP *um den Zah/* mit *um den Arzt*. *herum* zählt nicht mit in das Reparans, da *um* und *herum* eine feste Zirkumfigierung bilden und die Reparatur sich innerhalb dieser abspielt.

8. Beispiel

instructor_dipl	einen	runden	Kreis	äh	einen	Kreis
instructor_repair		rd		ir	rs	
instructor_subrep		d1			r1	r2