

Gdańsk 2023, Nr. 49

Elisabeth Scherr

(Karl-Franzens-Universität Graz)

ORCID: 0000-0002-4709-6068

Kotext-induzierte Variation aus quantitativer Sicht am Beispiel epistemischer Modalverben

<https://doi.org/10.26881/sgg.2023.49.01>

[I]t is about being the best among a choice of options,
not about being objectively perfect.
(McCarthy 2002:1)

Der Beitrag diskutiert die Frage der Disambiguierung (nicht-)epistemischer Interpretationsvarianten von finiten Modalverben anhand quantitativ-korpuslinguistischer Verfahren. Basierend auf einer allgemeinen Definition (epistemischer) Modalität, die auf deiktischen Relationen beruht, werden Eigenschaften festgelegt, die für die Interpretation der Verben relevant sind und an der sprachlichen Oberfläche semi-automatisiert analysiert werden. Anhand der Verben *dürfen* und *müssen* wird eine Methodik illustriert, die das Potenzial hat, dominierende Lesarten ebenso wie Zweifelsfälle erklären zu können.

Schlüsselwörter: Modalverben, epistemische Modalität, Deixis, quantitative Methoden, Korpuslinguistik

Cotext-Induced Variation from a Quantitative Perspective: The Example of Epistemic Modal Verbs – This article discusses the question of disambiguating (non)epistemic variants from a quantitative-corpus-linguistic perspective. After defining (epistemic) modality based on deictic relations, characteristics relevant for the interpretation of the verbs are discussed. Those features are annotated and analysed in the course of a semi-automatized corpus study. Using the example of *dürfen* and *müssen* a method is illustrated that has the potential of explaining dominating readings as well as cases of doubt.

Keywords: modal verbs, epistemic modality, deixis, quantitative methods, corpus linguistics

1. Hinführung¹

Es ist mittlerweile ein Gemeinplatz in der Modalitätsforschung, dass es sich bei Modalverben (auch, aber nicht nur im Deutschen) um formal opake Elemente handelt,

¹ Der vorliegende Beitrag versteht sich im Groben als komprimierte Zusammenschau der Ergebnisse aus Scherr (2019), wobei ein stärkerer Fokus auf die praktische Umsetzung der hier nur kurz skizzierten theoretischen Fundierung und auf das Erklärungspotenzial der vorgeschlagenen Methodik gelegt wird.

um „polyfunktionale [...] Modalausdrücke“ (Reis 2001: 310), die sich auf den ersten Blick durch eine intransparente Form-Bedeutungs-/Funktions-Zuweisung auszeichnen. Als finite Verben kongruieren sie mit dem Subjekt und verlangen in den meisten Fällen einen Infinitiv ohne *zu* (1. Status), drücken dabei aber höchst unterschiedliche Bedeutungen aus, die von Verpflichtung, Erlaubnis, Zwang, Möglichkeit oder Wunsch/Absichten einerseits (im Folgenden nicht-epistemische Modalverben) bis hin zum Ausdruck von zweifelhafter Faktizitätseinschätzung, reportativer oder Hearsay-Semantik reichen (im Folgenden epistemische Modalverben). Für prototypische Verwendungen vgl. (1) und (2).

- (1) *Der Student darf/kann/möchte/muss/soll/will die Prüfung ablegen.*
- (2) *Der Student dürfte/kann/mag/muss/soll/will die Prüfung nicht bestanden haben.*

Tatsächlich kann bereits in Anbetracht dieser kontrastiven Beispiele die Vermutung angestellt werden, dass es sehr wohl formale Hinweise gibt, die die Interpretation der Verben steuern könnten. So zeigt sich etwa, dass das Verb *mögen/möchten* im Konjunktiv eher zum Ausdruck einer Absicht/eines Wunsches verwendet wird, während die indikativische Form häufig auch eine zweifelhafte Faktizitätseinschätzung ausdrücken kann. Anders scheint die Lage bei *dürfen* zu sein, wo nur die Form im Konjunktiv die Faktizität eines Sachverhalts anzweifelt, während das Verb im Indikativ eine Erlaubnis-Lesart evoziert. Abgesehen von den Eigenschaften der flektierten Modalverben selbst zeigt auch der Infinitiv bestimmte Auffälligkeiten, so etwa im temporaldeiktischen Verweis: Während Infinitive mit terminativer Change-of-State Semantik wie *ablegen* offenbar die Bedeutung Erlaubnis, Möglichkeit, Zwang etc. stützen, führt ein Ersetzen des einfachen Infinitivs in (1) durch einen Infinitiv Perfekt zu einer Bedeutungsverschiebung, in Kombination mit *darf* sogar zu einem Beispiel mit fragwürdiger Grammatikalität (*?Der Student darf die Prüfung abgelegt haben.*). Dass die Interpretation der Verben unter anderem von solchen kotextuellen Faktoren abhängt, ist der Forschung auch nicht unbemerkt geblieben und wurde mehrfach ausführlich behandelt. So nennen bereits Brünner/Redder (1983: 52) die Eigenschaften des Subjekts, des Infinitivs sowie die syntaktische Form des Modalverbs selbst als Faktoren, die auf die Lesart wirken (vgl. auch Heine 1995: 3536). Ebenso sind es der Infinitiv Perfekt, die semantischen Merkmale des Infinitivs oder die Progressivform, die in den meisten einschlägigen Untersuchungen als interpretationssteuernd erkannt werden (vgl. z. B. Hentschel/Weydt 2003: 78; Engel 2004: 246; Zifonun et al. 1997: 1270f.).

Bevor jedoch die Merkmale der konkreten Realisierungsformen an der sprachlichen Oberfläche und deren Zusammenspiel in den Fokus genommen werden, soll im Folgenden zunächst der Frage nachgegangen werden, *warum* bestimmte formale Präferenzen vorliegen bzw. welche tiefgehenden funktionalen Unterschiede zwischen nicht-epistemischen und epistemischen Modalverbverwendungen festgemacht werden können (siehe Abschnitt 2). Basierend auf den Arbeiten u. a. von Abraham (2009), Diewald (1999) und Kotin (2010; 2005) wird die Funktion von Modalität und Modalverben insgesamt skizziert. Auf dieser Basis werden danach die in der Forschungsliteratur thematisierten sowie aus den theoretischen Annahmen logisch ableitbaren formalen Eigenschaften epistemischer (und nicht-epistemischer) Modalverben erklärt. Diese Merkmale fließen schließlich in eine dreiteilige quantitativ-empirische Studie ein (siehe Abschnitt 3), die aus einer

Korpusanalyse, einer statistischen Auswertung (multivariate Korrelationsanalysen und Regression) und einer Darstellung mittels Methoden der stochastischen Optimalitätstheorie besteht. Während insbesondere jüngere Arbeiten vermehrt auf Korpora und zumindest relative Häufigkeiten zurückgreifen (vgl. Baumann 2017), bleibt die genaue statistische Evaluation und insbesondere die Beziehung der einzelnen Merkmale untereinander unbeachtet. Die vorliegende Studie ist ein Beitrag zur Schließung dieser Forschungslücke, wobei für eine stärkere Verknüpfung von theoretischen Vorüberlegungen einerseits und quantitativ-statistischer Verfahren andererseits plädiert wird. Dabei geht es zu keinem Zeitpunkt um die Festlegung auf eine ‚objektive‘ Interpretation der Verben, sondern vielmehr um das Aufzeigen interpretativer Tendenzen und um deren Erklärung durch die quantitativ messbare Einflussstärke der einzelnen Faktoren.

2. Eine Frage der Modalität

Im Folgenden soll davon ausgegangen werden, dass sich alle Modalverbverwendungen wie auch andere Realisierungsformen von Modalität wie Modus, Adverbien, Partikeln etc. auf eine gemeinsame, funktional bestimmte Domäne rückführen lassen. Grundlegend für eine Definition eines solchen „Genus proprium“ (Abraham 2009: 251) ist ein Deixis-Begriff, der in Anlehnung an Diewald (1999; 2013) entwickelt wurde. Deixis wird hier in einem weiten Sinn verstanden, als Verbindung eines „Ausgangspunktes über eine bestimmte Strecke mit einem Zielpunkt“ (Diewald 1991: 25). Im Unterschied zu Diewald wird dieser grundlegende deiktische Wert für beide Domänen, sowohl für epistemische als auch für nicht-epistemische, als zentrale Funktion angenommen. Der Ausgangspunkt ist dabei immer eine spezifische hic-et-nunc-Situation mit den jeweiligen personellen und raum-zeitlichen Gegebenheiten und Wissensbeständen eines Handlungsraums. Modalität zeichnet sich einerseits durch eine Situationsgebundenheit aus, die einen denotativ-deiktischen Wert hat. Dieser Wert zeigt sich bei Modalität darin, dass das Modalitätsmittel eine Form von Notwendigkeit, Möglichkeit, Wahrscheinlichkeit etc., also bestimmte Gegebenheiten *denotiert*, die in der konkreten Sprechhandlungssituation vorliegen. Gleichzeitig wird jedoch die durch den deiktischen Prozess lokalisierte Entität, in Falle der Modalverben die Prädikation des Infinitivs, aus der konkreten Situation entrückt. Damit entsteht die Interpretation, dass die Faktizität einer modalisierten Äußerung als faktisch offen gilt, und dies betrifft sowohl die nicht-epistemischen als auch die epistemischen Teilspektren modaler Semantik. Kotin (2005: 249) beschreibt im folgenden Zitat die verlagerte deiktische Relation zwischen den beiden Bereichen und gleichzeitig den Bedeutungszusammenhang sämtlicher Modalverben im letzten Satz:

So wie z. B. ein Zurückliegen oder ein Bevorstehen in der Zeit, ein zeitlich-retrospektives resp. ein zeitlich-prospektives Dort, die temporale origo-exklusive Deixis kennzeichnet, drückt die origo-exklusive Deixis [d. i. hier epistemische Modalität, E. S.] epistemischer Modalverben das modale *D-rt* aus, das in diesem Kategorialbereich als ‚vielleicht‘, ‚wahrscheinlich‘, ‚kaum‘ u. dgl. umgedeutet wird. Origo-inklusive [d. i. hier nicht-epistemische Modalität, E. S.] sind dabei temporale und modale Sachverhalte, welche die Origo in ihre unmittelbare Wirkungssphäre einbezieht. Das temporale *Hier* ist das *Jetzt* und das modale *Hier* ist eine faktisch formulierte Aussage.

Mit Fokus auf die Interpretationsvarianten von Modalverben bedeutet dies, dass im Falle nicht-epistemischer Modalität die Prädikation des Infinitivs in einer temporaldeiktisch prospektivischen Zeit (also: in Zukunft) realisiert wird (bzw. *muss, soll, darf* etc. sie realisiert *werden*, siehe Beispiel (1), unten wiederholt als Beispiel 3). Bei epistemischer Modalität hingegen *wurde* die Verbalhandlung des Infinitivs bereits realisiert, der Faktizitätsstatus ist jedoch aus Perspektive der aktualisierten Origo nicht zugänglich (siehe Beispiel 4). Bei letzterer Verwendung handelt es sich also um einen temporaldeiktischen Verweis in eine nicht zugängliche „Welt“ vor dem Hintergrund der Sprech- und/oder Referenzzeit (Gleichzeitigkeitsrelation). Die Abbildungen (1) und (2) illustrieren den deiktischen Verweis, der die Verortung der



Abb. 1: Deiktische Verortung der Prädikation des Infinitivs bei nicht-epistemischer Modalverbverwendung

Prädikation des Infinitivs darstellt.

- (3) *Der Student darf/kann/möchte/muss/soll/will die Prüfung [in Zukunft!] ablegen.* Faktizität wird in Zukunft entschieden.



Abb. 2: Deiktische Verortung der Prädikation des Infinitivs bei epistemischer Modalverbverwendung

- (4) *Der Student dürfte/kann/mag/muss/soll/will die Prüfung [*in Zukunft!] nicht bestanden haben.*

Faktizität ist entschieden, aktuell aber nicht zugänglich.

Es zeigt sich also, dass der Unterschied zwischen epistemischer und nicht-epistemischer Modalität in der Verweisrichtung der deiktischen Relation liegt (potentielle Welt vs. Prospektivität). Der entscheidende Punkt ist nun, dass sich durch diese Verlagerung am Beispiel der finiten Modalverben eine Reihe von Eigenschaften erklären lassen, die sich an der sprachlichen Oberfläche zeigen und die bereits teilweise in der Forschung als interpretationslenkend diskutiert wurden. Der Fokus liegt dabei ausdrücklich auf overt Merkmalen, die im Rahmen einer korpuslinguistischen Studie (semi-)automatisiert untersucht werden können. Eigenschaften wie etwa (Negations-)Skopus oder kontextuell-semantische Einflussfaktoren, die nicht oder nur indirekt (z. B. durch vergleichsweise subjektive Tests und Fallentscheidungen) empirisch untersucht werden können, bleiben dabei explizit ausgeschlossen.

a. Eigenschaften des Subjekts

Sowohl bei epistemischen als auch bei nicht-epistemischen finiten Modalverben liegt formale Kongruenz mit dem Subjekt vor, die semantische Kongruenz besteht bei den epistemischen Interpretationsvarianten ausschließlich zwischen Subjekt und Infinitiv (vgl. Engel 1992: 72).² Weitere Eigenschaften, die mit der oben angenommenen verlagerten deiktischen Relation zu erklären sind, können quantitativ erfasst werden. Dazu zählt die Markierung der Person, wobei die erste und zweite Person vom Typ *Ich muss* oder *Du musst* meist innerhalb konkreter Handlungszusammenhänge verankert sind. Sie werden deshalb grundlegend als Betroffene eines direktiven, temporaldeiktisch-prospektivischen Bezugs gelesen (vgl. Plungian 2010: 4245 und Beispiel 5).

- (5) „*Wir müssen täglich Schutt vom Hof wegräumen*“, *klagt die Zwickauerin*. („Freie Presse“, 4.07.2012)

Weitere Eigenschaften des Subjekts, die ebenfalls an der sprachlichen Oberfläche annotiert werden können, sind Belebtheit/Agentivität. Wenn die Verbalhandlung des Infinitivs prospektiv verortet (also nicht-epistemisch) werden soll, braucht es agentive Entitäten (‘elements of will’), die eine bestimmte Notwendigkeit, Möglichkeit etc. auch realisieren können (vgl. Diebold 2001: 34). Es wird daher vermutet, dass unbelebte Subjekte eher eine epistemische Interpretation favorisieren, da sie in der Regel³ weder Quelle noch Ziel von Notwendigkeit, Wunsch, Wille etc. sind. Eine ähnliche Tendenz wird für indefinite Subjekte (etwa *man*-Subjekte oder jene mit expletivem *es*) angenommen (vgl. Hundt 2003: 274 und Beispiel 6).

- (6) *Bei einer der Personen soll es sich um einen dunkelhäutigen Mann gehandelt haben*. („Fränkischer Tag“, 25.2.2012)

b. Eigenschaften des Infinitivs

Die Darstellung des deiktischen Verweises, der bei epistemischer Modalität vorliegt, hat gezeigt, dass die Prädikation des Infinitivs *gleichzeitig* zu einer aktualisierten Sprech- und/oder Referenzzeit verortet wird. Dies spiegelt sich auch in semantischen und formalen Eigenschaften der Infinitive, die bei solchen Modalverbinterpretationen vorliegen: Formal sind es der Infinitiv Perfekt (siehe Beispiel 7) sowie der Infinitiv mit Zustandspassiv (Beispiel 8) als Resultativkonstruktion⁴, die eine epistemische Interpretation stützen können; semantisch sind es Verben mit statischer oder progressiver Aktionsart (vgl. Beispiel 9).

- (7) *Angesichts des gerade angefangenen Septembers und erwarteter frühherbstlicher Temperaturen könnte sich das Marktplatz-Problem für dieses Jahr obnehin erledigt haben*. („Augsburger Allgemeine“, 9.1.2012)

² Bei epistemisch interpretierten Modalverben handelt es sich daher um keine *root modals*, also Kontrollverben, sondern um α -Regentia mit dem gesamten Restsatz als einzige Leerstelle (*Es muss so sein, dass er die Prüfung nicht bestanden hat*. vs. **Er muss es*.)

³ Auf allfällige metaphorische Verwendungsweisen oder Belebtheit unbelebter Referenten (Personifizierung) kann an dieser Stelle nicht weiter eingegangen werden.

⁴ Ganz anders ist das Vorgangs- oder *werden*-Passiv zu interpretieren, das eine Change-of-State-Semantik evoziert und damit einen prospektivischen Verweis nahelegt.

- (8) *Aufgrund der am Unfallort gefundenen Teile lässt sich sagen, dass es sich bei dem Verursacherfahrzeug aller Wahrscheinlichkeit nach um einen Roller der Marke Peugeot handelt, der silberfarben oder teilweise silberfarben ist. Das Zweirad **müsste** im Frontbereich beschädigt sein.* („Neue Westfälische“, 15.8.2012)
- (9) *Der Grund ist offenbar, dass er neben der deutschen auch die türkische Staatsangehörigkeit besitzen soll.* („Der Tagesspiegel“, 6.3.2013)

c. Modus

Durch die gebotene Kürze des vorliegenden Beitrags wird insbesondere die Rolle des Konjunktivs, genauer jene des Konjunktivs II, bei der Interpretationssteuerung thematisiert. Fabricius-Hansen (1997: 18) hat gezeigt, dass eine der Kernaufgaben des Konjunktivs II die Markierung von Irrealis und unklarem Faktizitätswert darstellt, weshalb diesem Modus eine hohe Relevanz für die Interpretationssteuerung zugeschrieben wird. Die empirische Analyse zeigt, dass eine Konjunktivmarkierung höchst unterschiedliche Effekte auslöst, die modalverbspezifisch diskutiert werden sollten.

d. VP-nahe Adverbiale

Adverbiale mit terminativer Semantik können die Change-of-State-Semantik und damit nicht-epistemische Interpretationen stärken oder sogar auslösen. Einzelne Beispiele zeigen, dass Eigenschaften wie etwa ein Infinitiv Perfekt durch ein temporaldeiktisch prospektivisch zu interpretierendes Adverbial ‚überschrieben‘ werden kann (siehe Beispiel 10). Eine epistemische Interpretation kann andererseits durch entsprechende Adverbiale gestützt werden, wie die Beispiele (11) und (12) zeigen.

- (10) *Wer noch ein passendes Weihnachtsgeschenk sucht, **kann** ab sofort Karten im Vorverkauf im Verkehrsbüro der Stadt Lingen erwerben.* („Osnabrücker Zeitung“, 21.11.2011)
- (11) *Die Regierung **soll** angeblich über den Ankauf einer Luxemburg-Lichtensteiner Hundesteuer-Datei nachdenken.* („Lausitzer Rundschau“, 23.6.2012)
- (12) *Er **will** der Firma nach eigenen Angaben auch nur 694.000 Euro bezahlt haben.* („Kronen Zeitung“, 30.4.2011)

e. Komplexe Sätze

Angenommen wird, dass einzelne Konnektoren mit finaler Semantik (*auf dass, damit, dass* oder *um...zu*) einen prospektivischen Verweis stützen können. Anders gestaltet sich die Situation bei Konnektoren mit kausaler bzw. begründender Semantik (*von daher, da, nachdem*

oder *denn*), da sie inferenzielle Bezüge, wenn auch nicht automatisch auslösen, so doch stützen können (vgl. Blühdorn 2008: 219 und Beispiel 13).

- (13) *Von daher dürfte am Samstag für den Tabellensechsten schon ein knappes Ergebnis als ein Erfolg gewertet werden.* („Hessische/Niedersächsische Allgemeine“, 15.12.2011)

3. Daten und Methoden

Die Datenbasis für die empirische Analyse bildet ein Großkorpus (rund 600 Millionen Tokens), das dem Forschungsprojekt *Varietengrammatik des Standarddeutschen* (Austrian Science Fund (FWF): I716-G18, vgl. www.variantengrammatik.net) entstammt. Es besteht aus informationsorientierten Presstexten aus sämtlichen Voll- und Halbzentren des deutschen Sprachgebiets. Für jedes Modalverb wurden Subkorpora mit je 300 epistemischen und 300 nicht-epistemischen Interpretationsvarianten gebildet.⁵ In einem nächsten Schritt fand die Annotation des Datenmaterials anhand der zuvor skizzierten Eigenschaften statt. Tabelle 1 gibt einen Überblick über die relevanten Tags.

Tabelle 1: Annotations-Tags

Annotations-Tag	Erläuterung
KONJII	Die Flexionsform des Modalverbs steht im Modus Konjunktiv II.
*SUBJDEF	Das Subjekt ist indefinit.
*SUBJANI	Das Subjekt ist nicht animiert.
*SUBJPERS1/2	Das Subjekt steht nicht in der 1. oder 2. Person (Singular/Plural).
INFPROG	Beim Infinitiv liegt eine Progressiv-Konstruktion vor (<i>sein + am/beim/im</i>).
INFZUPASS	Der Infinitiv drückt ein Zustandspassiv aus.
INFKOP	Der Infinitiv wird von einer Kopulakonstruktion gebildet.
INFKONSTREPI	Es liegen syntaktische Eigenschaften des Infinitivs vor, die auf eine epistemische Interpretation hindeuten (Infinitiv Perfekt, Kopula, Zustandspassiv).
INFZUSTAND	Die Aktionsart des Infinitivs kann als statisch interpretiert werden.
*INFVORPASS	Der Infinitiv zeigt kein <i>werden</i> zur Realisierung eines Vorgangspassivs.
ADVETPI/AKT	Der betreffende Satz zeigt eine Adverbiale, die eine epistemische/nicht-epistemische Interpretation nahelegt.
KONJETPI	Es liegt ein epistemisch zu interpretierender Kon- oder Subjunktiv vor.

Die Annotation dieser Eigenschaften bildet die Ausgangsbasis für die weitere quantitative Untersuchung. Zunächst wird die Signifikanz der annotierten Eigenschaften im Datenmaterial (jeweils modalverbspezifisch, jedoch alle Features bei allen Verben⁶) durch die Anwendung des Chi-Quadrat-Tests überprüft; es werden ausschließlich jene Merkmale weiter

⁵ Die Validität dieser Auswahl wurde mittels der Methode des Annotator Agreements (vgl. Sudarikov/Bojar 2015: 91) geprüft. Die Berechnung des Kennwerts (Fleiss' Kappa) ergibt $\kappa = 0.787536551$. Da Werte ab 0.75 als gut bis ausgezeichnet betrachtet werden (vgl. Greve/Wentura 1997: 111), ist die Validität der Auswahl der Analysebeispiele hinreichend abgesichert.

⁶ Dies bedeutet, dass alle Eigenschaften mechanisch bei allen Verben gleich annotiert wurden, ohne auf einzelne Besonderheiten einzugehen, die bereits in der Forschungsliteratur als abgesichert bekannt sind (zum

berücksichtigt, die in signifikanter Häufigkeit im Korpus vorkommen.⁷ Danach zeigen multivariate Korrelationsanalysen die Stärke des Zusammenhangs zwischen der jeweiligen Interpretationspräferenz und den einzelnen Eigenschaften. Zur Darstellung dieser Ergebnisse und um insbesondere das Zusammenspiel der einzelnen Merkmale zu akzentuieren, werden Aspekte der stochastischen Optimalitätstheorie (vgl. Bresnan/Deo 2007) verwendet, die Eigenschaften als *Constraints* interpretiert, und im Rahmen von *Rankings* zueinander in Beziehung gebracht. Dieser Zugang ermöglicht nicht nur eine Beurteilung der unterschiedlichen Relevanz von binären Parameter für die Erklärung von Bedeutungsvarianten; es sind insbesondere auch die Darstellungsweisen der Optimalitätstheorie, die den Vorteil bieten, unterschiedlich gewichtete Einflussfaktoren sowie deren Zusammenspiel zu illustrieren. Welche Vorhersagekraft für die Interpretation des Modalverbs sich aus den definierten Faktoren ergibt, zeigen schließlich Regressionsanalysen separat für jedes Verb.

Aus Platzgründen werden die einzelnen Schritte anhand von lediglich zwei Modalverben, *dürfen* und *müssen*, gezeigt. Die Auswahl dieser Verben motiviert sich dadurch, dass einerseits deren stark divergierende Ergebnisse aus der empirischen Analyse die Modalverbspezifität der definierten kotextuellen Eigenschaften unterstreicht; andererseits bleibt die grundlegend funktionale Definition epistemischer Modalität, wie in Abschnitt 2 dargelegt, von diesen unterschiedlichen formalen Realisierungen unangetastet. Zudem werden die klar interpretierbaren Merkmale für *dürfen* von den weniger deutlichen Eigenschaften von *müssen* relativiert, um nicht ein trügerisches Bild von vermeintlich klaren Interpretationszuweisungen durch die präsentierte Methodik entstehen zu lassen.

4. Ergebnisse für *dürfen* und *müssen*

Nach der Durchführung des Chi-Quadrat-Tests und dem damit geprüften signifikanten Vorkommen der Merkmale im Korpus werden letztere zunächst im Zuge von multivariaten Korrelationsanalysen beurteilt. Damit kann geprüft werden, ob die Eigenschaften mit den Bewertungen (epistemisch/nicht-epistemisch) signifikant korrelieren. Dabei wurde für jede Eigenschaft (für jedes Modalverb separat) der Korrelationskoeffizient r ermittelt (vgl. Albert/Marx 2010: 121). Das Ergebnis liegt zwischen -1 und +1 und illustriert positive und negative Zusammenhänge, wobei $r = 1$ eine vollständige positive Korrelation zwischen einem Merkmal x und der jeweiligen Modalverbinterpretation bedeuten würde. In einer Gesamtauswertung wird darüber hinaus die Korrelation der definierten Eigenschaften untereinander dargestellt. Die Abbildungen 3 und 4 zeigen die Korrelationskoeffizienten, wobei die Korrelationen zwischen den Merkmalen und der Bewertung als epistemisches Modalverb (EMV) in Fettdruck hervorgehoben sind.⁸

Beispiel die Rolle des Konjunktivs bei *dürfte*). Eine solche Relevanz bestätigt sich gegebenenfalls ohnehin im Zuge der Analyse.

⁷ Aus Platzgründen wird dieser Analyseschritt im Kap. 4 nicht mehr gesondert dargestellt.

⁸ Die Farbgebung der einzelnen Zellen richtet sich nach der Interpretation der Korrelationswerte nach Albert/Marx (2010: 129):

0.9–1: sehr hohe Korrelation, sehr starke Beziehung

	EMV	ModKonjII	*SubjDef	*SubjAni	*SubjPers1/2	INFZustand	InfKonstrEpi
EMV	1						
ModKonjII	0,98019606	1					
*SubjDef	-0,1029705	-0,0901568	1				
*SubjAni	0,54671527	0,54709139	-0,0006337	1			
*SubjPers1/2	0,21282896	0,21712897	0,11631863	0,193637	1		
INFZustand	0,42734305	0,42345716	-0,0685243	0,23708992	0,13590313	1	
InfKonstrEpi	0,39857814	0,39030951	-0,0943775	0,15094167	0,08882092	0,4693965	1

Abb. 3: Korrelationskoeffizienten für *dürfen*

	EMV	ModKonjII	*SubjDef	*SubjAni	*SubjPers1/2	INFZustand	InfKonstrEpi
EMV	1						
ModKonjII	0,13060753	1					
*SubjDef	-0,0594686	-0,0332795	1				
*SubjAni	0,31387589	0,11536661	-0,0009091	1			
*SubjPers1/2	-0,1882361	-0,0841399	-0,1706513	-0,1840588	1		
INFZustand	0,81129524	0,04014268	-0,0608418	0,27838401	-0,1537753	1	
InfKonstrEpi	0,61917262	-0,0442447	-0,0646239	0,24702302	-0,1232408	0,65687558	1

Abb. 4: Korrelationskoeffizienten für *müssen*

Bei $r \geq 0.4$ wird generell von einer substantiellen Beziehung ausgegangen (vgl. Albert/Marx 2010: 129). Nun zeigt sich bei *dürfen* eindeutig, dass die Modusmarkierung mit $r = 0.98$ die höchste positive Korrelation mit epistemischen Interpretationsvarianten aufweist. Dies bedeutet mit anderen Worten, dass, wenn dieses Merkmal vorliegt, mit 98%iger Wahrscheinlichkeit davon ausgegangen werden kann, dass das Verb epistemisch gelesen wird. Im Rahmen der stochastischen Optimalitätstheorie liegt damit der Effekt der Neutralisation vor: Liegt ein solcher *Constraint* vor, werden andere Eigenschaften, die möglicherweise gegen eine epistemische Interpretation sprechen, neutralisiert. Konkrete Beispiele zeigen genau diesen Effekt: Auch wenn die Konjunktivierung von *dürfen* mit den Eigenschaften eines statischen und perfektivischen Infinitivs (INFZUSTAND und INFKONSTREPI) sowie eines nicht-animierten Subjektreferenten korrelieren sollte, setzt sich die epistemische Interpretation durch (siehe Beispiel 14 sowie für eine optimalitätstheoretische Darstellung Tabelle 2⁹).

- (14) *Die Vertragssumme für die Erschließung habe sich bei 791.000 Euro bewegt. „Nach Abschluss der Maßnahme dürfen wir bei 752.000 Euro ankommen“, erklärte er. („Fränkischer Tag“, 12.9.2012).*

	KONJII	*SUBJANI	INFZUSTAND	INFKONSTREPI
epistemisch ⁹		*	*	*
nicht-epistemisch	*!			

0.7–0.89: hohe Korrelation, ausgeprägte Beziehung

0.4–0.69: mäßige Korrelation, substantielle Beziehung

0.2–0.39: schwache Korrelation, definitive, aber geringe Beziehung

0–0.19: keine oder leichte Korrelation, Beziehung zu vernachlässigen

⁹ Die Reihenfolge der *Constraints* von links nach rechts illustriert ihre jeweilige korrelative Stärke, bemessen am Korrelationskoeffizienten r . Im Rahmen der stochastischen OT wird das Victory-Zeichen für sich durchsetzende Varianten verwendet.

Auf Basis der Stärke des Korrelationskoeffizienten kann nun ein *Ranking* entwickelt werden, das für *dürfen* sowohl die Dominanz von KONJII, die relativ gleich gewichtete Relevanz der anderen Eigenschaften (durch das Zeichen \geq) sowie die neutralisierende Wirkung von KONJII (durch die Klammersetzung) illustriert:

$$\text{KONJII} (> * \text{SUBJANI} \geq \text{INFZUSTAND} \geq \text{INFKONSTREPI})$$

Zusätzlich zur Korrelation, die die Basis für dieses *Ranking* bildet, kann nun noch berechnet werden, mit welcher Wahrscheinlichkeit die Prädikatoren (hier die signifikanten Merkmale für eine epistemische Interpretation) den Regressanden (die epistemische Interpretation) vorhersagen. Die Regressions-Statistik (siehe Tabelle 2) für *dürfen* zeigt, dass ein Bestimmtheitsmaß (R^2) von 0.96 vorliegt, was einer 96%igen Wahrscheinlichkeit dafür entspricht, dass die Merkmale die Interpretation des Modalverbs vorhersagen.¹⁰

Tabelle 2: Output der Regressionsstatistik für *dürfen*.

Multipler Korrelationskoeffizient	0.95058143
Bestimmtheitsmaß	0.96153994
Standardfehler	0.09879998
Beobachtungen	600

Zum Vergleich sollen nun die Ergebnisse von *müssen* präsentiert werden. Die Ergebnistabelle der multiplen Korrelationskoeffizienten in Abbildung 3 zeigt, dass bei diesem Modalverb ausschließlich die semantischen (INFZUSTAND) sowie syntaktischen Eigenschaften (INFKONSTREPI) der Infinitive eine Rolle spielen; alle anderen Merkmale nehmen einen Wert unter dem Schwellenwert von $r = 0.4$ an. Dieses auf den ersten Blick recht eindeutige Verhältnis zeigen auch die Belegbeispiele: Von 300 epistemischen Varianten ist der Infinitiv in 287 Fällen eindeutig als statisch zu beschreiben:

- (15) „*Ich wusste nicht, wie mir geschieht, der Wind nahm mir den Atem, ich kriegte überhaupt keine Luft. Jetzt könne er verstehen, wie sich Amerikaner bei einem Hurrikan fühlen müssen.*“ („Märkische Oderzeitung“, 25.8.2012)

	INFZUSTAND	INFKONSTREPI
epistemisch \wp		*
nicht-epistemisch	*!	

Mehr als die Hälfte der Sätze mit statischem Infinitiv haben zusätzlich eine perfektive oder resultative Konstruktionsweise (siehe Beispiel 16), sind also in diesem Sinn ‚doppelt‘ als nicht-prospektivisch markiert:

- (16) *Nur eine Erkenntnis gibt es: Die Täter gingen ungemein professionell vor und müssen eine ausgezeichnete Ortskenntnis gehabt haben.* („Kurier Wien“, 4.4.2013)

¹⁰ Das Bestimmtheitsmaß gibt an, wie viel Prozent der Varianz der abhängigen Variable (die Interpretation des flektierten Modalverbs) durch die definierten Merkmale erklärt werden. Auch der relativ geringe Standardfehler von 0,099 spricht für eine Verlässlichkeit der Ergebnisse.

	INFZUSTAND	INFKONSTRÉPI
epistemisch \wp		
nicht-epistemisch	*!	*

Umgekehrt zeigt sich, dass bei Verletzung der *Constraints* INFZUSTAND und INFKONSTRÉPI jedenfalls davon auszugehen ist, dass *keine* epistemische Interpretationspräferenz vorliegt:

- (17) *In der Folgezeit ebte der Angriffsschwung jedoch ab. Hier dürfte der Schlüssel zu einem möglichen Erfolg über die Wolfsburger Gäste liegen. Die VSSG muss mehr Druck aus dem Rückraum erzeugen.* („Hessische/Niedersächsische Allgemeine“, 1.12.2011)

	INFZUSTAND	INFKONSTRÉPI
epistemisch	*!	*
nicht-epistemisch \wp		

Im *Ranking* zeigt sich, dass INFZUSTAND das Merkmal INFKONSTRÉPI überschreiben kann, nicht jedoch umgekehrt: Im Korpus kommen sehr wohl etwa Kopula-Konstruktionen mit temporal prospektivischem Verweis vor (siehe Beispiel 18)¹¹.

- (18) *In Schaufenstern, Schaukästen und im Laden müsse der geforderte Preis für Waren oder Dienstleistungen sichtbar ausgehängt sein.* („Thüringer Allgemeine“, 9.6.2012)

	INFZUSTAND	INFKONSTRÉPI
epistemisch	*!	
nicht-epistemisch \wp		*

Das Beispiel zeigt, dass die Verletzung von INFZUSTAND eindeutig stärker ist als die Konstruktionsweise des Infinitivs. Für eindeutige Interpretationspräferenzen kann daher das folgende Ranking angenommen werden:

$$\text{INFZUSTAND} > \text{INFKONSTRÉPI}$$

Für *müssen* zeigt die Regressionsstatistik (siehe Tabelle 3), dass eine Vorhersagekraft von rund 69% aus den definierten Merkmalen abgeleitet werden kann:

Tabelle 3: Output der Regressionsstatistik für *müssen*.

Multipler Korrelationskoeffizient	0.8318419
Bestimmtheitsmaß	0.6919610
Standardfehler	0.2796113
Beobachtungen	600

¹¹ Die umgekehrte Situation, also ein Infinitiv mit statischer Aktionsart und prospektivischer Konstruktionsweise (z. B. mit Vorgangspassiv) ist im Korpus nicht belegt.

Im Vergleich zu *dürfen* ist bei *müssen* nicht nur die berechnete Vorhersagekraft geringer, auch der höhere Wert des Standardfehlers spricht dafür, dass nicht in allen Fällen mit einer verlässlichen Einschätzung der Interpretationspräferenz zu rechnen ist. Dies wird auch durch den Vergleich mit einer Kontrollgruppe von Annotator*innen (20 Studierende der Germanistik) im Zuge einer Fragebogenerhebung anhand einer zufälligen Stichprobe von 50 Beispielsätzen bestätigt:¹² Bei einer Reihe von Beispielen, bei denen nicht beide relevanten Merkmale (INFZUSTAND und INFKONSTREPI) realisiert sind und zusätzlich andere Eigenschaften gegen eine epistemische Interpretation sprechen (z. B. *SUBJANI in Beispiel (19)), treten Zweifelsfälle auf, die aber – und das ist der entscheidende Punkt – ebenfalls mit den definierten Eigenschaften bzw. in diesem Fall mit dem Nicht-Vorliegen dieser erklärt werden kann. Bei dem folgenden Beispiel stimmen die Annotator*innen zu zirka 50% überein, Eine nähere Betrachtung zeigt, dass es sich um die situationell gegebene Notwendigkeit handelt, die durch *müssen* ausgedrückt werden kann.

- (19) *Gewaltige Ausgaben bleiben Würzburg durch den Nahverkehr erspart. Ohne den öffentlichen Nahverkehr müssten dort 73 Hektar an zusätzlichem Parkraum entstehen.* („Osnabrücker Zeitung“, 28.2.2013).
 ‚Vermutlich würden dort 73 Hektar an zusätzlichem Parkraum entstehen.‘
 ‚Gäbe es den öffentlichen Nahverkehr nicht, würde die Notwendigkeit bestehen, dass dort 73 Hektar an zusätzlichem Parkraum entstehen.‘

Diese Variante steht semantisch-funktional zwischen einem Zwang einerseits und einer Vermutung andererseits (aufgrund zwingender situationeller Gründe), was auch in der Forschungsliteratur als Zweifelsfall erkannt wird.¹³ Da der Infinitiv des Beispiels auch nicht eindeutig eine epistemische Lesart kennzeichnet, ist die semantische Ambiguität des Verbs offenbar weder durch Annotator*innen noch durch semi-automatisierte Vergleiche zu disambiguieren. Die Rolle der Merkmale erklärt in diesem Fall keine Interpretationspräferenz, sondern unter Berücksichtigung der Korrelationsstärke vielmehr *warum* hier ein Zweifelsfall vorliegt.

5. Ausblick

Zusammenfassend lässt sich sagen, dass der Beitrag nur einen kleinen Einblick in die grundsätzliche Methodik bieten kann, die in Scherr (2019) vorgeschlagen wurde. Die Modalverbsemantik bewegt sich auf einem Differenzierungsgrad, der zu umfangreich und

¹² Dieser Schritt diene dazu, die berechneten Relevanzen der einzelnen Eigenschaften und die daraus abgeleiteten Rankings einer Gegenprüfung zu unterziehen. Diese bestand darin, die errechnete Bewertung von Beispielsätzen gemäß den Ergebnissen der quantitativen Analyse mit Einschätzungen durch kompetente Sprecherinnen und Sprecher zu vergleichen. Letzteres wurde im Rahmen der erwähnten kurzen Fragebogenerhebung erhoben und die Übereinstimmung mit den berechneten Ergebnissen abermals mittels Annotator Agreement geprüft. Fleiss' Kappa ergibt über alle Beispielsätze hinweg $\kappa = 0.91$, was als sehr guter Übereinstimmungswert gewertet werden kann.

¹³ Nicht ohne Grund werden die „Zwischenstufen“ der *Circumstantial* oder *Dynamic Modality* als separate semantische Gruppen diskutiert (vgl. z. B. Narrog 2012: 10).

zu modalverbspezifisch ist, um an dieser Stelle erschöpfend dargestellt werden zu können. Der Vergleich der beiden diskutierten Verben zeigt jedoch in groben Zügen den Wert, den eine umfassende theoretische Begriffsbestimmung sowie eine intersubjektiv überprüfbare Methodik für den Gegenstand haben können: Zum einen ist die Definition einen *genus proprium* unverzichtbar, um die Besonderheiten der semantischen Untergruppen von Modalität fassen zu können, ohne gleichzeitig die Zusammenfassung unter dieselbe Domäne ad absurdum zu führen. Zum anderen wird durch die Festlegung kotextueller Eigenschaften, die sich wiederum aus den theoretischen Überlegungen ableiten und motivieren, das empirische ‚Instrumentarium‘ geschaffen, das der Intersubjektivität der Analysen entgegenkommt. In diesem Sinne unterscheidet sich die vorliegende Studie von einem großen Anteil an Untersuchungen zur Modalität, die sich überwiegend auf die Diskussion von Einzelbeispielen oder auf die introspektive Beurteilung von Bedeutungsvarianten stützen (müssen). Dass das Bild trotz allem ein differenziertes bleibt, hat der Vergleich zwischen *dürfen* und *müssen* nahegelegt. Die Ergebnisse zeigen ein dynamisches Zusammenspiel der unterschiedlichen Faktoren, die sich zudem modalverbspezifisch höchst unterschiedlich zeigen. Zwar besitzen die einzelnen Features ein kleinstes Gemeinsames, indem sie eine epistemische Funktionalität im Sinne eines nicht-prospektivischen Verweisen stützen, einzeln betrachtet nehmen sie jedoch bei den Verben höchst unterschiedlichen Stellenwert ein: Während bei *dürfen* das Merkmal KONJII in den allermeisten Fällen als Indikator einer Verschiebung der deiktischen Verweisrichtung im Sinne eines Differenzmerkmals gewertet werden muss, hat Modus bei *müssen* nicht den gleichen Effekt. Bei letztem Verb sind es vornehmlich die Eigenschaften des Infinitivs, die die nicht-prospektivische Verweisrichtung ausdrücken. Funktional betrachtet tragen die Parameter insgesamt zur Verhinderung eines zukunftsdeiktischen Zeitbezugs bei, was direkt die hier präsentierte Definition von Epistemik (Abschnitt 2) reflektiert.

Die Diskussion der Ergebnisse hat zudem gezeigt, dass eine quantitative Analyse der kotextuellen Faktoren im Rahmen einer Korpusanalyse Interpretationspräferenzen sowie die Existenz von Zweifelsfällen erklären kann. Dabei wird klar, dass als Resultat niemals eindeutig ableitbare, objektiv festzulegende Bedeutungen stehen, sondern lediglich präferierte Varianten, die aber (zumindest mit statistischer Wahrscheinlichkeit) auf intersubjektive Faktoren zurückgeführt werden können.

Literaturverzeichnis

- Abraham, Werner (2009): Die Urmasse von Modalität und ihre Ausgliederung. Modalität anhand von Modalverben, Modalpartikeln und Modus. In: Werner Abraham, Elisabeth Leiss (Hg.): *Modalität. Epistemik und Evidentialität bei Modalverb, Adverb, Modalpartikel und Modus*. Tübingen: Narr, 215–302.
- Albert, Ruth / Marx, Nicole (2016): *Empirisches Arbeiten in Linguistik und Sprachlehrforschung. Anleitung zu quantitativen Studien von der Planungsphase bis zum Forschungsbericht* (3. Aufl.). Tübingen: Narr.
- Baumann, Carolin (2017): *Bedeutung und Gebrauch der deutschen Modalverben. Lexikalische Einheit als Basis kontextueller Vielfalt*. Berlin, Boston: de Gruyter.

- Blühdorn, Hardarik (2008): Epistemische Lesarten von Satzkonnectoren. Wie sie zustande kommen und wie man sie erkennt. In: Inge Pohl (Hg.): *Semantik und Pragmatik. Schnittstellen*. Frankfurt a. M.: Lang.
- Bresnan, Joan / Deo, Ashwini (2007): Typology in variation. A probabilistic approach to ‚be‘ and ‚n’t‘ in the ‚Survey of English Dialects‘. In: *English Language and Linguistics*. 11, 301–346.
- Brünner, Gisela / Redder, Angelika (1983): Modalverben im Diskurs. In: Gisela Brünner, Angelika Redder (Hg.): *Studien zur Verwendung der Modalverben. Mit einem Beitrag von Dieter Wunderlich*. Tübingen: Narr, 13–90.
- Diewald, Gabriele (2013): Modus und Modalverben. Kategorisierungsoptionen im grammatischen Kernbereich der Modalität. In: Werner Abraham, Elisabeth Leiss: *Funktionen von Modalität*. Berlin, Boston: de Gruyter, 77–110.
- Diewald, Gabriele (2001): Scheinen-Probleme. Analogie, Konstruktionsmischung und die Sogwirkung aktiver Grammatikalisierungskanäle. In: Reinmar Müller, Marga Reis (Hg.): *Modalität und Modalverben im Deutschen*. Hamburg: Buske, 87–110.
- Diewald, Gabriele (1999): *Die Modalverben im Deutschen. Grammatikalisierung und Polyfunktionalität*. Tübingen: Niemeyer.
- Engel, Ulrich (2004): *Deutschen Grammatik* (Neubearbeitung). München: iudicium.
- Greve, Werner / Wentura, Dirk (1997): *Wissenschaftliche Beobachtung. Eine Einführung*. Weinheim: Beltz.
- Heine, Bernd (1995): Agent-oriented vs. epistemic Modality. Some Observations on German Modals. In: Joan Bybee, Suzanne Fleischmann (Hg.): *Modality in grammar and discourse*. Amsterdam, Philadelphia: Benjamins, 17–53.
- Hentschel, Elke / Weydt, Harald (2003): *Handbuch der deutschen Grammatik* (3., völlig neu bearbeitete Auflage). Berlin u. a.: de Gruyter.
- Hundt, Markus (2003): Zum Verhältnis von epistemischer und nicht-epistemischer Modalität im Deutschen. Forschungspositionen und Vorschlag zur Neuorientierung. In: *Zeitschrift für Germanistische Linguistik*. 31, 343–381.
- Kotin, Michail (2010): Zur referentiellen Identität von Tempus- und Modusformen. In: Andrzej Katny / Anna Socka (Hg.): *Modalität/Temporalität in kontrastiver und typologischer Sicht*. Frankfurt a. M.: Lang, 29–38.
- Kotin, Michail (2005): *Die Sprache in statu movendi. Sprachentwicklung zwischen Kontinuität und Wandel*. Bd. 1: *Einführung – Nomination – Deixis*. Heidelberg: Winter.
- Kratzer, Angelika (1991): Modality. In: Arnim von Stechow, Dieter Wunderlich (Hg.): *Semantik. Ein internationales Handbuch zur zeitgenössischen Forschung*. Berlin, New York: de Gruyter, 639–650.
- McCarthy, John (2002): *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.
- Narrog, Heiko (2005): On defining modality again. In: *Language Sciences*. 27, 165–192.
- Plungian, Vladimir (2010): Types of verbal evidentiality marking: an overview. In: Gabriele Diewald, Elena Smirnova (Hg.): *Linguistic Realization of Evidentiality in European Languages*. Berlin, New York: de Gruyter, 15–58.
- Reis, Marga (2001): Bilden Modalverben im Deutschen eine syntaktische Klasse? In: Reinmar Müller, Marga Reis (Hg.): *Modalität und Modalverben im Deutschen*. Hamburg: Buske, 287–318.

- Scherr, Elisabeth (2019): *Die Opazität epistemischer Modalverben im Deutschen. Form, Funktion und empirische Fassbarkeit*. Berlin, Boston: de Gruyter.
- Sudarikov, Roman / Bojar, Ondřej (2015): Giving a Sense. A Pilot Study in Concept Annotation from Multiple Resources. In: Jakub Yaghob (Hg.): *Proceedings of the 15th conference ITAT 2015*. Prag: Charles University Prag, 88–94.
- Zifonun, Gisela / Hoffmann, Ludger / Strecker, Bruno (1997): *Grammatik der deutschen Sprache*. Berlin, New York: de Gruyter.