

MARKUS BADER / EMILIA ELLSIEPEN

EXPERIMENTELLE SYNTAX: EINE FALLSTUDIE ZUR VERBALKOMPLEXBILDUNG IM DEUTSCHEN

Abstract

This paper shows how experimental methods can advance syntactic description and syntactic theory. The empirical domain is the order of verbs in German verb clusters containing a modal verb in the perfect. Such clusters are special insofar as prescriptive grammar requires the finite verb to appear in cluster-initial instead of cluster-final position (e.g., *hat lesen müssen* 'has read must' instead of *lesen müssen hat* 'read must has'). Contrary to this requirement, experiments show that native speakers accept the auxiliary also in later positions as long as it precedes the modal verb. The acceptability data are corroborated by corpus data and experimental data from language production. The relevance of the experimental data for syntactic theory are discussed.

1. Einleitung

Bis vor nicht allzu langer Zeit waren Experimente innerhalb der Linguistik eine Sache der Psycho- und Neurolinguistik. Diese beiden linguistischen Teilbereiche beschäftigen sich mit Fragen nach der Verarbeitung von Sprache im menschlichen Geist bzw. Gehirn. Als Gegenstand der theoretischen Linguistik gilt dagegen das sprachliche Wissen, das der Verarbeitung von Sprache zugrunde liegt. Zur Erforschung des sprachlichen Wissens hat sich die theoretische Linguistik im Wesentlichen auf sprachliche Intuitionen gestützt. Im Falle der Syntax handelt es sich dabei insbesondere um Intuitionen bezüglich der Grammatikalität von Sätzen. Die meisten Sprecher des Deutschen haben beispielsweise die Intuition, dass der Satz in (1a) syntaktisch wohlgeformt ist, der Satz in (1b) dagegen nicht.

- (1) a. Ich glaube, dass Peter das Fahrrad hatte verkaufen wollen.
b. *Ich glaube, dass Peter das Fahrrad verkaufen wollen hatte.

Die Verwendung von Grammatikalitätsurteilen als Datengrundlage für die theoretische Syntax ist vielfach kritisiert worden (z.B. Edelman/Christiansen 2003; Ferreira 2005; Wasow/Arnold 2005; Gibson/Fedorenko 2010; Antworten auf diese Kritiken geben z.B. Phillips/Lasnik 2003; Phillips 2009; Culicover/Jackendoff 2010). Ein Teil dieser Kritik richtet sich nicht gegen Grammatikalitätsurteile als solche, sondern gegen die gängige Praxis, dass syntaktische Analysen auf Grammatikalitätsurteilen sehr weniger Personen beruhen – im Extremfall einer einzigen Person, nämlich dem Autor oder der Autorin der entsprechenden Analyse.

Beginnend mit den Arbeiten von Bard/Robertson/Sorace (1996) und Cowart (1997) hat sich die Situation grundlegend gewandelt. Syntaktische Fragestellungen mittels experimenteller Methoden zu untersuchen ist mittlerweile eine weitverbreitete Praxis. In zwei großen Studien haben Sprouse/Almeida (2012) und Sprouse/Schütze/Almeida (2013) gezeigt, dass die in einem verbreiteten Syntax-Lehrbuch (Adger 2003) bzw. in einer führenden linguistischen Fachzeitschrift (*Linguistic Inquiry*) veröffentlichten Bewertungen weitestgehend übereinstimmen mit den Urteilen von Muttersprachlern, die keinen Bezug zur Linguistik haben und deren Urteile deshalb nicht durch theoretische Voreingenommenheit beeinflusst sein können. Über die mangelnde empirische Fundierung syntaktischer Theorien zu lamentieren, ist deshalb nicht mehr notwendig. Die Frage ist dann, ob und in welcher Weise die Syntaxforschung vom Einsatz verschiedener experimenteller Methoden profitieren kann. Wir haben auf diese Frage keine generelle Antwort und werden stattdessen – frei nach dem Motto *The proof of the pudding is in the eating* – eine Fallstudie zur Syntax der Verbalkomplexe im Deutschen präsentieren, d.h. zu Sätzen wie in (1). Diese Fallstudie zeigt, so hoffen wir, dass mittels experimenteller Methoden fundiertere Analysen syntaktischer Phänomene möglich sind.

Aufgebaut ist unser Artikel wie folgt. In Abschnitt 2 stellen wir zwei ausgewählte experimentelle Methoden vor und geben zugleich einen kurzen Überblick über Experimente, die mittels dieser Methoden die Akzeptabilität von Verbalkomplexen untersucht haben. Abschnitt 3 skizziert eine syntaktische Analyse, die auf den experimentell gewonnenen Daten beruht. In Abschnitt 4 diskutieren wir die Frage, ob experimentelle Untersuchungen der Akzeptabilität von Sätzen angesichts der vielfachen Fortschritte im Bereich der Korpuslinguistik überhaupt noch notwendig sind. Dabei werden wir zu dem Schluss kommen, dass Akzeptabilitätsurteile weiterhin benötigt werden und nicht durch andere Daten – wie z.B. Frequenzdaten aus Korpusuntersuchungen – ersetzt werden können. Daraus folgt aber nicht, dass Daten aus der Sprachproduktion irrelevant sind für die theoretische Modellierung syntaktischer Strukturen. Wie Abschnitt 5 zeigt, können auch hier experimentelle Untersuchungen von zentraler Bedeutung sein. Der Artikel endet mit einem Resümee in Abschnitt 6.

2. Experimentelle Untersuchungen zur Akzeptabilität von Verbalkomplexen

Das Deutsche ist im Vergleich zu den anderen westgermanischen OV-Sprachen weitgehend konsistent, was die Platzierung der Komplemente betrifft. Ebenso wie die nicht-verbalen Komplemente (mit Ausnahme von Komplementsätzen und bestimmten PPs) werden verbale Komplemente links von dem

Element platziert, von dem sie selegiert werden. Dies zeigen die folgenden Beispiele. Dabei erhält das finite Verb als nicht-selegiertes Element den Index 1. Ein Verb, das vom finiten Verb selegiert wird, enthält den Index 2, usw.

- (2) a. ... dass Peter das Buch *liest*₁.
 b. ... dass Peter das Buch *gelesen*₂ *hat*₁.
 c. ... dass Peter das Buch *gelesen*₃ *haben*₂ *soll*₁.
- (3) a. ... dass das Buch anscheinend von Peter *gelesen*₂ *wird*₁.
 b. ... dass das Buch anscheinend von Peter *gelesen*₃ *worden*₂ *ist*₁.
 c. ... dass das Buch anscheinend von Peter *gelesen*₄ *worden*₃ *sein*₂ *soll*₁.

Wie die Beispiele in (2) und (3) zeigen, stehen die Verben im Verbalkomplex in absteigender Ordnung, d.h. V_N steht vor V_{N-1} . Dies gilt für andere westgermanische Sprachen nicht. Für das Niederländische beispielsweise ist die umgekehrte, aufsteigende Abfolge der Verben charakteristisch.¹

Bekanntlich gibt es eine Ausnahme zur strikten Abfolge V_N vor V_{N-1} . Für die Modalverben, die Perzeptionsverben und das Verb *lassen* gilt, dass bei zusammengesetzten Zeitformen das Auxiliar am Anfang des Verbalkomplexes platziert wird. Für das Perfekt gilt dies obligatorisch, für das Futur optional. Im Folgenden werden wir uns auf den im Sprachgebrauch am häufigsten vorkommenden Fall, das Perfekt der Modalverben, konzentrieren (für entsprechende Daten zu *lassen*; vgl. Bader 2014).

Die eigentlich zu erwartende Abfolge der Verben wird in (4) gezeigt. Diese gilt als ungrammatisch und stattdessen muss die Abfolge in (5) verwendet werden.²

- (4) a. *... dass Peter das Buch *lesen*₃ *müssen*₂ *hat*₁.
 b. *... dass das Buch anscheinend von Peter *gelesen*₄ *werden*₃ *müssen*₂ *hat*₁.
- (5) a. ... dass Peter das Buch *hat*₁ *lesen*₃ *müssen*₂.
 b. ... dass das Buch anscheinend von Peter *hat*₁ *gelesen*₄ *werden*₃ *müssen*₂.

Insgesamt gibt es für einen Verbalkomplex mit drei Verben sechs unterschiedliche Serialisierungen. Diese sechs Abfolgen lassen sich systematisch nach zwei

¹ Einen ausführlichen Überblick über die Verbalkomplexbildung im Westgermanischen bietet Wurmbrand (2004b 2017). Eine Auswahl unterschiedlicher syntaktischer Ansätze zur Analyse von Verbalkomplexen findet sich in Seuren/Kempen (Hg.) (2003).

² Eine weitere Besonderheit dieser Konstruktion ist der sog. Doppelinfininitiv oder IPP-Effekt (*infinitivus pro participio*): Das Modalverb erscheint hier als Infinitiv und nicht als Partizip, obwohl es sich um ein Perfekt handelt.

Faktoren klassifizieren: die Position des Auxiliars (1, 2 oder 3) und die Abfolge von lexikalischem Verb und Modalverb. Tabelle 1 zeigt die Verteilung der sechs verschiedenen Verbalkomplexe gemäß dieser beiden Faktoren.

	<i>Aux</i> = 1	<i>Aux</i> = 2	<i>Aux</i> = 3
<i>V</i> < <i>Mod</i>	Aux ₁ - <i>V</i> ₃ - <i>Mod</i> ₂ <i>hat</i> ₁ <i>lesen</i> ₃ <i>müssen</i> ₂	<i>V</i> ₃ - Aux ₁ - <i>Mod</i> ₂ <i>lesen</i> ₃ <i>hat</i> ₁ <i>müssen</i> ₂	<i>V</i> ₃ - <i>Mod</i> ₂ - Aux ₁ <i>lesen</i> ₃ <i>müssen</i> ₂ <i>hat</i> ₁
<i>Mod</i> < <i>V</i>	Aux ₁ - <i>Mod</i> ₂ - <i>V</i> ₃ <i>hat</i> ₁ <i>müssen</i> ₂ <i>lesen</i> ₃	<i>Mod</i> ₂ - Aux ₁ - <i>V</i> ₃ <i>müssen</i> ₂ <i>hat</i> ₁ <i>lesen</i> ₃	<i>Mod</i> ₂ - <i>V</i> ₃ - Aux ₁ <i>müssen</i> ₂ <i>lesen</i> ₃ - <i>hat</i> ₁

Tab. 1: Die sechs Permutationen eines Verbalkomplexes mit drei Elementen

Die Akzeptabilität von Verbalkomplexen wie in Tabelle 1 haben wir in zahlreichen Experimenten mittels Magnitude Estimation oder mittels binärer Grammatikalitätsurteile untersucht.³

Magnitude Estimation (ME) stammt ursprünglich aus der Psychophysik und wird dort verwendet, um den Zusammenhang zwischen physikalischen Größen und wahrgenommenen Größen zu untersuchen. Wenn beispielsweise ein Ton A die doppelte Intensität wie ein Ton B hat, kann gefragt werden, um wie viel lauter Ton A im Verhältnis zu Ton B wahrgenommen wird. Zur Beantwortung von Fragen dieser Art müssen die Versuchspersonen in einem ME-Experiment den Stimuli Zahlenwerte in Relation zu einem vorgegebenen Referenzstimulus zuweisen. Zuerst wird dem Referenzstimulus ein beliebiger Wert größer Null zugewiesen, beispielsweise der Wert 100. Anschließend werden die experimentellen Stimuli bewertet. Wenn, um bei unserem Beispiel zu bleiben, ein Ton als doppelt so laut wie der Referenzwert wahrgenommen wird, erhält er den Wert 200. Ein Ton, der dagegen nur als halb so laut wahrgenommen wird, erhält den Wert 50.

Die ME-Methode wurde von Bard/Robertson/Sorace (1996) und Cowart (1997) in die Linguistik eingeführt. Genau wie oben für ein Experiment aus der Psychophysik beschrieben, wird zunächst einem Referenzsatz eine Zahl größer Null zugewiesen. Der Referenzsatz, den wir in den weiter unten berichteten ME-Experimenten verwendet haben, wird in (6) gezeigt. Bei diesem Satz han-

³ Wir verwenden im Folgenden den Terminus 'Grammatikalitätsurteile' statt des Begriffs 'Akzeptabilitätsurteile', wenn Probanden explizit die Instruktion erhalten haben, die Grammatikalität von Sätzen zu beurteilen. Es hat nie Schwierigkeiten bereitet, dieses Konzept den Probanden anhand von Beispielen klar zu machen (z.B. *Peter haben verschlafen* ist ein ungrammatischer, d.h. 'inkorrekt' Satz). Vielen unserer Probanden ist dieser Begriff auch eher vertraut, da moderne Textverarbeitungsprogramme typischerweise eine Grammatikprüfung anbieten, aber keine Akzeptabilitätsprüfung.

delt es sich um eine leicht abgewandelte Version des Referenzsatzes aus Experiment 2 von Keller (2000).⁴

(6) Ich glaube, dass den Bericht der Chef in seinem Büro gelesen hat.

Die Versuchsperson weist dem Referenzsatz einen beliebigen Wert zu und beurteilt anschließend alle weiteren Sätze durch Zuweisung eines Zahlenwertes, der die Akzeptabilität des Satzes in Relation zum Referenzsatz widerspiegelt. Die dabei gewonnenen Werte können nur relativ zum Referenzwert interpretiert werden. Hat ein experimenteller Satz den Wert 50 zugewiesen bekommen, so bedeutet dies „doppelt so akzeptabel wie der Referenzsatz“, wenn dem Referenzsatz eingangs der Wert 25 zugewiesen worden ist. Hat der Referenzsatz dagegen den Wert 100 erhalten, bedeutet ein Wert von 50 „halb so akzeptabel wie der Referenzsatz“. Es gibt verschiedene rechnerische Verfahren, um die Abhängigkeit der Rohdaten vom Referenzwert bei der Auswertung zu berücksichtigen. Die ME-Daten, die im Folgenden präsentiert werden, wurden aus den Rohdaten gewonnen, indem jeder Datenpunkt durch den Referenzwert geteilt wurde. Ein Satz, der als gleich akzeptabel wie der Referenzsatz empfunden wird, hat nach dieser Transformation den Wert 1. Um die anschließende statistische Analyse zu ermöglichen, wird von den dabei resultierenden Werten zusätzlich der Logarithmus genommen, so dass die Daten zumindest näherungsweise einer Normalverteilung entsprechen. In den unten präsentierten Daten bedeutet damit der Wert 0, dass ein Satz als gleichermaßen akzeptabel bewertet wurde wie der Referenzsatz (da 0 der Logarithmus von 1 ist). Sätze mit Akzeptabilitätswerten größer 0 wurden folglich als akzeptabler als der Referenzsatz bewertet, Sätze mit Werten kleiner 0 dagegen als weniger akzeptabel.

Experimentelle Untersuchungen zur Akzeptabilität von Sätzen lassen sich auch mit Hilfe von binären Grammatikalitätsurteilen durchführen, wie sie aus der theoretischen Syntax hinlänglich bekannt sind. Binäre Grammatikalitätsurteile sind, wie bereits eingangs erwähnt wurde, ziemlich umstritten. Ein Teil dieser Kritik bezieht sich auf den Umstand, dass es sich typischerweise um die Urteile einer sehr kleinen Personengruppe handelt (häufig nur eine einzige Person), die eine sehr kleine Anzahl von Sätzen betreffen, welche die strukturellen Konfigurationen von Interesse aufweisen (ebenfalls häufig nur ein einziger Satz). Diese Kritikpunkte entfallen, wenn binäre Urteile wie andere experimentelle Daten unter kontrollierten Bedingungen erhoben werden. Dies bedeutet u.a., dass eine größere Anzahl von Sätzen von einer größeren Anzahl von Versuchspersonen beurteilt wird, so dass auch eine statistische Absicherung der Daten möglich ist.

⁴ Der Originalreferenzsatz von Keller (2000) lautet wie folgt:

(i) Ich behaupte, dass den Bericht der Chef im Büro liest.

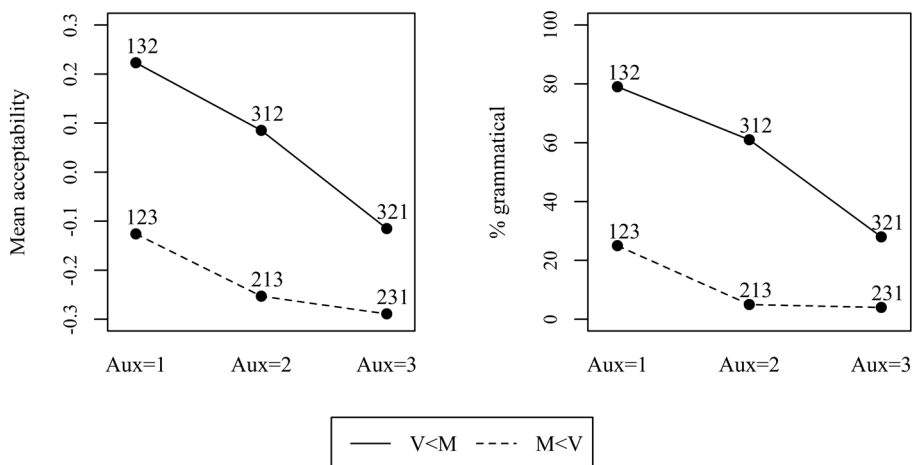


Abb. 1: Akzeptabilität gemessen mittels Magnitude Estimation (links) und mittels binärer Grammatikalitätsurteile (rechts) für Verbalkomplexe mit drei Verben

Repräsentative Daten für Verbalkomplexe mit drei Verben, die mit den beiden oben eingeführten Methoden gewonnen wurden, finden sich in Abbildung 1. Die linke Grafik zeigt Daten aus einem ME-Experiment, die rechte Grafik zeigt die Ergebnisse eines Experiments mit binären Grammatikalitätsurteilen.⁵ Die beiden Grafiken zeigen das gleiche Muster. Die beste Bewertung erhalten Sätze, bei denen wie in (7a) der Verbalkomplex in der standarddeutschen Abfolge AuxVMod auftritt.

- (7) a. ... dass Peter das Buch *hat*₁ *lesen*₃ *müssen*₂.
 b. ... dass Peter das Buch *lesen*₃ *hat*₁ *müssen*₂.

Die restlichen fünf Abfolgen, die im Standarddeutschen allesamt als ungrammatisch gelten, werden aber keineswegs im gleichen Ausmaß als ungrammatisch zurückgewiesen. Insbesondere werden Verbalkomplexe wie in Satz (7b), in denen das Auxiliar an zweiter Stelle steht und das lexikalische Verb dem Modalverb vorangeht, im ME-Experiment als akzeptabler als der Referenzsatz bewertet und im Experiment mit binären Grammatikalitätsurteilen in der überwiegenden Zahl der Fälle als grammatisch beurteilt. Alle anderen Abfolgen können dagegen als unakzeptabel betrachtet werden.

Bevor wir experimentelle Resultate für Verbalkomplexe mit vier Verben präsentieren, wollen wir noch einmal auf die experimentellen Methoden zurückkommen. Die Erhebung von binären Grammatikalitätsurteilen unter kontrol-

⁵ Die binären Daten entsprechen Tabelle 2 von Bader/Schmid (2009b). Die Magnitude-Estimation-Daten entsprechen Tabelle A3 aus Bader/Häussler (2010). Die ME-Resultate beruhen auf 24 Versuchspersonen und 30 Sätzen. Die Resultate der binären Grammatikalitätsurteile beruhen auf 30 Versuchspersonen und ebenfalls 30 Sätzen.

lierten experimentellen Bedingungen entkräftet einen wesentlichen Einwand gegen solche Urteile. Es bleibt aber noch ein weiterer Kritikpunkt gegen binäre Grammatikalitätsurteile, der die Binarität der Urteile selbst betrifft. Im Gegensatz zu ME-Urteilen oder anderen numerischen Urteilen, z.B. auf einer Punkteskala von 1-7, hat eine Versuchsperson im Falle binärer Grammatikalitätsurteile nicht die Möglichkeit, graduelle Abstufungen vorzunehmen. Trotz dieser offensichtlichen Limitation zeigen die Grafiken ein überraschend hohes Maß an Übereinstimmung zwischen den ME-Resultaten und den Resultaten der binären Grammatikalitätsurteile. Dies liegt zum einen daran, dass die Resultate für Experimente mit binären Grammatikalitätsurteilen in Prozenten ausgedrückt werden und damit einen beliebigen Wert zwischen 0 und 100 annehmen können. So werden aus binären Urteilen durch Mittlung kontinuierliche Daten. Dies allein garantiert aber noch nicht, dass die derart gewonnenen Mittelwerte in so auffälliger Weise übereinstimmen mit den Mittelwerten, die mittels der ME-Prozedur gewonnen wurden. Diese Übereinstimmung – die auch für andere Konstruktionen gefunden wurde (z.B. Bader/Häussler 2010; Fukuda et al. 2012) – deutet darauf hin, dass die Bewertung von Sätzen auf einer numerischen Skala und die Bewertung auf einer binären Skala nicht auf grundlegend unterschiedlichen mentalen Prozessen beruhen. Ein Modell, das die beiden Urteilstypen zueinander in Beziehung setzt, wird in Bader/Häussler (2010) vorgeschlagen.

Ein letzter Punkt, den es noch zu bedenken gilt, ist, dass durch die Mittlung über Versuchspersonen hinweg Ergebnismuster entstehen könnten, die nicht repräsentativ für individuelle Sprecher sind. Wenn für eine bestimmte Konstruktion beispielsweise ein Mittelwert von 60% gefunden wurde, so könnte dieser dadurch zustande gekommen sein, dass 60% der Versuchspersonen die Konstruktion zu 100% als grammatisch bewertet haben, während sie von den restlichen 40% in 0% der Fälle als grammatisch bewertet wurde. Ein solches Szenario könnte beispielsweise dann auftreten, wenn in der Gesamtheit der Versuchspersonen zwei unterschiedliche Dialekte im entsprechenden Mischungsverhältnis repräsentiert sind. Diese Möglichkeit muss bei jedem Experiment, das Akzeptabilitätsurteile erhebt, in Betracht gezogen werden. Eine Analyse der experimentellen Daten bezüglich dieser Fragestellung wird in Abbildung 2 illustriert.

In dem zur Diskussion stehenden Experiment hat jede Versuchsperson 30 Sätze bewertet, die sich gleichmäßig auf die 6 verschiedenen Permutationen verteilt haben. Pro Permutation gab es damit 5 Urteile für jede Versuchsperson, was 6 verschiedenen Prozentwerten entspricht (0%, 20%, 40%, 60%, 80%, 100%).⁶

⁶ Diese Werte kommen wie folgt zustande: 0% – keiner der fünf Sätze wurde als grammatisch bewertet; 20% – einer der fünf Sätze wurde als grammatisch bewertet; 40%, 60%, 80%, 100% – 2, 3, 4, 5 der fünf Sätze wurden als grammatisch bewertet.

Die Histogramme in Abbildung 2 zeigen für jede Permutation des Verbalkomplexes, wie häufig jeder Prozentwert bei den insgesamt 30 Versuchspersonen aufgetreten ist. Für Verbalkomplexe vom Typ VAuxMod (Aux=2, V vor Mod) reicht die Verteilung von 0% bis 100%, wobei die Verteilung bei ca. 60% ihren höchsten Wert besitzt. Dies bedeutet, dass der Mittelwert von ca. 60% repräsentativ für die Gesamtheit der Population ist, wobei wie zu erwarten Schwankungen nach oben und unten auftreten.

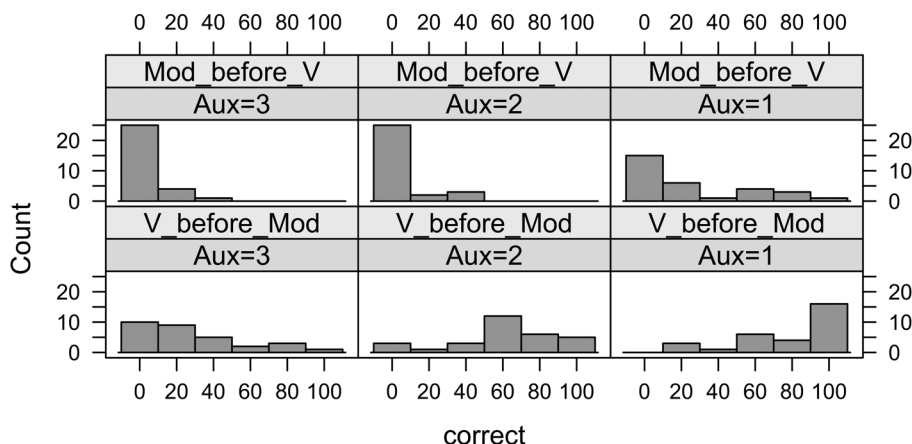


Abb. 2: Histogramm der Prozentwerte der 30 Versuchspersonen aus Experiment 1 von Bader/Schmid (2009b)

Eine genauere Analyse der Ergebnisse der einzelnen Versuchspersonen zeigt damit, dass das in Abbildung 1 gezeigte Akzeptabilitätsmuster für die Mehrheit der Versuchspersonen gilt und die gemittelten Ergebnisse keine Artefakte der Mittlung sind. Dass dem nicht immer so ist, zeigt die Untersuchung der Tun-Periphrase von Bader/Schmid (2006). Sätze des Typs *Peter tut das Buch bestimmt nicht lesen* wurden von einem Teil der Versuchspersonen mehrheitlich abgelehnt, von einem anderen Teil dagegen mehrheitlich akzeptiert.

Entsprechende Akzeptabilitätsdaten liegen auch für Verbalkomplexe mit vier Verben vor. Bei diesen Verbalkomplexen handelt es sich um Modalverben im Perfekt, die ein passiviertes Vollverb einbetten, d.h. Verbalkomplexe wie *hat₁ gelesen₄ werden₃ müssen₂*. Da die 24 unterschiedlichen Permutationen, die für vier Elemente möglich sind, den Rahmen eines Experimentes sprengen würden, wurden nur 8 Serialisierungen untersucht. Wie in Tabelle 2 zu sehen ist, folgt in allen Fällen das Passivauxiliar unmittelbar auf das Partizip des lexikalischen Verbs. Variiert werden lediglich die Position des Auxiliars (1, 2, 3 oder 4) sowie die Stellung des Modalverbs relativ zu dem Komplex aus Partizip und Passivauxiliar (V+Pass < Mod oder Mod < V+Pass).

	V+Pass < Mod	Mod < V+Pass
Aux = 1	Aux ₁ V ₄ Pass ₃ Mod ₂ <i>hat₁ gelesen₄ werden₃ müssen₂</i>	Aux ₁ Mod ₂ V ₄ Pass ₃ <i>hat₁ müssen₂ gelesen₄ werden₃</i>
Aux = 2	V ₄ Aux ₁ Pass ₃ Mod ₂ <i>gelesen₄ hat₁ werden₃ müssen₂</i>	Mod ₂ Aux ₁ V ₄ Pass ₃ <i>müssen₂ hat₁ gelesen₄ werden₃</i>
Aux = 3	V ₄ Pass ₃ Aux ₁ Mod ₂ <i>gelesen₄ werden₃ hat₁ müssen₂</i>	Mod ₂ V ₄ Aux ₁ Pass ₃ <i>müssen₂ gelesen₄ hat₁ werden₃</i>
Aux = 4	V ₄ Pass ₃ Mod ₂ Aux ₁ <i>gelesen₄ werden₃ müssen₂ hat₁</i>	Mod ₂ V ₄ Pass ₃ Aux ₁ <i>müssen₂ gelesen₄ werden₃ hat₁</i>

Tab. 2: Die acht experimentell untersuchten Permutationen eines Verbalkomplexes mit 4 Verben

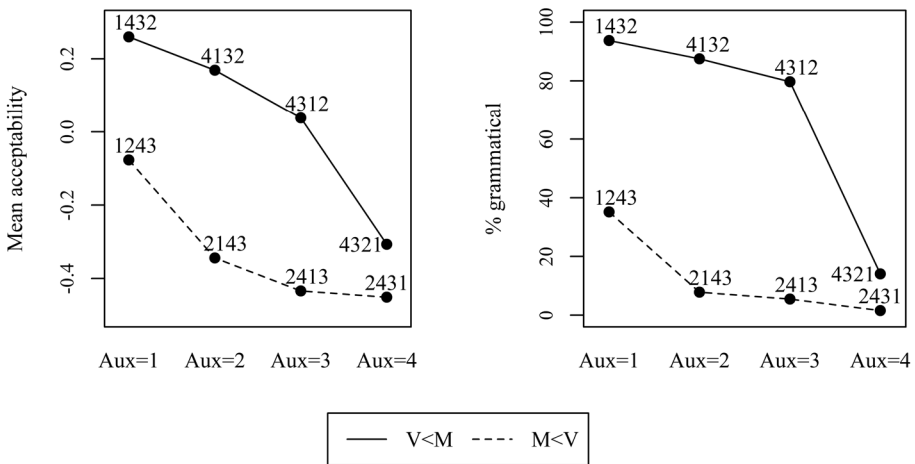


Abb. 3: Akzeptabilität gemessen mittels Magnitude Estimation (links) und mittels binärer Grammatikalitätsurteile (rechts) für Verbalkomplexe mit vier Verben

Die Ergebnisse aus zwei Akzeptabilitätsexperimenten werden in Abbildung 3 gezeigt.⁷ Wiederum zeigt sich ein hohes Maß an Übereinstimmung zwischen den ME-Daten und den Daten, die mittels binärer Grammatikalitätsurteile gewonnen wurden. Wie bei Verbalkomplexen mit drei Verben erhält auch bei Verbalkomplexen mit vier Verben die standarddeutsche Abfolge (*hat₁ gelesen₄ werden₃ müssen₂*) die beste Bewertung. Auch die beiden anderen Abfolgen, in denen das Modalverb dem Komplex aus Partizip und Passivauxiliar folgt und das finite Perfektauxiliar dem Modalverb vorangeht, d.h. an zweiter oder dritter Position steht (*gelesen₄ hat₁ werden₃ müssen₂* und *gelesen₄ werden₃ hat₁ müssen₂*),

⁷ Die binären Daten entsprechen Tabelle 10 von Bader/Schmid (2009b). Bei den Magnitude Estimation-Daten handelt es sich um unveröffentlichte Daten. Die ME-Resultate beruhen auf 40 Versuchspersonen und 32 Sätzen. Die Resultate der binären Grammatikalitätsurteile beruhen auf 32 Versuchspersonen und ebenfalls 32 Sätzen.

schneiden ziemlich gut ab. Im Experiment mit binären Grammatikalitätsurteilen erreichen diese beiden Abfolgen Akzeptabilitätswerte von über 80%. Die restlichen fünf Abfolgen schneiden dagegen relativ schlecht ab und werden bei binärer Beurteilung meist als ungrammatisch zurückgewiesen.

Von diesen fünf Abfolgen erhielt die Abfolge mit initialem Auxiliär und dem Modalverb in zweiter Position (Aux₁ Mod₂ V₄ Pass₃ – *hat₁ müssen₂ gelesen₄ werden₃*) mit ca. 40% Akzeptanz noch die beste Bewertung. Diese Abfolge wird in Bech (1983), dem Standardwerk zur Verbalkomplexbildung im Deutschen, als grammatisch betrachtet. In seiner Untersuchung des Schlesischen führt Dubenion-Smith (2011, S. 316) zwei Belege für Vierverbkomplexe an, von denen einer genau diese Verbabfolge besitzt (*bis Korn wieder hat₁ müssen₂ gesät₄ werden₃*). Im heutigen Deutsch scheint diese Abfolge dagegen nicht mehr akzeptiert zu werden.

Die bisher diskutierten Ergebnisse sind in (8) zusammengefasst. Standarddeutsch meint dabei das Deutsch, wie es in präskriptiven Grammatiken des Deutschen kodifiziert ist. Umgangsdeutsch dagegen ist das Deutsch, wie es von unseren Probanden, bei denen es sich generell nicht um Dialektsprecher handelt, tatsächlich akzeptiert und verwendet wird.⁸

- (8) Position des Perfektauxiliars in Verbalkomplexen vom Typ „Modalverb im Perfekt“
- a). Standarddeutsch und Umgangsdeutsch: Alle Verben außer dem Auxiliär stehen in absteigender Reihenfolge
 - b). Standarddeutsch: Das Auxiliär steht an erster Position
 - c). Umgangsdeutsch: Das Auxiliär steht vor dem Modalverb

Die Generalisierung in (8) wurde von Bader/Schmid (2009b) auf der Basis von Verbalkomplexen mit drei und vier Verben aufgestellt. Wenn diese Generalisierung stimmt, dann sollten für Verbalkomplexe mit fünf Verben sogar vier Auxiliärpositionen als akzeptabel bewertet werden. Diese Vorhersage wurde von Bader/Schmid/Häussler (2009) anhand von Verbalkomplexen wie in (9) überprüft. Das Perfektauxiliär sollte in initialer Position (in (9) gezeigt) sowie in den durch + markierten Positionen akzeptiert werden, während die finale Position zur Ungrammatikalität führen sollte.

⁸ Umgangs sprachliches Deutsch ist selbst eine Vereinfachung, da beispielsweise die Existenz von Regiolekten nicht berücksichtigt wird. Dass diese im Bezug auf den Verbalkomplex zu berücksichtigen sind, legt die Analyse des regionalen Hintergrundes der Versuchspersonen von Bader/Schmid (2009b) nahe. Es zeigte sich zwar das gleiche Muster unabhängig vom regionalen Hintergrund der Versuchspersonen, die numerischen Abstufungen weisen aber auf regionale Einflüsse hin. Aufgrund der sehr unterschiedlichen Anzahl von Versuchspersonen pro Regiolekt konnte dies statistisch nicht überprüft werden, so dass hier weitere Forschungen notwendig sind.

- (9) ... dass das Fahrrad **hätte**₁ repariert + worden + sein + müssen.

Die sechs Live-Musiker saßen bzw. standen vor einem Werbeplakat, welches in New York hätte fotografiert worden sein können. Quelle: www.musicalfotojournalismus.com/?cat=4&m=2014				
Daran war größtenteils meine Angst schuld, dass diese Serie, die ich so sehr liebe, zu sehr verhunzt hätte worden sein können. Quelle: www.digitalnippon.de/anime-manga-japan/focus-asia/asian-movies-tv/rurouni-kenshin_t18669/drucken/				
Den Anfang bestritt mit Werner Dannemann ein phantastischer Gitarrist, der, wüsste man nicht, dass er ein Schwabe durch und durch ist, auch direkt aus einem englischen Club der späten 60er- und frühen 70er-Jahre für diesen Abend entliehen worden hätte sein können. Quelle: www.swp.de/heidenheim/lokales/heidenheim/Guitar-Festival-Die-hohe-Schule-des-Blues-Rock;art1168893,2321196				
Noch härter ist ohne Zweifel „Neverlight“, das auch von großen Metal-Ikonen geschrieben worden sein hätte können. Quelle: www.mindbreed.de/rezensionen/tarja-colours-in-the-dark/				
Prozentsatz als grammatisch bewerteter Sätze				
Aux = 1	Aux = 2	Aux = 3	Aux = 4	Aux = 5
79	73	79	63	8

Tab. 3: Korpusbeispiele und Akzeptabilitätsdaten für Verbalkomplexes mit fünf Verben

Verbalkomplexe dieser Art sind extrem selten, kommen aber tatsächlich vor. Durch Suche mit Google konnten wir eine Reihe von Belegen finden. Je ein Beispiel für die Auxiliarpositionen 1-4 wird in Tabelle 3 gezeigt. Einen Beleg mit dem Auxiliar in finaler Position haben wir nicht gefunden, was natürlich nicht ausschließt, dass es solche Belege gibt. Zusätzlich zeigt Tabelle 3 die experimentellen Ergebnisse von Bader/Schmid/Häussler (2009).⁹ Die Ergebnisse zeigen einen drastischen Unterschied zwischen den Verbalkomplexen mit dem Auxiliar vor dem Modalverb und dem Auxiliarkomplex mit dem Auxiliar in finaler Position: Während erstere häufiger akzeptiert als zurückgewiesen werden, werden letztere fast immer als ungrammatisch bewertet. Trotz der relativ hohen Komplexität von Verbalkomplexen dieser Art hatten die Versuchspersonen also eine ziemlich klare Intuition darüber, was erlaubt ist und was nicht.

⁹ Die Daten in Tabelle 3 stimmen nicht genau mit den Daten von Bader/Schmid/Häussler (2009) überein, da nach Fertigstellung dieses Manuskripts weitere Versuchspersonen getestet wurden.

3. Vom Experiment zur Theorie und zurück

Um die Verzahnung von Experiment und Theorie zu illustrieren, geben wir in diesem Abschnitt einen kurzen Überblick über die von Bader/Schmid (2009b) vorgeschlagene syntaktische Analyse der Verbalkomplexe im Deutschen, wobei wir uns auf Verbalkomplexe mit Modalverben im Perfekt beschränken. Die Analyse von Bader/Schmid (2009b) beruht auf der Theorie von Williams (2003), die wiederum zur Familie der kategorialegrammatischen Analysen der Syntax von Verbalkomplexen gehört (siehe Bader/Schmid 2009b für weitere Referenzen).

Williams (2003) definiert die formale Sprache CAT, die den Aufbau syntaktischer Strukturen mittels der Kombinationsregel in (10) erlaubt.

$$(10) \quad \text{Rule of Combination} \\ X_{_Y} + Y_{_Z} \rightarrow [X + Y]_{X_{_Z}}$$

Diese Regel ist wie folgt zu lesen. 'X_{_Y}' ist eine syntaktische Einheit der Kategorie X, die eine syntaktische Einheit der Kategorie Y subkategorisiert. 'Y_{_Z}' ist entsprechend eine syntaktische Einheit der Kategorie Y, die eine syntaktische Einheit der Kategorie Z subkategorisiert. Wenn 'X_{_Y}' und 'Y_{_Z}' mittels der Rule of Combination kombiniert werden, entsteht eine neue Einheit, die die syntaktische Kategorie X besitzt und Z subkategorisiert. In anderen Worten: Die komplexe Kategorie 'X_{_Z}' ist vom gleichen Typ wie der Kopf 'X_{_Y}' und besitzt die gleichen Subkategorisierungseigenschaften wie das Komplement 'Y_{_Z}'.

In kategorialegrammatischen Begriffen ausgedrückt erlaubt die Rule of Combination sowohl funktionale Applikation (wenn das Subkategorisierungsmerkmal Z von Y in (10) leer ist) als auch funktionale Komposition (wenn Z nicht leer ist). Dies wird in (11) illustriert.

- (11) a. Funktionale Applikation b. Funktionale Komposition



Die Grammatik einer Einzelsprache wird in CAT durch Festlegung von Subkategorisierungsrahmen für die lexikalischen Elemente der Sprache spezifiziert. Spezifiziert werden können u.a. die Richtung der Selektion (nach links oder nach rechts, im Folgenden mit \leftarrow bzw. \rightarrow symbolisiert) sowie die syntaktische Kategorie des Komplements. Zur Illustration werden in (12) einige Lexikoneinträge für Verben des Deutschen gezeigt.

- (12) Auswahl lexikalischer Einträge für Verben im Deutschen
- a. Lexikal. Verben — $V_{\text{Main}}: \text{DP} \leftarrow \mid V_{\text{Main}}: \text{PP} \leftarrow \mid V_{\text{Main}}: \text{DP PP} \leftarrow \mid \dots$
 - b. Modal- und Auxiliärverben — $V_{\text{Mod} \mid \text{Aux}}: \text{V} \leftarrow$

Um Verbalkomplexe mit Modalverben im Perfekt analysieren zu können, ist noch ein weiteres Subkategorisierungsmerkmal notwendig. Die Modalverben im Deutschen können nicht nur mit Verben kombiniert werden, sondern auch direkt mit nicht-verbalen Komplementen, teilweise mit NPs (*ein Bonbon wollen*) oder genereller mit PPs. Wie in (13a) gezeigt wird, verhalten sich Modalverben ohne verbales Komplement im Standarddeutschen so wie alle Verben auch, was die Bildung des Perfekts betrifft. Nur zusammen mit einem verbalen Komplement wie in (13b) erfolgt die Voranstellung des Perfektauxiliars.

- (13) a. ... das Peter nach Paris (gewollt hat | *hat wollen).
 b. ... das Peter nach Paris (*fahren gewollt hat | hat fahren wollen).

Aus diesem Grund hat Williams (2003) ein Komplexitätsmerkmal eingeführt, das in (14) definiert wird (unter einer anderen Bezeichnung als bei Williams, der von „stem“ bzw. „non-stem“ spricht).

- (14) Komplexitätsmerkmal für Verbalkomplexe
- a. Verbalkomplex: [VV]
 - b. Einfaches Verb: [V]

Mit Hilfe des Komplexitätsmerkmals können wir jetzt die Subkategorisierungsrahmen für das Perfektauxiliar festlegen, wenn dieses ein Modalverb subkategorisiert. In Abhängigkeit von der Spezifikation des Komplexitätsmerkmals erhalten wir die drei Systeme in (15).

(15)

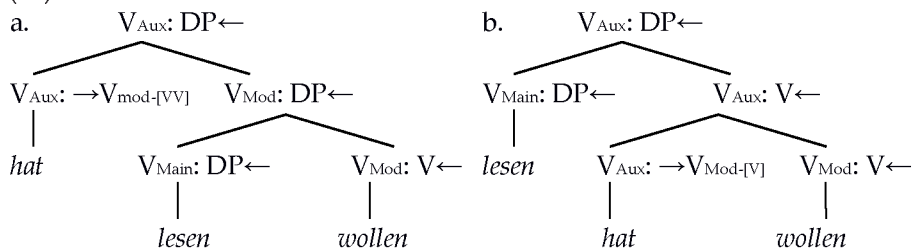
	System 1: → $V_{\text{Mod}}\text{-[VV]}$	System 2: → $V_{\text{Mod}}\text{-[V]}$	System 3: → V_{Mod}
AUX V MOD	√	–	√
V AUX MOD	–	√	√

In System 1 selegiert das Perfektauxiliar komplexe Modalverben nach rechts. Dieses System entspricht dem Standarddeutschen, wo sich Modalverben ohne verbales Komplement wie alle anderen Verben verhalten (*gewollt hat*), während Modalverben mit verbalem Komplement Voranstellung des Perfektauxiliars erfordern (*hat lesen wollen*). System 2 fordert die Voranstellung des Perfektauxiliars vor ein nicht-komplexes Modalverb. Dies entspricht dem von Louden (1990) beschriebenen Pennsylvania-Deutsch, das nur Verbalkomplexe vom Typ

lesen hat wollen zulässt.¹⁰ System 3 schließlich ist unterspezifiziert bezüglich des Komplexitätsmerkmals. Dieses System entspricht dem Umgangsdeutschen, das Verbalkomplexe vom Typ *hat lesen wollen* ebenso zulässt wie Verbalkomplexe vom Typ *lesen hat wollen*.

Für Verbalkomplexe des Typs AuxVMod und VAuxMod werden in (16) die syntaktischen Strukturen gezeigt, die gemäß der in diesem Abschnitt referierten Analyse erzeugt werden.

(16)



Die hier vorgestellte syntaktische Analyse wurde auf der Grundlage experimenteller Ergebnisse entwickelt. Sie macht aber auch Vorhersagen, die experimentell überprüft werden können. Eine Vorhersage betrifft das Verhältnis von Verbalkomplexen, in denen ein Modalverb ein lexikalisches Verb selegiert, und solchen, wo es dies nicht tut. Wie oben anhand von Beispiel (13) erläutert, ist im Standarddeutschen Auxiliarinversion im Falle eines Modalverbs ohne verbalem Komplement nicht erlaubt. Dies wird in der syntaktischen Analyse durch das Komplexitätsmerkmal VV im Subkategorisierungsrahmen sichergestellt. Da dieses Merkmal im Umgangsdeutschen nicht vorhanden ist, ist es hier möglich, ein vorangestelltes Perfektauxiliar mit einem Modalverb zu kombinieren, auch wenn das Modalverb kein lexikalisches Verb selegiert.

Wir erhalten damit die Vorhersage, dass unabhängig von der Größe des Verbalkomplexes im Standarddeutschen niemals die Sequenz AuxMod zulässig ist, im Umgangsdeutschen dagegen immer. Diese Vorhersage haben Bader/Schmid (2009a) getestet, indem sie Versuchspersonen sowohl Verbalkomplexe vom Typ VAuxMod als auch Verbalkomplexe vom Typ AuxMod beurteilen ließen. Dabei zeigte sich eine signifikante Korrelation. Versuchspersonen, die VAuxMod-Komplexe als akzeptabel beurteilt haben, haben auch AuxMod-Komplexe akzeptiert, während Versuchspersonen, die VAuxMod-Komplexe

¹⁰ Die partiell invertierte Verbabfolge VAuxMod ist laut Louden (1990) die einzig grammatisch zulässige Abfolge. Dies unterscheidet das Pennsylvania-Deutsch von anderen Varianten des Deutschen, in denen diese Abfolge zwar präferiert wird, aber zusätzlich andere Abfolgen als grammatisch gelten, wie das Bairische (siehe Weiß 1998) und bestimmte Varietäten des Österreichischen (siehe Wurmbbrand 2004b).

als eher wenig akzeptabel beurteilt haben, dazu tendierten, AuxMod-Komplexe als ungrammatisch zurückzuweisen. Auf der Ebene individueller Akzeptabilitätsdaten zeigt sich damit der vorhergesagte Zusammenhang zwischen VAuxMod und AuxMod.

Wie andere kategorialgrammatische Analysen des Verbalkomplexes ist die hier vorgestellte Analyse sehr flexibel. Angesichts der bekanntermaßen hohen Variation, die die westgermanischen Sprachen in diesem Bereich der Syntax zeigen, kann dies durchaus als ein Vorteil gesehen werden. Es kann aber auch nicht von vornherein ausgeschlossen werden, dass der kategorialgrammatische Ansatz zu flexibel ist, indem er Abfolgen zu erzeugen erlaubt, die aus grundsätzlichen Gründen ausgeschlossen sind. Beispielsweise können alle Abfolgen eines Verbalkomplexes mit drei Verben erzeugt werden, aber es ist behauptet worden, dass die Abfolge V2-V1-V3 (*müssen hat lesen*) in keiner der westgermanischen Sprachen zu finden sei (z.B. Barbiers 2005). Dieser Behauptung widersprechen Daten, die Schmid/Vogel (2004) sowie Salzmann (2013) anführen. Sollten diese Daten einer genaueren Prüfung standhalten, würde dies bedeuten, dass alle sechs möglichen Abfolgen eines Verbalkomplexes mit drei Verben belegt wären (für weitere Diskussion siehe Abels 2016; Salzmann 2016; Wurmbrand 2017).

Eine weitere Eigenschaft von Verbalkomplexen, hinsichtlich derer der kategorialgrammatische Ansatz zu flexibel sein könnte, betrifft das Auftreten nicht-verbalen Materials im Verbalkomplex, das sog. Verb Projection Raising (VPR). In Verbalkomplexen vom Typ AuxVMod darf im Deutschen bekanntlich nicht-verbales Material zwischen Auxiliar und lexikalischem Verb stehen, wenn dieses Material in einer besonders engen Beziehung zum lexikalischen Verb steht. Dies wird in (17) illustriert.

- (17) ... dass Peter ohne den Lottogewinn niemals *hätte*₁ nach Paris *fahren*₃ *können*₂.

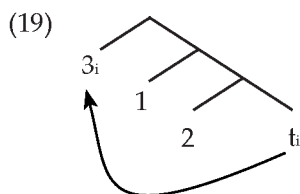
Im Standard-Niederländischen existiert diese Möglichkeit nicht. In Haider (2003) wird deshalb für das Deutsche eine grundlegend andere syntaktische Struktur des Verbalkomplexes vorgeschlagen als für das Niederländische, so dass der Unterschied zwischen den beiden Sprachen bezüglich VPR aus prinzipiellen Gründen folgt. Mittels lexikalischer Spezifikation der phrasalen Ebene im Sinne der X-bar-Theorie kann der oben vorgestellte Ansatz den Unterschied ebenfalls erfassen. Im Deutschen können Modalverben sowohl mit einem bloßen Verb als auch mit einer verbalen Projektion kombiniert werden (unspezifiziertes Ebenenmerkmal), im Niederländischen dagegen nur mit einem bloßen Verb (Ebenenmerkmal V°). Im Gegensatz zu Haiders Analyse ist dies aber kein prinzipieller syntaktischer Unterschied, sondern einer, der lexikalisch stipuliert werden muss. Die Frage ist dann, wie grundlegend der Unterschied

zwischen Deutsch und Niederländisch bezüglich VPR ist. Im Gegensatz zum Standard-Niederländischen sind bestimmte niederländische Dialekte deutlich toleranter, was nicht-verbales Material im Verbalkomplex betrifft (siehe Hendriks 2014 und die dort zitierte Literatur). Dies kann als Indiz gewertet werden, dass die Eigenschaft, nicht-verbales Material im Verbalkomplex zuzulassen, als kontinuierliche Eigenschaft verstanden werden kann, mit schweizerdeutschen Dialekten als besonders freizügigen Systemen auf der einen Seite, Standarddeutsch ungefähr in der Mitte, und Standard-Niederländisch mit besonders restriktiven Regeln auf der anderen Seite. In diesem Falle würde die Motivation für prinzipiell unterschiedliche Grammatiksysteme wegfallen.

Eine restriktivere syntaktische Analyse von Verbalkomplexen, die wie die hier vorgestellte Analyse auf der kategorialgrammatischen Sprache CAT von Williams (2003) basiert, wird in Wurmbrand (2004a) vorgeschlagen. Für die Basisgenerierung lässt Wurmbrand (ebd.) nur funktionale Applikation zu, so dass die hierarchische Struktur eines Verbalkomplexes die zugrundeliegenden Selektionsbeziehungen direkt widerspiegelt. Die lineare Abfolge ist dagegen variabel. Unter diesen Annahmen lassen sich die vier in (18) gezeigten Abfolgen erzeugen.

(18) a. [1 [2 3]] b. [1 [3 2]] c. [[2 3] 1] d. [[3 2] 1]

Nicht erzeugen lassen sich die beiden Abfolgen, in denen das bezüglich der Selektion höchste Element 1 zwischen den beiden tieferen Elementen 2 und 3 steht. Hierbei handelt es sich zum einen um die Abfolge V2-V1-V3 (*müssen hat lesen*), deren Existenz, wie oben erwähnt wurde, umstritten ist. Zum anderen kann aber auch die weitverbreitete Abfolge V3-V1-V2 (*lesen hat müssen*) nicht erzeugt werden. Hier kommt nun eine weitere Annahme ins Spiel, die besagt, dass die Abfolgevariation, die durch CAT erzeugt wird, keine semantisch-pragmatischen Konsequenzen hat und deshalb als Teil der PF-Komponente der Grammatik zu betrachten sei. Eine zusätzliche Quelle der Variation ist syntaktische Bewegung, die Konsequenzen für die Interpretation hat. Mittels Bewegung kann die Abfolge V3-V1-V2, wie in (19) gezeigt, erzeugt werden.



Der interpretative Effekt, den die Bewegung des lexikalischen Verbs an den Anfang des Verbalkomplexes auslöst, besteht laut Wurmbrand (2004a) darin, dass V3 in der Abfolge V3-V1-V2 eng fokussiert sein muss. In Abschnitt 5

werden wir zwei Experimente diskutieren, die diesen Vorschlag überprüft haben. Zuvor werden wir in Abschnitt 4 einen kurzen Blick auf das Verhältnis von Akzeptabilität und Korpusfrequenz werfen.¹¹

4. Zum Verhältnis von Korpusdaten und Akzeptabilitätsurteilen

Abbildung 4 zeigt unveröffentlichte Korpusdaten für Verbalkomplexe mit drei und vier Verben, wie sie oben experimentell untersucht wurden. Diese Daten beruhen auf einer Analyse des deWac-Korpus, einer großen Sammlung von Internettexten (Baroni et al. 2009). Die Korpusdaten zeigen ein anderes Bild als die experimentellen Daten in Abschnitt 2. Die akzeptabelste Verbabfolge, d.h. die standardsprachliche Abfolge, tritt mit weitem Abstand am häufigsten auf. Die nur geringfügig weniger akzeptable Abfolge mit dem Auxiliar an zweiter Position ((20a) und (20b)) tritt ebenfalls auf, aber nur mit einem sehr geringen Anteil. Im Falle von Verbalkomplexen mit vier Verben tritt darüber hinaus gelegentlich auch die Abfolge (20c) mit dem Auxiliar in dritter Position auf. Alle anderen Abfolgen wurden gar nicht oder mit verschwindend geringer Frequenz gefunden.

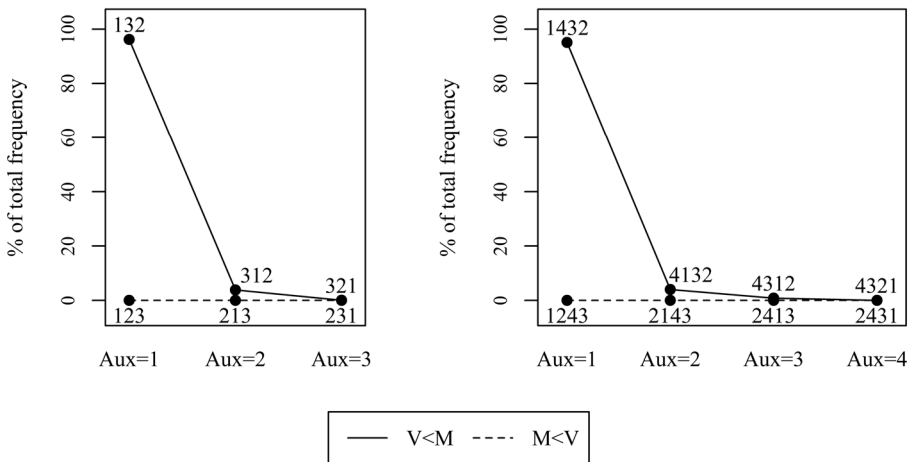


Abb. 4: Korpusdaten für Verbalkomplexe mit drei Verben ($n = 20213$) (links) und für Verbalkomplexe mit vier Verben ($n = 3195$) (rechts)

¹¹ Die Sprache CAT spielt in dieser Analyse nur insofern eine Rolle, als sie eine flexible Linearisierung erlaubt. Wird generell von flexibler Linearisierung ausgegangen, wird der Bezug auf CAT überflüssig. In neueren Arbeiten wie Abels (2016) und Wurmbrand (ersch. demn.) wird dann diese Art von Analyse auch ohne Bezug auf CAT und ohne informationstrukturellen Trigger für Bewegung weitergeführt.

- (20) a. ... dass Peter das Buch *lesen*₃ *hat*₁ *müssen*₂.
 b. ... dass das Buch anscheinend von Peter *gelesen*₄ *hat*₁ *werden*₃ *müssen*₂.
 c. ... dass das Buch anscheinend von Peter *gelesen*₄ *werden*₃ *hat*₁ *müssen*₂.

Angesichts der recht geringen Korpusfrequenzen für alle Abfolgen außer der standarddeutschen Abfolge könnte man vermuten, dass es sich bei den nicht-standarddeutschen Belegen um Fehler, beispielsweise Verschreiber, handelt. Gerade da das analysierte Korpus ein Korpus von Internettexten ist, die in nicht unerheblicher Zahl ohne redaktionelle Bearbeitung veröffentlicht worden sein dürften, ist mit Fehlern zu rechnen. Ein Vergleich mit den experimentellen Daten spricht aber dafür, dass wir es mit regulär produzierten Belegen zu tun haben. Das Muster, das sich zeigt und das auch für andere Konstruktionen gefunden wurde (Featherston 2005; Kempen/Harbusch 2008), lässt sich wie folgt zusammenfassen: (i) Ein geringfügiger Akzeptabilitätsunterschied zwischen zwei absolut gesehen akzeptablen Strukturen führt zu einem starken Frequenzunterschied. (ii) Strukturen, die in unterschiedlichem Ausmaß als unakzeptabel bewertet werden, tendieren dazu, überhaupt nicht produziert zu werden.

Die Diskrepanz zwischen Akzeptabilitätsbewertungen und Frequenzdaten schließt nicht aus, dass es eine gesetzmäßige Beziehung zwischen Akzeptabilität und Frequenz gibt. Was ausgeschlossen wird, ist lediglich, dass es sich dabei um eine lineare Beziehung handelt. Theorien, die mit gewichteten Beschränkungen arbeiten (z.B. Featherston 2005; Pater 2009), können das beobachtete Muster erklären, wenn sie annehmen, dass Produktionsfrequenzen auf einem Wettbewerb zwischen alternativen Kandidaten beruhen, Akzeptabilitätsurteile dagegen nicht. Die Akzeptabilität eines Kandidaten – wobei ein Kandidat in unserem Falle eine bestimmte Serialisierung eines Verbalkomplexes ist – wird ausschließlich dadurch bestimmt, welche syntaktischen Beschränkungen ein Kandidat erfüllt oder verletzt. Bei der Sprachproduktion dagegen wird aus der Menge aller Kandidaten derjenige mit der besten Bewertung ausgewählt. Wenn diese Auswahl nicht völlig deterministisch ist, sondern ein gewisses Maß an zufälligem Rauschen einschließt, wird der beste Kandidat meistens, aber nicht immer, gewinnen – zumindest dann, wenn konkurrierende Kandidaten vorhanden sind, die nur geringfügig weniger akzeptabel sind als der beste Kandidat.

Aus Platzgründen müssen wir darauf verzichten, anhand von konkreten Simulationen zu überprüfen, inwieweit gewichtete Beschränkungen sowohl die Akzeptabilitätsdaten als auch die Korpusdaten modellieren können, wenn wir von einem Wettbewerb bei der Produktion, aber nicht beim Bewerten von Sätzen ausgehen. Zum Abschluss dieses Abschnitts über das Verhältnis von

Akzeptabilität und Korpusfrequenz wollen wir deshalb lediglich noch kurz die Frage aufwerfen, was zuerst kommt, Akzeptabilität oder Korpusfrequenz. Geht man beispielsweise von gewichteten Beschränkungen aus, lässt sich die Frage umformulieren zu: Reflektieren die Gewichte primär die Akzeptabilität und die beobachtbaren Korpusfrequenzen sind eine sekundäre Konsequenz der Gewichte, oder ist es umgekehrt, und die Gewichte werden aus Korpusfrequenzen gelernt und bestimmen dann anschließend die Akzeptabilität der jeweiligen Abfolgen?

Empirische Untersuchungen anderer Phänomene (z.B. Arppe/Järvikivi 2007) deuten darauf hin, dass es nicht möglich ist, Akzeptabilität generell aus Korpusfrequenzen oder Korpusfrequenzen generell aus Akzeptabilität abzuleiten. Inwieweit das eine oder das andere im Falle der Verbalkomplexbildung möglich ist, muss als eine offene Frage betrachtet werden.

5. Experimentelle Untersuchungen der Sprachproduktion

Die bislang diskutierten Experimente haben alle die Akzeptabilität von Sätzen untersucht. Wie wir in diesem Abschnitt zeigen wollen, können auch Experimente zur Sprachproduktion zu einem besseren Verständnis syntaktischer Phänomene beitragen. Die konkrete Frage, um die es dabei gehen soll, betrifft den Einfluss prosodischer Faktoren auf die Serialisierung der Verben im Verbalkomplex.

In Abschnitt 3 haben wir den Vorschlag von Wurmbbrand (2004a) diskutiert, dass Verbalkomplexe vom Typ VAuxMod nur dann grammatisch sind, wenn V eng fokussiert ist. Evidenz für den Einfluss prosodischer Faktoren im Allgemeinen und diesen Vorschlag im Speziellen findet sich in Untersuchungen zur Sprachgeschichte (z.B. Ebert 1981; Sapp 2011) sowie in Untersuchungen zur Syntax verschiedener deutscher Dialekte (z.B. Schmid/Vogel 2004; Dubenion-Smith 2011). Der größte Teil der vorliegenden Evidenz betrifft dabei Verbalkomplexe aus zwei Verben, die in früheren Stufen des Deutschen sowie in verschiedenen Dialekten weitaus variabler hinsichtlich der Abfolge der Verben waren als dies im heutigen Deutsch der Fall ist. Im Falle des Perfekts war damit beispielsweise sowohl die Abfolge VAux als auch die Abfolge AuxV möglich, wie in Beispiel (21) illustriert (aus Gründen der Einfachheit in standardsprachlicher Form).

- (21) a. ... dass Peter einen ONKEL hat₁ besucht₂ | besucht₂ hat₁.
 b. ... dass Peter den Onkel hat₁ BESUCHT₂ | BESUCHT₂ hat₁.

In mehreren Korpusstudien wurde gefunden, dass die Abfolge AuxV eher verwendet wird, wenn dem Verbalkomplex ein indefinites Objekt vorangeht, als wenn ihm ein definites Objekt vorangeht (Dubenion-Smith 2011; Ebert 1981; Sapp 2011). Da, wie ebenfalls in (21) illustriert, der Hauptakzent im Falle eines

indefiniten Objekts typischerweise auf das Nomen des Objekts fällt, im Falle eines definiten Objekts dagegen auf das folgende lexikalische Verb, wird dieser Befund von den zitierten Autoren als Evidenz für eine rhythmische Präferenz interpretiert, die bereits von Behagel postuliert worden ist.

Effekte dieser Art wurden aber nicht konsistent gefunden, insbesondere nicht für Verbalkomplexe mit drei Verben. Da dies vermutlich an der zu schmalen Datenbasis für Verbalkomplexe mit mehr als zwei Verben liegt, bietet es sich an, dies experimentell zu untersuchen, da dadurch Daten in ausreichendem Maße erzeugt werden können. Wir haben folglich zwei Experimente durchgeführt, in denen der Einfluss der Informationsstruktur durch die unterschiedliche Position einer Fokuspartikel manipuliert wurde.¹² Die Fokuspartikel konnte entweder direkt vor dem Objekt stehen, wie in (22a), oder direkt vor dem Verbalkomplex, wie in (22b).

- (22) a. ... dass Peter sogar das BUCH hat₁ lesen₃ müssen₂ | lesen₃ hat₁ müssen₂.
 b. ... dass Peter das Buch sogar hat₁ LESEN₃ müssen₂ | LESEN₃ hat₁ müssen₂.

Sätze von der Struktur in (22) wurden digital aufgenommen, wobei der Hauptakzent in (22a) auf dem Nomen und in (22b) auf dem lexikalischen Verb lag. In (22a) haben wir es folglich mit Objektfokus, in (22b) mit Verbfokus zu tun. Die aufgenommenen Sätze wurden in zwei Experimenten mit unterschiedlichen Aufgabenstellungen Probanden vorgespielt. In einem ersten Experiment mussten die Probanden die gehörten Sätze unmittelbar nach Satzende als grammatisch oder ungrammatisch klassifizieren. Dieses Experiment ist damit identisch zu den oben besprochenen Experimenten, außer dass die Satzpräsentation auditiv erfolgte.

Die Ergebnisse, die in Abbildung 5 zu sehen sind, zeigen den erwarteten Effekt der Auxiliarposition. Wenn das Auxiliar am Anfang des Verbalkomplexes stand, wurden die Sätze zu ca. 90% als grammatisch bewertet; stand es dagegen an zweiter Position, war die Akzeptanzrate mit ca. 80% etwas niedriger, aber immer noch ausgesprochen hoch.¹³ Im Gegensatz zur Position des Auxiliars hatte die Position der Fokuspartikel – und damit die Position des

¹² Den Einfluss von Rhythmus auf die Stellung des Auxiliars in Verbalkomplexen wurde von Vogel et al. (2015) untersucht, wobei ähnlich wie hier ein Akzeptabilitätsexperiment mit einem Produktionsexperiment kombiniert wurde.

¹³ In diesem Experiment gab es keine Verbalkomplexe mit finalem Auxiliar, diese waren aber in einem weiteren Experiment mit auditiver Präsentation enthalten. Solche Verbalkomplexe wurden nur zu etwas mehr als 20% als grammatisch bewertet. Auch bei auditiver Satzpräsentation zeigt sich damit ein scharfer Unterschied zwischen grammatischen und ungrammatischen Verbalkomplexen.

Hauptakzent – keinen Effekt auf die Akzeptabilität der Sätze. Dieser Befund repliziert einen früheren Befund von Bader/Schmid (2009b), allerdings mit auditiver statt mit visueller Satzpräsentation. Die Interpretation der Ergebnisse von Bader/Schmid (ebd.) beruhte auf zwei Hypothesen, deren Korrektheit nicht als a priori gegeben betrachtet werden kann. Die erste Hypothese besagt, dass beim stillen Lesen eine prosodische Struktur berechnet wird, die der prosodischen Struktur beim Sprechen entspricht (vgl. Fodor 2002). Die zweite Hypothese besagt, dass beim Lesen der Hauptakzent auf das Nomen bzw. auf das lexikalische Verb gelegt wird. Dies ist nicht zwingend so, da beispielsweise auch der Artikel anstatt des Nomens oder das Modalverb anstatt des lexikalischen Verbs betont werden kann. Die Tatsache, dass das Experiment mit auditiver Stimuluspräsentation sehr ähnliche Ergebnisse erzielt hat wie das Experiment mit visueller Stimuluspräsentation, liefert starke Evidenz für die Korrektheit dieser beiden Hypothesen.

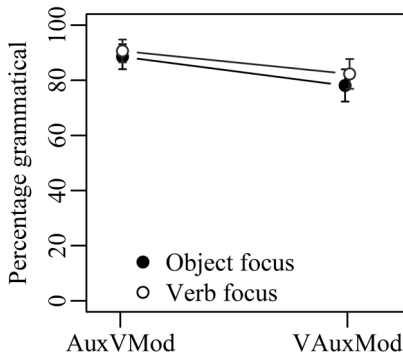


Abb. 5: Akzeptabilitätsergebnisse für Sätze mit Fokuspartikel (unveröffentlichte Daten)

In einem zweiten Experiment, das dieselben Sprachaufnahmen wie das soeben beschriebene Akzeptabilitätsexperiment verwendet hat, wurde der Einfluss von Fokus auf die Sprachproduktion untersucht. Ein generelles Problem bei der Erforschung der Prozesse, die beim Produzieren von Sätzen für die Berechnung syntaktischer Strukturen verantwortlich sind, ist, wie man Versuchspersonen dazu bringt, Sätze von der Art zu produzieren, die für die Fragestellung des Experiments von Interesse sind. In einfachen Fällen kann man Versuchspersonen Bilder beschreiben lassen. Diese Methode stößt aber schnell an ihre Grenzen, wenn z.B. Inhalte verbalisiert werden sollen, die sich nicht bildlich darstellen lassen. Als eine alternative Methode hat sich die Satz-wiedergabe aus dem Gedächtnis (*production from memory*) als eine Standardmethode zur Untersuchung syntaktischer Prozesse bei der Sprachproduktion etabliert. In einem Experiment dieser Art hören Versuchspersonen einen oder mehrere Sätze. Sie müssen sich diese Sätze merken und anschließend aus dem Gedächtnis wiedergeben.

Wenn man einen nicht allzu langen Satz hört und direkt wiederholen muss, so ist normalerweise eine wortwörtliche Wiedergabe problemlos möglich (Fletcher 1994). Das Gedächtnis für den genauen Wortlaut schwindet allerdings schnell, wenn die Wiedergabe mit einer gewissen Verzögerung erfolgt, indem z.B. eine kurze Zusatzaufgabe zwischen Hören und Wiedergabe eingefügt wird oder indem erst nach der Präsentation mehrerer Sätze mit der Wiedergabe begonnen wird. In diesem Fall muss ausgehend von der Satzbedeutung, die am wenigsten vom Vergessen betroffen ist, die Oberflächenstruktur neu generiert werden (Potter/Lombardi 1990). In diesem Sinne handelt es sich bei der Satz-wiedergabe aus dem Gedächtnis um eine Produktionsaufgabe.

McDonald/Bock/Kelly (1993) ist eine der klassischen Studien, die von den Versuchspersonen Satz-wiedergabe aus dem Gedächtnis verlangt haben. In dieser Studie mussten sich Versuchspersonen jeweils 12 kurze Texte aus zwei Sätzen anhören und merken. Anschließend wurde nur noch der erste Satz vorgelesen und die Versuchspersonen mussten den zweiten aus dem Gedächtnis wiedergeben. Mittels dieser Methode wurde u.a. untersucht, welchen Einfluss die Belebtheit der Argumente auf ihre Linearisierung hat. Zur Illustration werden in (23) und (24) einige ausgewählte Beispieltex-te gezeigt.

- (23) *Prompt:* Appliances were rare in rural America until after World War II. What occasioned a lot of talk in Deadwood, South Dakota, one week in March, 1940?
- | | | |
|----|--|-----|
| a. | <i>Zielsatz:</i> A farmer purchased a refrigerator. | 0% |
| b. | <i>Zielsatz:</i> A refrigerator was purchased by a farmer. | 32% |
- (24) *Prompt:* After investigating the loud rumbling in the hallway, the elementary school teacher returned to find her entire class under their desks. Why?
- | | | |
|----|---|----|
| a. | <i>Zielsatz:</i> The sound frightened the students. | 4% |
| b. | <i>Zielsatz:</i> The students were frightened by the sound. | 5% |

Jeder Text bestand aus einer Aufforderung (Prompt) und einem Zielsatz, der entweder im Aktiv oder im Passiv formuliert war. Die Prozentwerte hinter den Zielsätzen geben an, wie häufig das Genus Verbi des Zielsatzes bei der Wiedergabe geändert wurde. Ein Wert sticht dabei hervor: Sätze wie (23b), d.h. Passivsätze mit unbelebtem Subjekt und belebter by-Phrase, wurden zu 32% als Aktivsatz wiedergegeben, d.h. in der Form von (23a). Hierbei handelt es sich laut McDonald/Bock/Kelly (1993) um einen Effekt der Belebtheit auf die Argumentserialisierung. Aus einem Passivsatz mit unbelebtem vor belebtem Argument machen die Versuchspersonen einen Aktivsatz mit belebtem vor unbelebtem Argument.

Kommen wir damit zurück zu unserem Produktionsexperiment zur Serialisierung der Verben im Verbalkomplex. Wie die obige Diskussion klargemacht hat, muss man bei einem Experiment, das sich der experimentellen Methode der Wiedergabe aus dem Gedächtnis bedient, dafür sorgen, dass die Sätze nicht einfach wortwörtlich aus dem Arbeitsgedächtnis wiedergegeben werden können. Andererseits muss die Gedächtnisrepräsentation hinreichend detailliert bleiben, so dass die Wiedergabe alle Elemente enthält, die für die geplante Auswertung notwendig sind. In unserem Fall muss sichergestellt sein, dass ein Verbalkomplex mit drei Verben beispielsweise nicht – bzw. nicht allzu häufig – zu einem Verbalkomplex mit zwei Verben vereinfacht wird. Wird nämlich ein Verbalkomplex wie *hat lesen wollen* wiedergegeben als *lesen wollte*, kann die entscheidende Frage nach der Position des Auxiliars nicht beantwortet werden. Wir haben uns deshalb für eine Variante der Satz wiedergabe entschieden, die weniger Anforderungen an das Arbeitsgedächtnis stellt als die Variante von McDonald/Bock/Kelly (1993). Die Versuchsperson hört einen Satz und merkt sich ihn. Danach löst die Versuchsperson eine kleine Rechenaufgabe (Addition zweier Zahlen zwischen 1 und 50). Direkt im Anschluss gibt sie den Satz wieder.

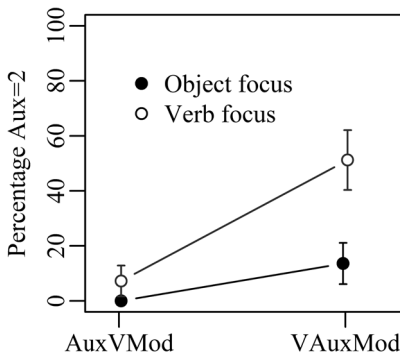


Abb. 6: Produktionsergebnisse für Sätze mit Fokuspartikel (unveröffentlichte Daten)

Die Ergebnisse für das beschriebene Produktionsexperiment werden in Abbildung 6 dargestellt. Im Gegensatz zur Beurteilung der Grammatikalität zeigt sich bei der Sprachproduktion ein klarer Fokuseffekt. Wenn die Versuchspersonen einen Satz mit dem Auxiliar in erster Position hörten, d.h. mit standardsprachlicher Abfolge der Verben, haben sie diesen bei Fokus auf dem Objekt niemals mit geänderter Stellung des Auxiliars wiedergegeben. Bei Fokus auf dem Verb kommt es in einem geringen Prozentsatz von 7% zu einem Stellungswechsel des Auxiliars, so dass das lexikalische Verb in die initiale Position gelangt. Wenn das lexikalische Verb bereits im gehörten Satz in initialer Position stand und das Auxiliar damit in zweiter Position, wurde bei Objektfokus mit 86%

fast immer zur Standardabfolge mit initialem Auxiliar gewechselt. Bei Verbfokus dagegen kam es bei der Wiedergabe nur in etwa der Hälfte der Fälle zu einer Änderung der Verbabfolge. Bei Verbfokus wurde damit ungefähr jeder zweite Satz, der mit der Abfolge VAuxMod gehört wurde, anschließend mit eben derselben Abfolge VAuxMod produziert.

Das Muster, das die in Abbildung 6 gezeigten Produktionsdaten aufweisen, stimmt in einer wesentlichen Hinsicht überein mit dem Muster der von McDonald/Bock/Kelly (1993) erzielten Daten. In beiden Fällen wird die unmarkierte Struktur – wobei unmarkiert im Sinne von Haspelmath (2006) als die am häufigsten auftretende Struktur verstanden werden kann – fast immer unverändert wiederholt, sogar dann, wenn eine Änderung der Struktur durch die experimentelle Manipulation favorisiert wird. Im Falle des Experiments von McDonald/Bock/Kelly (1993) sind dies Aktivsätze mit unbelebtem Subjekt und belebtem Objekt, im Falle der Verbalkomplexe sind dies Sätze mit Verbalkomplexen vom Typ AuxVMod und fokussiertem lexikalischen Verb. In beiden Fällen wurde in weniger als 10% der Fälle zur markierten Struktur gewechselt, obwohl die experimentellen Bedingungen einen solchen Wechsel nahegelegt haben. Die Ergebnisse für die markierte Struktur stimmen dagegen nur partiell überein. Sowohl im Experiment von McDonald/Bock/Kelly (ebd.) als auch in unserem Produktionsexperiment wurde die Ausgangsstruktur deutlich häufiger geändert, wenn sie den untersuchten Präferenzregeln widersprochen haben. Passivsätze wurden im Experiment von McDonald/Bock/Kelly (ebd.) in 32% der Fälle zu Aktivsätzen, wenn dadurch eine Übereinstimmung mit der Präferenz „belebt vor unbelebt“ erzielt wurde, aber nur in 4% der Fälle, wenn diese Präferenz bereits im Passivsatz erfüllt war. Für die Verbalkomplexe kam es analog zu einer Änderung von VAuxMod zu AuxVMod in 86% der Fälle, wenn der Fokus auf dem Objekt lag, aber nur in 49% der Fälle, wenn der Fokus auf dem Verb lag. Der offensichtliche Unterschied zwischen den beiden Experimenten ist, dass im Experiment von McDonald/Bock/Kelly (ebd.) generell die Wechselrate wesentlich geringer war als in unserem Experiment. Aufgrund dieses Unterschiedes können die Ergebnisse von McDonald/Bock/Kelly (ebd.) in dem Slogan „Ändere eine markierte Ausgangsstruktur nur dann, wenn dies zu einer präferierteren Struktur führt!“ zusammengefasst werden, während unsere Ergebnisse zusammengefasst werden können zu „Behalte eine markierte Ausgangsstruktur nur dann bei, wenn eine Änderung zu einer weniger präferierten Struktur führt“.

Wir haben es hier vermutlich nicht mit einem prinzipiellen Unterschied zu tun. Beispielsweise haben Tanaka et al. (2011) in Anlehnung an McDonald/Bock/Kelly (1993) untersucht, welchen Effekt die Belebtheit auf die Anordnung der Argumente im Japanischen hat. Im Gegensatz zum Englischen, aber in Übereinstimmung mit dem Deutschen, können die Argumente im Japanischen

ohne Änderung des Genus Verbi umgeordnet werden. OS-Sätze wurden in 55% der Fälle als SO-Sätze wiedergegeben, wenn dadurch ein Satz mit der Abfolge „belebt vor unbelebt“ entstand. Aber auch wenn das Gegenteil erreicht wurde, wurden OS-Sätze noch mit einem Prozentsatz von etwas mehr als 30% als SO-Sätze wiedergegeben. Was genau dafür verantwortlich ist, wie häufig markierte Strukturen spontan als unmarkierte Strukturen wiedergegeben werden, ist unbekannt. Dass dies im Falle von Verbalkomplexen besonders häufig vorkommt, ist wohl kein Zufall. Erstens kommt die Abfolge VAuxMod sehr selten vor im Vergleich zur Abfolge AuxVMod. Zweitens ist die damit verbundene Änderung besonders gering. Sie betrifft ausschließlich die oberflächenstrukturelle Realisierung. Wie das Ausbleiben eines Akzeptabilitätsunterschieds in Abhängigkeit von der Auxiliarposition in dem oben beschriebenen Experiment zum Zusammenhang von Fokus und Abfolge im Verbalkomplex zeigt, hat diese Änderung keinen Effekt auf die semantisch/pragmatische Interpretation eines Satzes. Drittens schließlich führt die Abfolge VAuxMod zu einem Zusammenstoß zweier lexikalisch betonter Wörter (ein sog. ‘stress clash’), wenn dem Verbalkomplex direkt ein Nomen vorausgeht, was unter rhythmischen Gesichtspunkten ungünstig ist. Es ist mehrfach in der Literatur festgestellt worden, dass unter diesen Umständen präferiert das Auxiliar zwischen Nomen und lexikalisches Verb gestellt wird, so dass der Betonungszusammenstoß vermieden wird (siehe Sapp 2011 für Diskussion und weitere Referenzen).

Abschließend können wir festhalten, dass die in diesem Abschnitt diskutierten Daten zeigen, dass die Informationsstruktur eines Satzes – wie in der einschlägigen Literatur vorgeschlagen – einen Einfluss auf die Serialisierung der Verben in einem Verbalkomplex hat. Dieser Einfluss betrifft nicht die Akzeptabilität der zur Diskussion stehenden Verbalkomplexe, sondern ihre Auftretenshäufigkeit. Dieser Befund widerspricht den theoretischen Annahmen von Schmid/Vogel (2004) und Wurmbrand (2004a), da diese davon ausgehen, dass die Abfolge VModAux in der Grammatik nur dann lizenziert ist, wenn das lexikalische Verb oder das Modalverb eng fokussiert sind. Unsere Daten weisen aber darauf hin, dass wir es hier nicht mit einem grammatischen Effekt zu tun haben, sondern mit einem Effekt, der die Auswahl aus zwei grammatischen Alternativen bei der Sprachproduktion betrifft, also einem Effekt, den man traditionellerweise als Performanzeffekt bezeichnen würde. Diese Schlussfolgerung wird unterstützt durch die oben zitierten Daten aus der Sprachgeschichte und der Dialektsyntax. In den Korpusdaten von Ebert (1981), Sapp (2011) und Dubenion-Smith (2011) kommen die beiden Abfolgen AuxV und VAux sowohl unter engem Objektfokus als auch unter engem Verbfokus vor, lediglich die prozentuale Verteilung unterscheidet sich in Abhängigkeit vom Fokus.

Wie der Einfluss des Fokus auf die Abfolge im Verbalkomplex im Rahmen eines Sprachproduktionsmodells zu modellieren ist, ist momentan eine offene

Frage. Beispielsweise ist noch zu klären, ob der hier beschriebene Effekt ein rein prosodischer Effekt ist, der auch auftritt, wenn sich Sätze nicht hinsichtlich ihrer Fokusstruktur unterscheiden. Dass dem so ist, darauf deuten Daten hin, die die Wortart des Wortes berücksichtigen, das dem Verbalkomplex unmittelbar vorangeht. Insbesondere wurde in der oben zitierten Literatur gefunden, dass ein Pronomen im Gegensatz zu einem Nomen die Voranstellung des Auxiliars favorisiert, wie in (25) illustriert (siehe auch Swerts/van Wijk 2005 fürs Niederländische).

- (25) a. ... dass Peter seinen ONKEL hat₁ besuchen₃ müssen₂ | besuchen₃ hat₁ müssen₂.
 b. ... dass Peter ihn hat₁ BESUCHEN₃ müssen₂ | BESUCHEN₃ hat₁ müssen₂.

Da die Datenlage auch hier nicht eindeutig ist, bietet sich eine Untersuchung mittels experimenteller Methoden an.

6. Resümee

Beginnend mit den einflussreichen Veröffentlichungen von Bard/Robertson/Sorace (1996) und Cowart (1997) werden zunehmend experimentelle Methoden eingesetzt, um syntaktische Fragestellungen zu untersuchen. Diese Methoden sind nicht abhängig von speziellen syntaktischen Theorien und sie sind potenziell sowohl für die Grammatiktheorie und Grammatikographie von Interesse. In unserem Beitrag haben wir das Zusammenspiel von experimentellen Daten und syntaktischer Analyse anhand einer Fallstudie illustriert, die sich mit der Abfolge der Verben in Verbalkomplexen mit einem Modalverb im Perfekt beschäftigt hat. Abschließend wollen wir auf die eingangs aufgeworfene Frage zurückkommen, ob und in welcher Weise Grammatiktheorie und Grammatikographie vom Einsatz experimenteller Methoden profitieren können.

Im Falle des hier zur Diskussion stehenden Teilbereichs der deutschen Syntax war das Hauptergebnis unserer Experimente, dass die Abfolge mit partieller Voranstellung des Perfektauxiliars von der von uns untersuchten Versuchspersonengruppe als relativ akzeptabel beurteilt wurde. Für Verbalkomplexe mit drei Verben bedeutet dies beispielsweise, dass Verbalkomplexe vom Typ „lesen hat wollen“ zwar weniger akzeptabel sind als Verbalkomplexe mit der standardsprachlichen Abfolge „hat lesen wollen“, aber immer noch so akzeptabel, dass sie als prinzipiell grammatisch einzustufen sind. Aus Studien zur historischen Syntax und zur Dialektsyntax war bereits bekannt, dass Verbalkomplexe in bestimmten Sprachstufen des Deutschen und in bestimmten Dialekten präferiert mit partieller Auxiliarvoranstellung realisiert wurden

bzw. werden. In der theoretischen Syntax dagegen war der Status der partiellen Auxiliarvoranstellung umstritten. Während Haider (2003) Verbalkomplexe vom Typ „lesen hat wollen“ als grammatisch bewertet, werden sie von Sternefeld (2006, S. 645) als ungrammatisch zurückgewiesen. Diese Diskrepanz führt Sternefeld darauf zurück, dass Autoren, für die Verbalkomplexe vom Typ „lesen hat wollen“ grammatisch sind, nicht die Verhältnisse im Standarddeutschen wiedergeben, sondern die Verhältnisse in den süddeutschen Dialekten, die diese Abfolge erlauben oder sogar präferieren. Nun ist aber Haider (2003) – ebenso wenig wie Sternefeld (2006) – als ein Beitrag zur Dialektsyntax intendiert. Natürlich ist nicht ausgeschlossen, dass dies unbeabsichtigt dennoch der Fall ist. Die von uns gefundenen Daten weisen aber einen anderen Ausweg aus diesem scheinbaren Widerspruch.

Die in diesem Aufsatz diskutierten Akzeptabilitätsdaten beruhen alle auf Experimenten, die an der Universität Konstanz durchgeführt wurden. Mit seiner Lage in Südwestdeutschland gehört Konstanz nicht zu den Spracharealen des Deutschen, in denen Verbalkomplexe mit partieller Auxiliarinversion besonders stark im Sprachgebrauch vertreten sind (Südostdeutschland sowie Österreich; siehe Niehaus 2014). Des Weiteren ergab eine Analyse des regionalen Hintergrunds der Versuchspersonen durch Bader/Schmid (2009b), dass das grundlegende Akzeptabilitätsmuster unabhängig vom regionalen Hintergrund zu finden ist. Bader/Schmid (2009b) haben daraus den Schluss gezogen, dass die relativ hohe Akzeptanz von Verbalkomplexen mit partieller Auxiliarinversion eine Eigenschaft des umgangssprachlichen Deutsch ist. Neuere Arbeiten kommen zu demselben Schluss (Krasselt 2013; Niehaus 2014).

Der scheinbare Widerspruch zwischen Haider (2003) und Sternefeld (2006, S. 645) lässt sich also dahingehend auflösen, dass Haider (2003) Umgangdeutsch analysiert, Sternefeld (2006, S. 645) dagegen Standarddeutsch. Einer der Vorteile der oben vorgestellten syntaktischen Analyse von Bader/Schmid (2009b) ist es, dass der Zusammenhang zwischen Standarddeutsch und Umgangdeutsch auf einfache Weise erfasst werden kann. Einige der Probleme, die dieser Ansatz aufwirft, werden ausführlich in Salzmann (2013) diskutiert.

Die in diesem Beitrag diskutierten Methoden können dazu dienen, grammatische Phänomene besser zu verstehen. Die damit erzielten Ergebnisse sind für Grammatiktheorie und Grammatikographie gleichermaßen relevant. Dass wir uns auf die Implikationen für die Grammatiktheorie konzentriert haben, ist ausschließlich dem Umstand geschuldet, dass die von uns referierten Arbeiten entsprechend ausgerichtet waren. Die wichtigste Implikation für die Grammatikographie betrifft den Status der Abfolge mit partieller Auxiliarinversion (AuxVMOD) als weitverbreitete umgangssprachliche Abfolgeoption (siehe auch Krasselt 2013; Niehaus 2014). Dass in bestimmten Fällen auch fein-

körnigere Aussagen möglich sind, zeigt die Studie von Bader (2014) zur Verbalkomplexbildung mit *lassen* statt mit Modalverb. Bezüglich *lassen* machen verschiedene deskriptiv orientierte Grammatiken einander widersprechende Aussagen, die anhand von experimentell gewonnenen Daten überprüft werden können.

Abschließend können wir festhalten, dass die experimentelle Untersuchung von Verbalkomplexen gezeigt hat, dass es für das Verständnis der Syntax von Verbalkomplexen nicht ausreicht, zwischen Standarddeutsch auf der einen und Dialekten auf der anderen Seite zu unterscheiden. Dazwischen ist zumindest noch eine Ebene des Umgangsdeutsch anzunehmen. Umgangssprachlich weisen Verbalkomplexe mehr Variation auf als standardsprachlich. Ein wichtiges Ergebnis der diskutierten Experimente ist es, gezeigt zu haben, dass die beobachtbare Variation keinesfalls regellos ist, sondern sich genau eingrenzen lässt: Während im Standarddeutschen das Perfektauxiliar an der ersten Stelle stehen muss, sind im Umgangsdeutschen alle Positionen erlaubt, die dem Modalverb vorangehen. Dies gilt unabhängig von der Anzahl der Verben, die im Verbalkomplex vorkommen.

Literatur

- Abels, Klaus (2016): The fundamental left-right asymmetry in the Germanic verb cluster. In: *The Journal of Comparative Germanic Linguistics* 19, 3, S. 179-220.
- Adger, David (2003): *Core syntax: A minimalist approach*. Oxford.
- Arppe, Antti/Järvikivi, Juhani (2007): Every method counts: Combining corpus-based and experimental evidence in the study of synonymy. In: *Corpus Linguistics and Linguistic Theory* 3, 2, S. 131-159.
- Bader, Markus (2014): Das Verb *lassen* im Deutschen: Zum Zusammenspiel von Auxiliarinversion und Ersatzinfinitiv. In: *Zeitschrift für Sprachwissenschaft* 33, 1, S. 1-44.
- Bader, Markus/Häussler, Jana (2010): Toward a model of grammaticality judgments. In: *Journal of Linguistics* 46, 2, S. 273-330.
- Bader, Markus/Schmid, Tanja (2006): An OT-analysis of *do*-support in modern German. <http://roa.rutgers.edu/files/837-0606/837-BADER-0-0.PDF> (Stand: 4.1.2018).
- Bader, Markus/Schmid, Tanja (2009a): CAT meets GO: Auxiliary inversion in German verb clusters. In: van Cranenbroeck, Jeroen (Hg.): *Alternatives to cartography*. Berlin/Boston, S. 203-244.
- Bader, Markus/Schmid, Tanja (2009b): Verb clusters in colloquial German. In: *The Journal of Comparative Germanic Linguistics* 12, 3, S. 175-228.
- Bader, Markus/Schmid, Tanja/Häussler, Jana (2009): Optionality in verb-cluster formation. In: Winkler, Susanne/Featherston, Sam (Hg.): *The fruits of empirical linguistics*. Bd. 2: Product. (= *Studies in Generative Grammar* 102). Berlin, S. 37-58.

- Barbiers, Sjef (2005): Word order variation in three-verb clusters and the division of labour between generative linguistics and sociolinguistics. In: Cornips, Leonie/Corrigan, Karen P. (Hg.): *Syntax and variation: Reconciling the biological and the social.* (= Amsterdam Studies in the Theory and History of Linguistic Science 265). Amsterdam/Philadelphia, S. 233-264.
- Bard, Ellen G./Robertson, Dan/Sorace, Antonella (1996): Magnitude estimation of linguistic acceptability. In: *Language* 72, 1, S. 32-68.
- Baroni, Marco et al. (2009): The WaCky Wide Web: A collection of very large linguistically processed web-crawled corpora. In: *Language Resources and Evaluation* 43, 3, S. 209-226.
- Bech, Gunnar (1983): *Studien über das deutsche verbum infinitum.* 2., unveränd. Aufl. (= Linguistische Arbeiten 139). Tübingen.
- Cowart, Wayne (1997): *Experimental syntax: Applying objective methods to sentence judgments.* Thousand Oaks, CA u.a.
- Culicover, Peter W./Jackendoff, Ray (2010): Quantitative methods alone are not enough: Response to Gibson and Fedorenko. In: *Trends in Cognitive Science* 14, 6, S. 234-235.
- Dubenion-Smith, Shannon A. (2011): Der Verbalkomplex im Schlesischen. In: *Zeitschrift für Dialektologie und Linguistik* 78, 3, S. 281-320.
- Ebert, Robert P. (1981): Social and stylistic variation in the order of auxiliary and non-finite verb in dependent clauses in Early New High German. In: *Beiträge zur Geschichte der deutschen Sprache und Literatur* 103, S. 204-237.
- Edelman, Shimon/Christiansen, Morten H. (2003): How seriously should we take minimalist syntax? In: *Trends in Cognitive Science* 7, 2, S. 60-61.
- Featherston, Sam (2005): The decathlon model of empirical syntax. In: Kepsers, Stephan/Reis, Marga (Hg.): *Linguistic evidence. Empirical, theoretical and computational perspectives.* (= Studies in Generative Grammar 85). Berlin, S. 187-208.
- Ferreira, Fernanda (2005): Psycholinguistics, formal grammars, and cognitive science. In: *Linguistic Review* 22, 2, S. 365-380.
- Fletcher, Charles R. (1994): Levels of representation in memory for discourse. In: Gernsbacher, Morton A. (Hg.): *Handbook of psycholinguistics.* San Diego, S. 589-607.
- Fodor, Janet D. (2002): Prosodic disambiguation in silent reading. In: *Proceedings of the North East Linguistic Society* 32, 1, S. 113-132.
- Fukuda, Shin et al. (2012): Is magnitude estimation worth the trouble. In: Choi, Jaehoon et al. (Hg.): *Proceedings of the 29th West Coast Conference on Formal Linguistics.* Somerville, MA, S. 328-336.
- Gibson, Edward/Fedorenko, Evelina (2010): Weak quantitative standards in linguistics research. In: *Trends in Cognitive Science* 14, 6, S. 233-234.
- Haider, Hubert (2003): V-clustering and clause union: Causes and effects. In: Seuren/Kempen (Hg.), S. 91-126.

- Haspelmath, Martin (2006): Against markedness (and what to replace it with). In: *Journal of Linguistics* 42, 1, S. 25-70.
- Hendriks, Lotte (2014): Variation in verb cluster interruption. In: *Linguistics in the Netherlands* 31, 1, S. 53-65.
- Keller, Frank (2000): *Gradience in grammar: Experimental and computational aspects of degrees of grammaticality*. Edinburgh.
- Kempen, Gerard/Harbusch, Karin (2008): Comparing linguistic judgments and corpus frequencies as windows on grammatical competence: A study of argument linearization in German clauses. In: Steube, Anita (Hg.): *The discourse potential of underspecified structures*. Berlin/Boston, S. 179-192.
- Krasselt, Julia (2013): Zur Serialisierung im Verbalkomplex subordinierter Sätze. Gegenwartssprachliche und frühneuhochdeutsche Variation. In: *Jahrbuch für germanistische Sprachgeschichte* 4, S. 128-143.
- Louden, Mark L. (1990): Verb raising and the position of the finite verb in Pennsylvania German. In: *Linguistic Inquiry* 21, 3, S. 470-477.
- McDonald, Janet L./Bock, Kathryn/Kelly, Michael H. (1993): Word and world order: Semantic, phonological, and metrical determinants of serial position. In: *Cognitive Psychology* 25, 2, S. 188-230.
- Niehaus, Konstantin (2014): Kontinuität im Neuhochdeutschen 'von unten' und 'von oben'. Ein variationslinguistisches Nutzungsszenario. In: *Jahrbuch für germanistische Sprachgeschichte* 5, 1, S. 299-313.
- Pater, Joe (2009): Weighted constraints in generative linguistics. In: *Cognitive Science* 33, 6, S. 999-1035.
- Phillips, Colin (2009): Should we impeach armchair linguists? In: Iwasaki, Shoichi (Hg.): *Japanese/Korean Linguistics*. Bd. 17. Stanford, CA, S. 49-72.
- Phillips, Colin/Lasnik, Howard (2003): Linguistics and empirical evidence: Reply to Edelman and Christiansen. In: *Trends in Cognitive Sciences* 7, 2, S. 61-62.
- Potter, Mary C./Lombardi, Linda (1990): Regeneration in the short-term recall of sentences. In: *Journal of Memory and Language* 29, 6, S. 633-654.
- Salzmann, Martin (2013): New arguments for verb cluster formation at PF and a right-branching VP: Evidence from verb doubling and cluster penetrability. In: *Linguistic Variation* 13, 1, S. 81-132.
- Salzmann, Martin (2016): Displaced morphology in German: Evidence for post-syntactic morphology. In: Barnickel, Katja et al. (Hg.): *Replicative processes in grammar*. (= *Linguistische Arbeitsberichte* 93). Leipzig, S. 401-446.
- Sapp, Christopher D. (2011): *The verbal complex in subordinate clauses from Medieval to Modern German*. (= *Linguistik aktuell* 173). Amsterdam.
- Schmid, Tanja/Vogel, Ralf (2004): Dialectal variation in German 3-verb clusters. A surface-oriented optimality theoretic account. In: *Journal of Comparative Germanic Linguistics* 7, 3, S. 235-274.
- Seuren, Pieter A. M./Kempen, Gerard (Hg.) (2003): *Verb constructions in German and Dutch*. (= *Current Issues in Linguistic Theory* 242). Amsterdam.

- Sprouse, Jon/Almeida, Diogo (2012): Assessing the reliability of textbook data in syntax: Adger's Core Syntax. In: *Journal of Linguistics* 48, 3, S. 609-652.
- Sprouse, Jon/Schütze, Carson T./Almeida, Diogo (2013): A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001-2010. In: *Lingua* 134, S. 219-248.
- Sternefeld, Wolfgang (2006): *Syntax. Eine morphologisch motivierte generative Beschreibung des Deutschen. Bd. 2.* (= Stauffenburg Linguistik 31). Tübingen.
- Swerts, Marc/Wijk, Carel van (2005): Prosodic, lexico-syntactic and regional influences on word order in Dutch verbal endgroups. In: *Journal of Phonetics* 33, 2, S. 243-262.
- Tanaka, Mikihiro N. et al. (2011): Conceptual influences on word order and voice in sentence production: Evidence from Japanese. In: *Journal of Memory and Language* 65, 3, S. 318-330.
- Vogel, Ralf et al. (2015): Function words in rhythmic optimisation. In: Vogel, Ralf/Vijver, Ruben van de (Hg.): *Rhythm in cognition and grammar: A Germanic perspective.* (= Trends in Linguistics. Studies and Monographs 286). Berlin, S. 255-274.
- Wasow, Thomas/Arnold, Jennifer (2005): Intuitions in linguistic argumentation. In: *Lingua* 115, 11, S. 1481-1496.
- Weiß, Helmut (1998): *Syntax des Bairischen. Studien zur Grammatik einer natürlichen Sprache.* (= Linguistische Arbeiten 391). Tübingen.
- Williams, Edwin (2003): *Representation Theory.* Cambridge, MA.
- Wurmbrand, Susi (2004a): Syntactic vs. post-syntactic movement. In: Burelle, Sophie/Somesfalean, Stanca (Hg.): *Proceedings of the 2003 Annual Conference of the Canadian Linguistic Association (CLA).* Montreal, S. 284-295.
- Wurmbrand, Susi (2004b): West Germanic verb clusters: The empirical domain. In: Kiss, Katalin É./Riemsdijk, Henk van (Hg.): *Verb clusters: A study of Hungarian, German and Dutch.* (= Linguistik aktuell 69). Amsterdam, S. 43-86.
- Wurmbrand, Susi (2017): Verb clusters, verb raising, and restructuring. In Everaert, Martin/Riemsdijk, Henk van (Hg.): *The Blackwell companion to syntax.* 2. Aufl. Bd. 5. Oxford, S. 1-109.