



MASTERARBEIT | MASTER'S THESIS

Titel | Title

Das *gehören*-Passiv in der österreichischen Standard(schrift)sprache.
Eine Analyse im Austrian Media Corpus (amc)

verfasst von | submitted by

Claudia Mattes BA BA

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Master of Arts (MA)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt |
Degree programme code as it appears on
the student record sheet:

UA 066 817

Studienrichtung lt. Studienblatt | Degree
programme as it appears on the student
record sheet:

Masterstudium Deutsche Philologie

Betreut von | Supervisor:

Univ.-Prof. Dr. Alexandra Lenz

Danksagung

Ich bedanke mich bei

Meinen Eltern Gabriele und Georg für die finanzielle und seelische Unterstützung, die es mir erst möglich gemacht hat, dieses Studium bestreiten zu können.

Meinem Partner David für die jahrelange Partnerschaft und Begleitung während des Studiums.

Meiner Familie, meinen Freund*innen und meinen Brüdern Daniel und Florian für das Anfeuern und die (manchmal nötige) Ablenkung gleichermaßen.

Hannes Pirker für die Beratung zum amc und die Hilfe bei den CQL-Abfragen.

Stefan Resch für die Beratung bezüglich Python und spaCy.

Amelie Dorn für die vielzähligen Diskussionen zu meinem Thema.

Wolfgang Koppensteiner und Katherina Korecky-Kröll für das Beantworten von Fragen in ihren jeweiligen Expertisen.

Allen Kolleg*innen des SFB „Deutsch in Österreich“, Standort Wien.

Meiner Betreuerin Univ.-Prof. Alexandra Lenz für das Vertrauen in mich (angefangen bei der Themenwahl), für die Unterstützung und Möglichkeiten.

Inhaltsverzeichnis

Abkürzungsverzeichnis	6
1. Einleitung	8
2. Theoretische Grundlagen: Standardsprache – Passiv – <i>gehören</i>	11
2.1 Allgemeines zur österreichischen Standardsprache.....	11
2.2 Grammatikalisierung und Passiv-Entstehung.....	16
2.3 Forschungsstand zu <i>gehören</i> im Allgemeinen und zum <i>gehören</i> -Passiv.....	23
3. Methode.....	43
3.1 Allgemeines zum Austrian Media Corpus (amc).....	43
3.2 Vorgehen bei der Zusammenstellung des Korpus.....	46
3.2.1 Vorüberlegungen.....	46
3.2.2 Finale Auswahl der Datenextraktion und weitere Herausforderungen.....	52
3.2.3 Datenextraktion, -reinigung und -filterung.....	59
3.3 Annotation der <i>gehören</i> -Formen.....	60
4. Das <i>gehören</i> -Passiv im amc.....	64
4.1 Quantitative Verhältnisse der <i>gehören</i> -Passiv-Konstruktionen	65
4.2 Morphologie des <i>gehören</i> -Verbs.....	76
4.3 Semantik der VVPPs	79
4.4 Transitivität der Verben.....	84
4.5 Subjekte	84
4.6 Präpositionalphrasen.....	88
4.6.1 Vorhandensein einer lokativen Präpositionalphrase	88
4.6.2 Vorhandensein eines Agens-Anschluss.....	89
4.7 Aspekt der Mündlichkeit	90
5. Diskussion der Ergebnisse.....	94
5.1 Grammatik des <i>gehören</i> -Passivs.....	95
5.2 Semantik des <i>gehören</i> -Passivs.....	96
5.3 Weitere Aspekte der Konstruktion.....	106
5.4 Grammatikalisierung des <i>gehören</i> -Passivs	110
6. Konklusion	114

7. Literaturverzeichnis.....	119
7.1 Verwendete Links	123
7.2 Verwendete Packages	123
8. Abbildungsverzeichnis	124
9. Tabellenverzeichnis	125
10. Anhang	127
Anhang A: Github Repository	127
Anhang B: Medien-/Normalisierungstabelle	127
Anhang C: qualitatives Sample	128
Anhang D: quantitatives Korpus	131
11. Abstracts	134

Abkürzungsverzeichnis

amc	Austrian Media Corpus
ÖAW	Österreichische Akademie der Wissenschaften
APA	Austria Presse Agentur
ACDH-CH	Austrian Centre of Digital Humanities and Cultural Heritage
CQL	Corpus Query Language
DWB	Deutsches Wörterbuch
KWIC	KeyWord in Context
NLP	Natural Language Processing
OTS	Originaltext-Service
PoS	Part of Speech
PP	Präpositionalphrase
STTS	Stuttgart-Tübingen-Tagset
VVPP	Vollverb Perfektpartizip (nach STTS)
VWB	Variantenwörterbuch des Deutschen

1. Einleitung

Das gehört doch mal gemacht! Mit diesem Gedanken trat ich an das Thema des *gehören*-Passivs heran, nachdem ich im Zuge einer Lehrveranstaltung Stathi (2010) Artikel „Is German *gehören* an auxiliary? The grammaticalization of the construction *gehören* + participle II“ vorstellte, der einige Fragen offenließ, und in Folge bei einer ersten Recherche kaum weitere Forschungsliteratur zu diesem spezifischen Thema fand.

Stathi (2010) selbst betrachtet auf diachroner und synchroner Ebene die Entwicklung von *gehören*, im Speziellen die Grammatikalisierung als Passiv. Außer ihr hat sich Szatmári (2002) eingehender mit dem Phänomen des *gehören*-Passivs beschäftigt, ansonsten finden sich in der Forschungslandschaft großteils nur Erwähnungen mit unterschiedlichen Termini und Umschreibungen für die Konstruktion. Dieser Umstand wird in Kapitel 2.3 besprochen. Aus diesem Grund wird eine datenfokussierte Untersuchung dieser spezifischen Verwendung des Verbs vorgenommen, wobei der Fokus auf seine Ausformung in Österreich gerichtet ist.

Die Wahl des Korpus fiel hierbei auf das Austrian Media Corpus (amc), das der Untersuchung von sprachwissenschaftlichen Themen in der österreichischen Zeitungslandschaft dient und Texte aus allen möglichen Medien seit 1984 versammelt (vgl. amc: Das amc!; amc: amc in Zahlen). Um die Datenmenge bereits im Vorhinein etwas einzuschränken, wurde eine zeitliche Begrenzung zwischen 2010 und 2022 vorgenommen.

Durch dieses Vorgehen sollen sowohl weitere Erkenntnisse über die Konstruktion im Allgemeinen gewonnen als auch erstmals ein Blick auf das *gehören*-Passiv in einem dediziert österreichischen Korpus geworfen werden.

Das grundlegende Forschungsinteresse bezieht sich auf die zwei folgenden überdachenden Forschungsfragen:

Wie gestaltet sich die Grammatikalisierung des *gehören*-Passivs im Korpus (in Abgleich mit der bisherigen Forschungsliteratur)?

Welche Regularitäten sind bei der Konstruktion in Bezug auf Morphologie und Semantik zu beobachten?

Bei der Untersuchung wird zudem auf Punkte eingegangen, die im Besonderen bei Szatmári (2002) und Stathi (2009; 2010) eine Rolle spielen, sowie auf Aspekte, die während der Beschäftigung mit dem Thema und den Daten auftraten. Dies sind im Speziellen die Verbformen, in denen *gehören* auftaucht, die Semantik der Perfektpartizipia, mit denen das Passiv gebildet wird, aber auch die Betrachtung des Patiens-Subjekts (vgl. Szatmári 2002: 182), der

lokativen Präpositionalphrasen (vgl. Stathi 2010: 327–328) und des Agens-Anschlusses (vgl. Stathi 2010: 338).

Der Aufbau der Arbeit setzt sich aufgrund der bereits angesprochenen Punkte wie folgt zusammen:

Die theoretischen Grundlagen der Arbeit im zweiten Kapitel sind in drei Stützpunkte aufgeteilt; die Standardsprache, besonders in Österreich – das Passiv im Allgemeinen – das Verb *gehören* mitsamt seiner Verwendung als Passiv.

Zuerst geht es in Kapitel 2.1 um Begrifflichkeiten zur und das Konzept der (schriftlichen) Standardsprache mit Fokus auf die Verwendung in Österreich. Ein Großteil der verwendeten und erklärten Termini stammen von Ammon (1995; 2017) und dem Variantenwörterbuch (VWB) (2016), die Einschätzung zur aktuellen Lage der Standardsprache in Österreich von Koppensteiner (2023). Dabei geht es primär um die Schaffung eines Kontextes für das Korpus aus österreichischen Medien und die dadurch erforschte Varietät des Deutschen in der vorliegenden Untersuchung, die nach Betrachtung der Daten in Kapitel 5.3 noch einmal eingehender besprochen wird.

Das zweite Unterkapitel (2.2) behandelt die Grammatikalisierung und Passiv-Entstehung im Allgemeinen; hier geht es ebenfalls um grundlegende Konzepte des Passivs und den Begriffsdschungel der Grammatikalisierungstheorie seit den 90er-Jahren sowie um unterschiedliche Passivformen

In Kapitel 2.3 wird der Forschungsstand zum *gehören*-Verb im Allgemeinen und zum *gehören*-Passiv im Speziellen aufgearbeitet. Die Passivkonstruktion mit *gehören* ist seit Ende der 70er in der Diskussion vertreten, jedoch bleibt sie häufig eine Nebenbemerkung, während anderen Passivformen in der Forschung in den letzten Jahren mehr Aufmerksamkeit geschenkt wurde; um einige ausgewählte Beispiele zu nennen: das Rezipienten-Passiv mit *bekommen/erhalten/gehören* bei Askedal (2005), *give*- und *get*-Passive bei Lenz (2018) oder gar *kommen*-Passiv-Verwendungen bei Gaeta (2018).

Der Fokus liegt deshalb auf Szatmári (2002) und Stathi (2009; 2010), ihren Vorgehensweisen und nicht zuletzt auch den Beispielen und Belegen, mit denen sie gearbeitet haben.

Dies führt zu Kapitel 3, in dem die Methode dargelegt wird. Zuerst wird das Austrian Media Corpus (amc) vorgestellt (Kapitel 3.1), wobei vor allem auf grundlegende Informationen zum Korpus und den Metadaten wie Jahr und Medien eingegangen wird.

Ein nicht unwesentlicher Teil der Arbeit floss in methodische Überlegungen hinein, wie das *gehören*-Passiv extrahiert werden soll, denen hier ebenfalls Rechnung getragen wird, samt einem Einblick in die Probleme und Herausforderungen mit den Daten. Im Anschluss

wird die endgültig gewählte Vorgehensweise vorgestellt, die sich auf eine Kombination aus CQL-Code, der die Daten aus dem amc extrahiert, und eine Weiterbearbeitung mithilfe des Python Packages spaCy stützt (Kapitel 3.2). Das Vorgehen bezüglich Datenextraktion, -reinigung und -filterung bis hin zur Normalisierung der Belege wird in Kapitel 3.2.4 genauer dargelegt.

Kapitel 3.3 befasst sich mit der Annotation der *gehören*-Passiv-Belege. Da die Untersuchung selbst in eine quantitativ orientierte größere Datensammlung und ein kleineres, qualitativ gestaltetes Sample geteilt ist, wird hierbei vor allem auf die Kategorien der manuellen Annotation in zweiterem eingegangen.

Die Ergebnisse werden in Kapitel 4 gezeigt, angefangen bei der Entwicklung der Belegzahlen im Zuge der Datenextraktion und -bearbeitung.

Im Anschluss wird auf die Verhältnisse des *gehören*-Passivs nach verschiedenen Metadaten eingegangen, das Jahr, in welchem Medium sie erschienen sind, zu welchem der drei Medientypen dies gehört, und welcher Region innerhalb Österreichs das zugeteilt wird.

Der Fokuspunkt liegt im weiteren Verlauf aufgrund der Forschungsfragen auf den Formen von *gehören* und des Verbs als VVPP (Vollverb Perfektpartizip)¹, um Regularitäten bezüglich Morphologie und Semantik der Konstruktion ausfindig zu machen.

Durch mehr qualitativ in Augenschein genommene Belege wird außerdem ein Blick auf weitere Fragestellungen und Aspekte geworfen, die sich in der quantitativen Datensammlung aus technischen oder methodischen Gründen nicht derart untersuchen lassen.

Dies betrifft zum einen eine detailliertere Aufschlüsselung der Morphologie bezüglich Person, Anzahl, Tempus und Modus, und andererseits Punkte, aber auch semantische Gesichtspunkte der Patiens-Subjekte, das Vorkommen von Präpositionalphrasen – einmal in ihrer lokativen Variante und einmal als Agens-Anschluss – sowie Aspekte der konzeptionellen Mündlichkeit.

Eine spannende Frage, die nämlich während der Betrachtung des Themas auftauchte, ist die nach der möglichen mündlichen Konnotation der Konstruktion, auch in einem schriftlichen (Medien-)Korpus. Darauf wird ein besonderes Augenmerk geworfen, unter Einbezug der Ergebnisse aus der quantitativen Datensammlung, die Aufschluss über diese

¹ In dieser Arbeit wird auf das Perfektpartizip des Verbes mit der Abkürzung VVPP Bezug genommen. Diese wurde von der Annotationskonvention geliehen, mit der Vollverben im Perfektpartizip nach dem STTS annotiert werden (vgl. STTS 1999: 7). Dies dient einerseits dazu, bei einer kohärenten Kurzform zu bleiben, die auch bei der Bearbeitung des Korpus verwendet wird, und andererseits den Unterschied zur Präpositionalphrase aufrecht zu erhalten, die häufig mit PP abgekürzt wird.

Mündlichkeit geben können, aber auch durch konkrete Annotationen vor allem zu direkten und indirekten Reden im Text, die im qualitativen Sample betrachtet werden.

In Kapitel 5 werden die Ergebnisse vor dem Hintergrund der Forschungsliteratur und den Vorannahmen noch einmal diskutiert, eine Einschätzung über die Grammatikalisierung und Gestaltung der Konstruktion abgegeben sowie die weiters untersuchten Aspekte – vor allem zur Mündlichkeit – kontextualisiert.

Die Arbeit schließt mit einer Konklusion ab, die die wichtigsten Punkte und Ergebnisse zusammenführt, aber auch einen ausführlichen Ausblick gibt, der auf weiterführende Forschungsdesiderata bezüglich des *gehören*-Passivs eingeht und im Besonderen auf weitere mögliche Methoden und thematische Anknüpfungen verweist.

2. Theoretische Grundlagen: Standardsprache – Passiv – *gehören*

In diesem Kapitel werden die theoretischen Grundlagen dargelegt, die zur Untersuchung des *gehören*-Passivs im Rahmen der Arbeit relevant sind. Zuerst werden die Standard(schrift)sprache, im Besonderen auf ihre Ausprägung in Österreich, und die theoretischen Rahmenbedingungen des Themenfeldes beleuchtet.

Im Anschluss wird auf das Passiv im Allgemeinen, die Grammatikalisierung desselben sowie einige passivische Konstruktionen eingegangen.

Zu guter Letzt schließt der Theorieteil mit der Betrachtung des Verbs *gehören* und der damit gebildeten Passiv-Form ab. Der Fokus liegt dabei auf den Artikeln von Szatmári (2002) und Stathi (2009; 2010), den unterschiedlichen Theorien und Konzepten, die im Zuge der Grammatikalisierung der Konstruktion bei ihnen genannt werden, sowie Aspekte des Passivs, die seine Gestalt und Verwendung prägen.

2.1 Allgemeines zur österreichischen Standardsprache

In diesem Kapitel geht es um jenen Teil des Titels der Arbeit, der auch in die Kapitelüberschrift einfließt, nämlich die „österreichische(...) Standard(schrift)sprache“.

Zuerst sei an dieser Stelle gesagt, dass die Formulierung als Attribution mit Adjektiv vor allem aus stilistischen Gründen für den Titel bevorzugt wurde; es könnte genauso gut „Standardsprache in Österreich“ heißen. Diese Begriffe werden austauschbar verwendet, ohne auf die Diskussionen genauer einzugehen, die in der Vergangenheit in Bezug darauf geführt wurde.²

² Als Teil der Diskussion zum Begriff der Standardsprache wurde in meiner Bachelorarbeit mit Fokus auf unterschiedliche Konzepte (Plurizentrik, Pluriarealität, Sprachdynamik) das Verhältnis der unterschiedlichen Standardvarietäten des Deutschen diskutiert, wobei im Besonderen die Varietät in Österreich

Deutsch ist die offizielle Amtssprache in Österreich, aber nicht nur in diesem Land, was das Deutsche als plurizentrische Sprache platziert (vgl. Ammon 1995: 12).³ Grundsätzlich meint ‚plurizentrisch‘ eben, dass es mehrere Länder gibt, in denen die Sprache gesprochen wird. Je nach ihrem Status in dem jeweiligen Land wird dabei von Voll-, Halb- oder Viertelzentren gesprochen (vgl. Ammon/Bickel/Lenz 2016: XXXIX–XL; Ammon 2017: 18).



Abbildung 1: Die nationalen Zentren der deutschen Sprache (VWB 2016: XLII)

Eines der ersten Werke, das sich ausführlich mit Standardvarietäten des Deutschen im Kontext der Plurizentrik – oder Plurizentrizität – beschäftigte und dieses Konzept auch problematisierte, ist Ammon (1995). Bis hin zur Tätigkeit am Variantenwörterbuch des Deutschen (siehe Ammon/Bickel/Lenz 2016) nimmt Ammon (2017) immer wieder auf den

beleuchtet wurde und dabei die teilweise stark emotionalisierte Debatte, auch zur „richtigen“ Benennung, problematisiert (siehe Mattes (2021), unter anderem auf Glauninger (2015b) stützend, sowie die weiterführende Literatur zur Debatte und ihrer Aufarbeitung in Fußnote 3).

³ Die Thematik rund um das Konzept der Plurizentrik/Plurizentrizität wird hier nur verkürzt dargestellt, da es grundsätzlich nur um österreichische Standardsprache bzw. Standard in Österreich geht – also in einem der DACH-Länder, nicht in allen und kein direkter Vergleich angestrebt wird. Zudem soll die gesamte Diskussion rund um die umstrittene Beschreibung der unterschiedlichen Standardvarietäten ausgeblendet werden (Stichwort: österreichisches Deutsch vs. Deutsch in Österreich), da eine Ausführung zu diesem Thema den Rahmen sprengen würde und nicht zwingend relevant ist. Zur Diskussion siehe (beispielhaft und nicht extensiv) Muhr/Schrodt/Wiesinger (1995), Scheuringer (1996), Auer (2014), Wiesinger (2014), Glauninger (2013; 2015a; 2015b), Ammon/Bickel/Lenz (2016: XXXIX–XLII), Muhr (2017), Koppenssteiner/Lenz (2017; 2021: 394–397), De Cecilia/Ransmayer (2019), Dollinger (2019), Koppenssteiner (2023).

Begriffskomplex rund um die ‚Plurizentrik‘ der deutschen Sprache Bezug, genau die verwendeten Termini, die eben auch die theoretische Basis für das VWB (Ammon/Bickel/Lenz 2016) darstellen, und aus diesem Grund nach wie vor für variationslinguistische Untersuchungen relevant sind.

Die postulierten ‚(nationale[n]) Zentr[en] einer Sprache‘ (Ammon 1995: 42) im Zuge der Plurizentrität unterliegen einer langen Diskussion, die mindestens seit Mitte des letzten Jahrhunderts besteht (siehe Kapitel "Eigentliche Forschungsgeschichte" in Ammon 1995: 42–49) und sich über eine besonders heiße Phase in den 90ern und frühen 2000ern teilweise bis heute zieht.⁴

Es sind grundsätzlich bei der Betrachtung der deutschen Sprache Standardvarietäten und Nonstandard-Varietäten voneinander zu unterscheiden (vgl. Ammon 1995: 3–4)⁵, was weiters zur Folge hat, dass, wenn von ‚nationale[n] Variante[n]/Varietät[en] einer Sprache‘ (Ammon 1995: 42) die Rede ist, so eben auch von ‚(nationale[n]) Standardvarietät[en]‘ gesprochen wird (vgl. Ammon 1995: 69);⁶ also auf die einzelnen „Zentren“ der Sprache bezogene Standards.

Ein wichtiger Schritt der Definition einer (nationalen) Standardvarietät ist die Kodifizierung derselben (vgl. Ammon 1995: 74), zu welchem Zweck Ammon (1995, 80) ein einfaches Modell entwirft, dass das ‚soziale Kräftefeld einer Standardvarietät‘, in dem diese Aushandlung geschieht, darstellen soll:

⁴ Die österreichische Ausformung des Deutschen nahm dabei eine besondere Stellung ein, unter anderem, weil die Diskussion häufig keine linguistische, sondern ideologische ist (siehe beispielsweise Scheuringer 1996; Glauninger 2013, 2015a, 2015b).

⁵ Auch wenn sich hier – wie so oft – die Frage stellt, wie sehr dies in Bezug auf tatsächliche Sprachverwendung möglich ist, siehe Konzept des Gebrauchsstandards, wie es beispielsweise bei Elspaß/Dürscheid/Ziegler (2017) behandelt wird.

⁶ Zur Diskussion und Problematisierung des Begriffes ‚Nation‘, im Allgemeinen, aber auch besonders in einer ‚plurinationalen‘ Auslegung der Plurizentrik, siehe unter anderem Ammon (1995) selbst (Kapitel A 2 Zum Begriffsfeld ‚Sprachgemeinschaft‘ und ‚Nation‘, s. 18–34) und Glauninger (2013).

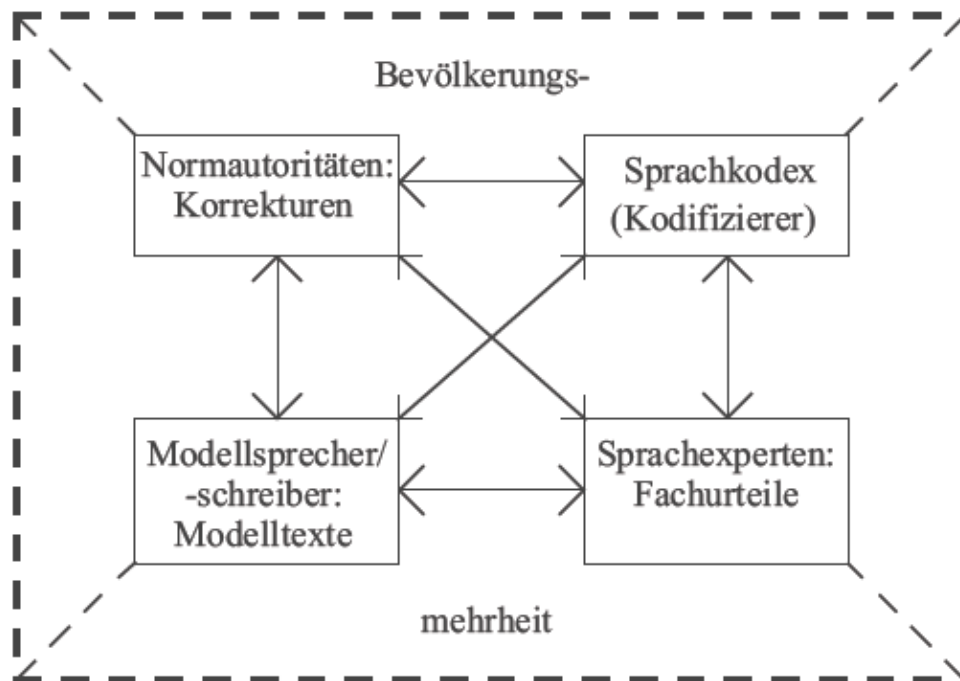


Abbildung 2: Soziales Kräftefeld einer Standardvarietät (Ammon 1995: 80; 2017: 11)

Unterschiedliche Gruppen sind hierbei beteiligt: Normautoritäten, die Korrekturen durchführen, beispielweise Lehrpersonen; Kodifizierende, die für die Erstellung von Sprachkodizes zuständig sind; Sprachexpert*innen, die die Verwendung von Sprache beobachten und einschätzen, zum Beispiel Wissenschaftler*innen der Philologien. Eingebettet ist dieser Diskurs in die Sprachverwendung der Bevölkerungsmehrheit (vgl. Ammon 1995, 78–79; 2017, 11–18).

Für die Betrachtung des *gehören*-Passivs in dieser Arbeit ist vor allem das Element der Modellschreiber*innen und -sprecher*innen und ihrer Modelltexte von Relevanz. Bei diesen handelt es sich um Personen, die dies in den meisten Fällen beruflich machen und im Zuge ihrer Tätigkeit schriftliche und mündliche Texte produzieren, die für eine öffentliche Publikation gemacht sind. Man spricht von Modelltexten, da diese sprachlichen Dokumente eine gewisse Vorbildwirkung für die Kodifizierung von verwendeten Neuerungen und Änderungen haben. Sprachexpert*innen und Kodifizierer*innen beobachten die Sprachverwendung in den öffentlichkeitswirksamen Texten, untersuchen diese und stützen neue Normen und Regelungen darauf (vgl. Ammon 1995: 79; 2017: 11–12).

Zwischen den Beteiligten gibt es Rückkopplungsprozesse, die über die entgegenlaufenden Pfeile dargestellt werden. Diese stellen die gegenseitige Beeinflussung der unterschiedlichen Aspekte des Modells dar, da diese keine in sich abgegrenzten, unabhängig

operierbaren Instanzen sind, sondern sich eben bei der Ausbildung der Standardvarietät(en) gegenseitig stützen (vgl. Ammon 1995: 79).

Eben diese Kodifizierung der Konstruktion ist augenscheinlich bereits im Gange: In der 9. Auflage des Grammatik-Dudens wird *gehören* + Partizip II als Alternative zum *werden*-Passiv angeführt (vgl. Duden 2016: 562).

Nicht nur wer über dreißig ist, gehört respektiert. (,Nicht nur wer über dreißig ist, muss/sollte respektiert werden.‘) Unpersönlich: Dir gehört geholfen. (Duden 2016: 562)

Hier wird es als ‚modale Passivvariante‘ beschrieben und als prominentes Beispiel wird *Das gehört gehört!*, ein Werbespruch für einen Radiosender, genannt (vgl. Duden 2016: 434).⁷

Ammon (1995: 82) spricht in Bezug auf die Setzung des Standards ebenfalls das Problem der Abgrenzung desselben vom Nonstandard an. Eine beispielhafte „Normsetzung“, die von Sprachexpert*innen kritisiert wurde, ist die Aufnahme von Nonstandard-Varianten in der 35. Auflage des Österreichischen Wörterbuches (ÖWB) im Jahre 1979. Derartige Fälle zeigen, wie die Grenzen nicht nur für die durchschnittlichen Sprachnutzer*innen verwischt sind, sondern auch bei Kodizes Grenzfälle des Standards Gang und Gäbe sind (vgl. Ammon 2017: 14).⁸

Ab wann etwas als Standard gelten kann, ist also ebenfalls ein Aushandeln zwischen den einzelnen Gruppen des Kräftefeldes, eine Beurteilung der Verwendung unterschiedlicher Konstruktionen, Wörter und Grammatikalisierungen durch die Sprecher*innen einer Sprache (vgl. Ammon 1995: 82–83).

Ähnliche Fragestellungen rund um dieses Spannungsfeld sind bei Koppensteiner (2023) zu finden, der sich mit der attitudinal-perzeptiven Sichtweise auf ‚Standard in Österreich‘ beschäftigt, wobei ein besonderes Augenmerk auf die Perspektive von linguistischen Lai*innen gelegt wird.

Koppensteiner (2023: 12) reflektiert weiters über die Schwierigkeit bei der Ad-hoc-Definition des – oder der – Standards in Österreich, vor allem aus dieser konkreten Sichtweise. Dessen oder deren Existenz wird nicht angezweifelt, allerdings herrscht Unklarheit über die genaue Gestalt der Standardvarietät(en) (vgl. Moosmüller 1991: 22), was eben auch mit

⁷ Bei dem angesprochenen Radiosender handelt es sich entweder um Ö1 (aus Österreich) oder SWR1 (aus Deutschland), die beide eine Version des Spruches als Werbung verwend(et)en. (vgl. Radioszene (2012)). Der Slogan stammt ursprünglich aus der Feder von Wolf Haas, in der Version „Ö1 gehört gehört“ (vgl. Standard (2020); Presse (2020)).

⁸ Weitere Diskussion unter anderem bei Wiesinger (2014: 231–258) Kapitel: Zum „Österreichischen Wörterbuch“. Aus Anlaß der 38. neubearbeiteten Auflage (1997).

Abgrenzung zum Nonstandard in dem fließenden Dialekt-Standard-Kontinuum zu tun hat (vgl. Soukup 2009: 39).

Seine Untersuchungen beziehen sich in einem Teilaspekt im Speziellen auf die *Sprecher*innen* des Standards in Österreich (vgl. Koppensteiner 2023: 13) und die Ergebnisse heben hierbei in Österreich vor allem die *ORF-Nachrichtensprecher*innen* und *Nachrichtensendungen* hervor. Die professionelle sprachliche Schulung derselben, aber die dennoch hörbare (staatsbezogene) sprachliche Herkunft werden als Merkmale angeführt. Standardsprachliche Sprachäußerungen, die eben gar keine regionalen Marker aufweisen, sind quasi unerreichbare Ideale, für die es keine Sprechenden gibt (vgl. Koppensteiner 2023: 380, 390–391).

„Standard in Österreich“ scheint keine solche idealisierte Sprachform darzustellen, sondern eine mit **prononciert gesprochen sprachlicher Relevanz**, wiewohl diese auf spezifische Kontexte beschränkt wird. (Koppensteiner 2023: 391) (Hervorhebung im Original, C.M.)

Gleichzeitig ist in der grundsätzlichen Konzeption des (sprechsprachlichen) Standards ein Bezug auf die Schriftsprache gegeben, gerade wenn es um Ab-/Vorlesesituationen geht, wie eben Nachrichtensendungen oder Vorträge (vgl. Koppensteiner 2023: 379).

Bei der Betrachtung des *gehören*-Passivs in einem österreichischen Zeitungskorpus werden aufgrund der Präsuppositionen, die sich auf die Herausbildung der Standardsprache nach Ammon (1995) beziehen, vor allem standardsprachliche Texte von Modellschreiber*innen erwartet. Durch die weitere Kontextualisierung in Koppensteiner (2023) im Besonderen auf den österreichischen Standard – auch wenn diese sich vor allem auf den gesprochenen Standard beziehen – muss jedoch bereits hier eingehakt werden. Wie sich diese Vorannahmen dann tatsächlich in Bezug auf das Phänomen verhalten, wird in Kapitel 5.3 diskutiert, in dem es primär um die potenzielle Mündlichkeit der Konstruktion und die Bedeutung der Ergebnisse für die hier untersuchte (und allgemein im Korpus reproduzierte) Varietät geht.

2.2 Grammatikalisierung und Passiv-Entstehung

Das *gehören*-Passiv wird im gleichen Atemzug mit anderen Passiv-Formen genannt (vgl. Szczepaniak 2011: 153)⁹. Um also Kontext für die Diskussion der Grammatikalisierung der Konstruktion zu schaffen, wird auf einige Grundkonzepte eingegangen sowie ein kurzer Abriss der bereits ausführlicher beforschten Passive gegeben.

⁹ In Kapitel 2.3 zum Forschungsstand zu *gehören* wird kurz auf die Diskussion eingegangen, inwieweit die *gehören*-Konstruktion als Passiv bezeichnet wird bzw. werden kann.

Zur Grammatikalisierungstheorie, vor allem zum Passiv, wurden besonders in den 90ern und 2000 viel Forschungsliteratur produziert; sie bleibt aber nach wie vor relevant wie Artikel aus dem letzten Jahrzehnt zeigen (siehe beispielsweise Szczepaniak (2011) oder Ausgabe 43 (2) der Zeitschrift Sprachwissenschaft aus dem Jahre 2018, die sich mit unterschiedlichen Aspekten des Passivs beschäftigt; für diese Arbeit beispielhaft herangezogen: Lenz (2018) und Gaeta (2018)). Einige grundlegende Begriffe und Konzepte stammen aus Werken, die zu dieser Zeit der Grammatikalisierung entstanden sind und nach wie vor angewandt werden.

Ein Problem bei der Betrachtung von Grammatikalisierung, besonders von Passiv-Formen, scheint zu sein, dass jede*r Forschende seine bzw. ihre präferierten Theorien und Begrifflichkeiten hat, die teilweise mit anderen Konzepten übereinstimmen, aber durch die unterschiedlichen Bezeichnungen ist ein Überblick durchaus schwierig. In dieser Arbeit wird dennoch versucht, durch den Dschungel der Begriffe, Termini und Modelle zu navigieren, um die Grammatikalisierungstheorie produktiv für die Betrachtung des *gehören*-Passivs nutzen zu können.

Zuerst wird der Begriff und das Konzept der Grammatikalisierung genauer betrachtet. Hopper/Traugott (2012: XV) definieren dies kurz und knapp als “the change whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions and, once grammaticalized, continue to develop new grammatical functions.”

Lexikalische Items werden also in bestimmten Kontexten zu grammatikalischen, dies geht mit unterschiedlichen strukturellen Veränderungen einher.

Laut Askedal (2005: 214) ist der Terminus Grammatikalisierung aus funktionaler Sicht doppeldeutig: Einerseits kann er resultatbezogen „als funktionale Eingliederung in ein grammatisches (Sub-)System (beispielsweise das der Tempus- oder Passivhilfsverben)“ verstanden werden, andererseits eben auch als der Prozess der einzelnen Vorgänge, die eben zu dieser Eingliederung führen.

Askedal (2005: 213–214) stellt zudem die Grammatikalisierungsprinzipien nach Hopper (1990) vor, fünf an der Zahl:

Prinzip I: *Accumulation/Layering* (Akkumulation/Schichtung)

Prinzip II: *Divergence* (Divergenz)

Prinzip III: *Specialization* (Spezialisierung)

Prinzip IV: *Persistence* (Persistenz)

Prinzip V: *Decategorialization* (Dekategorialisierung)

In seinem Artikel zum *bekommen/kriegen/erhalten*-Passiv nimmt er eine Problematisierung bzw. Erweiterung von Hoppers Theorie vor. Wichtige Punkte sind hierbei die Ergänzung des Terminus Dekategorialisierung um die „Endstation“ Rekategorialisierung, die Erklärung, dass Rekategorialisierung je nach Grammatikalisierungsprozess als Auxiliarisierung, Partizipialisierung, usw. zu bezeichnen ist, und die Frage, ob Wortklassenwechsel generell als Grammatikalisierungsprozesse aufzufassen sind (vgl. Askedal 2005: 214).

Zudem werden Hoppers (1990) Begriffe in ein Baumdiagramm gebracht, um die Prozesshaftigkeit bzw. ihre Abfolge im Zuge der Grammatikalisierung darzustellen anstatt sie parallel nebeneinander stehen zu lassen (vgl. Askedal 2005: 215).

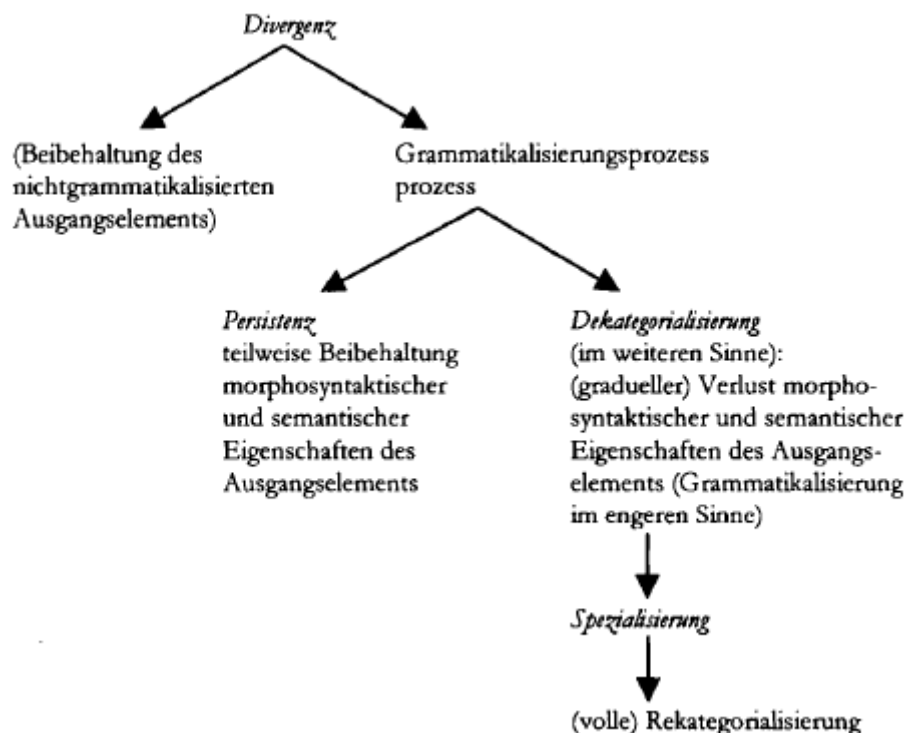


Abbildung 3: Anpassung der Theorie Hoppers (1990) als hierarchisches Baumdiagramm (Askedal 2005: 215)

Der Grammatikalisierungsprozess besteht also aus einer Abwechslung von Persistenz und Dekategorialisierung, solange bis die Spezialisierung eintritt und das Item als (vollständig) rekategorialisiert betrachtet werden kann (vgl. Askedal 2005: 214–215).

Die Grundprozesse der Grammatikalisierung werden auch bei Szatmári (2002: 183–184) auf Basis von Wegener (1998: 37, 43) dargestellt:

1. Verlust an phonologischer Substanz
2. Verlust an semantischer Substanz (*semantic bleaching*) (Kursivsetzung im Original, C.M.)
3. Verlust an syntaktischer Freiheit
4. Gewinn an pragmatischer/illokutiver/metakommunikativer Stärke

Diese Abfolge an Schritten ist nicht direkt in dieser Reihenfolge aneinandergeknüpft oder in dieser Ausformung bei jedem grammatikalisierten Item vollständig ausgeprägt vorzufinden, aber ähnlich der Aufzählung bei Askedal (2005: 212–213) sind dies Eigenschaften von sich in der Grammatikalisierung befindenden Elementen.

Szczepaniak (2011) beschäftigt sich in ihrem Werk zur Grammatikalisierung im Deutschen mit diversen Formen davon, beispielsweise den Plural- oder Pronominalformen (siehe Szczepaniak 2011: 55–80), aber auch der Passiv-Entstehung (siehe Szczepaniak 2011: 143–150, 152–158). Grundbegriffe werden hierbei ebenfalls anhand des *bekommen*-Passivs gezeigt (vgl. Szczepaniak 2011: 36–37):

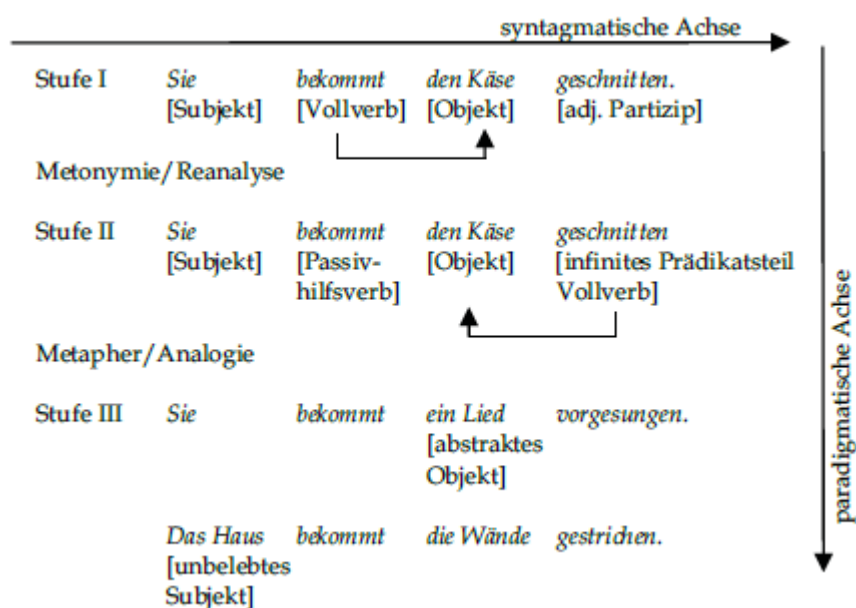


Abbildung 4: Stufen der Grammatikalisierung (Szczepaniak 2011: 37)

Zwischen Stufe I und II passiert eine Reanalyse, die auch als Metonymie bezeichnet wird, indem die syntaktische Struktur anders interpretiert wird als zuvor und eine neue Lesart entsteht. In weiterer Folge zu Stufe III wird die Struktur analog auf weitere Kontexte ausgeweitet, die zuvor in Stufe I nicht möglich waren. Dies wird auch als Metapher bezeichnet (vgl. Szczepaniak 2011: 37). Beim Beispiel des Rezipienten-Passivs führt Szczepaniak (2011: 36–37) abstrakte Objekte und unbelebte Subjekte an, um dies zu zeigen.

Wenn man die Grammatikalisierung von Items betrachtet, gibt es eine Vielzahl von Begriffen und Konzepten, die man heranziehen kann, um die Vorgänge zu beschreiben; hier wurden vor allem jene angeführt, die für die Betrachtung des *gehören*-Passivs am relevantesten sind, auch weil sich die bereits existierende Forschungsliteratur ebenfalls darauf stützt.

Im Anschluss werden nun die bereits üblichen Passiv-Konstruktionen dargestellt, angefangen bei den Prototypen mit *sein* und *werden*. Das *sein*- und das *werden*-Passiv verbindet eine gemeinsame Entwicklung, da heutzutage zwar die *werden*-Konstruktion als das Vorgangspassiv gilt, während das mit *sein* gebildete Passiv einen Zustand ausdrückt (vgl. Szczepaniak 2011: 141). So ist *Das Auto wird gewaschen.* als Vorgangspassiv mit *werden* demselben Beispiel mit *sein* *Das Auto ist gewaschen.* gegenüberzustellen (vgl. Szczepaniak 2011: 141).

Dies war jedoch im Althochdeutschen noch umgekehrt:

min tohter ubilo fon themo tiuuale giuueigit ist (T 85, 2) (Vorgangspassiv)

'meine Tochter wird schrecklich vom Teufel gequält' (Szczepaniak 2011: 141).

Im Gegenzug kann auch das *werden*-Passiv nach wie vor nicht nur für den Ausdruck des Vorgangs, sondern auch des Zustands verwendet werden. „Alle umliegenden Gebäude werden von dem neuen Hochhaus überragt.“ (Szatmári 2002: 176).

Sein und *werden* als Zustands- und Vorgangspassiv ist demnach eine grobe Einteilung, die je nach Verwendung im Kontext auch divergieren kann, da diese Unterscheidung historisch gewachsen ist und die jeweils andere Bedeutung nicht vollständig verdrängt wurde.

Askedal (2005) zeigt einige Konzepte der Grammatikalisierungstheorie anhand des sogenannten Dativ- oder Rezipienten-Passivs, das eine Grammatikalisierung der drei Verben *kriegen/bekommen/erhalten* mit sich zieht (vgl. Askedal 2005: 211; siehe oben).

Lenz (2018) widmet ihren Artikel dem weniger häufig behandelten *geben*-Passiv (in dem vorliegenden Werk als GIVE-Passive bezeichnet), das vor allem im Luxemburgischen sowie einigen wenigen westmitteldeutschen Dialekten in Deutschland belegt ist (vgl. Lenz 2018: 118).

In Luxemburg ist zudem auch das GET-Passive (zuvor bereits als Dativ- oder Rezipienten-Passiv vorgestellt) verbreitet, hier jedoch vor allem mit dem Verb *kriegen* (auf Luxemburgisch *kreien*) (vgl. Lenz 2018: 208). Allgemein sind *kriegen* und *bekommen* im deutschsprachigen Raum häufiger anzutreffen als *erhalten*, letzteres Verb scheint in der Grammatikalisierung noch weniger weit fortgeschritten zu sein (vgl. Lenz 2018: 197–198).

Ein weiterer Artikel, der passivische Formen behandelt, die regional eingeschränkt sind, ist Gaetas (2018) „Im Passiv sprechen in den Alpen“: *kommen*-Passiv (vgl. Gaeta 2018: 227), *gehen*-Passiv (vgl. Gaeta 2018: 230–232) sowie eine spezielle Form des *sein*-Passivs (vgl. Gaeta 2018: 229).

Eine weitere am Rande erwähnte Passiv-Variante ist das *bleiben*-Passiv, das bei Vogel (2006: 3) gemeinsam mit *gehören* + Perfektpartizip erwähnt, aus ihrer Untersuchung zum unpersönlichen Passiv jedoch ausgeschlossen wird, weil es nur eine persönliche Form gibt.

Formen wie das *geben*-, *kommen*- und *gehen*-Passiv, die entweder mündliche und/oder räumlich eingeschränkte substandardsprachliche Varianten zu sein scheinen, werden ähnlich selten beforscht wie das *gehören*-Passiv.

Die „üblicheren“ Passiv-Formen werden bei Szatmári (2002: 175) in einer Baumstruktur dargestellt. Sie nennt das Modell *Passiv-Mikrofeld 1 [-Modalfaktor]*, in Abgrenzung zu passivischen Konstruktionen mit Modalfaktor, die im folgenden Kapitel 2.3 zum *gehören*-Passiv kurz angesprochen werden.

Die Passive werden in ‚vorgangsorientiert‘ und ‚zustandsorientiert‘, ‚ohne jedwede Zentrierung‘ und ‚mit Nichtagens-Zentrierung‘ eingeteilt, wobei letztere Kategorie wiederum in ‚mit Patiens-Zentrierung‘ und ‚mit Rezipienten-Zentrierung‘ unterteilt wird (Szatmári 2002, 175).

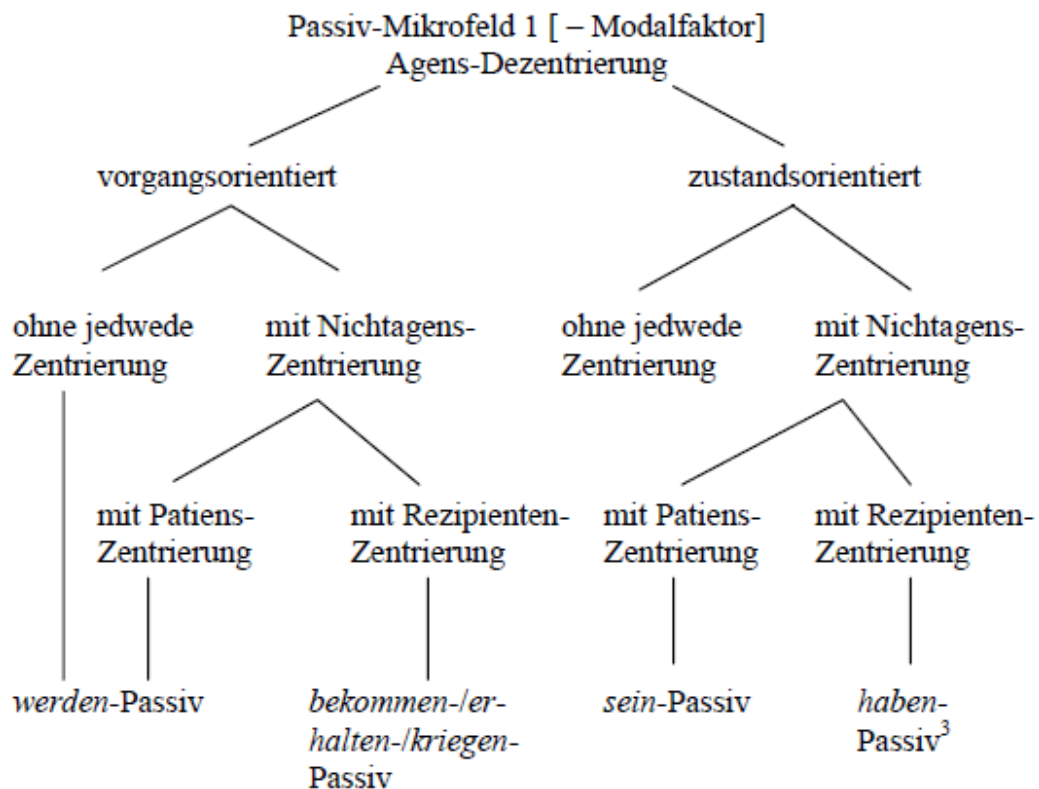


Abbildung 5: *Passiv-Mikrofeld 1 [-Modalfaktor]* (Szatmári 2002: 175)

Das angenommene *haben*-Passiv wird von Szatmári (2002: 175) in der Fußnote kurz angeführt, sie stützt sich dabei vor allem auf Helbig (1989). Es handelt sich hierbei um eine Konstruktion mit *haben* plus Perfektpartizip, bei dem jedoch das Subjekt nicht als Agens gilt: *Er*

hat die Wunde verbunden (Helbig 1989: 219, nach Szatmári 2002: 175, Fußnote 3).¹⁰ Es handelt sich dabei um eine eingeschränkte, eher seltene Konstruktion, die in Zusammenhang mit *wollen* etwas üblicher zu sein scheint, in Sätzen wie beispielsweise *Sie will das laut und deutlich vorgelesen haben* (vgl. Szatmári 2002: 175, Fußnote 3).

Bei den Passiv-Grammatikalisierungen handelt es sich um die spezielle Form der Auxiliarisierung (siehe Szatmári 2002: 183; Askedal 2005: 214), also ein lexikalisches Item – in dem konkreten Fall ein Verb – wird zum grammatikalischen Part des Prädikats.

Bei Szatmári (2002) wird eine Definition von Auxiliarverben nach Amrhein (1995) angegeben, die im Kontext der Passiv-Konstruktionen und besonders der „Nebenformen“ von Relevanz ist:

Auxiliarverben („Nebenverben“) sind alle Verben, die mit einem weiteren Element das Prädikat des Satzes bilden, wobei das Auxiliar Träger der Verbflexion ist (nach Person, Numerus, Tempus, Modus) und das zweite Element zwar keine Flexionsmorpheme aufweist, aber eine vom Verb abgeleitete Form darstellt (im Deutschen sind das Infinitiv, Partizip und Nominalisierung), die als Träger des verbsemantischen Kerns im komplexen Prädikat fungiert. Die funktionale Leistung des Auxiliars kann auf grammatische Kernfunktionen (wie Passiv und Tempus) beschränkt sein, sie kann aber auch stärker „eigensemantisch“ sein, indem sie die Satzaussage modalisiert, aspektualisiert oder sogar situativ und pragmatisch modifiziert. (Amrhein 1995: 86-87, nach Szatmári 2002: 183)

Hilfsverben sind nicht die alleinigen Verben eines Satz-Prädikats, sondern der primär grammatikalische Teil, während der infinite Verbteil einen Großteil der Semantik trägt, jedoch wie bei Amrhein (1995: 86-87, nach Szatmári 2002: 183) auch angemerkt wird, kann eben auch das Auxiliar Bedeutung tragen.

Die Bemerkung, dass auch ein Auxiliarverb eine gewisse Eigensemantik haben kann, ist ein Aspekt, der im Besonderen im nächsten Kapitel noch einmal diskutiert wird, nämlich konkret auf das *gehören*-Passiv bezogen.

¹⁰ Die Gleichsetzung mit dem *bekommen*-Passiv und die Rezipienten-Zentriertheit könnten auch nahelegen, dass es sich hier um mehr als nur strukturell verwandte Konstruktionen handelt. Der Beispielsatz *Er hat die Wunde verbunden* würde durch ein angeschlossenes *bekommen*, also eine Bildung des Rezipientenpassivs im Perfekt, ebenfalls die grundlegende Semantik behalten. Auch beim Beispiel mit Modalverb wäre das *haben* durch *bekommen* ersetzbar. Möglicherweise handelt es sich hier um eine (sprechsprachliche) Reduktion der Konstruktion (beispielsweise aus Gründen der Sprachökonomie), wenn durch den Kontext eine passivische Lesart ohne die Explikation bereits deutlich ist. Dies spricht wiederum für das Argument der konversationellen Implikatur/pragmatischen Inferenz, die bei der Grammatikalisierung des Passivs beteiligt ist (siehe Szatmári 2002: 181; Szczepaniak 2011: 36–37).

2.3 Forschungsstand zu gehören im Allgemeinen und zum gehören-Passiv

Das folgende Kapitel widmet sich konkret dem Verb *gehören* und der damit gebildeten Passivform. Ein Großteil der Ausführungen stützt sich auf Szatmári (2002) und Stathi (2009; 2010), die es ausführlicher behandeln.

Um in dieser Hinsicht weitere Beobachtungen anstellen zu können, wird nun kurz *gehören* im Allgemeinen betrachtet, angefangen bei der Etymologie und der Bedeutung des Verbes.

gehören swV. (< 8. Jh.). Das Wort war weit verbreitet als Intensiv-Form zu hören. Die heutige Bedeutung entwickelte sich aus ‚auf jmd. Hören‘ = ‚zu ihm gehören‘, geht also ursprünglich von Personen aus und ist dann auf Sachen verallgemeinert worden. [...] (Hervorhebung im Original, C.M.) (Kluge/Seebold 1999: 307)

Diese Definition findet sich auch bei Stathi (2010: 326), die es als Ausgangspunkt für die diachrone Betrachtung von *gehören* und der passivischen Konstruktion nahm, die im späteren Verlauf des Kapitels noch einmal aufgegriffen wird.

Szatmári (2002: 187) gibt in ihrem Artikel unterschiedliche Konstruktionen mit *gehören* an, die außer der Passivvariante in der Gegenwartssprache zu finden sind:

Variante 1: gehören<Sn, Edat>

(Ausdruck einer Haben-Relation: *Das Buch gehört mir. Dem Kind gehört ihre ganze Liebe.*)

Variante 2: gehören<Sn, Eprp>

[(a) Ausdruck des Erforderlichseins, der Notwendigkeit, der Zuordnung: *der Kranke gehört ins Bett, dazu gehört viel Mut, zu diesem Rock gehört eine elegante Bluse;*
(b) Ausdruck der Teil-von-Relation: *er gehört zu meinen Freunden, Gartenarbeit gehört nicht zu meinen Pflichten*]

Variante 3: gehören<Sn, Edir >

Wohin gehört dieses Buch?

Variante 4: gehören<Sn, Edat, Eprp>

Für deine Frechheit gehören dir ein paar Ohrfeigen!

Variante 5: sich gehören<Sn>

(Ausdruck eines Erfordernisses im Sinne von ‚sich (ge)ziemen‘, ‚sich schicken‘: *wie es sich gehört*) (Szatmári 2002: 187)

Für die hier vorliegende Untersuchung sind vor allem die Varianten 2, 4 und 5 von Interesse. Bei ersterer ist es die Bedeutung a) des ‚Erforderlichseins‘, das sich als deontische Modalität bzw. *obligation* auch bei Stathis (2010: 327) diachroner Betrachtung wiederfindet und hier auch ohne dem Perfektpartizip vorkommt (vgl. Szatmári 2002: 187).

Zu Punkt 4 und 5 sind folgende Überlegungen in den Sinn gekommen: Hier könnte man die Ellipse eines *gehören*-Passivs annehmen, beispielsweise mit dem Verb *verpassen* (*Für deine Frechheit gehören dir ein paar Ohrfeigen (verpasst)!.*¹¹ Umgekehrt wäre es auch denkbar, dass derartige Konstruktionen über die Pragmatik des Ausdruckes die Tür für das Passiv mit *gehören* geöffnet haben, einerseits über die Notwendigkeit, die durch *gehören* ausgedrückt wird, andererseits aber auch, dass der*die Sprechende sich in dieser (gewalttätigen) Handlung nicht als Agens hinstellen will (siehe Kapitel 5.2 Diskussion rund um die Semantik des Ausdrucks, auch im Zusammenhang mit der passivtypischen Agenstilgung).

Bei Variante 5 wird von Szatmári (2002: 187) selbst eine parallele Entwicklung des *gehören*-Passivs mit dem *gehören*-Reflexiv vermutet, da eine ähnliche Semantik vorliegt. Dies ist anhand des ‚Ausdruck[s] eines Erfordernisses‘ erkennbar.

Ähnliche Punkte finden sich auch im Duden, der ebenfalls die derzeitige Gegenwortsprache abzubilden versucht; hier werden teilweise auch ähnliche oder gar dieselben Beispiele angegeben, wenn auch etwas anders eingeteilt:

- 1.) jemandes Besitz, jemandes Eigentum sein
 - a. das Buch gehört mir
 - b. 〈in übertragener Bedeutung:〉 (dichterisch) dir will ich gehören (in Liebe verbunden sein)
 - c. 〈in übertragener Bedeutung:〉 (gehoben) ihr Herz gehört einem andern (sie liebt einen andern)
 - d. 〈in übertragener Bedeutung:〉 dieser Tag gehört der Familie (wird der Familie gewidmet)
- 2.) Glied oder Teil eines Ganzen sein, zu etwas zählen
 - a. er gehört schon ganz zu unserer Familie
 - b. dieser Wein gehört zur Spitzenklasse
- 3.) (an einer bestimmten Stelle passend) am Platze sein
 - a. das Fahrrad gehört nicht in die Wohnung
 - b. das gehört nicht hierher
 - c. die Kinder gehören um sieben Uhr ins Bett (sollten um sieben Uhr im Bett sein)
- 4.) für etwas erforderlich, Voraussetzung sein
 - a. es gehört viel Mut dazu, diese Aufgabe zu übernehmen
 - b. es gehört schon einiges dazu (man muss schon beherzt sein), so etwas zu wagen
 - c. dazu gehört nicht viel (sind keine besonderen Eigenschaften oder Fähigkeiten erforderlich)
- 5.) den Regeln des Anstands, den Normen der Sittlichkeit entsprechen, sich schicken: Grammatik **sich gehören**

¹¹ An dieser Stelle sei auf Fußnote 10 hingewiesen, in der eine ähnliche Beobachtung bei der Betrachtung eines möglichen *haben*-Passivs gemacht wurde.

- a. das gehört sich nicht [für dich]!
 - b. benimm dich, wie es sich gehört!
- 6.) für jemanden angebracht sein, jemandem gebühren: Gebrauch **landschaftlich, besonders süddeutsch, österreichisch**
- a. ihm gehört eine Ohrfeige
 - b. (meist in Verbindung mit einem 2. Partizip:) der gehört eingesperrt! (man sollte ihn einsperren)
- (Hervorhebungen im Original, C.M.) (Online-Duden: *gehören*)

Punkt 3, besonders c, zeigt Verbindungen zu Szatmáris (2002: 187) Punkt 2 des Erforderlich-Seins, wobei auffällig ist, dass diese mit lokativen Phrasen in Verbindung stehen (vgl. Online-Duden: *gehören*).

Punkt 4 knüpft ebenfalls an die Erforderlichkeit an, jedoch ohne jegliche lokativen Komplemente, wodurch die Konstruktion wie eine übertragen gemeinte Variante von Punkt 1 und 2 wirkt. Im Allgemeinen werden vor allem an diesem Punkt gewisse Unterschiede zwischen der semantischen Klassifikation im Online-Duden (vgl. Online-Duden: *gehören*) und bei Szatmári (2002: 187) deutlich.

Punkt 5 entfällt wiederum auf das bereits bekannte *gehören*-Reflexiv (vgl. Online-Duden: *gehören*).

Punkt 6 entspricht eben Punkt 4 bei Szatmári (2002: 187); hier wird ebenfalls die Verbindung zur Konstruktion mit Perfektpartizip gezogen, wenn auch hierbei ein neues Beispiel gewählt wird (vgl. Online-Duden: *gehören*).

Interessant ist an dieser Stelle auch die Anmerkung zum Gebrauch dieser Bedeutung, in der die Konstruktion in beiden Varianten (mit und ohne Perfektpartizip) als landschaftlich, mit regionaler Markierung für süddeutsch und österreichisch gekennzeichnet ist. Das Beispiel wird zudem als Aktivsatz paraphrasiert, anstatt *gehören* wird ein Modalverb im Konjunktiv angeführt (vgl. Online-Duden: *gehören*).¹²

Bevor es nun tiefer in die Untersuchungen zum *gehören*-Passiv hineingeht, wird auf die Diskussion der Frage „Ist es überhaupt ein Passiv?“ eingegangen. Sczcenepaniak (2011, 153) bezeichnet Bildungen mit *gehören* als passivähnlich. Auch in den Artikeln, die sich direkt mit dieser Konstruktion beschäftigen,¹³ wird die Bezeichnung von *gehören* + Perfektpartizip als tatsächliches Passiv diskutiert. Zu sehen ist dies bereits an Stathis (2010: 232)

¹² In weiterer Folge könnte man kontrastiv eine Reihe von *gehören*-Konstruktionen bezüglich ihrer Semantik in den Fokus nehmen, um hier tiefergehende Einblicke zu gewinnen; siehe Kapitel 5.4 zur Grammatikalisierung des *gehören*-Passivs.

¹³ Stathi (2009) präsentiert dieselbe Untersuchung zum *gehören*-Passiv, jedoch als „Case Study“ für „The Origin of Semantic Change in Discourse Tradition“.

Titel: „Is German *gehören* an auxiliary?“ und dem kurzen Exkurs bei Szatmári (2002: 179), in dem unterschiedliche Einordnungen verschiedener Forscher zu dieser Thematik präsentiert werden.

Szatmári (2002: 179) kommt zu dem Schluss, dass die Konstruktion mit *gehören* passivtypische Merkmale aufzeigt, jedoch die Modalität einen Problemfaktor darstellt, der die Akzeptanz als Passiv in der Forschungsliteratur in Frage stellt.

Unter dieser Prämisse werden für die Konstruktion mit *gehören* bereits seit den 70er-Jahren unterschiedliche Begrifflichkeiten verwendet, die sich thematisch in Passiv und Modalität einreihen. Es folgt eine nicht-extensive, aber umfassende Liste, um einen kurzen Überblick über die Bezeichnungen zu geben:

- *gehören*-Variante des Passivs (Reis 1976: 70)¹⁴
- Nebentyp des Passivs (Ágel 1996: 82)
- modalisierte Passivkonstruktion mit *gehören* als Auxiliar (Eroms 2000: 405–406)
- modal-passivisches *gehören* + Partizip II (Askedal 2005, 221)
- Modales Passiv (Vogel 2006: 3)
- *gehören*-Passiv (Engel 2009: 240)
- passivähnlich (Szczepaniak 2011: 153)
- modale Partizipialkonverse mit *gehören* (Zifonun/Hoffmann/Strecker 2011: 1792)¹⁵
- deontische Konstruktion mit der Halbkopula *gehören* (Gaeta 2018: 245)

Die Frage, die sich durch die angeschnittene Diskussion in der Forschungsliteratur ergibt, ist: Wie „passivisch“ muss eine Konstruktion sein, welche Aspekte muss sie (bzw. darf sie nicht) haben, um als Passiv gelten zu können, und ist der Modalfaktor der *gehören*-Konstruktion einer dieser „Dealbreaker“?

Laut Zifonun/Hoffmann/Strecker (2011: 1792) ist eine ihrer vorgeschlagenen Bedingungen des Passivs¹⁶ nicht erfüllt, da aufgrund der Modalität, die durch die Konstruktion mit *gehören* ausgedrückt wird, keine Entsprechung zum einfachen Aktivsatz besteht:

¹⁴ Bei Reis (1976) findet sich eine der ersten Erwähnungen des *gehören*-Passivs als solches – wenn nicht eventuell sogar die erste. Dies geschieht in einer Abhandlung über Hilfsverben und ihre Grammatikalisierung.

¹⁵ In „Grammatik der deutschen Sprache“ wird *gehören* mitsamt der ‚modalen Partizipialkonverse‘ dreimal mit Begrifflichkeiten aus dem Passivfeld und Modalität erwähnt: *Gehören* wird dabei zusätzlich zur Peripherie der Hilfsverben, die ein Passiv bilden können (vgl. Zifonun/Hoffmann/Strecker 2011: 52), und zur abgestuften Peripherie von Modalverben i.w.S. (vgl. Zifonun/Hoffmann/Strecker 2011: 1252) gezählt.

¹⁶ Die angesprochene Bedingung, die die Nummer (vi) trägt, lautet wie folgt:

„Sätze mit der nach [i] bis [v] aufgebauten periphrastischen Form und entsprechende Aktivsätze, die eine in Tempus und Modus übereinstimmende Form des Verbs enthalten, sind - von skopussensitiven Fällen abgesehen - in ihren

Und es war vor einiger Zeit im Gemeinderat eine heftige Debatte, ob der Münsterplatz endgültig gesperrt werden soll oder nicht. Nach meiner Meinung gehört er gesperrt. (YBA, 19)
(...) Nach meiner Meinung muß er gesperrt werden. (Zifonun/Hoffmann/Strecker 2011: 1792)

Auch bei Befürworter*innen der Bezeichnung als Passiv wird die Modalität der Konstruktion thematisiert. Ágel (1996, 82) schreibt davon, dass es als Nebentyp des *werden*-Passivs gesehen werden kann und durch die Wahl von *gehören* (anstatt eines anderen Auxiliars) eben der modale Aspekt hinzukommt.

Als Schlussfolgerung der kurz angeschnittenen Diskussion und unter der Prämisse, dass die Konstruktion in den Artikeln, die sich bereits eingehender mit ihr beschäftigen, als Passiv bezeichnet wird, wird in der hier vorliegenden Arbeit ebenfalls vom *gehören*-Passiv gesprochen. Im Kapitel 5, in dem die Ergebnisse der Untersuchung besprochen werden, wird nochmal vor dem Hintergrund der Grammatikalisierung als Passiv auf das Thema eingegangen.

Viele der Konzepte, die bei der Betrachtung des *gehören*-Passivs in diesem Kapitel von Relevanz sind, wurden bereits im vorangegangenen Kapitel (2.2) besprochen, jedoch sind dadurch noch nicht alle Begriffe abgedeckt, die in der Forschungsliteratur explizit mit dem *gehören*-Passiv in Verbindung gebracht werden.

Im Gegensatz zu den anderen üblicheren und auch „ungewöhnlicheren“ Passiv-Konstruktionen¹⁷ gibt es zu *gehören* augenscheinlich nur zwei (bzw. drei, siehe Fußnote 13) Artikel, die sich extensiver mit dieser spezifischen Form befassen, nämlich, wie bereits mehrmals erwähnt, Szatmári (2002) und Stathi (2009; 2010). Bevor nun auf die detaillierteren Ausführungen in dieser Forschungsliteratur eingegangen wird, wird ein konkreter Blick auf die Konstruktion geworfen.

Bei Stathi (2010: 324) wird das Beispiel mit einem entsprechenden *werden*-Passiv mit einem zusätzlichen Modalverb umschrieben, ein *müssen* oder *sollen*, das entweder im Indikativ oder Konjunktiv steht:

- a. Das Ladenschlussgesetz gehört abgeschafft.
- b. Das Ladenschlussgesetz soll/sollte/muss/müsste abgeschafft werden.
- c. Das Ladenschlussgesetz ist abzuschaffen. (Stathi 2010: 324)

Wahrheitsbedingungen gleich. Im Hinblick auf solche Sätze kann auch von zwei verschiedenen Diathesen gesprochen werden.“ (Zifonun/Hoffmann/Strecker 2011: 1790)

Laut ihnen ist also zwischen der *gehören*-Konverse und einem paraphrasierenden Aktivsatz keine direkte Entsprechung vorhanden, da *gehören* auf andere Weise ausgedrückt werden muss als im Passiv.

¹⁷ Mit Ausnahme des *gehen*- und *kommen*-Passivs bei Gaeta (2018); diese sind bei der Recherche zum Passiv nur in diesem Artikel vorgekommen und dadurch augenscheinlich noch weniger häufig beforscht als das *gehören*-Passiv.

Die letzte Entsprechung in c) ist eine ähnliche Konstruktion mit *sein* + Infinitiv, auf die nicht weiter eingegangen wird.

Szatmári (2002: 180–181) gibt in ihrem Artikel ebenfalls Beispiele für Umschreibungen, jedoch stellt sie den Passivformen in d) und e) einen Aktivsatz gegenüber.

d. Das ganze System gehört reformiert. (NEWS 26/94: 46)

e. Man muss/müsste/soll/sollte das ganze System reformieren.

f. Das ganze System muss/müsste/soll/sollte reformiert werden. (Szatmári 2002: 180–181)

Dieser wird mit dem unpersönlichen *man* gebildet, da im Passivsatz der Agens nicht explizit erwähnt wird. Es wird hier in den Umschreibungen in b) und e) gezeigt, wie das *gehören*-Passiv mit dem *werden*-Passiv plus einem Modalverb paraphrasiert wird.

Wie die Verben *sein*, *werden*, *bekommen/erhalten/kriegen*, *haben* durchläuft das Verb *gehören* einen Grammatikalisierungsprozess, der mit der Dekategorialisierung und Desemantisierung des Verbs einhergeht und dem „abgespaltenen“ *gehören* grammatische Funktionen zuweist. (Szatmári 2002: 182)

Mit dieser Vorannahme findet bei Szatmári (2002) von vornherein eine synchrone Betrachtung des *gehören*-Passivs statt, die auf die Perspektivierung im Zuge der Konstruktion (vgl. Szatmári 2002: 171–172) und daher beispielsweise auf die Semantik der Subjekte eingeht (vgl. Szatmári 2002: 182; siehe weiter unten), stellt Stathi (2010: 326–329) dem eine diachrone Betrachtung der möglichen Entstehung voran, die aus einem zuvor erschienenen Artikel hervorgeht, in dem *gehören* als Case Study zum Thema „The Origin of Semantic Change in Discourse Tradition“ behandelt wird (vgl. Stathi 2009: 291). Hier werden Beispiele ab dem 14. Jahrhundert angeführt, die *gehören* mit einer lokativen Präpositionalphrase zeigen:

die holz, wiesen, wasser und weiden mit gebrauchen [...], **gehoren aufs gericht**, sollen uf den rechten glockenschlag volgen. (14th century) (Hervorhebungen im Original, C.M.) (Stathi 2009: 293; 2010: 326)

Die Bedeutung von *gehören* mit lokativer Präpositionalphrase wird im DWB bereits wie folgt ausgezeichnet, nämlich dass die Mitgliedschaft bei einer Gruppe, in den belegten Beispielen meist eine damalige Rechtseinheit wie ein Gericht, zu diesem eine Verpflichtung ausdrückt, wobei die Mitgliedschaft dadurch obligatorisch wird (vgl. Stathi 2010: 327).¹⁸

¹⁸ DWB = Deutsches Wörterbuch von Wilhelm und Jakob Grimm, das umfangreichste Wörterbuch der deutschen Sprache, das sich auf historische Quellen bezieht, zwischen 1854 und 1961 in 32 Bänden erschien und von der Berlin-Brandenburgischen Akademie der Wissenschaften (BBAW) zum Teil

‘Membership in a group (assembly, court)’ > ‘obligatory membership in a group’ (Stathi 2010: 327)¹⁹

Stathi (2010, 327) greift diesen Punkt auf und führt aus, dass über diesen Kontext und die pragmatische Inferenz²⁰ aus diesem die weitere allgemeine Bedeutung der deontischen Modalität, also der ‚Verpflichtung‘ (*obligation*) und ‚Notwendigkeit‘ (*necessity*) herbeigeführt wird:

Summarizing, the historical data suggests that the verb developed the meaning ‘obligation’ and ‘necessity’ in the fourteenth century in a particular context, namely law and administration. This new sense arose from the ‘membership’ sense via pragmatic inferencing and is first attested with locative prepositional phrases. (Stathi 2010: 327)

Die Bedeutung des *gehören*-Verbs als Zugehörigkeit in manchen Bedeutungen entwickelte sich also zu einer Bedeutung der Obligation, die Erforderlichkeit in der Verwendung bloß mit lokativer Präpositionalphrase (vgl. Stathi 2009: 293; 2010: 327). Diese Entwicklung in einem bestimmten Kontext wird von Stathi (2009: 299) angelehnt an Moore (2006) als ‚genre-specific grammaticalization‘ bezeichnet.²¹ Sie zeigt allerdings auch Beispiele, die im DWB angeführt werden, die bereits im 16. Jahrhundert auch außerhalb expliziter Gesetzes- und Administrationstexte vorkommen (vgl. Stathi 2010: 327–328).

Die Verbindung von *gehören* mit Perfektpartizip scheint jünger zu sein (vgl. Stathi 2009: 293) Im DWB selbst wird die Konstruktion nicht angeführt, aber im Archiv finden sich zwei historische Beispiele. Stathi (2010: 328) zeigt die mögliche strukturelle Veränderung anhand der beiden folgenden Sätze, zwischen denen etwas mehr als 200 Jahre vergangen sind:

- a) Die Knäblein **gehören in eine besondere Kammer gelegt**: die Mägdlein auch in eine sondere. 1643 (Hervorhebung im Original, C.M.)
- b) Zwischen dem Gamsarchen und der Hochscharte **gehören** die Aichenköpfe **aufgeführt**. 1874 (Stathi 2010: 328) (Hervorhebung im Original, C.M.)

überarbeitet wurde (vgl. BBAW: DWB). Eine Version des Wörterbuchs ist online bei Wörterbuchnetz zu finden (siehe Wörterbuchnetz: DWB).

¹⁹ Stathi (2010, 327) führt diesen Punkt an, der eine andere Schlussfolgerung bei Heine/Kuteva (2002: 245) widerspiegelt, bei der ein POSSESSION-Schema der OBLIGATION von *gehören*, also dem Markieren einer deontischen Modalität, vorangestellt wird. Dies wird als ‚belong to‘ übersetzt, jedoch in einem Beispielsatz *Das Buch gehört mir*: im Sinne des Besitzes (vgl. Heine/Kuteva 2002: 245) und nicht der Mitgliedschaft (vgl. Stathi 2010: 327). Der Endpunkt der Grammatikalisierung scheint derselbe zu sein.

²⁰ Stathi (2010: 327) weist weiterführend zu dem Konzept des ‚pragmatic inferencing‘ auf Traugott (1988); Traugott/König (1991) hin, die dieses im Zuge der Grammatikalisierung besprechen.

²¹ Moore (2006; nach Stathi 2009: 299) zeigt diese Genre-spezifische Grammatikalisierung anhand eines lateinischen Ausdrucks *videlicet*, der in englischsprachigen „slander depositions“ als Diskursmarker grammatikalisiert wurde. Dieser Kontext und dieses Genre wirken sehr stark begrenzt, stärker begrenzt als die Umgebungen, in denen *gehören* vorkommt.

In a) wird das Partizip neben einer lokativen Präpositionalphrase angeführt, die ein Komplement zu demselben darstellt, während bei b) das Verb kein lokatives Komplement verlangt. Aus dem Kontrast dieser beiden Beispiele sieht man ebenfalls den Unterschied zwischen der Formulierung mit lokativer Präpositionalphrase und einem fakultativen Adverbial (vgl. Stathi 2010: 328).

Stathi (2010: 329) zieht aus ihren Beobachtungen der diachronen Belege Schlüsse, die sie auf den folgenden möglichen Grammatikalisierungspfad schließen lassen; dieser ist aber als Hypothese zu verstehen, da nur wenige historische Belege vorhanden sind, wie die oben angeführten (vgl. Stathi 2010: 328).

I. With locative prepositional phrase

- *Er gehört ins Gefängnis* ‚He should be in prison‘

II. With locative prepositional phrase + participle II (reinforcement)

- *Er gehört ins Gefängnis eingesperrt* ‚He should be shut in prison‘

III. With participle II (omission of PP, inferred from context or verb meaning)

- *Er gehört eingesperrt* ‚He should be imprisoned (shut)‘

IV. With participle II of transitive verbs (host class expansion to verbs that do not require a locative complement)

- *Die Bergspitze gehört erwähnt* ‚The mountain peak should be mentioned‘ (Stathi 2010: 329)

Interessant ist hierbei auch die gleichbleibende Bedeutung der Obligation, die an den entsprechenden Übersetzungen erkennbar ist, bei Veränderung der Verwendung der lokativen Präpositionalphrase, die zuerst ohne Perfektpartizip, dann gemeinsam mit dem Verbsuffix verwendet wird und schließlich weggelassen werden kann. In weiterer Folge wird die Konstruktion auch mit transitiven Verben verwendet, die kein derartiges Komplement nehmen (können), wie auch in dem oben angeführten Beispiel (vgl. Stathi 2010: 329).

Bei Szatmári (2002: 171) ist die Präpositionalphrase bereits im für den Titel gewählten *gehören*-Passiv-Beispiel zu finden: *Das gehört nicht vom Tisch gewischt*. (Unterstreichung, C.M.) Diese wird jedoch im Artikel nicht weiter betrachtet.

In der Untersuchung des *amc* wird ebenfalls ein Blick auf (lokative) Präpositionalphrasen im Zusammenhang mit dem *gehören*-Passiv geworfen (siehe Kapitel 3.3 zur Annotation des *gehören*-Passivs und 4.6.1 zur Untersuchung der lokativen Präpositionalphrasen).

Wie im vorhergehenden Kapitel zur Grammatikalisierung des Passivs spricht auch Stathi (2010: 329) von Desemantisierung (*desemantization*), also Bedeutungsverlust (*semantic*

bleaching) (siehe Kapitel 2.2 und Wegener (1998: 37, 43) nach Szatmári (2002: 183–184)) und Dekategorialisierung (siehe Hopper (1990), Askedal (2005)).

Um auf diese Aspekte weiter einzugehen, wird die Entwicklung von *gehören* anhand der sieben Stufen der Auxiliarisierung nach Heine (1993) besprochen, die die Entwicklung eines Vollverbes zum grammatikalischen Marker zeigen (vgl. Stathi 2010: 330), wobei aufgrund der Relevanz nur vier der insgesamt sieben Stufen vorgestellt werden.

Stufe A stellt die Ausgangssituation dar, ein Verb mit gefestigter lexikalischer Bedeutung und einem Komplement, das häufig auf ein konkretes Objekt oder auch Adverbial verweist. Laut Stathi (2010: 330) entspricht das der Bedeutung der Zugehörigkeit, die bereits in dieser Situation den semantischen Shift zur obligatorischen Mitgliedschaft vollzogen hat, mitsamt der lokativen Präpositionalphrase.

Stufe B als dynamische Situation, bereits hier werden nicht-finite Verbteile hinzugezogen, gemeinsam mit der bereits etablierten Konstruktion (vgl. Stathi 2010: 330).

In Stufe C befinden sich Verb und nicht-finites Verb in einem Prädikat und werden gleich flektiert, wodurch das Ursprungssitem zu einem ‚quasi-auxiliary‘ oder ‚semi-auxiliary‘ wird und keine eigenständige lexikalische Bedeutung mehr trägt (vgl. Stathi 2010: 330–331).

Als Semi- bzw. Halb-Auxiliar mit einem Perfektpartizip befindet sich *gehören* auf dem Weg der Grammatikalisierung, in der Dekategorialisierung, indem es nicht mehr die Ausgangsbedeutung ausdrückt, sowohl grammatisch als auch semantisch (vgl. Heine 1993: 61, nach Stathi 2010, 331).

In this sense, *gehören* expresses only one dimension of the TAM complex²², as predicted by the model. While *gehören* marks (TA)M distinctions,²³ the participle carries the semantic load of the predicate, i.e. the participle expresses lexical content and assigns semantic roles to the arguments. (Stathi 2010: 331)

Der Fokus liegt also auf der Modalität; aus diesem Grund wird ein expliziter Blick darauf geworfen:

Modalität als Konzept für die Sprachwissenschaft wurde aus der Philosophie entlehnt. Bereits Kant (1787/2003: 144, nach Schümann 2010: 180) beschreibt, dass dadurch keine „Inhalte“ formuliert, sondern die Notwendigkeit oder das nur mögliche Zutreffen derselben

²² Der sogenannte TAM-Komplex (oder auch TMA-Komplex) meint mehr oder minder das finite Verb, das im Prädikat für die Markierung der Kategorien Tempus, Aspekt und Modus zuständig ist, beziehungsweise steht der Begriff allgemein für dieses System, das im Deutschen häufig Überschneidungen aufweist (vgl. Heinold 2015: 11–13).

²³ Das Deutsche markiert typischerweise Aspekt nicht explizit und oft auch nicht verbal (siehe Heinold 2015: 60–61).

ausgedrückt werden. Aufbauend auf dieser (verhältnismäßig einfachen) Idee wurde Modalität in unterschiedlichen sprachwissenschaftlichen Werken weiterentwickelt und produktiv gemacht, einerseits durch die Einführung von Subkategorien (vgl. Wright 1951, nach Schümann 2010: 180), andererseits durch die Erhebung zur grammatischen Kategorie (vgl. Palmer 2001, nach Schümann 2010: 180).²⁴

Dies geht wiederum so weit, dass später auch Zweifel an der einheitlichen Darstellung des Begriffes der Modalität sowie seiner grammatischen Ausdrücke aufkommen (siehe Bybee/Perkins/Pagliuca 1994: 176, nach Schümann 2010: 180), bis hin zur Infragestellung der Kategorie im Allgemeinen (siehe Nuyt 2006, nach Schümann 2010: 180).

Eine der Unterkategorien ist die deontische Modalität, wobei *deon* gr. ‚das Nötige‘, ‚das Angemessene‘, ‚die Pflicht‘ bedeutet. Damit wird „eine objektive Notwendigkeit, eine Verpflichtung oder eine Erlaubnis bzw. ein Verbot ausgedrückt“ (Schümann 2010: 181).

Den Begriff der objektiven Modalität findet man auch bei Szczepaniak (2011: 165–166), die diese in Bezug auf die Modalverben beschreibt und folgendes Beispiel anführt: „*Der*

²⁴ Schümann (2010: 180) fügt an dieser Stelle hinzu, dass Modalität grammatisch nicht einheitlich ausgedrückt wird, „[i]m Gegensatz zu Tempus und Aspekt“. Es wird dabei nicht genauer ausgeführt, ob sich diese Aussage auf Sprachen im Allgemeinen oder auf das Deutsche im Speziellen bezieht, jedoch ist sie – meiner Ansicht nach – zumindest zu kurz gegriffen (wenn vielleicht auch aus manchen grammatiktheoretischen Perspektiven nicht falsch). Besonders in Bezug auf die deutsche Sprache, die – davon kann ausgegangen werden – in einer Grammatik zur deutschen Sprache betrachtet wird, hätten die Beispiele für „einheitlich grammatische“ Strukturen besser gewählt werden können. Tempus wird im Deutschen primär (zumindest aus schulgrammatischer, latein-strukturierter Sicht) durch das Verb ausgedrückt, jedoch zeigen beispielsweise Formen des Präsens (futuristisches Präsens: *Ab morgen gehe ich zur Schule*. Oder historisches Präsens: *1912 bricht der erste Weltkrieg aus.*), das Präteritum als allgemeine Erzählform (*Das kleine Mädchen ging in den Wald.*), das trotz des formalen Ausdrucks der Vergangenheit Aktualität ausdrückt, sowie die Variation zwischen Perfekt und Plusquamperfekt anstatt des Präteritums (Perfekt-Periphrase) nach regionalen Aspekten unterschiedliche Ausdrucksformen der Kategorie. Hier könnte man anhand der Präsens-Beispiele argumentieren, dass es sich bei den Tempus-Verhältnisse ausdrückenden Temporaladverbia bzw. -adverbiale und Präpositionen ebenfalls um einheitliche grammatische Strukturen handelt, allerdings wird der Begriff der Einheitlichkeit bereits aufgeweicht. Ähnliche Problematisierungen und weitere Ausführungen zu anderen Tempora finden sich bei Eisenberg (2020), im Kapitel 4.3 zum Tempus (Eisenberg 2020: 109–119) und Heinold (2015), besonders Kapitel 5.2.1 zum Präsens (Heinold 2015: 82–87).

Im Falle des Aspekts ist die Situation im Deutschen noch uneindeutiger, da Progressivität und Abgeschlossenheit keinesfalls einheitlich ausgedrückt und unterschiedliche Ausdrücke von Aspekt seit Jahren diskutiert werden (siehe beispielsweise Heinold (2015), die in ihrem Studienbuch neben Tempus auch Modus und Aspekt thematisiert). Während im Englischen zu fast jeder Tempusbildung eine progressive *-ing* Form existiert, gibt es bei der deutschen Verbkonjugation keine entsprechenden Strukturen. Beispiele für Progressivität wären: *Ich habe seit gestern nicht geschlafen.*, bei der sich die progressive Lesart aus der Kombination der Präpositionalphrase mit *seit* und dem Perfekt (das Perfekt zwischen Tempus und Aspekt siehe beispielsweise Heinold (2015: 102–111) ergibt, oder *Ich bin gerade am Telefonieren*, das aufgrund des nominalisierten Verbs in der Präpositionalphrase mit *am* sowie dem Adverb *gerade* Progressivität ausdrückt (*am*-Progressiv, siehe beispielsweise Szczepaniak (2011: 158–164); Heinold (2015: 60–61)). Die Frage, die sich an dieser Stelle wiederum stellt – ähnlich der Situation der passivischen Konstruktionen, die nicht als Passiv per se anerkannt werden, weil sie aus dem üblichen Schema fallen –, ist, ob die Notwendigkeit besteht, grammatisch bis ins letzte Detail einheitlich zu sein, um eine Kategorie zu bilden bzw. damit einzelne Items zu einer Kategorie zu gehören.

Junge muss den Abwasch machen.“ (Szczepaniak 2011: 165) Ein Umstand – auch als modale Quelle bezeichnet – führt dazu, dass das Subjekt verpflichtet ist, etwas zu tun. Dabei wurde diese Tätigkeit noch nicht umgesetzt und muss auch nicht eintreten; der Aussagesatz gilt somit als nicht-faktisch. Je nach Wahl des Modalverbs, nach Wahl des Ausdrucks, wird eben die Art der modalen Quelle festgelegt, ob es sich um einen Zwang, eine Möglichkeit, eine Empfehlung etc. handelt (vgl. Szczepaniak 2011, 165).²⁵

Die Modalität der Konstruktion spielt eben auch bei der expliziten Betrachtung des *gehören*-Passivs bei Szatmári (2002) und Stathi (2010) eine wichtige Rolle.

Ausgehend vom bereits in Kapitel 2.2 vorgestellten Passivfeld stellt Szatmári (2002: 178) ein Pendant mit Modalität bzw. – wie sie es ausdrückt – mit Modalfaktor gegenüber:

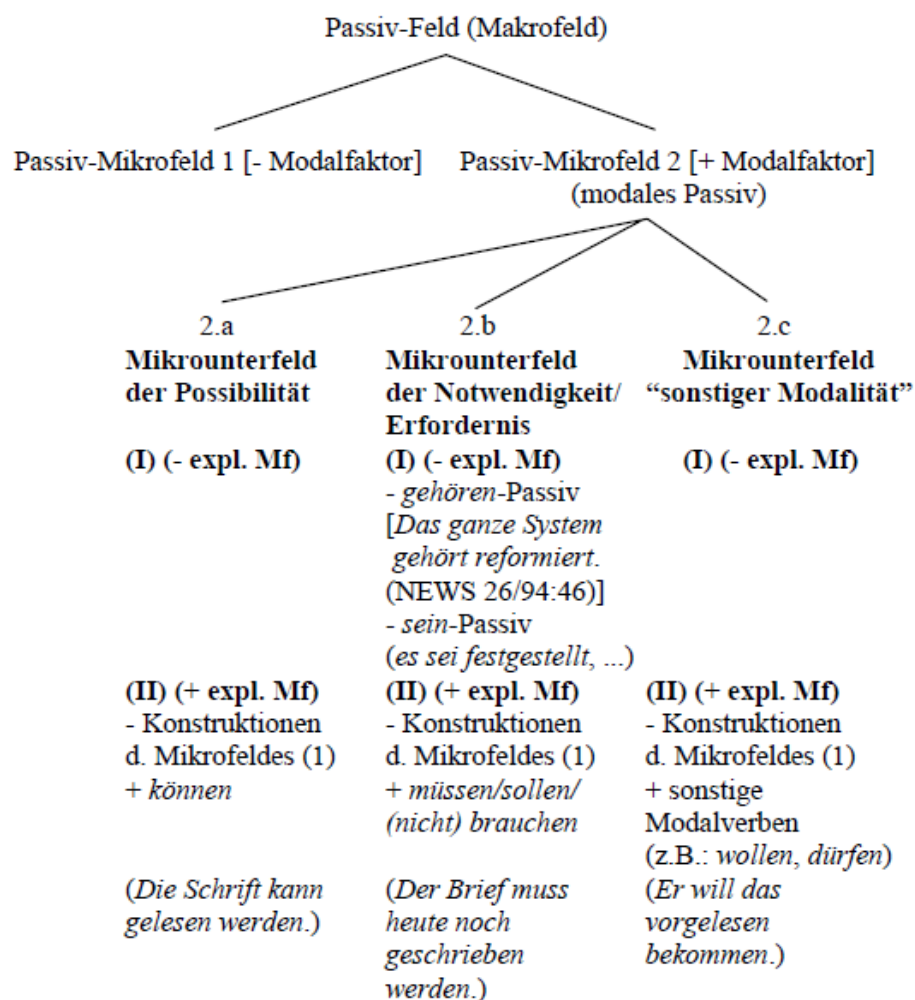


Abbildung 6: Passiv-Mikrofeld [+Modalfaktor] im Passiv-Feld (Grobgliederung) (Szatmári 2002: 178)

²⁵ Weiteres zum Konzept der modalen Quelle bei Diewald (1999).

Das *gehören*-Passiv wird hier 2.b dem Mikrounterfeld der Notwendigkeit/Erfordernis unterstellt, das keinen expliziten Modalfaktor hat (-expl. Mf). Die Modalität der Konstruktion wird also durch kein Modalverb ausgedrückt, wie es bei den anderen Beispielen mit Modalfaktor (+expl. Mf) der Fall ist (vgl. Szatmári 2002: 178).²⁶

Szatmári (2002: 181) geht in weiterer Folge weniger explizit auf die Grammatikalisierung der Modalität im Zuge der Bildung des Passivs ein, sondern bezieht sich hauptsächlich auf die Eigensemantik des (potenziellen) Auxiliars und die Entstehung einer Implikation von Modalität, die bei der Konstruktion mit *gehören* entsteht und eventuell sprachökonomischer ist als die Ausformulierung über das *werden*-Passiv mit einem Modalverb (vgl. Szatmári 2002: 181).

Das Konzept der deontischen bzw. objektiven Modalität wurde bei Stathi (2010) über die Verbindung zum „Verb-to-TAM“-Weg im Zuge der Grammatikalisierung dargelegt. Diese Etablierung als TAM-Marker ist dabei auch bei der Differenzierung der Stufen nach Heine (1993) relevant, denn dies entspricht Stufe C (vgl. Stathi 2010: 331).

Die letzte Stufe, die im Artikel angesprochen wird (die allerdings nicht die letzte mögliche Stufe im Zuge der Auxiliarisierung ist), ist die vierte, Stufe D, bei der das Verb nur mehr ein „defektives“ Paradigma im Gegensatz zu seinem Ursprung aufweist.

Bezüglich der Grammatik des Verbes beschäftigt sich Stathi (2010: 331–332) in ihren Ausführungen mit dem eben diesem, indem sie die Formen von *gehören* betrachtet, die in Verbindung mit der Gesamtkonstruktion vorkommen:

²⁶ Auffällig ist, dass bei den meisten der in der Graphik aufgezeigten Passivformen der Modalfaktor/die Modalität explizit ausgedrückt wird. Nur 2b zur Notwendigkeit, zu dem auch das *gehören*-Passiv gehört, ist nicht-explizit. Die andere Konstruktion ist ein *sein*-Passiv, bei dem die modale Komponente durch den Konjunktiv bei *sein* ausgedrückt wird (vgl. Szatmári 2002: 178).

Person / Number	Indicative Present	Occurrences in corpus	Subjunctive I	Occurrences in corpus
Sg 1	<i>gehöre</i>	2	<i>gehöre</i>	–
2	<i>gehörst</i>	6	<i>gehörst</i>	–
3	<i>gehört</i>	254	<i>gehöre</i>	121
Pl 1	<i>gehören</i>	–	<i>gehören</i>	–
2	<i>gehört</i>	4	<i>gehört</i>	–
3	<i>gehören</i>	205 ¹⁶	<i>gehören</i>	–
Person / Number	Indicative Simple Past	Occurrences in corpus	Subjunctive II	Occurrences in corpus
Sg 1	<i>gehörte</i>	–	<i>gehörte</i>	–
2	<i>gehörtest</i>	–	<i>gehörtest</i>	–
3	<i>gehörte</i>	3 (?)	<i>gehörte</i>	30
Pl 1	<i>gehörten</i>	–	<i>gehörten</i>	–
2	<i>gehörtet</i>	–	<i>gehörtet</i>	–
3	<i>gehörten</i>	3 (?)	<i>gehörten</i>	110

Abbildung 7: *gehören*-Formen im Korpus (Stathi 2010: 332)

Das defektive Paradigma wird vor allem damit begründet, dass Perfektformen, Futurkonstruktionen, Nominalisierungen und dergleichen grundsätzlich nicht gebildet werden oder als ungrammatisch gelten (vgl. Stathi 2010: 331) sowie gewissen Personen/Anzahl-Kombinationen nicht vorkommen (vgl. Stathi 2010: 332).

Stathi (2010: 331) legt in Fußnote 15 dar, dass es 738 Vorkommnisse von *gehören* + participle II gibt. Diese wurden für das obere Graphik annotiert. *Gehört* als Perfektpartizip ist daher bei Stathi (2010: 331–332) in dieser Aufzählung des Paradigmas ausgeblendet, da relevante analytische Tempusformen mit einem weiteren Auxiliar genau die umgekehrte Aneinanderreihung (*participle* II + *gehören* als VVPP) vornehmen würden.

Dafür wird eine weitere Anfrage aus participle II und *gehört* (Wortform) im Korpus gesucht, die 376 Hits produziert und insgesamt zehn Fälle des Konjunktiv II Plusquamperfekt zeigt (vgl. Stathi 2010, 332).

Es sind Informationen gehütet worden, von Mann zu Mann, die **ausgetauscht gehört hätten**. (Süddeutsche Zeitung 07.03.1994, p. 6) (Hervorhebungen im Original, C.M.) (Stathi 2010, 332)

Die Frage stellt sich, ob nur diese Form des periphrastischen Konjunktivs aufscheint oder ob andere Vergangenheitsformen mit analytischer Bildung zu finden sind. In Kapitel 3, 4 und 5 dieser Arbeit wird von den methodischen Überlegungen bis hin zur Diskussion der Ergebnisse über die Probleme gesprochen, das Phänomen in einem Korpus zu untersuchen; ein

Versuch, weitere Formen zu extrahieren, die über die Abfolge *gehören* und Partizip II hinausgehen, wird dabei angestellt.

Futurformen werden ebenfalls ausgeschlossen, was dadurch begründet wird, dass die Konstruktion eine Zukunftsgewandtheit innehat (vgl. Stathi 2010: 332).²⁷ Eine Formulierung wie „*Wenn sich das bewahrheitet, wird er eingesperrt gehören.*“ scheint möglich, also ist eine derartige Formulierung zwar unwahrscheinlich anzutreffen, jedoch nicht zwingend ungrammatisch.²⁸

Gehören allein macht nämlich kein Passiv, der andere Teil – nämlich das VVPP – ist ebenfalls von Interesse in den Untersuchungen: Sowohl bei Szatmári (2002: 180) als auch Stathi (2010: 335) wird ein Augenmerk auf die Semantik der Perfektpartizipia geworfen. Hier werden die neun häufigsten Verben angeführt:

Participles	Occurrences in corpus
<i>abgeschafft</i> ‘abolished’	88
<i>verboten</i> ‘forbidden’	55
<i>bestraft</i> ‘punished’	23
<i>erschossen</i> ‘shot’	10
<i>eingesperrt</i> ‘put in prison’	8
<i>geändert</i> ‘changed’	7
<i>abgeschoben</i> ‘deported’	6
<i>abgewählt</i> ‘voted out’	6
<i>beseitigt</i> ‘removed, eliminated’	6

Abbildung 8: Verben als Perfektpartizip in der Konstruktion (Stathi 2010: 335)

Des Weiteren wird eine Aufzählung aller „typischen Partizipia“ (Übersetzung C.M.) ohne genauere Darlegung der Belegzahlen angeführt (vgl. Stathi 2010: 335). Eine ähnliche Liste findet sich im vorangegangenen Artikel zum „semantic change“ ebenfalls, hier werden die

²⁷ Diese Bemerkung spielt in die Argumentation bei Heinold (2015: 12) hinein, dass innerhalb des TAM-System Überschneidungspunkte vorliegen. Durch die Modalität, die über *gehören* ausgedrückt wird, schwingt auch eine mögliche temporale Komponente mit, die in weiterer Folge für eben die Beschränkungen gesorgt haben könnte – bei Futur- und Perfektformen, die Stathi (2010: 330–331) hier eben auch anführt.

²⁸ Die gesamte Diskussion des defektiven Paradigmas wirft die Frage auf, inwiefern andere ähnliche Konstruktionen diese aufzeigen. Auxiliare wie *sein*, *haben* und *werden* sind voll grammatikalisiert und haben als solche zwar Restriktionen in den jeweiligen Konstruktionen, aber sind innerhalb der Paradigmata nicht defektiv (volle Rekategorialisierung). Das Dativ-Passiv mit *bekommen/kriegen/erhalten*, wie es beispielsweise bei Askedal (2005) besprochen wird, scheint beurteilbar anhand der angeführten Beispiele ebenfalls keine derart starken Beschränkungen zu haben (nur die zweite Person Singular und Plural kommen in den ausgewählten Belegen nicht vor, ebenso werden unterschiedliche Tempora gebildet sowie eine Infinitiv-Konstruktion) (vgl. Askedal 2005: 219–220, 221–222).

Verba allerdings nochmal genauer in semantische Unterkategorien eingeteilt (vgl. Stathi 2009: 296–297).

Bei Szatmári (2002: 180) handelt es sich um eine vollständige Aufzählung der Verben in ihrem betrachteten Kleinstkorpus, nämlich 18 an der Zahl.

abnehmen, abnehmen (= wegnehmen), abschaffen, abstellen,
auftauen, ausmisten, behandeln, bestrafen, entsorgen, feiern,
legalisieren, machen, reformieren, stoppen, trennen, überprüfen,
verfilmen, verhängen

Abbildung 9: Liste der im Kleinkorpus aufscheinenden Vollverben (Szatmári 2002: 180)

Auffällig ist, dass zwei Verben – *abschaffen* und *bestrafen* – in beiden Untersuchungen vorkommen und es sich dabei um zwei der Top 3 der häufigsten Verben bei Stathi (2010: 335) handelt.

Stathi (2010: 337) sieht die im Allgemeinen übergreifende Semantik der meisten Verben, mit denen ein *gehören*-Passiv gebildet wird:

A. *gehören* + [Participle II [PUBLIC AFFAIRS, PUBLIC OPINION]]

B. *gehören* + [Participle II [UNPLEASANT ACTION ON PATIENT (PUNISHMENT)]]

Das Ausdiskutieren in der Öffentlichkeit ist der grundlegende Aspekt, der laut Stathi (2010: 337) im Zuge des *gehören*-Passivs ausgedrückt wird (‘Public Affairs, Public Opinion’)²⁹, und es finden sich eben auch Beispiele, die dies ohne die ‘unpleasant action’ machen. Auf diese entfallen die anderen 27 Prozent der Verben (vgl. Stathi 2010: 336).

²⁹ Sie stützt sich dabei auf Eroms (2000: 405–406), der das *gehören*-Passiv kurz bei den modalisierten Passivkonstruktionen anführt:

(38) und (39) sind derzeit noch regional gebunden. Sie sind auch semantisch stark restringiert, als damit nicht nur eine einzige Modalisierung, nämlich durch *müssen*, erzielt wird, sondern dies auch nur für eine bestimmte Äußerungsweise gilt, für öffentlichkeitsbezogene Verhaltensweisen. Immerhin ist damit eine sehr kompakte, syntaktisch geregelte Konstruktion ermöglicht.

(38) Eine solche Äußerung gehört vom Gesetz verboten.

(39) Dies gehört sofort geändert.

Die Ableitung erfolgt analog den anderen Passivtypen. Das Verb *gehören* wird dabei als Auxiliar aufgefasst. (Eroms 2000: 405–406)

Die Punkte der starken Restriktion sowie der regionalen Gebundenheit sind hierbei hervorzuheben. Dass eine Modalisierung nur über *müssen* möglich ist, wird bei den Paraphrasierungen durch Szatmári (2002: 180–181) und Stathi (2010: 324) widerlegt (siehe oben), auch wenn die Hinzunahme von *sollen* und den jeweiligen Konjunktivformen diese nur aufweicht. Im Allgemeinen bleibt zu fragen, ob die Beschränkung auf diesen Modalitätskontext als Restriktion im semantischen Sinne einzuordnen ist oder ob es sich dabei schlichtweg um die grundlegende Bedeutung der Konstruktion handelt. Ebenso ist der einzige Äußerungskontext für ‘öffentlichkeitsbezogene Verhaltensweisen’ in Frage zu stellen (siehe Diskussion in Kapitel 5 und die zwei Beispiele, die auch bei Stathi (2010: 337) zu finden sind).

Laut ihren Untersuchungen ist die Semantik der Verben stark von der negativen Bedeutung für den Patiens geprägt, nämlich in ungefähr zwei Drittel der Belege (73%), wobei konkrete Beispiele sich auf die Verben *kontrollieren*, *überwachen*, *konfrontieren*, (*vor die Tür*) *setzen* und *wegjagen* beziehen und weitere Verben ohne ihren Kontext angeführt werden (vgl. Stathi 2009: 294–296; 2010: 335–336). Ihre Betrachtungen der Semantik der Konstruktion wird wie folgt auf den Punkt gebracht: „The common semantic core of most participles is the notion of punishment or criticism and destruction (according to the speaker).” (Stathi 2010: 337)

Im Anschluss an die Feststellung wird weiter auf das Konzept der *semantic prosody* nach Louw (1993)³⁰ eingegangen. Als abschließende Bemerkung wird dadurch *gehören* plus VVPP als „form-meaning pairing in the sense of Construction Grammar („Form-Bedeutungs-Paar im Sinne der Konstruktionsgrammatik“ (Übersetzung C.M.))“ bezeichnet (vgl. Stathi 2009: 298; 2010: 337–338).

Es werden ebenfalls zwei weitere Beispiele genannt, die die Bedeutung des Ausdiskutierens in der Öffentlichkeit nicht haben.³¹

³⁰ *Semantic prosody* als grundlegendes Konzept wird nicht in die Arbeit übernommen; es soll hier dennoch kurz erklärt werden, den Ausführungen in Stathi (2009) folgend: „A consistent aura of meaning with which a form is imbued by its collocates is referred to [...] as a semantic prosody. Semantic prosodies have been largely inaccessible to human intuition about language and they cannot be retrieved reliably through introspection.” (Louw 1993: 157)

Wie die Definition durch Louw (1993) selbst, der diesen Ausdruck geprägt hat, zeigt, ist das Konzept ein Versuch, Kollokationen in kontextueller Umgebung zu verbinden und dadurch Regularitäten in der Verwendung bestimmter Konstruktionen über die Semantik des umliegenden Kontextes („consistent aura of meaning“) zu erklären. Prosodie wird dabei aus der Phonetik übernommen, nämlich dass phonetische Eigenschaften auf die benachbarten Laute Einwirkung haben. Ein Wort und seine syntagmatische Umgebung beeinflussen sich also gegenseitig, wodurch manche Restriktionen entstehen können (vgl. Stathi 2009: 298).

Ein weiterer Begriff, der hier hinzukommt, ist die *semantic association*, die bei Hoey (2005) erklärt wird: „Semantic association exists when a word or word sequence is associated in the mind of a language user with a semantic set or class, some members of which are also collocates for that user.” (Hoey 2005: 24)

Er beschreibt ein ähnliches Prinzip, das die Konklusion mit sich zieht, dass das *gehören* mit Partizip II als Konstruktion im Sinne der Konstruktionsgrammatik (siehe Goldberg 1995; nach Stathi 2009: 298) beschrieben wird, die mit „negative-impact participles“ konventionalisiert ist (vgl. Stathi 2009: 298).

Ohne nun zu tief in die Thematik einzusteigen, soll auf die ausführliche Kritik des Konzeptes durch Whitsitt (2005) hingewiesen werden und ein Passiv-adjazentes Beispiel für einen der semantischen Assoziation widersprechenden Umstand gegeben werden, nämlich dass beim Rezipienten-Passiv entgegen der Benennung auch Verben damit vorkommen können, die explizit ausdrücken, dass dem Dativ-Objekt etwas genommen wird, wie eben stehlen, entzogen, ausschlagen, etc. (vgl. Askedal 2005: 223).

³¹ Dies demonstriert die Schwierigkeit, die zahlenmäßigen Verhältnisse genau zu beurteilen, da diese zwei Beispiele über die von Stathi (2010: 334, 336–337) als 100 Prozent geführten Belege für die Bedeutung der Public Opinion, Public Matters (27%) and der Unpleasant Action on the Patient (73%) (siehe auch Stathi 2009: 294) hinausgehen. Des Weiteren ist unklar, ob es sich hierbei um prozentuelle Angaben zu den einzelnen Verben oder zu den betrachteten Gesamtbelegen handelt, also ob es 73 Prozent der n=738 oder der einzelnen Partizipia sind.

Von diesen drei als möglich betrachteten Bedeutungen wird nun jeweils ein Beleg betrachtet, um einen Einblick in die konkrete Gestaltung vorzunehmen:

- a) „Solche Kollegen **gehören vor die Tür gesetzt**“, kommentierte zufrieden ein Gewerkschaftssekretär, als das Berliner Landesarbeitsgericht im Mai 1991 die Kündigung eines Metzgermeisters bestätigte. (die tageszeitung, 05.03.1992) (Hervorhebungen im Original, C.M.) (Stathi 2009: 295–296; 2010: 335–336)
- b) Auch offen verkaufte Eier, gerade offen verkaufte (sic) **gehören datiert!** (die tageszeitung, 27.04.1991) (Hervorhebungen im Original, C.M.) (Stathi 2009: 297; 2010: 336)
- c) Philosophische Probleme **gehören** demnach **naturalisiert** [...]. (Frankfurter Allgemeine, 08.03.2000, p. N6) (Hervorhebungen im Original, C.M.) (Stathi 2009: 299; 2010: 337)

Beleg a) zeigt die Semantik von B. *Jemanden vor die Tür setzen* drückt diese ‚unpleasant action on the patient‘ aus, in diesem Fall sind es *solche Kollegen*. Auffällig ist auch, dass dieses Beispiel von den dreien dasjenige mit einem belebten Subjekt ist.

Datieren in Beispiel b) per se scheint kein Verb zu sein, dass nur in öffentlich wirksamen Kontexten verwendet werden kann; diese Bedeutung kommt durch die Verwendung mit *gehören* und dem Bezug auf ‚offen verkaufte Eier‘ zustande (und dadurch eine Verknüpfung zu einem Kontext geschaffen wird, der für die Öffentlichkeit potenziell von Interesse ist).

Ein Beispiel wie in c) zeigt ebenfalls, wie schwierig die Bewertung der Relevanz für die Öffentlichkeit ist, die entweder durch weiteren Kontext oder durch den grundsätzlichen Umstand, dass dieses Gespräch in einer öffentlichkeitswirksamen Tageszeitung – der Frankfurter Allgemeinen in diesem Fall – geführt wird, bestimmt werden könnte. Zur Quellenreflexion und Bedeutung der ‚Public Matters‘ bzw. Öffentlichkeitsrelevanz siehe auch Diskussion in Kapitel 5.2 und 5.3.

Ein Aspekt, der bei Szatmári (2002: 179–180) beleuchtet wird, jedoch bei Stathi (2009; 2010) keine Beachtung findet, ist die Semantik des Patiens auf der Subjektposition, auf das sich das *gehören*-Passiv bezieht, also mit dem ‚etwas gemacht gehört‘. Dies ist augenscheinlich der Theorie der Perspektivierung geschuldet, die ein wichtiges Thema bei Szatmári (2002: 172–174) darstellt. Sie bezieht sich nämlich auf die Veränderung der Rollen im Zuge der Passivierung, was ihren Fokus auf die Patiensrolle im Subjekt lenkt.

Dies ist ein anderer Forschungsrahmen als Stathi (2010: 330–331), die die Passiv-Grammatikalisierung, wie oben bereits geschildert, über ein allgemeines Konzept der Auxiliarisierung schildert, die sich auf das Verb und seine Komplemente bezieht und nicht genauer auf die Rollen eingeht.

Die Beispiele, die Szatmári (2002: 179–180) genauer untersucht, weisen Patiens-Lexeme auf, die sie in die Merkmale [abstrakt] und [human] einteilt. [Abstrakt] scheint hierbei auf

konkrete oder abstrakte Dinge zu referieren (,diese Speise‘ gegenüber ,die Ungerechtigkeit‘), [human] auf tierische oder menschliche Lebewesen (,die Adler‘ gegenüber ,Unruhestifter‘) (vgl. Szatmári 2002: 182).

[± abstrakt]	[± human]
das, Dummheit, etwas, dieses Gesetz, Liebe, der Prozess, das „Spiel im ...“, diese Speise, das ... System, die Theorien ..., die Ungerechtigkeit,	die Adler, Dr. Thoma ein unbescholtener Bürger, Unruhestifter,

Abbildung 10: Semantische Merkmale der Patiens-Lexeme des Kleinkorpus (Szatmári 2002: 182)

Auffällig ist das einzelne Vorkommen der tierischen Lebewesen, das in folgendem Beleg zu finden ist:

Dem Mann gehören die Adler abgenommen und ein Tierhalteverbot verhängt, dann wird für alle Zeit Ruhe sein. (Die ganze Woche 10/96: 3) (Szatmári 2002: 180)

Die Adler als Patiens des Aktiv-Satzes ist durch die Kennzeichnung mit Nominativ als Subjekt des Passiv-Satzes erkennbar; im Vorfeld steht jedoch das Dativ-Objekt *dem Mann*, der thematisch hervorgehoben wird.³²

Aufgrund der zahlenmäßigen Verhältnisse – noch einmal sei betont: eines kleinen Korpus – wird der Schluss gezogen, dass die Konstruktion derart gelesen wird,

wie man in einer bestimmten Situation eine bestimmte Angelegenheit erledigen soll (sollte) bzw. muss (müsste). Eine solche Entscheidung ist „Unbelebtem“ gegenüber weitaus einfacher zu treffen als belebten Wesen gegenüber, die selbst über einen „Willen“, d.h. über die Fähigkeit zu entscheiden, verfügen. (Szatmári 2002: 182)

Dies wirft einige Fragen auf, nämlich ob die Passiv-Perspektivierung primär einen Fokus auf das Patiens darstellt und nicht eher einen mehr oder weniger dedizierten Nicht-Fokus auf das Agens, unabhängig vom „Willen“ des Sachverhalts, des Dinges oder des Lebewesens, mit dem ,etwas gemacht gehört‘.

³² Bei dem vorliegenden Beispiel ist eine Paraphrase durch das Dativ-Passiv möglich, bei dem allerdings die durch das *gehören*-Passiv ausgedrückte Modalität nicht ausgedrückt wird. Diese müsste durch ein hinzugefügtes Modalverb expliziert werden: *Der Mann bekommt die Adler weggenommen*. Aktiv-Satz: *Sie nehmen dem Mann die Adler weg/ Sie soll(t)en dem Mann die Adler wegnehmen*. Dies wird durch die dreiwertige Valenz des Verbes *abnehmen* ermöglicht, welche bei den Dativ-Passiven gefragt ist (vgl. Askedal 2005: 223). Dies hebt ebenfalls die Schwierigkeit hervor, mit der die Restriktionen auf das *gehören*-Passiv ermittelt werden, denn es ist unklar, warum diese Konstruktion auch in der mit *gehören* modalisierten Variante funktioniert.

In diese Annahme spielt auch hinein, dass in manchen Werken der Forschungsliteratur festgehalten wird, dass das *gehören*-Passiv auch eine unpersönliche Form hat – in anderen jedoch die Existenz desselben verneint wird. Vogel (2006: 3) nimmt die Konstruktion aus *gehören* und Perfektpartizip eben aus diesem Grund nicht in die Untersuchung der unpersönlichen Passiv-Formen auf. Diese Annahme wird allerdings später in einem Artikel zum Passiv revidiert bzw. relativiert: Die unpersönliche Form des Passivs mit *gehören* (und *bleiben*) ist ‚eventuell‘ möglich, aber durchaus unüblich, da kaum belegt (vgl. Vogel 2010: 229).

In Kapitel 2.1 wurde bereits die Erwähnung der Konstruktion im Grammatik-Duden dargestellt, denn dort ist von einer unpersönlichen Variante des *gehören*-Passivs (vgl. Duden 2016: 562). Dadurch ist sie zumindest am Rande kodifiziert.

„Dir gehört geholfen.“ (Duden 2016: 562). Das Vorfeld muss dennoch durch einen Satzteil gefüllt werden, weshalb im Beispiel aus dem Duden das Dativ-Pronomen *dir* hineingezogen wird (vgl. Duden 2016: 562).

Bei Szatmári (2002: 182) wird ein Beispiel aus Ágel (1996) angeführt, bei dem der Satz mit *gehören*-Passiv kein Subjekt besitzt:

Hier gehört gründlich ausgemistet. (Ágel 1996: 82)

- a) Man mistet den Stall aus. – Der Stall wird ausgemistet.
- b) Der Stall gehört ausgemistet.
- c) Man mistet hier aus. – Hier wird ausgemistet.
- d) Hier gehört ausgemistet. (Szatmári 2002: 182)

Alternativ zu d) wäre auch eine Variante mit *es* als formales Subjekt denkbar: *Es gehört ausgemistet*. Hier wird das transitive Verb *ausmisten* (mit Patiens *der Stall*) intransitiv verwendet, das die Bildung des subjektlosen Passivs in diesem Fall möglich macht (vgl. Szatmári 2002: 182).

Dieser Umstand findet auch Eingang in Szatmáris (2002: 189) Konklusion, in der sie das Passiv-Mikrofeld 2 mit Modalfaktor auf das *gehören*-Passiv bezogen ein weiteres Mal darstellt.

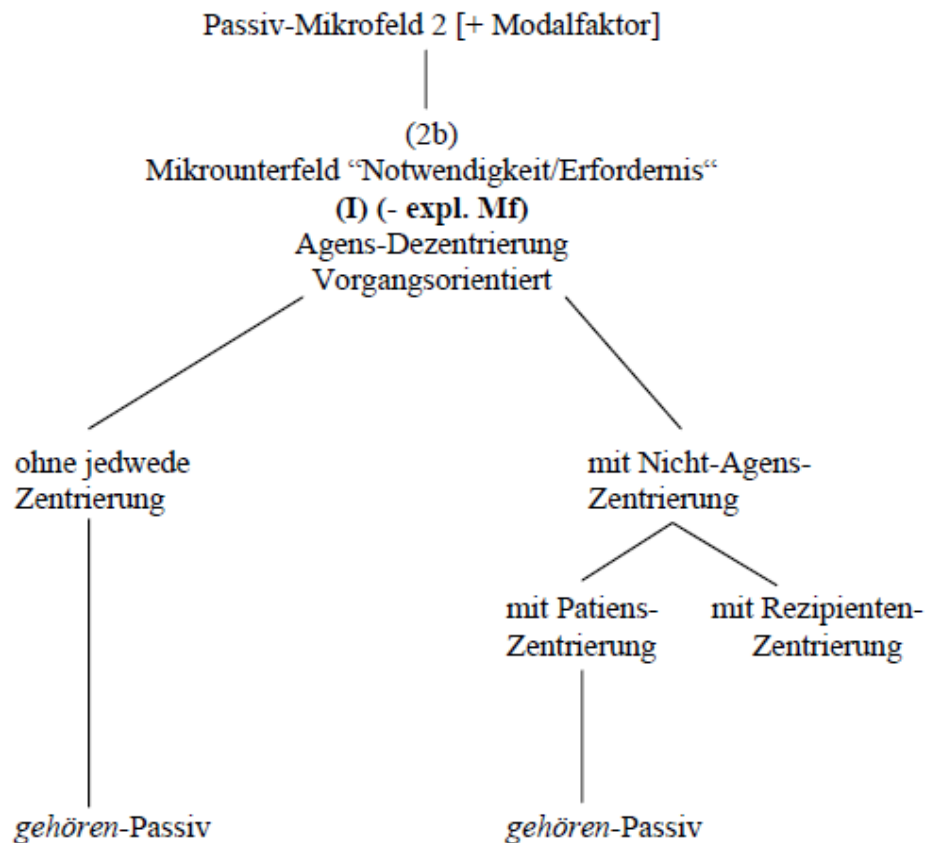


Abbildung 11: Einordnung des *gehören*-Passivs ins Passiv-Mikrofeld 2 [+Modalfaktor] (Szatmári 2002: 189)

Auffällig ist hierbei, dass die *gehören*-Konstruktion in dieselben Zweige des Baumdiagramms eingefügt wird, wie das *werden*-Passiv in der Darstellung des Passiv-Mikrofelds 1. Beide sind vorgangsorientiert, einmal mit Patiens-Zentrierung und einmal ohne Zentrierung (also unpersönliches Passiv) (vgl. Szatmári 2002: 175, 189).

Ein ähnlicher Aspekt wird in Engels (2009) Deutsche Grammatik angesprochen, nämlich dass *gehören* zu allen Verben gebildet werden kann, die ein *werden*-Passiv erlauben (vgl. Engel 2009: 240).

Grundsätzlich werden durch die Darlegungen in der Forschungsliteratur so gut wie ausschließlich transitive Verben erwartet; eventuelle Ausnahmen in Form der intransitiven Verwendung könnten bei unpersönlichen Passivformen auftauchen (weitere Einblicke hierzu in Kapitel 4.4 *Transitivität der Verben* und 4.5 *Subjekte*).

Das *gehören*-Passiv wird nicht in der gesamten Forschungsliteratur tatsächlich als Passiv gesehen, jedoch grundlegend im Passivfeld verortet. Besonders die Artikel von Szatmári (2002) und Stathi (2009; 2010) sehen an der passivischen Verwendung von *gehören* keine Zweifel, betonen aber die wichtige modale Komponente, die durch das Verb als Auxiliar selbst ausgedrückt wird.

Die aufgeworfenen Aspekte ihrer Untersuchungen fließen in der hier vorliegenden Arbeit in die Betrachtung der Daten – besonders in den qualitativen Teil – ein, teilweise auch in verbindender Weise, zum Beispiel wenn es um die Frage geht: Sind es bestimmte Patiens-Subjekte, denen hier die negativ-bestrafenden Bedeutungen zukommen? Weiteres dazu folgt in Kapitel 3.3 zur Annotation des *gehören*-Passiv-Belege.

3. Methode

Die Untersuchung des *gehören*-Verbs, insbesondere der Passiv-Form, wird über eine Analyse des amc (Austria Media Corpus) gestaltet.

Durch Datenextraktion wird aus den Belegen des *gehören*-Passivs ein eigenes kleines Korpus erstellt, dass Aufschlüsse darüber geben soll, wie sich die Verwendung der Konstruktion in der österreichischen Medienlandschaft verhält.

Die Schwierigkeit hierbei ist, dass sich das Phänomen an der Schnittstelle zwischen Morphologie und Syntax befindet und aufgrund der deutschen Wortstellung des finiten und infiniten Teils von analytischen Verbformen das Herausfiltern valider Belege eine Herausforderung darstellt.

Aus diesem Grund wird hier ausführlicher zu den getesteten Methoden, den im Zuge der Datenextraktion auftretenden Herausforderungen und Problemen sowie der gewählten Vorgehensweise berichtet.

3.1 Allgemeines zum Austrian Media Corpus (amc)

Als Grundlage für die Datenbasis dient das Austrian Media Corpus (amc), das die österreichische Medienlandschaft verzeichnet und für sprachwissenschaftliche Zwecke zur Verfügung stellt (vgl. amc: Das amc!). Die hier verwendete Version des amc – 4.2 – stammt vom Jänner 2023 und umfasst über 12 Milliarden Tokens (vgl. amc: Versionsarchiv 4.2).³³

Die Erstellung und Betreuung des amc findet durch das ACDH-CH an der ÖAW statt, in Zusammenarbeit mit der APA, die für die Bereitstellung der Inhalte zuständig ist (vgl. Ransmayr/Mörth/Đurčo 2017: 28).

³³ Inzwischen gibt es eine neue Version des amc, 4.3, das die Daten des letzten Jahres miteinbezieht (vgl. amc: amc in Zahlen). Die Analysen, die hier besprochen werden, entstanden jedoch bereits vor dessen Erscheinen.

Zur Abfrage von Items wird die sogenannte „NoSketch-Engine“-Software verwendet, die „Sprache“, in der diese formuliert werden, nennt man Corpus Query Language (CQL), also wortwörtlich ‚Korpus-Abfrage-Sprache‘.

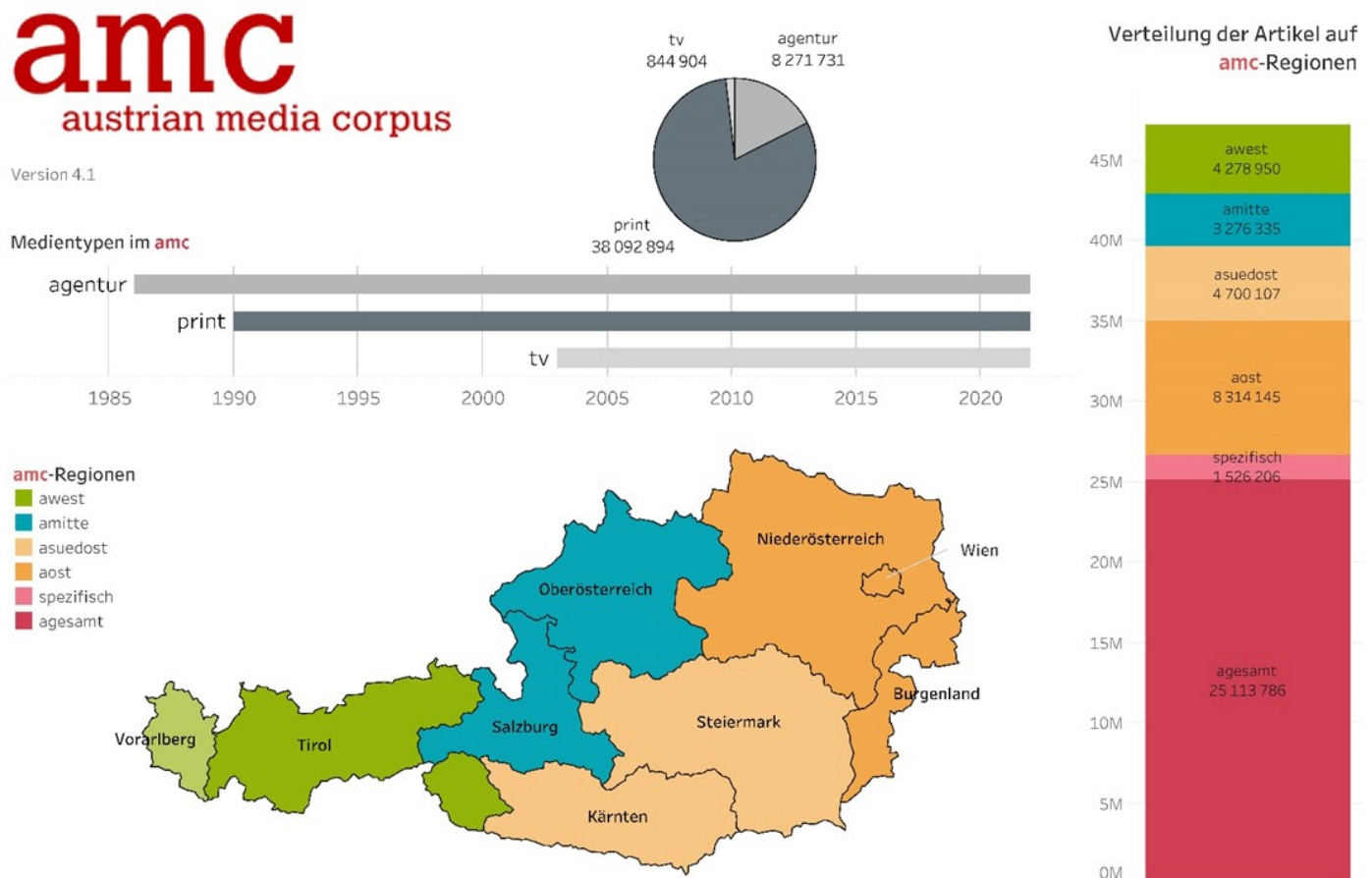


Abbildung 12: Grafiken zum amc (Dorn et al. 2023: 45); CC-BY Theresa Ziegler

Den Großteil der Dokumente im Korpus machen die Texte der österreichischen Printmedienlandschaft aus, dazu gehören sowohl täglich und wöchentlich erscheinende Zeitungen als auch Magazine und thematisch spezifischere Zeitschriften. Zudem sind die Pressemitteilungen und -aussendungen durch die APA (Kürzel APA und OTS³⁴) inbegriffen. Auch die TV- und Radiolandschaft Österreichs wird in kleinem Maße aufgegriffen, Nachrichtensendungen österreichischer Medien sind als Transkripte verfügbar (vgl. Dorn et al. 2023: 44).

Die einzelnen Medien sind dabei den Regions-Metadaten des amc zugeteilt, die sich mit den Kategorien des VWB decken (vgl. Dorn et al. 2023: 44).

³⁴ OTS: Originaltext-Service.

doc.region	doc.docsrc	Notiz
aost	FaltER, BVZ, HEUTE, NOEN	Wien, Niederösterreich, Burgenland
amitte	OOEN, NVB, SBGW, SVZ	Oberösterreich, Salzburg
asuedost	KLEINE, KTZ, GRAZER, KTNMONAT, STMONAT	Kärnten, Steiermark
awest	NVT, OBERRUND, TT, TTKOMP, VN	Tirol, Vorarlberg
agesamt	APA, DATUM, FORMATDB, FURCHE, KRONE, KURIER, MEDIANET, NEWS, OEREICHE, OTS, PRESSE, PROFIL, SN, STANDARD, WZ, WIBLATT, ATVVOLL, PRO7VOLL, PULSVOLL, SAT1VOLL	Medien, die regional nicht zuzuordnen sind
spezifisch	ACADEMIA, ARBEITW, AUGUSTIN, BAUERNZT, DIEWIR, ECHO, EMEDIA, GEWINN, HOR, IM, KONSUM, KW, SOLI, SPORTZTG, TREND, TVMEDIA, WIENER, WIENERIN, WOMAN, MWVOLL	Medien, die regional nicht zuzuordenbar und/oder thematisch sehr spezifisch sind

Abbildung 13: Zuteilung der einzelnen Medien zu den Regionskategorien (vgl. amc: Korpusinhalt-Attribute #doc.region)

Zu **agesamt** werden überregionale Zeitungen wie Kronen Zeitung, Kurier, Die Presse oder Der Standard gezählt, aber auch die TV-Sender ATV, Pro7, Puls4, Sat1. Überraschenderweise finden sich auch die Salzburger Nachrichten (SN) und die Wiener Zeitung (WZ) in dieser Kategorie wieder; dies ist vermutlich dadurch zu erklären, dass diese auch außerhalb der über den Namen suggerierten Region erhältlich sind.

Der einzige TV- bzw. Radio-Sender, der zu **spezifisch** gezählt wird, ist der ORF unter dem Kürzel MWVOLL.³⁵

Im Allgemeinen haben **agesamt** und **spezifisch** die meisten Medien unter sich versammelt, die einzelnen kleineren Regionen nur jeweils vier bis fünf Zeitschriften (von täglichem bis monatlichem Erscheinungszyklus).³⁶

Seit dem Erscheinen des Artikels von Ransmayr/Mörth/Đurčo (2017) zum amc sind einige Forschungsarbeiten mithilfe des Korpus publiziert worden. Diese sind auf der Seite des amc aufgelistet (siehe amc: Publikationen). Bei einem Blick auf die Titel fällt auf, dass die meisten von ihnen auf lexikalischer oder morphologischer Ebene forschen.

³⁵ Zur Aufschlüsselung der Medien nach ihren Kürzeln siehe amc: amc_4.2 Medienliste und Tabelle in Anhang B.

³⁶ Die Metadaten zu den Regionen werden der Vollständigkeit halber betrachtet, grobe Unterschiede bei den Regionen **awest**, **amitte**, **asüdost**, **aost** werden nicht erwartet. Interessanter sind wohl die übergreifenden Regionen **agesamt** und **spezifisch**.

In Dorn et al. (2023: 46–47) werden als wichtige Forschungsprojekte und -werke, die mit den Daten arbeiten, „(flexions)morphologische Arbeiten“ (siehe Korecky-Kröll 2020; Ziegler 2021) angegeben sowie der Fokus auf die lexikalischen Ebene (beispielsweise in Koppensteiner 2015; Lenz/Dorn/Ziegler 2021) betont. Außerhalb dieser Bereiche wird noch ein attitudinal-perzeptives Forschungsdesideratum angegeben (siehe Höll 2020).

Die Beispiele, die bei Dorn et al. (2023: 47–53) genauer dargestellt werden, sind wie bei Ransmayr/Mörth/Đurčo (2017) auf Orthographie- und Wortbildungswandel und auf morphologisch kompakte Wörter und Phrasen beschränkt, die sich durch RegEx-Abfragen in Bezug auf ein Token (im Falle von *Acht geben* auf zwei: `[word="(Ala)chtgeben"]` bzw. `[word="(Ala)cht"] [word="geben"]`)³⁷ extrahieren lassen.

Der Hinweis auf diese Forschungsarbeiten soll nicht zeigen, dass es nicht möglich ist, auf morphologisch-syntaktischer Ebene, über die Wortgrenze hinaus, zu arbeiten, sondern um deutlich zu machen, dass derartige Untersuchungen der Recherche nach noch nicht stattgefunden haben – vor allem auf disjunkte Kollokationen bezogen (mehrteiliges Prädikat im Vergleich zu adnominaler Attribution) –, weshalb in Bezug auf das Korpus nicht auf Erfahrungswerte zurückgegriffen werden konnte, um ein Phänomen in diesem Bereich zu erforschen.

In den nächsten Kapiteln wird die Vorgehensweise bei der Extraktion der Daten und der Zusammenstellung des Korpus bei der hier vorliegenden Arbeit dargelegt sowie auf die Herausforderungen und Probleme, die dabei entstanden sind, eingegangen.

3.2 Vorgehen bei der Zusammenstellung des Korpus

3.2.1 Vorüberlegungen

Die größte Herausforderung der vorliegenden Arbeit ist die Filterung der Belege des *gehören*-Passivs aus dem amc. Ein beträchtlicher Teil der Arbeitszeit ist eben in die Suche nach Lösungen für dieses Problem geflossen, unter anderem weil die eigene Expertise in der quantitativen Korpuslinguistik noch eingeschränkt war, aber auch weil bislang keine derartigen Themen im Umfeld des amc beforscht wurden, wodurch auf keine bereits vorhandenen Optionen zurückgegriffen werden konnte.

³⁷ CQL-Abfragen werden zur deutlichen Kennzeichnung und besseren Leserlichkeit durch Änderung der Schriftart von Fließtext und Zitaten abgehoben.

Wie im vorherigen Kapitel bereits beschrieben, arbeitet das amc mit der NoSketch-Engine und Abfragen sind über eine bestimmte CQL möglich. Dies war der erste Schritt, um eine Übersicht über das Verb *gehören* zu bekommen.³⁸

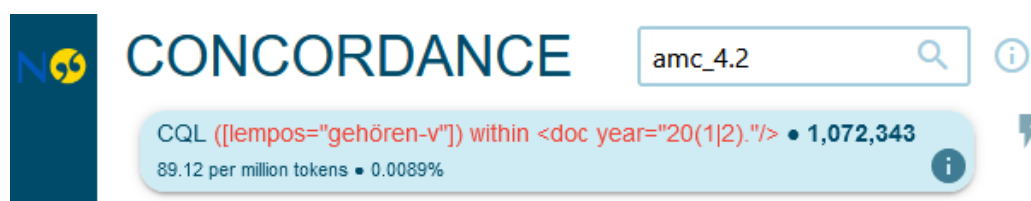


Abbildung 14: Abfrage des *gehören*-Verbs über die NoSketch-Engine des amc

Um die Anzahl der Belegzahlen im Vorhinein grob einzuschränken und die Datenmengen im Rahmen dieser Arbeit bearbeitbar zu halten, wurde beschlossen, bloß Treffer nach 2010 miteinzubeziehen. Aus diesem Grund sind alle CQL-Abfragen mit dem Zusatz `within <doc year="20(1|2)."/>` versehen, der die Filterung nach Jahr vornimmt. Das schränkt bereits die Treffer für das Verb *gehören* auf etwas über eine Million Hits in der Concordance ein (siehe Abbildung 14).

Eine der frühen Überlegungen ging in die Richtung, dass im Zuge der Grammatikalisierung vorkommende *gehören*-Konstruktionen im Gesamtkorpus des Verbs untersucht werden sollen, also die unterschiedlichen Bildungen, die mit dem Verb üblich sind. Davon wurde später Abstand genommen, um den Fokus auf dem *gehören*-Passiv allein zu belassen.

Die erste Version der CQL-Abfrage, um das *gehören*-Passiv zu extrahieren, lautete:

```
([lempos="gehören-v"][]* [posTT="VVPP"] | [posTT="VVPP"] []*[lempos="gehören-v"] ) within <s/> within <doc year="20(1|2)."/>
```

Die wichtigen Teile der Abfrage sind eben einerseits `lempos = „gehören-v“`, das die Verben filtert, die dem Lemma *gehören* zugeschrieben werden, um eben möglichst keine Formen von *hören* sowie vom Substantiv *Gehör* in den Ergebnissen zu haben.

Der andere Teil mit dem Tag `VVPP`, das nach dem Stuttgart-TübingenTag-Set (STTS) annotiert ist, bezieht sich auf alle Vollerben, die Perfektpartizipia sind, unabhängig von ihrer semantischen Bedeutung (vgl. STTS 1999: 7).

Dieses Vorgehen unterscheidet sich von den im vorherigen Unterkapitel vorgestellten vor allem dadurch, dass primär nicht nach RegEx, sondern nach PoS-Tags und Lemmatisierung gefiltert wird. Die Anordnung der einzelnen CQL-Parts folgt wiederum einem RegEx-Schema.

³⁸ Zugriff auf das Korpus: Austria Media Corpus (amc), Version 4.2, zugänglich über <http://hdl.handle.net/21.11115/0000-000C-4F08-4>, zuletzt abgerufen am 15.12.2023.

Die Doppelung der beiden Teilcodes stammt daher, dass damit zu rechnen ist, das VVPP in der Satzreihenfolge sowohl nach als auch vor *gehören* anzutreffen; etwa wenn das *gehören*-Passiv in einem Nebensatz steht und dadurch das finite *gehören* am Ende der rechten Verbklammer steht. Aus diesem Grund wird nach dem VVPP sowohl vor als auch nach *gehören* gesucht.

Diese Abfrage lieferte jedoch über 270.000 Belege, die nur bedingt das *gehören*-Passiv darstellen, da die Partizipia auch zu anderen Konstruktionen im Satz gehören könnten (Beispiele hierzu werden in Kapitel 3.2.2 präsentiert). Im Anschluss wurde deshalb nach Möglichkeiten gesucht, die Daten aufzubereiten, um ein möglichst zielführendes, kohärentes Korpus zu erhalten.

Durch die Zuhilfenahme des spaCy-Packages, das in der Programmiersprache Python funktioniert, habe ich einen Code geschrieben, mit dessen Hilfe das *gehören*-Passiv extrahiert werden konnte. Diese Vorgehensweise wird dem Natural Language Processing, kurz NLP, zugeordnet.

Python ist eine gängige Programmiersprache, die vor allem in den Data Sciences für unterschiedlichste Forschungen verwendet wird und im Allgemeinen viele Teilbereiche der Bearbeitung von Daten abdeckt, unter anderem durch verschiedenste sogenannte Packages oder Libraries, die vorgefertigte Sets an Functions bereitstellen, besonders auch im Bereich der Textver- und -bearbeitung (vgl. Srinivasa-Desikan 2018: 26).

Zu den Vorteilen von Python zählen nicht nur die Implementation dieser Libraries, sondern auch die Einfachheit, mit der eigene Skripte gebaut und ausgeführt werden können, sowie die Flexibilität und (menschliche) Lesbarkeit der Programmiersprache (vgl. Srinivasa-Desikan 2018: 26).

Für NLP gibt es verschiedene Packages; eines davon ist spaCy, das Algorithmen implementierende Pipelines für unterschiedliche Bearbeitungen von Text, wie die weiter oben beschriebenen Annotationen, zur Verfügung stellt (siehe spaCy; vgl. Srinivasa-Desikan 2018: 33–34). Für die Untersuchung wird im Script Python Version 3.11.3 sowie spaCy Version 3.7.2 mit dem Modell `de_dep_news_trf` (3.5.0) verwendet.

Unter Natural Language Processing versteht man die Bearbeitung „natürlicher“ Sprachdaten, also mündliche oder – häufiger üblich – schriftliche Äußerungen von Sprachproduzent*innen. Dabei werden unterschiedliche Tools und Methoden der Computertechnologie verwendet, um die Daten zu bearbeiten und zumeist in höherer Quantität zu untersuchen (vgl. Malik et al. 2022: 124).

Vorgänge wie Tokenisierung, Lemmatisierung und PoS-Tagging werden automatisiert durch sogenannte Parser durchgeführt.

Tokenisierung teilt die Daten in Bedeutungssegmente, also Token, ein (vgl. Srinivasa-Desikan 2018: 41–42; Malik et al. 2022: 125) und Lemmatisierung beschreibt die Rückführung der einzelnen Token auf ihre Lemmata, was die Auffindung unterschiedlicher Wortformen, die eben zum selben Lemma gehören, erleichtert (vgl. Altinok 2021: 100–103; Malik et al. 2022: 127).

PoS-Tagging ist das Tagging von Part of Speech (PoS), eine Form der Annotation, bei der die Teile einer Äußerung nach ihrer Funktion im Satz den einzelnen Wortkategorien zugeordnet werden (vgl. Altinok 2021: 128; Malik et al. 2022: 127). Hier gibt es unterschiedliche Genauigkeiten, ob nur Wortarten zugeteilt (Nomen, Verb, Adjektiv, etc.) werden (unter dem `token.pos_-Feature`) oder ob genauer eingeteilt wird, um welche Unterkategorie es sich handelt (finites Verb, Eigennamen, persönliches Pronomen) (mit dem `token.tag_-Feature`) (vgl. Altinok 2021: 134). Ein detaillierterer PoS-Tag, der bereits erwähnt wurde und in der hier vorliegenden Untersuchung von großer Relevanz ist, ist das VVPP (siehe oben).

Eine Überlegung war die Arbeit mit Dependency Parses³⁹. Diese werden auch für das amc durchgeführt, jedoch wurden sie bislang nicht für Untersuchungen verwendet, weshalb keine Informationen über die Qualität bestehen (siehe persönliche Mitteilung Hannes Pirker, 7.11.2022). Das weitere Problem mit den Dependency Parses ist, dass das *gehören*-Passiv von den bereits etablierten Parsern nicht als solches erkannt wird, höchst wahrscheinlich aus dem Grund, dass dieses im Goldstandard-Korpus des Parsers nicht als mögliche Passivform annotiert wurde. *Gehören* ist deshalb nur als einfaches Verb, nicht aber als Auxiliar filterbar (die PoS-Tagger und Parser treffen hierbei keine Unterscheidung) und wird deshalb auch nicht mit dem VVPP in Verbindung gebracht.

Aus diesem Grund wurde dieser Ansatz verworfen und eine rudimentärere Herangehensweise ins Auge gefasst: Die Sätze in der Concordance werden in möglichst kleine Satzteile gesplittet und dann im Anschluss danach gefiltert, ob sowohl ein *gehören*-Verb als auch ein beliebiges VVPP in diesen kleinstmöglichen Satzteilen vorkommt. Dabei handelt es sich wie

³⁹ Dependency Parses referieren hier auf die Annotation der syntaktischen Dependenz, also vereinfacht der Abhängigkeit unterschiedlicher Satzteile voneinander (vgl. spaCy: Linguistic Features – Dependency Parse; Srinivasa-Desikan 2018: 44).

bei CQL um Rule-Based Matching⁴⁰. Die Daten werden über die Parser des NLP-Packages spaCy annotiert, wodurch im Anschluss nach den Annotationen gefiltert werden kann.

Es stellt sich an dieser Stelle die Frage, wie erfolgreich dieses Vorgehen tatsächlich ist, um das *gehören*-Passiv zu extrahieren, weshalb hier einige Überlegungen diesbezüglich angestellt werden.

Aus dem Subkorpus der *gehören*-Verben, die mithilfe einer CQL-Abfrage aus dem Gesamtkorpus gefiltert wurden, wurde ein randomisiertes Sample von 200 Belegen extrahiert, das als Testkorpus diente, um ein geeignetes Script zu programmieren.

Ein Problem mit der ersten Version des Codes, der die *gehören*-Verben, die kein VVPP sein dürfen, und ein weiteres VVPP filterte, war, dass *gehören*-Passiv-Belege wegfielen, bei denen *gehören* valide oder fälschlicherweise als VVPP gekennzeichnet wurde, zum Beispiel bei *gehören* als finites Verb im Verb-Letzt-Satz, wie anhand der beiden folgenden Beispiele ersichtlich:

...das gesamte Krisenmanagement der Regierung und sprechen mittlerweile von Corona-Wahnsinn, der dringend gestoppt gehört. MWVOLL_2020_0428
...montieren, auf den Straßen gedealt wird, Sprayer die Hauswände verunstalten und "etwas gegen die Ausländer gemacht gehört". PROFIL_2016_0523

Bei beiden Treffern, die durch den Code aus den Ergebnissen entfernt wurden, handelt es sich um *gehören*-Passiv-Formen, die in einem Nebensatz mit Verb-Letzt-Stellung vorkommen. Aus diesem Grund wurde die Condition, dass das VVPP kein *gehören*-Verb sein darf, revidiert.

Um einschätzen zu können, wie erfolgreich die Filter sind, wurde eine simple Art der quantitativen Evaluation herangezogen. Es handelt sich hierbei um ein binäres Klassifikationssystem – in diesem Fall: ‚ist ein *gehören*-Passiv‘ oder ‚ist kein *gehören*-Passiv‘. Hierzu folgt eine kurze Erklärung der wichtigsten Begriffe, die bei der Evaluation eine Rolle spielen.

Der – vorhin bereits erwähnte – Goldstandard oder die Goldannotation stellt eine Annotation bzw. ein Labelling durch einen Menschen dar; diese kann von Person zu Person divergieren, es können auch Fehler unterlaufen, aber grundsätzlich werden diese Daten als

⁴⁰ Rule-Based Matching ist das grundlegende Vorgehen traditioneller NLP-Methoden, bei dem Regeln aufgestellt werden, die von den Sprachdaten erfüllt werden müssen. In einer Feedbackschleife werden dann die Regeln durch einen Menschen angepasst, um das gewünschte Phänomen herauszufiltern (vgl. Malik et al. 2022: 128–129). Dadurch entscheidet sich das Rule-Based Matching von Machine-Learning-Ansätzen, in denen Algorithmen über die Analyse eines Trainingssets „lernen“, die Sprachdaten zu „verstehen“ (vgl. Malik et al. 2022: 129–130).

Referenz verwendet, um die Ergebnisse von Algorithmen abzugleichen (vgl. Dalianis 2018: 47, 51).

Table 6.1 Confusion matrix: predicted annotation is what the algorithm retrieves or annotates and gold annotation is what was marked up or annotated by a human

		Predicted annotation	
		Positive	Negative
Gold annotation	Positive	True positive (tp)	False negative (fn)
	Negative	False positive (fp)	True negative (tn)

Abbildung 15: Confusion matrix (Dalianis 2018: 47)

Die Klassifikation dieser Einschätzung passiert dann über die sogenannte Confusion Matrix (siehe Abbildung 15). True Positives sind jene Belege, die sowohl vom Algorithmus als auch durch menschliche Annotator*innen als korrekt eingestuft wurden, genauso wie true Negatives von beiden als inkorrekt annotiert werden. False Positives werden durch den algorithmischen Prädiktor extrahiert, aber vom Menschen als nicht valid annotiert. Bei false Negatives tritt der umgekehrte Fall ein, laut Mensch wäre der Beleg valide, der Algorithmus annotiert ihn allerdings nicht derart (vgl. Dalianis 2018: 47).

Wenn nun die Belege einer Daten-Extraktion derart klassifiziert wurden, können unterschiedliche Metrics – ‚Werte‘ – berechnet werden, anhand derer evaluiert werden kann, wie gut der Algorithmus funktioniert und das Gewünschte extrahiert.

Die Performance wird über Precision and Recall bewertet. Precision ist jener Wert, der beschreibt, wie viele korrekte Hits unter allen durch den Algorithmus extrahierten Belegen vorhanden sind. Recall beschreibt das Verhältnis der true Positives im Verhältnis zu allen korrekten Belegen, auch jenen, die nicht durch den Algorithmus erkannt werden (vgl. Dalianis 2018: 47). Der F1-Score ist das harmonische Mittel der beiden Werte (vgl. Dalianis 2018: 47).

$$Precision : P = \frac{tp}{tp + fp}$$

$$Recall : R = \frac{tp}{tp + fn}$$

Abbildung 16: Berechnung der Precision- und Recall-Werte (vgl. Dalianis 2018, 47)

Accuracy ist die Anzahl aller korrekten Vorhersagen durch die Gesamtanzahl des untersuchten Samples: $(tp+tn)/(tp+tn+fp+fn)$ (vgl. Dalianis 2018: 48).

Für die 200 Belege wurden eben diese Scores bestimmt, auch um einen ersten Einblick in die zahlenmäßigen Verhältnisse zu erhalten und die Ergebnisse zu evaluieren.

Daten	Anzahl
true positives	11
false positives	0
true negatives	187
false negatives	2

Metrics	Scores
Accuracy	0,99
Precision	1
Recall	0,85
F1-score	0,91

Accuracy, Precision, Recall und F1-Score des Samples zeigen, dass diese Vorgehensweise in diesem ersten Test die meisten der Belege korrekt extrahiert, deshalb wird diese Methode weiterverfolgt. Diese Metrics werden im nächsten Kapitel ebenfalls angewandt, um unterschiedliche Optionen zu vergleichen.

Aus diesem Datensatz wurde ebenfalls versucht, eine Abschätzung darüber aufzustellen, wie viele *gehören*-Passiv-Belege im amc voraussichtlich vorkommen, um in Folge auch Aussagen darüber treffen zu können, welcher Anteil der Belege untersucht wurde. Der Prozentsatz der validen *gehören*-Passiv-Belege, also true Positives, entspricht 0,055 Prozent in dem randomisierten Sample von 200, das aus dem Gesamtvorkommen aller *gehören*-Verben ab 2010 mit insgesamt 1.072.343 Hits stammt⁴¹. Daraus ergibt sich eine geschätzte erwartete Anzahl von 58.979.

3.2.2 Finale Auswahl der Datenextraktion und weitere Herausforderungen

An dieser Stelle tritt ein weiteres Problem auf, nämlich das Urheberrecht des Korpus'. Aus diesem Grund wurde die CQL-Anfrage nochmal überarbeitet, um mit den Bestimmungen in Einklang zu sein. Dies geschah durch die Hilfe von Hannes Pirker, der an der ÖAW für die Pflege des amc zuständig ist. Im Gespräch konnten wir eine Lösung finden, die sowohl eine größere Datenmenge als auch die Vorgaben der Urheberrechtsbestimmungen berücksichtigen.

Die CQL-Anfrage wurde von Hannes Pirker verfeinert und lieferte danach n=44.252 Treffer:

```
(( [lempos="gehören-v"] [word!=" , | - | - | - | - ; | : | \ " | „ | “ | \ ( | \ ) | und | oder | - . * " & postTT!="V.*"] * [postTT="VPPP"] | [postTT="VPPP"] [word!=" , | - | - | - | - ; | : | \ " | „ | “ | \ ( | \ ) | und | oder | - . * " & postTT!="V.*"] * [lempos="gehören-v"] ) within <s/> within <doc year="20(1|2)."/>
```

⁴¹ Abfragesyntax CQL: ([lempos="gehören-v"]) within <doc year="20(1|2)."/>

Diese Treffer wurden mit einem Minimum an Concordance als xlsx-Datei heruntergeladen, das heißt, es ist nur der Treffer von *gehören* plus VVPP ohne den linken und rechten Kontext für diese über 45.000 Belege.

Wie anhand des Codes zu sehen ist, lautet die Bezeichnung des Parsers, mit dem gesucht werden sollte, posTT. In der Dokumentation des amc lässt sich herausfinden, dass es sich hierbei um den sogenannten Treetagger handelt, ein NLP-Parser für die deutsche Sprache, der bereits seit 1994 verwendet wird (vgl. amc: Korpusinhalt-Attribute). Dieser ist allerdings nicht der einzige mögliche Parser. Es ist ebenfalls möglich, einen spaCy-Tagger für die CQL-Abfrage im amc zu verwenden, mit dem Kürzel posTI (vgl. amc: Korpusinhalt-Attribute).

Da bereits eine sehr starke Filterung über die rule-based CQL-Abfrage besteht, wurde danach eine minimalere Filterung über den Python-Code vorgenommen. Hierbei geht es dann vor allem um die automatische Aufteilung der Informationen aus dem Korpus, die in einer CSV-Datei zwischengespeichert sind, in eine für die Arbeit sinnvolle und übersichtliche Form zu bringen.

So liegt beispielsweise die Meta-Information, woher der Beleg stammt, in der Form ‚Zeitungsnamen‘_‘Datum im Jahr-Monat-Tag-Uhrzeit-Format‘ vor und diese wird in zwei separate Spalten mit Zeitungsnamen und Datum aufgeteilt. Die übrigen Informationen sind nicht weiter relevant und fallen deshalb weg.

Durch diese Vorgehensweise ist unklar, wie viele false Negatives, das heißt eigentlich positive Treffer, die durch die Filter nicht erkannt werden, wegfallen, einerseits aufgrund der Art der CQL-Abfrage, andererseits wegen der wiederholten Filterung durch den Python-Code. Um Abschätzungen dazu aufstellen zu können, werde ich durch die kleineren Samples daraus extrapolierte Schätzungen errechnen.

Die gezwungenermaßen auferlegte – und leider auch kompliziertere – Extraktion öffnet die Tür für eine Mannigfaltigkeit an Problemen, die sich nur bedingt umgehen lassen und in die hier ebenfalls ein Einblick gegeben werden soll.

Bereits der Ausschluss von *und* und *oder* ist nicht unproblematisch, jedoch gezwungenermaßen notwendig, da ansonsten die Beschränkung auf einzelne Satzteile kaum möglich wäre.

(K18756)⁴² Die Aufgaben gehören kompetenzorientiert und nicht nach parteipolitischem Proportz aufgeteilt. OTS_2014_0515⁴³

Zudem werden durch andere Mittel der koordinierenden Satzbildung ebenfalls Satzteile verbunden, die bei der Splittung nicht beachtet werden; zum Beispiel *weder ... noch*. Dies lässt sich wiederum nicht dazwischen ausschließen, da dies wiederum Belege wie den folgenden verdrängen würde:

(K26051) An welchen Schrauben gehört noch gedreht, um Sozialberufe attraktiver zu machen? STANDARD_2020_0219

Zudem wird *gehören* selbst in manchen Fällen als Perfektpartizip interpretiert, teilweise zu recht, in den meisten Fällen jedoch fälschlicherweise: Wenn man *gehören* als VVPP filtern würde, fallen einige Ergebnisse wiederum weg, unter anderem auch alle Belege, in denen das *gehören*-Passiv z.B. im Plusquamperfekt gebildet wird und das *gehören* selbst dadurch als VVPP gebildet wird, sowie alle Ergebnisse, in denen *gehört* (3P Sgl Präsens Indikativ) in einem Nebensatz steht und dadurch von den Parsern zu Unrecht als VVPP annotiert wird. Aufgrund des beschränkten Kontextes, da nur das KWIC (KeyWord in Context) extrahiert wurde, lässt sich für die einzelnen Fälle nicht nachvollziehen, welche der beiden Möglichkeiten vorliegt.

Zudem können auch Partizipformen von Verben extrahiert werden, die nicht Teil des Prädikats sind. Der häufigste Fall nicht-valider Belege stellt das Vorkommen des als adjektivisch-verwendeten VVPP in einer angrenzenden Nominalphrase wie im folgenden Beispiel:

(K45336) Vollmundige und kräftige Weine gehören zu den später gelesenen. BVZ_2022_0825

Die zweite Möglichkeit ist als partizipialer Kopf einer anderen Phrase, die nicht Teil des Prädikats darstellt, beispielsweise als Kopf einer Adverbialphrase; dies ist ebenfalls nicht über diese strukturelle Vorgehensweise zu disambiguieren.

Man vergleiche: (K42005) „Stark gefördert gehören sowohl schwächere Schüler, teils mit Migrationshintergrund, als auch die überdurchschnittlich begabten Schüler.“ (SBGW_2019_0523) mit (S443) „EU-weit gesehen gehört Österreich im Vergleich zu den

⁴² Da in beiden Korpora den Belegen IDs zugeordnet werden, die bei 1 anfangen, wird bei den Angaben zwischen K (quantitatives Korpus) und S (qualitatives Sample) unterschieden, auch wenn über die Zahl 500 hinausgehende Belege nur aus ersterem stammen können.

⁴³ Dieses Beispiel findet sich vermutlich deshalb dennoch im Korpus, weil *kompetenzorientiert* bei der Extraktion aus dem amc als VVPP interpretiert wurde. Nach der Bearbeitung durch das spaCy-Script war es eines der fehlerhaften Beispiele, die händisch nachkorrigiert wurden, weil dieses eben *kompetenzorientiert* nicht als VVPP deutet.

Ländern mit den höchsten Lohnnebenkosten (...).“ (OBERRUND_2013_1204)“. Bei ersterem handelt es sich tatsächlich um ein *gehören*-Passiv, das VVPP *gefördert* steht im Vorfeld (vermutlich aufgrund der thematischen Hervorhebung), es handelt sich dennoch um den infiniten Teil der analytischen Passivkonstruktion. Beim zweiten Beispiel befindet sich ebenfalls ein VVPP im Vorfeld, dieses gehört allerdings nicht zum Prädikat, sondern gemeinsam mit *EU-weit* fungiert *gesehen* als Kopf dieser Phrase und als Modaladverbial des Satzes und wird dabei thematisch hervorgehoben.

Ob es sich bei der Existenz von *gehören* und VVPP innerhalb eines Satzteiles im Endeffekt um eine Passivkonstruktion handelt, ist in manchen Fällen nicht mal durch die Grobstruktur des Satzes beurteilbar, sondern erst über genauere Analysen der Satzglieder bzw. aufgrund der Semantik, die ausgedrückt wird. Satzgliedanalyse und semantische Beurteilung können allerdings nicht durch die hier angewandten Methoden bewerkstelligt werden, weshalb solche Fälle wie S443 zu Unrecht im Korpus auftauchen (also ein false Positive sind), da eine strengere Filterung noch mehr eigentlich tatsächliche Belege nicht extrahieren würde (z.B. keine VVPPs vor den finiten Verben betrachten).

Unterschiedliche Satzzeichen verursachen ebenfalls Probleme, unabhängig davon, ob sie bei der Textnormalisierung entfernt oder drinnen gelassen werden. Durch Rule-Based Algorithmen und RegEx-Codes müsste man hier minutiös einzelne Fälle angeben; hier kann es dann wiederum passieren, dass durch die strengere Filterung eigentlich valide Belege ausgeschlossen werden; je nachdem werden mehr oder weniger false Positives oder false Negatives in den Ergebnissen sein.

Anhand einiger ausgewählter Beispiele, die bei der manuellen Durchsicht der fehlerhaften Belege aufgefallen sind, sollten diese Schwierigkeiten kurz demonstriert werden. Problemfälle bereiten beispielsweise „–“-Zeichen, die manchmal anstatt von Gedankenstrichen, manchmal als tatsächlicher Bindestrich verwendet werden:

(K5121) Schwere Unfälle gehören auf Österreichs Straßen zum traurigen Alltag–doch was am Wochenende im oberösterreichischen Adlwang geschehen ist, lässt niemanden kalt. KLEINE_2012_1016 (Unterstreichungen C.M.)

(K17910) Stattdessen gehört nicht nur die Scharfmacher-Rhetorik aus dem Iran als verblendet verurteilt (...). OTS_2011_1103 (Unterstreichungen C.M.)

Bereits beim Download aus der Engine dürfte beim ersten Beleg ein Fehler passiert sein, da „Alltag–doch“ als eine Zeichensequenz aufscheint. Würde man den Gedankenstrich nicht splitten, wäre dieser Beleg fälschlicherweise in den Daten vorhanden. Zweiteres (K17910)

ist ein tatsächlich valider Beleg, bei dem das Splitten jedoch dafür gesorgt hat, dass er bei den fehlerhaften Belegen vorkommt.

Satzzeichen wie Punkte und Beistriche bereiten ebenfalls Probleme: Teilweise fehlen diese im Ausgangsmaterial, wodurch Belege fälschlicherweise ihren Weg in das Korpus finden.

(K11064) Die Travel Factory, als ein wichtiger Dertour Partner, gehört zu einem der ersten Büros (sic) in dem das Konzept Travel EXPert realisiert wird. KRONE_2012_06101 (Unterstreichungen C.M.)

Dies gilt ebenso bei anderen Fällen von Punkten bei Abkürzungen (*bzw.*) oder wie im Beispiel K5594 als Form der Zensur.

(K5594) Es gehört sehr vielen in den A... getreten, auch bei unseren Politikern. OEREICHE_2010_0614 (Unterstreichungen C.M.)

Zudem gibt es noch weitere Zeichen wie „/“ und einfache Anführungszeichen, die für Probleme sorgen:

(K40971) Frauen und Männer entscheiden miteinander, Arbeit gehört neu/gerecht bewertet, bezahlte und unbezahlte Arbeit gerecht verteilt. OBERRUND_2018_1003 (Unterstreichungen C.M.)

(K17687) Allen voran gehört die ‚Hacklerregelung‘ endlich abgeschafft (...). OTS_2011_0318 (Unterstreichungen C.M.)

Aufgrund der analytischen Bildung und der Seltenheit bzw. Ungewöhnlichkeit der Konstruktion gibt es viele Probleme bzw. Herausforderungen mit den Daten, die teilweise von den Daten selbst stammen, teilweise aufgrund der Parser entstehen, und sich leider bei diesem Thema nicht vollständig vermeiden lassen. Egal für welche Vorgehensweise man sich entscheidet, es wird immer entweder auf false Negatives (valide Belege, die aber nicht als solche erkannt werden) oder false Positives (nicht-valide Belege, die aber bei den Ergebnissen aufscheinen) Einfluss genommen; dies ist ein allgemeines Problem der Korpuslinguistik und der Arbeit mit Sprachdaten. Im Besonderen gibt es hier viele „Stolpersteine“ bei der Extraktion der Daten, da durch die doppelte Filterung in den jeweiligen Bearbeitungen einzeln Probleme auftreten, aber auch durch das Zusammenspiel der beiden mehr Probleme entstehen. Wie man es dreht und wendet, eine (möglicherweise) bessere Bearbeitung würde nur durch ein anderes Vorgehen, eventuell mit Machine-Learning-Ansätzen bzw. der Annotation eines eigenen Parsers funktionieren, die den Rahmen der hier vorliegenden Arbeit sprengen

würden. Deshalb habe ich beschlossen, mich anhand der unterschiedlichen, nun bereits etablierten Codes für eine Vorgehensweise zu entscheiden.

Um die Entscheidung für die eine oder andere Variante mit Berechnungen abzusichern, wurden für unterschiedliche Kombinationen aus CQL-Abfragen und spaCy-Filtern die Accuracy, Precision, Recall und F1-Scores ermittelt.

```
posTT-Abfrage: (([lempos="gehören-v"][word!=",-|-|-|-|
|;|:|\\"|„|“|\\(|\\)|und|oder|-.*" & posTT!="V.*"]* [posTT="VVPP"] |
[posTT="VVPP"][word!=",-|-|-|-|-|;|:|„|“|\\(|\\)|und|oder|-.*" &
posTT!="V.*"]*[lempos="gehören-v"] ) within <s/>) within <doc
year="20(1|2)."/>
```

```
posTI-Abfrage: (([lempos="gehören-v"][word!=",-|-|-|-|
|;|:|\\"|„|“|\\(|\\)|und|oder|-.*" & posTI!="V.*"]* [posTI="VVPP"] |
[posTI="VVPP"][word!=",-|-|-|-|-|;|:|„|“|\\(|\\)|und|oder|-.*" &
posTI!="V.*"]*[lempos="gehören-v"] ) within <s/>) within <doc
year="20(1|2)."/>
```

```
Kombi-Abfrage: (([lempos="gehören-v"][word!=",-|-|-|-|
|;|:|\\"|„|“|\\(|\\)|und|oder|-.*" & posTI!="V.*" & posTT!="V.*"]*
[posTI="VVPP" & posTT="VVPP"] | [posTI="VVPP" & posTT="VVPP"][word!=",-|-|-|-|-|;|:|„|“|\\(|\\)|und|oder|-.*" &
posTI!="V.*" &
posTT!="V.*"]*[lempos="gehören-v"] ) within <s/>) within <doc
year="20(1|2)."/>
```

Von diesen wurde dann ein Sample von 500 randomisiert ausgewählt (über das Interface der NoSketchEngine), um diese zu analysieren und über statistische Werte zu ermitteln, welche der Optionen wohl „die besten“ Ergebnisse liefert.

Insgesamt belaufen sich diese Untersuchungen auf sechs verschiedene Samples aus den Kombinationen der unterschiedlichen Parser und Filter, einerseits die zwei Parser im amc (Treetagger, spaCy und eine Kombination der beiden) und andererseits die zwei Filteroptionen im Python-Skript (Filter nur mit *gehören* (leicht) und Filter mit *gehören* und *VVPP* (streng)).

Name	Treetagger_leicht	Treetagger_streng	spaCy_leicht	spaCy_streng	Kombi_leicht	Kombi_streng
Tagger (amc)	TT	TT	TI	TI	TTTI	TTTI
Anzahl (vor Filter)	500	500	500	500	500	500
valide Belege	489	489	484	484	482	482
nicht valide Belege	11	11	16	16	18	18
spaCyfilter	nur gehören	gehören + VVPP	nur gehören	gehören + VVPP	nur gehören	gehören + VVPP
Anzahl (nach Filter)	483	472	490	476	487	469
TP	476	466	472	466	474	460
FP	7	6	18	10	13	9
TN	4	5	3	6	5	9
FN	13	23	7	18	8	22
Accuracy	0.96	0.942	0.95	0.944	0.958	0.938
Precision	0.985507246	0.987288136	0.963265306	0.978991597	0.973305955	0.980810235
Recall	0.973415133	0.952965235	0.985386221	0.962809917	0.98340249	0.954356846
F1-Score	0.979423868	0.969823101	0.974200206	0.970833333	0.978328173	0.967402734

Tabelle 1: Accuracy, Precision, Recall und F1-Score der unterschiedlich bearbeiteten Daten

Allgemein ist erkennbar, dass alle hier angeführten Daten-Extraktionen im allgemeinen gute Ergebnisse liefern: unter 0,938, also 93,8-prozentiger – in diesem Fall Accuracy – sind keine Scores zu finden. Bei den anderen drei Kategorien sind alle Faktoren über 0,95. Das heißt, mit allen Methodenvarianten würde man einen großen Anteil verlässlicher Daten erhalten.

Die Kombination als Treetagger und spaCy-Filter liefert ebenfalls gute Ergebnisse, jedoch wirkt die Verwendung von zwei unterschiedlichen Parsern auf den ersten Blick methodisch problematisch. Aus diesem Grund wird dieses Vorgehen eher hintenangestellt.

Insgesamt ist auch zu beobachten, dass der strengere Filter mit *gehören* und VVPP nur in der Kategorie Precision bessere Ergebnisse liefert als sein Counterpart, das nur nach *gehören* filtert.

Bei den ‚streng‘-Extraktionen sind ebenfalls weniger fehlerhafte Belege zu finden als bei den ‚leicht‘-Counterparts. Im Gegensatz dazu erhält man ebenfalls mehr false Negatives.

Name	Treetagger_leicht	Treetagger_streng	spaCy_leicht	spaCy_streng	Kombi_leicht	Kombi_streng
Tagger (amc)	TT	TT	TI	TI	TTTI	TTTI
Anzahl (vor Filter)	44252	44252	45874	45874	39218	39218
spaCyfilter	nur gehören	gehören + VVPP	nur gehören	gehören + VVPP	nur gehören	gehören + VVPP
Anzahl (nach Filter)	41296	39859	43206	41696	36879	35895
fehlerhafte Belege (VVPP fehlt)	1827	390	1953	444	1211	227
Prozent TP	0.985507246	0.987288136	0.963265306	0.978991597	0.973305955	0.980810235
Prozent FP	0.014492754	0.012711864	0.036734694	0.021008403	0.026694045	0.019189765
geschätzte valide Belege	38896.98551	38967.27542	39737.58367	40385.36134	34715.8768	34983.53945
geschätzte nicht-valide Belege	572.0144928	501.7245763	1515.416327	866.6386555	952.1232033	684.4605544

Tabelle 2: Absolute Zahlen der Extraktionsvarianten

In der Arbeit mit den Daten wurde versucht, eine annehmbare Balance zu finden, um einerseits möglichst viele Belege zu erhalten, andererseits möglichst viele nicht-valide Ergebnisse aus den Daten zu entfernen, während gleichzeitig nicht zu viele weitere Variablen die Bearbeitung beeinflussen sollen. Während der Bearbeitung wurden – wie oben erwähnt – jedoch

Bedenken geäußert, dass die Verwendung zweier Parser problematisch sein könnte, auch in serieller Verwendung. Aus diesem Grund wurde nur ein Parser anstatt zwei und die leichtere Filterung verwendet, da die CQL-Abfrage bereits sehr stark einschränkend ist. Die Entscheidung fiel auf die Kombination der spaCy-Extraktion mit der leichten Filterung im spaCy-Python-Script (dritte Spalte ‚spaCy_leicht‘ in Tabelle 1 und 2).

Die erwähnten Herausforderungen haben das erschwert und auch die übrig gebliebenen Belege sind nicht fehlerfrei – nach der eingehenderen Analyse der Daten ist klar, dass auch diese Vorgehensweise nicht ohne Probleme ist –, jedoch ist dies das mit dem Korpus und den mir möglichen Methoden erzielbare Endergebnis, und anhand der damit extrahierten Belege kann und soll dennoch das *gehören*-Passiv untersucht werden.

3.2.3 Datenextraktion, -reinigung und -filterung

Das konkrete Vorgehen der Korpuszusammenstellung wird nun kurz dargelegt.

Zuerst wird im Interface des amc, der NoSketch-Engine, die CQL-Abfrage gestellt:

```
(([lempos="gehören-v"] [word!="",|-|-|-|-|;|:|\"|„|\"|\\(|\\)|und|oder|-.*" &
posTI!="V.*"]* [posTI="VPPP"] | [posTI="VVPP"] [word!="",|-|-|-|-|
|;|:|„|\"|\\(|\\)|und|oder|-.*" & posTI!="V.*"]*[lempos="gehören-v"] )
within <s/> within <doc year="20(1|2)."/>
```

Die Ergebnisse derselben werden als KWIC ohne den Kontext (aufgrund der bereits erwähnten Urheberrechts-Regelungen) als Excel-Datei downgeloadet.

Die Spalte mit der Ressortzuteilung wurde zuerst händisch bearbeitet und der Einfachheit halber nur die Erstnennung übernommen.

Dieses Dokument wurde mit Pandas als Dataframe eingelesen und die Textspalten einer einfachen Datenreinigung unterzogen, die auch als Textnormalisierung bezeichnet wird, bevor dieser Text in kleinere Teilsätze gesplittet wird. Dieser Schritt war in vorhergehenden Tests noch wichtiger, da längere Sätze als Input dienen, wie es auch beim qualitativen Sample der Fall ist.

Die Satzteile werden anschließend mit dem Parser automatisch annotiert, konkret ist für die hier vorliegende Arbeit vor allem die Lemmatisierung und das POS-Tagging von großer Bedeutung, da diese Pipelines ermöglichen, die Daten zu filtern und so nur die Textteile mit *gehören* zu betrachten.

Aus diesen werden zudem *gehören* in seiner konkret im Beleg verwendeten Form sowie das VVPP in eigene Spalten extrahiert, um diese separat von ihrem Äußerungskontext zu betrachten; bei einer hohen Anzahl an Belegen ist es sinnvoll, auch dies zu automatisieren. Dieser Schritt ist der Grund, warum nicht bloß auf die Ergebnisse der doch recht genauen

CQL-Abfrage zurückgegriffen wurde, da ohne ihn keine gesonderte Betrachtung der Formen möglich (im amc wird die gesamte Abfragesyntax als KWIC zusammengruppiert) und dadurch auch die Quantifizierbarkeit dieser Aspekte nicht gegeben wäre.

Das Dataframe mit diesen gefilterten Teilsätzen, den herausgefilterten *gehören*- und VVPP-Formen und den entsprechenden Metadaten wurde dann wieder in eine Excel-Datei überführt. Anhand dieser Daten wurden Tabellen und Grafiken erstellt, die zur Präsentation der Ergebnisse im folgenden Kapitel dienen.

Der Code, der hierfür verwendet wurde, ist in einem Github-Repositoryum versammelt (siehe Anhang A), ohne die Daten aus dem amc beizufügen, da diese den Urheberrechtsbestimmungen desselben unterliegen. Bei einem Zugang zum Korpus, mithilfe der CQL-Abfrage und des Python-Scripts sollten die Ergebnisse reproduzierbar sein.

Die Daten werden im Anschluss auch normalisiert. Dieser Schritt ist wichtig, da die absoluten Zahlen einen falschen Eindruck vermitteln würden, im Besonderen, wenn man unterschiedlich große Teilbereiche innerhalb einer Metadaten-annotierten Kategorie untersucht. Entsprechend der Terminologie wird hier dann von PPM – Parts per Million – gesprochen; im Falle der untersuchten Konstruktion bezieht sich das konkret auf die Bedeutung ‚jene *gehören*-Verben, die zur Bildung des Passivs verwendet werden, per Million Token‘.

Zur Normalisierung der Belege wird ebenfalls auf der Webseite ein Dokument mit den entsprechenden Daten zur Verfügung gestellt (vgl. amc: Normalisierung). Die Tabelle der Berechnungen ist in Anhang B der Arbeit beigelegt.

Ressorts werden nicht normalisiert, da bei der quantitativen Analyse immer nur die Erstnennung extrahiert wird und eine Normalisierung die Ergebnisse wiederum verzerren würde.

3.3 Annotation der *gehören*-Formen

Da durch diese Bearbeitung über die gewählte Vorgehensweise eine nicht unbeträchtliche Menge an Daten wegfällt und nicht alle für die Untersuchungen benötigten Informationen über NLP-Methoden extrahiert werden können, wird die quantitative Arbeit mit dem Korpus durch eine qualitative Analyse eines kleineren Samples von randomisiert ausgewählten Belegen gestützt, um zu gewissen Aspekten der Fragestellung detailliertere Aussagen treffen zu können, die sich nicht über einfache automatisierte Parser herauslesen lassen.⁴⁴

⁴⁴ Extensivere Arbeit mit dem Sample und weiterem Coding – eventuell mit Machine Learning Methoden – wirken vielversprechend, genauere und ausführlichere Ergebnisse zu erhalten. Der Zeitaufwand, der hierfür jedoch notwendig wäre, ist im Rahmen der Masterarbeit nicht durchführbar, unter anderem, weil höchstwahrscheinlich eine Gold-Standard-Annotation eines kleinen Korpus und das Schreiben eigener Parser vonnöten wären. Die Möglichkeiten der digitalen Tools sind bei der Betrachtung dieser Passiv-Konstruktion(en) nicht endgültig erschöpft.

Die Analyse lässt sich dabei in drei Fokuspunkte aufteilen:

Der erste Teil beschäftigt sich genauer mit dem Prädikat der Sätze, einerseits *gehören* als finites Verb und dessen grammatikalische Formen (Person, Numerus, Modus (Indikativ, Konjunktiv). Hierzu wird besonders auf Disambiguierung der Formengleichheit geachtet (beispielsweise 3. Person Singular und 2. Person Plural oder 3. Person Singular Konjunktiv und 1. Person Singular Indikativ). Andererseits geht es nochmal genauer um die Semantik der Verben, die in infinitiver Form vorkommen, sowie um die Transitivität derselben, um in dieser Hinsicht im Theorieteil vorgestellte Hypothesen für die Möglichkeit der Verwendung des Passivs zu überprüfen.

Rein theoretisch hätte die Option bestanden, die morphologische Verbbestimmung ebenfalls von den Parsern machen zu lassen, jedoch stößt man dabei wieder auf dieselben Probleme. Es gibt die Möglichkeit, direkt in der NoSketch-Engine diese Bestimmungen anzufügen und mit den Daten downzuloaden, jedoch wird dies dann für die gesamte Concordance gemacht, was einerseits die Datenmenge beträchtlich erhöht und andererseits im Nachgang die Textnormalisierung wiederum beträchtlich erschwert. Mit einem Parser im Nachhinein nur *gehören* derart zu taggen, ist wiederum nicht möglich, da nie der vollständige Satz, sondern nur das KWIC extrahiert wurde, und dadurch eventuell wichtige Satzteile (wie beispielsweise das Subjekt) fehlen, die die Bestimmung über die Parser erst möglich machen würden.

Der zweite Teil betrachtet einige der potenziell vorkommenden Subjekte und Objekte. Beim Subjekt wird untersucht, ob es vorhanden ist oder ob ein unpersönliches Passiv gebildet wurde. Die grammatikalische Form wird annotiert und, wenn es vorhanden ist, angelehnt an Szatmári (2002, 182) in die allgemeine Semantik der Patiens-Lexeme eingeordnet (siehe Kapitel 2.3.2), jedoch hauptsächlich auf belebt/unbelebt bezogen (während bei Szatmári (2002: 182) zwischen [$\pm$ abstrakt] und [$\pm$ human] unterschieden wird).

Bei den Objekten werden Präpositionalphrasen genauer unter die Lupe genommen: Zum einen geht es dabei um die lokativen Präpositionalphrasen, die bei Stathi (2010: 327–329) im Zuge der Grammatikalisierung thematisiert wurden (siehe Kapitel 2.3 zur Theorie des *gehören*-Passivs), zum anderen um den fakultativen Agens-Anschluss mit „von“. Wie häufig kommen diese Fälle bei der Verwendung des *gehören*-Passivs vor?

Zu guter Letzt wird ein Blick auf den Aspekt der Mündlichkeit geworfen, der in der Forschungsliteratur zum Thema bislang kaum bzw. eigentlich gar nicht beachtet wurde. Das untersuchte Korpus ist ein schriftsprachliches, eine Vermutung ist jedoch, dass es sich eventuell um eine eher sprechsprachliche Formulierung handelt, die eben Einzug in die

Schriftsprache erhält; durch die Verwendung in Zeitungstexten dringt es eben, wie in Kapitel 2.1 zur österreichischen Standardschriftsprache dargelegt, in den Bereich des Standards ein. Um nun eine Aussage zum Aspekt der Mündlichkeit treffen zu können, wird konzeptionelle Mündlichkeit in den Belegen genauer untersucht: Wird das *gehören*-Passiv in direkter oder indirekter Rede formuliert?

Als kurze Beispiele, wie diese Analyse angedacht ist, wird noch einmal auf die beiden Belege eingegangen, die bereits in Kapitel 3.2.1 im Zuge des Vorgehens beim Zusammenstellen des Korpus vorgestellt wurden.

...das gesamte Krisenmanagement der Regierung und sprechen mittlerweile von Corona-Wahnsinn, der dringend gestoppt gehört. MWVOLL_2020_0428

...montieren, auf den Straßen gedealt wird, Sprayer die Hauswände verunstalten und „etwas gegen die Ausländer gemacht gehört“. PROFIL_2016_0523

Der erste Beleg stammt aus MWVOLL, das heißt aus den Transkripten des ORF (Radio-/TV-Sender), und ist zudem in der Umgebung des Verbs *sprechen* vorzufinden, weshalb man sowohl eine direkt mündliche Äußerung als auch eine mögliche indirekte Rede darin sehen könnte.

Der zweite Beleg beinhaltet eine direkte Rede oder kenntlich gemachtes Zitat, das über die Anführungszeichen ausgedrückt wird. Dadurch wird Referenz auf den ursprünglich gesprochenen Kontext hergestellt, der die Aussage in den mündlichen Bereich rückt.

Zudem wird ein Blick auf die Ressort-Metadaten geworfen, um potenzielle Regularitäten ausfindig zu machen. Die Ressorts sind Zuteilungen nach Angaben der einzelnen Zeitungen und Zeitschriften, die für das Korpus bereinigt wurden. Es ist jedoch unklar, ob die Kategorien in unterschiedlichen Medien tatsächlich dasselbe meinen (vgl. amc: Korpusinhalt-Attribute. doc.ressort2).

Von den angeführten Analysekatégorien ist die (konzeptionelle) Mündlichkeit am schwierigsten zu beurteilen, jedoch wollte ich sie trotz möglichen Uneindeutigkeiten nicht exkludieren, da zu diesem Aspekt kein Versuch der Einordnung existiert und die Frage bei der Entwicklung der Forschungsfragen auftauchte.⁴⁵

Der Umstand, dass dies in der Forschungsliteratur bislang weniger beachtet wurde, ist auch der Grund, warum an dieser Stelle ein kurzer Exkurs zum Begriff der konzeptionellen Mündlichkeit stattfinden sollte. Er wurde von Koch/Österreicher (1985; 1990; 1994) ab dem

⁴⁵ Vielen Dank an dieser Stelle an Susanne Schmalwieser, die bei der ersten Präsentation meines Themas den Denkanstoß für die Entwicklung dieser Frage gab.

Ende der 80er geprägt, aufbauend auf Ausführungen von Söll (1974; 1985). Sie beschäftigen sich mit dem Verhältnis von Schriftlichkeit und Mündlichkeit in Kommunikation, einerseits auf medialer, andererseits auf konzeptioneller Ebene.

Medial meint hierbei die Realisation als gesprochene Sprache oder als graphematisch festgehaltener Text (vgl. Thaler 2007: 149), während konzeptionell zwischen unterschiedlichen Charakteristiken von Nähe (mündlich) und Distanz (schriftlich) unterscheidet, sowohl in tatsächlicher Distanz und den gegebenen technischen Kommunikationskanälen als auch die Verhältnisse von Mono- zu Dialogizität oder Privatheit und Öffentlichkeit (vgl. Elspaß 2005: 25–26; Thaler 2007: 155). Die Art der Kommunikation wird also an Kommunikationsbedingungen geknüpft.

Da Medialität einfach geändert werden kann, indem Gesagtes transkribiert oder Niedergeschriebenes vorgetragen wird⁴⁶, wird der konzeptionellen Mündlichkeit und Schriftlichkeit oft Vorrang gegeben (vgl. Elspaß 2005: 25).

Ihr Modell wurde in den Jahren darauf häufiger kritisiert, konkretisiert oder auch weiterentwickelt (vgl. Thaler 2007: 148).⁴⁷ Es bleibt jedoch in vielen Fällen ein Ansatzpunkt, um über die Verhältnisse von Schriftlichkeit und Mündlichkeit zu diskutieren, da es sich als „anregend und konstruktiv“ erwiesen hat (vgl. Elspaß 2005: 25).

Wenn also in dieser Arbeit von fraglicher Mündlichkeit die Rede ist, wird potenziell auf unterschiedliche Aspekte und Ausprägungen dieses Konzepts referiert, die aufgrund des gewählten Korpus und der *gehören*-Passiv-Konstruktion in einem komplexen Zusammenspiel von fixierter medialer Schriftlichkeit der Belege, der postulierten Standardsprachlichkeit (die eher dem Feld der Distanz zugeordnet wird) und ihrer tatsächlichen Gestaltung beziehungsweise vermuteten Verwendung in unterschiedlichen Situationen stehen. Dies wird in Kapitel 5.3 zu weiteren Aspekten der Konstruktion weiter diskutiert.

Zusammengefasst werden konkret folgende Kategorien für das qualitative Sample annotiert:

⁴⁶ Es scheint dennoch einen Unterschied zu machen, ob die Medialität der Aussage geändert wurde, da eine Übertragung potenziell mit Informationsverlust einhergeht (ein Interview wird vermutlich nicht im IPA übertragen werden), aber auch mit Informationsanreicherung (z. B. über den*die Sprecher*in). Besonders hier stellt sich auch die Frage nach der Einordnung von schriftlich zitierter Rede in mediale und potenziell auch konzeptionelle Schriftlichkeit und welche Wechselwirkungen sich entfalten.

⁴⁷ Dies geschah besonders auch im Hinblick auf computergestützte Kommunikationstechnologien (vgl. Thaler 2007: 148). Seit diesen Forschungswerken sind nun gute 20 Jahre vergangen, die Arten der Kommunikation haben sich durch die Weiterentwicklung der Technologie wiederum verändert und die durch das Modell vorgestellten Mittel reichen eventuell nicht, um diesen gerecht zu werden.

Siehe weiterführend Stein (2001), Dürscheid (2003), Thaler (2003: 31–45), Ágel/Hennig (2006; 2007).

- Finites Verb: Person, Numerus, Modus, im Besonderen der Fokus auf Formen-gleichheit
- Transitivität der Verben: transitiv oder intransitiv?
- Form des Subjekts: Nominalphrase, Nebensatz, formales Subjekt, ...; Semantik der Patiens-Lexeme: unbelebt oder belebt, menschlich oder tierisch
- Präpositionalobjekt: Nimmt das *gehören*-Passiv zusätzlich eine Präposition-alphrase? Ist diese statisch oder direktiv? Ist die Bedeutung metaphorisch?
- Vorhandensein des Agens-Anschlusses: Welche Präposition wird dafür verwendet?
- Aspekt der Mündlichkeit: direkte oder indirekte Rede, Anführungszeichen, Ressorts

Aufgrund der Vorgehensweise bei der Extraktion aus dem amc in Abstimmung mit den Urheberrechten muss das kleinere Sample wiederum über die NoSketch-Engine extrahiert werden. Dies wird mithilfe der bereits bekannten finalen CQL-Abfrage und der Randomisierungs-Funktion des Interfaces bewerkstelligt. Diese 500 Treffer werden mitsamt rechtem und linkem Kontext heruntergeladen, um die beschriebenen Analysen durchführen zu können. Für die Weiterbearbeitung werden die Daten wiederum in den Python-Code eingelesen und wie in 3.2.4 bearbeitet.

Zudem wurde das Sample händisch validiert: Handelt es sich bei dem einzelnen Beleg tatsächlich um ein *gehören*-Passiv oder wurde es durch die CQL-Abfrage fälschlicherweise als ein solches ausgeworfen? Von den 500 extrahierten Belegen sind 472 tatsächlich valide Passiv-Belege, die auch vom Script als solches erkannt wurden (also true Positives). Diese Zahlen wurden dann verwendet, um eine Abschätzung zu geben, wie viele *gehören*-Passive und wie viele false Positives im quantitativen Korpus zu finden sind.

4. Das *gehören*-Passiv im amc

In diesem Kapitel werden die Ergebnisse aus dem quantitativen Subkorpus und dem qualitativen Sample von 472 Belegen präsentiert. Angefangen bei den quantitativen Verhältnissen im Laufe der Extraktion und Datenbearbeitung sowie den manuellen Nachbesserungen hin zu den unterschiedlichen Analysekategorien zu Metadaten und grammatischen und semantischen Aspekten der Konstruktion werden die Daten in Zahlen und Grafiken dargestellt, und dann weiter durch Ergebnisse aus dem Sample kontextualisiert, einerseits durch zusammenfassende Belegzahlen, andererseits durch einzelne Beispiele, die einen tieferen Einblick in Fragestellungen geben sollen.

4.1 Quantitative Verhältnisse der gehören-Passiv-Konstruktionen

Die Untersuchungen der Konstruktion teilen sich auf zwei Teile auf, die jedoch mit derselben CQL-Abfrage aus dem amc extrahiert wurden – das quantitative Korpus (K) und das qualitative Sample (S). Sie unterscheiden sich in Bezug auf Samplegröße und Detailgenauigkeit, die Beispiele des qualitativen Samples sind aber ebenfalls in der größeren quantitativen Datensammlung aufzufinden.

Es wird zuerst auf die Entwicklung der Korpus-Größe im Zuge der Datenbearbeitung eingegangen, um diesbezüglich einen Überblick zu geben und danach auf die mengenmäßigen Verhältnisse der Belegzahlen nach unterschiedlichen Gesichtspunkten eingegangen, die sich in beiden Teilen abzeichnen.

Nach der Extraktion aus dem amc, unter Verwendung der CQL-Abfrage mit dem spaCy-Tagger, sind es 45.874 KWIC. Diese werden durch mein spaCy-Script in die vorgesehene Form als Excel-Spreadsheet mit folgenden Spalten gebracht: sentence_id (Identifizier), Zeitung (obwohl es besser Medium heißen hätte sollen), Datum (in Jahr), Ressort (nur Erstnennung), Mediatype (agentur, print, tv), Region (agesamt, aost, asuedost, awest, amitte, spezifisch), text (gesamter Text des KeyWords in Context (KWIC)), splitund (Text nach der Bearbeitung mit dem spaCy-Script), gehören (aus ‚splitund‘ extrahierte Form von *gehören*), VVPP (Vollverb Perfektpartizip aus ‚splitund‘).

Schritt in der Datenbearbeitung	Anzahl Hits
Extraktion aus dem amc (mit spaCy-Tagger)	45.874
nach meinem spaCy-Script	43.206
davon fehlerhafte Belege	1.953
von diesen valide <i>gehören</i> -Passive	1.255
von diesen nicht-valide Belege	698
Gesamtbelege	42.508
Davon false Positives (anhand des Precision-Werts)	1.562
Davon true Positives (anhand des Precision-Werts)	40.946

Tabelle 3: Entwicklung der Belegzahlen der quantitativen Datensammlung

Die 1.953 „fehlerhaften“ Belege rühren größtenteils daher, dass entweder *gehören* oder die VVPPs im Zuge des spaCy-Scripts nicht korrekt in die separaten Spalten extrahiert wurden. Aus diesem Grund wurden manuelle Nachbesserungen durchgeführt, und im Zuge dessen auch offensichtlich falsche Belege aussortiert (n=698).

Eine Fehlerquelle war hierbei die Doppelung des *gehören* in der dafür vorgesehenen Spalte, zumeist weil das VVPP zu *hören* als *gehören* gefiltert wurde. Dies wurde angepasst.

(K24011) Die 14 neuen Songs dürfen im Plattenregal – ja, diese Musik gehört auf Vinyl gehört – neben die Meisterwerke „Astral Weeks“, „Moondance“ und „Hymns to the Silence“. SN_2019_1111

(K22344) Doch auch nach 1995 verwendete die Rundfunkanstalt den Slogan "Ö1 gehört gehört" weiter. PRESSE_2020_0120⁴⁸

In manchen Fällen ist die Concordance nicht vollständig oder das VVPP fehlerhaft, vermutlich aufgrund eines Fehlers im Dokument, das im amc vorliegt.

(K11031) Das gehört abge-DORIS VETTERMANNstellt (sic) (...) KRONE_2012_0511

(K37652) „Die Bevölkerung gehört eingebuden (sic)“, regte Treiber eine Volksbefragung an. NOEN_2012_0508 (Unterstreichung, C.M.)

Ein weiteres Problem sind ungewöhnliche, nicht erkannte VVPPs, wobei unklar ist, ob dies durch die Version des Parsers verursacht wird oder ob die Unvollständigkeit der Sätze im Kontext Probleme bereitet.

Teilweise war es aufgrund des beschränkten Kontextes schwierig zu beurteilen, ob es sich nun um einen validen Beleg handelt. Im Zweifelsfall wurde der jeweilige Satz auch im amc kontrolliert, vor allem bei ungewöhnlichen Verben oder potenziell adjektivisch verwendeten Perfektpartizipia.

Überdacht scheint relativ häufig im Korpus auf (n=171), jedoch erkennt der Parser dieses Wort nicht, in keiner der beiden Bedeutungen (als VVPP von *überdachen* ‚mit einem Dach versehen‘, aber auch von *überdenken* ‚noch einmal überlegt‘):

(K28605) Hierzu gehörten Pkw-Parkflächen wie vor Supermärkten konsequent mit PV-Anlagen überdacht TT_2021_0814

(K29220) Das gesamte Konzept gehörte nochmals überdacht. OOEN_2011_0412

Ebenso bringen englische Verben (wie *to update* ‚aktualisieren‘), die deutsch flektiert werden, sowie Verbneuschöpfungen, meist aus etablierten Substantiven (*Notstrom*) zusammen mit einem Verb (*versorgen*) (K9569)⁴⁹, oder Eigennamen (*Attersee*, K2521), Probleme mit sich.

(K1852) (FOLDER) mitgeschickt am 29.06.2016 (gehört upgedatet!) APA_2019_0527

⁴⁸ Dies stammt eben aus dem Artikel zum Slogan, der bereits in Kapitel 2.1 erwähnt wurde.

⁴⁹ An dieser Stelle ist unklar, ob es sich um eine Komposition aus Substantiv plus Verb handelt oder ob das komponierte Substantiv Notstromversorgung in ein Verb umgewandelt wurde, vor allem wenn man Beispiele wie (K20740) „Der österreichische Sport gehört endlich ‚entparteipolitisiert‘.“ OTS_2021_0521 genauer betrachtet. *Politisieren* wird als lexikalisches Verb im Duden angeführt (vgl. Online-Duden: *politisieren*).

(K2521) Das gehört auch nochmal atterseeisiert. KURIER_2010_0822
 (K9569) Dass die BAWZ dringend notstromversorgt gehört, war außerdem ein Ergebnis der Übung ‚Combined Success‘. KLEINE_2019_0404.
 (K44944) Eßling, Aspern, Breitenlee, Stadlau, Hirschstetten und die Siedlungen entlang des Biberhaufenweges gehören verkehrsberuhigt. WZ_2022_0513
 (K12275) Wiewohl manche von den Behältnissen einfach zerdeppert gehören.
 KRONE_2015_0608

Diese Belege wurden manuell validiert und nachgebessert, sie stellen Ausnahmen dar.

Bei *verstärkt* (n=84) ist es nicht klar, ob es sich dabei um das Partizip in der Verwendung als Partizip oder als Adverb handelt, deshalb ist es schwierig, anhand der verkürzten Beispiele zu beurteilen, ob es sich nun um *gehören*-Passiv-Belege oder die adverbiale Verwendung handelt. Ein ähnlicher Fall trifft auf *gefühlt* zu, auch in geringerer Anzahl (n=9).

Nach der manuellen Validierung der fehlerhaften Belege verbleiben 42.508 Datenpunkte, von denen aufgrund der Hochrechnung der Precision-Werte schätzungsweise 1.562 false Positives und 40.946 true Positives sind.

Diese Belege werden im Folgenden nach den Metadaten Jahr, Medium, Medientyp und Region genauer beleuchtet, sowohl in absoluten als auch normalisierten Zahlen.

Jahr	Anzahl <i>gehören</i> -Belege	Frequenz (PPM)
2010	3537	7,473
2011	3517	7,052
2012	3571	7,139
2013	3879	7,873
2014	3508	7,334
2015	3333	7,186
2016	3163	6,994
2017	3269	7,456
2018	2953	6,618
2019	3156	7,038
2020	2693	7,078
2021	2899	7,255
2022	3030	7,648

Tabelle 4: Anzahl der *gehören*-Passiv-Belege nach Jahr seit 2010 (absolut und normalisiert)

Insgesamt wurde das *gehören*-Passiv über den Zeitraum 2010-2022 untersucht.

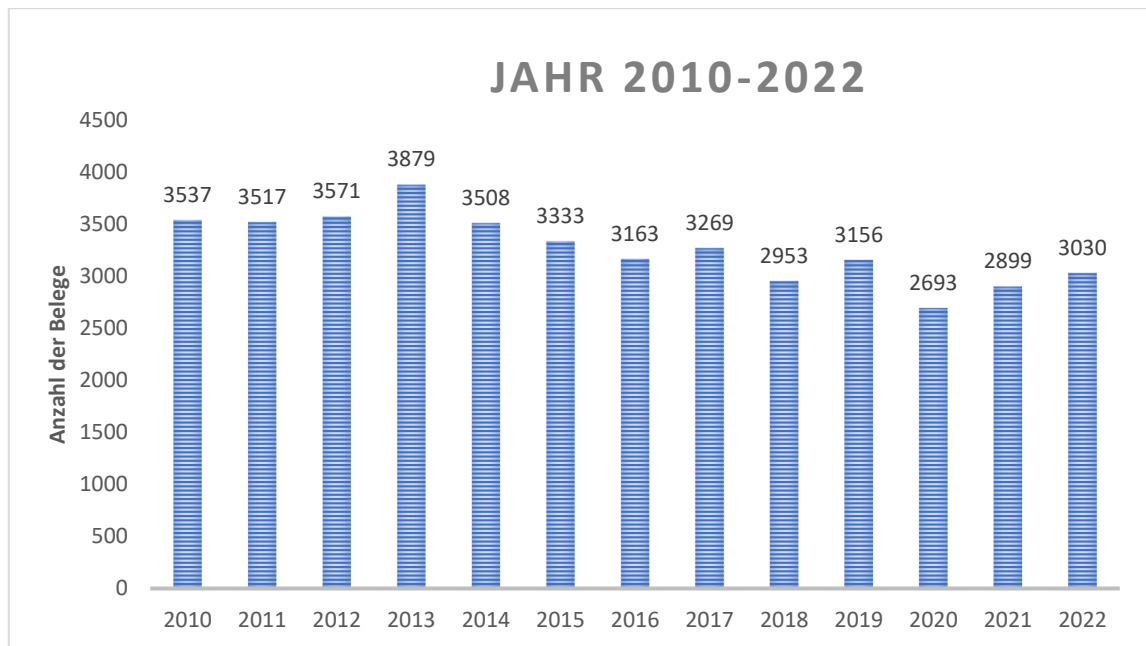


Abbildung 17: Anzahl der gehören-Passiv-Belege nach Jahren (absolut)

Die absoluten Zahlen der Passiv-Belege nach Jahr bewegen sich zwischen 2693 und 3879.

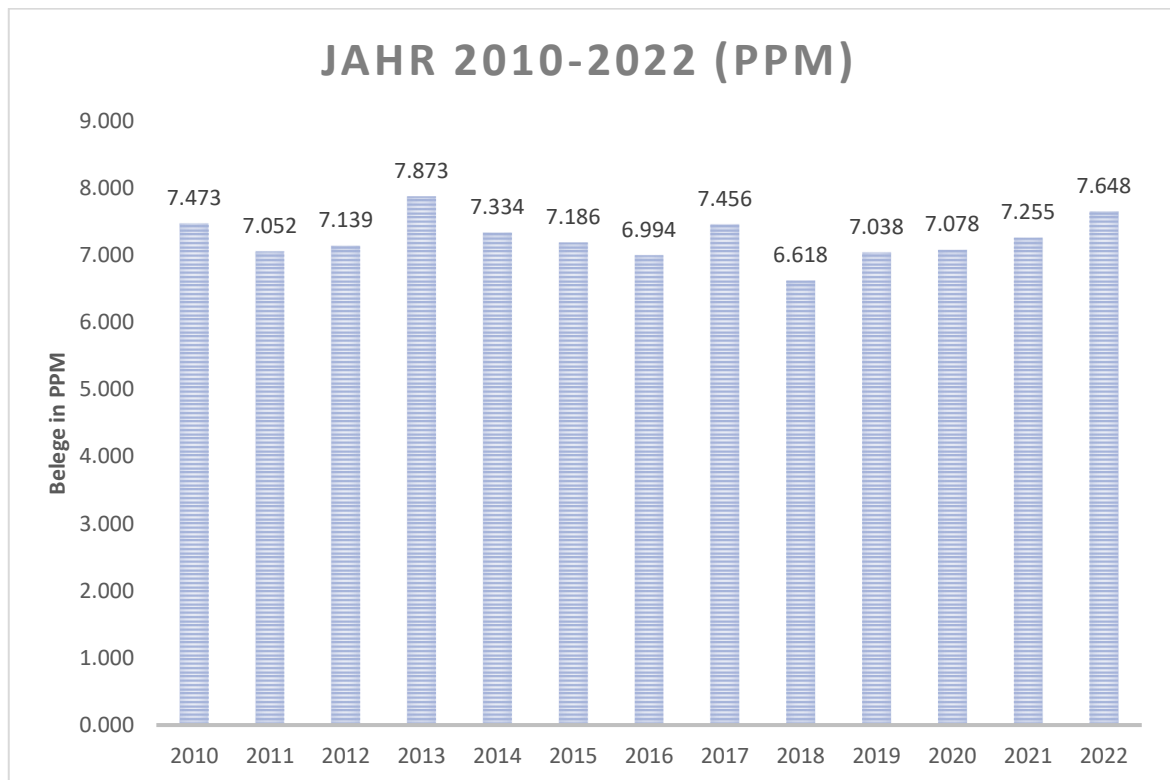


Abbildung 18: Anzahl der gehören-Passiv-Belege nach Jahren (normalisiert)

Als normalisierte Werte sind diese jedoch noch näher beieinander zwischen 6,62 PPM und 7,87 PPM, mit einem Mittelwert von 7,26 PPM.

Sowohl die absoluten Zahlen als auch die normalisierten Werte zeigen nur minimale Abweichungen, keinen bestimmten Auf- oder Abwärtstrend. Die Konstruktion scheint zumindest in den letzten 12 Jahren stabil geblieben zu sein.

An dieser Stelle ist anzumerken, dass die meisten der herangezogenen Medien, aber nicht alle, über den gesamten Zeitraum publiziert haben. Der Neue Kärntner Monat erschien von 2010 bis 2013, die Kärntner Tageszeitung von 2010-2014, der Steirer Monat von 2010 bis 2012, die Salzburger Volkszeitung 2010 bis 2014 und das Wirtschaftsblatt von 2010 bis 2016. Außerdem wurden die Fernsehsender (mit Ausnahme des ORF) nur bis 2020 im Korpus aufgenommen (siehe Anhang B, Medien-Tabelle). Insgesamt belaufen sich diese neun Medien auf eine Beleganzahl von 762, wobei die Kärntner Tageszeitung mit 244 und das Wirtschaftsblatt mit 303 *gehören*-Passiv-Konstruktionen den Großteil beisteuern.

Eine genauere Betrachtung der einzelnen Medien folgt nun, mit einem besonderen Blick auf die Zeitungen, Zeitschriften und anderweitigen Medien, in denen sich das *gehören*-Passiv besonders oft findet. Denn hier zeigen sich ebenfalls Unterschiede in der Häufigkeit, die im Vergleich von absoluten und normalisierten Belegzahlen noch einmal kontextualisiert werden.

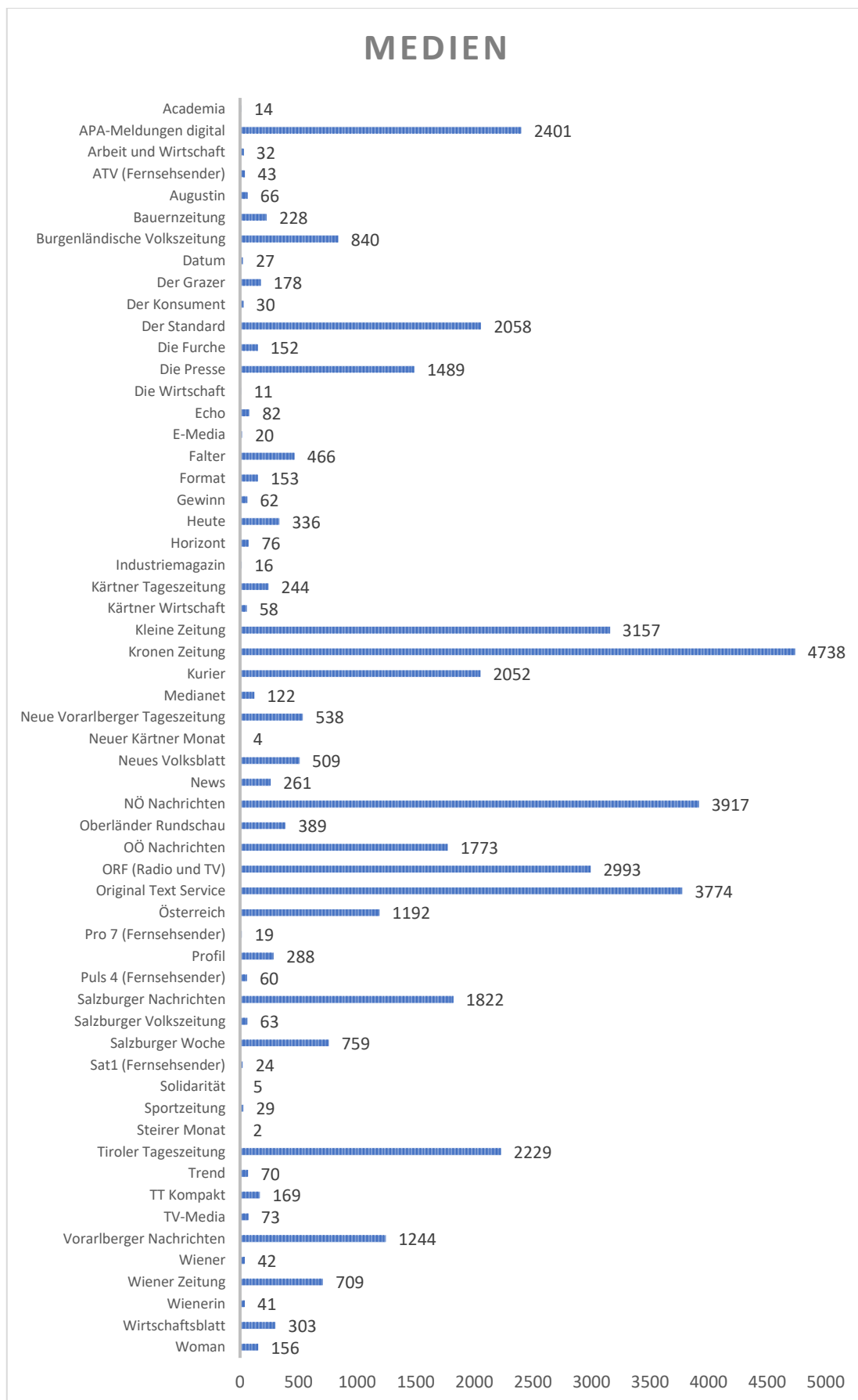


Abbildung 19: Anzahl der gehören-Passiv-Belege in den einzelnen Medien des Korpus (absolut)

Bei den absoluten Belegzahlen sind die ersten fünf Medien – Kronen Zeitung mit $n=4738$, NÖ Nachrichten mit $n=3917$, Original Text Service (OTS) mit $n=3774$, Kleine Zeitung mit $n=3157$ und ORF (Radio und TV) mit $n=2993$ – welche, die sehr häufig produzieren, da es sich entweder um Tageszeitungen, den Presstextservice der APA oder die größte Rundfunkanstalt des Landes handelt.

Es gibt im betrachteten Zeitraum keine Medien im amc, die nicht in der Liste vorkommen, auch wenn manche nur sehr wenige Belege aufweisen, drei davon sogar unter 10: Der Steirer Monat weist zwei *gehören*-Passive auf, der Neue Kärntner Monat vier und Solidarität insgesamt fünf. Die ersteren zwei Medien tauchen allerdings auch nur zwei bzw. drei Jahre im Untersuchungszeitraum auf, was die geringe Beleganzahl erklären könnte.

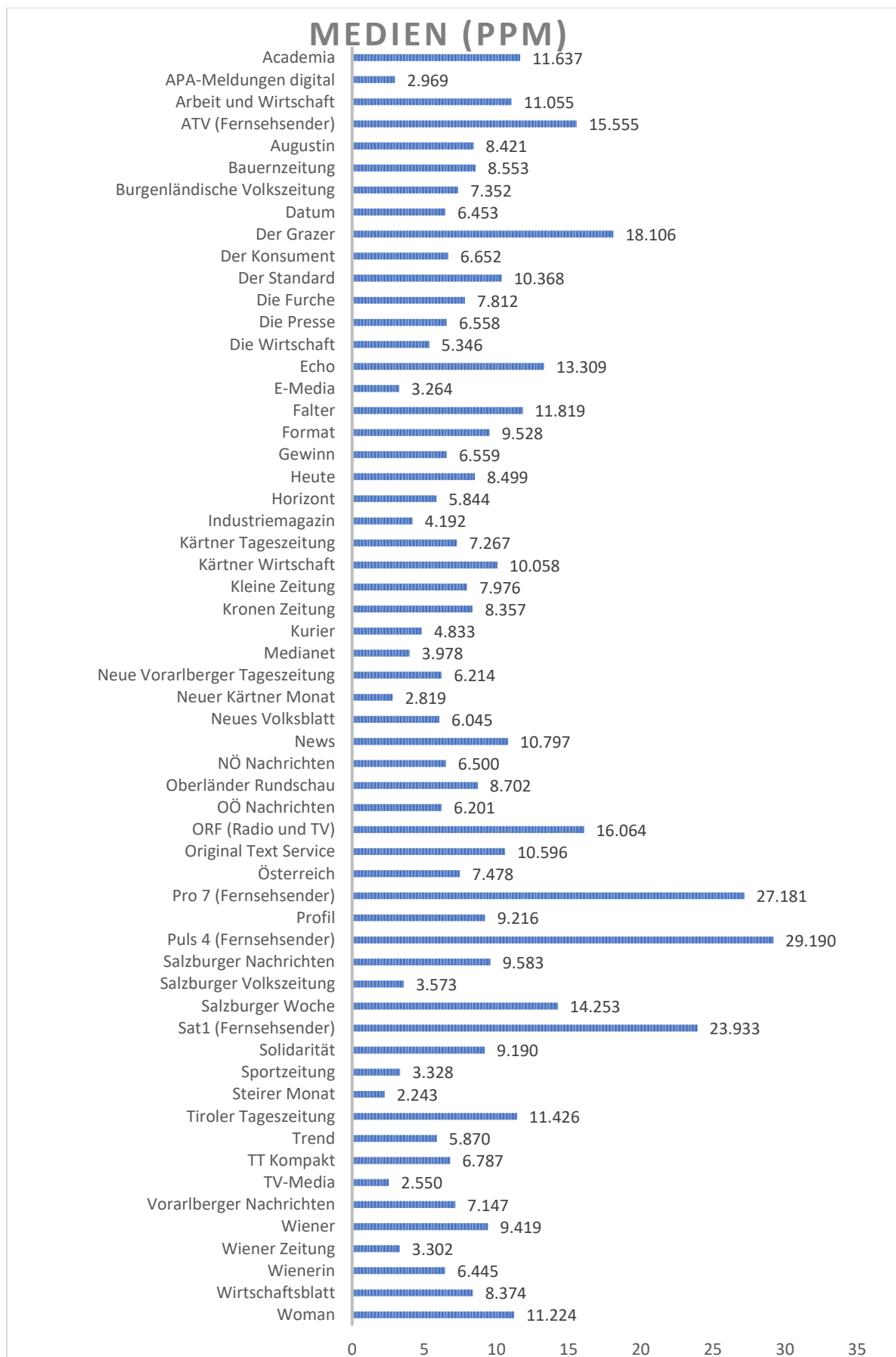


Abbildung 20: Anzahl der gehören-Passiv-Belege in den einzelnen Medien (normalisiert)

Bei den 5 häufigsten Medien normalisiert sind alle TV-Sender dabei, also der explizit mündlich produzierte und transkribierte Teil des Korpus: Puls 4 mit 29,19 PPM, Pro7 mit 27,18 PPM, Sat1 mit 23,93 PPM und etwas weiter dahinter ORF (Radio und TV) mit 16,06 PPM. Das einzige Printmedium ist der Grazer, der in PPM mehr *gehören*-Passiv-Belege aufweist als der ORF (n=18,11). Es sind also verstärkt die mündlich produzierten und transkribierten Medien, die hier hervorstechen, die bei den absoluten Zahlen nicht ins Gewicht fallen, weil die Tokenanzahl im Gesamtkorpus nur einen kleinen Teil ausmacht.

Der ORF als TV- und Radiosender ist sowohl bei den absoluten als auch normalisierten Zahlen auf Platz 5 aufzufinden.

Bei den übrigen Medien, die absolut unter den Top 5 vorkommen, wird die Gewichtung bei den normalisierten Zahlen relativiert: Kronenzeitung (8,537), NÖ Nachrichten (6,5), OTS (10,596) und Kleine Zeitung (7,976).

Diese Verhältnisse, die sich hier anhand einzelner Medien abzeichnen, spiegeln sich auch bei der Betrachtung des Medientyps wider, wobei bei den absoluten Zahlen der **print**-Teil des Subkorpus wie für das gesamte Korpus überwiegt.

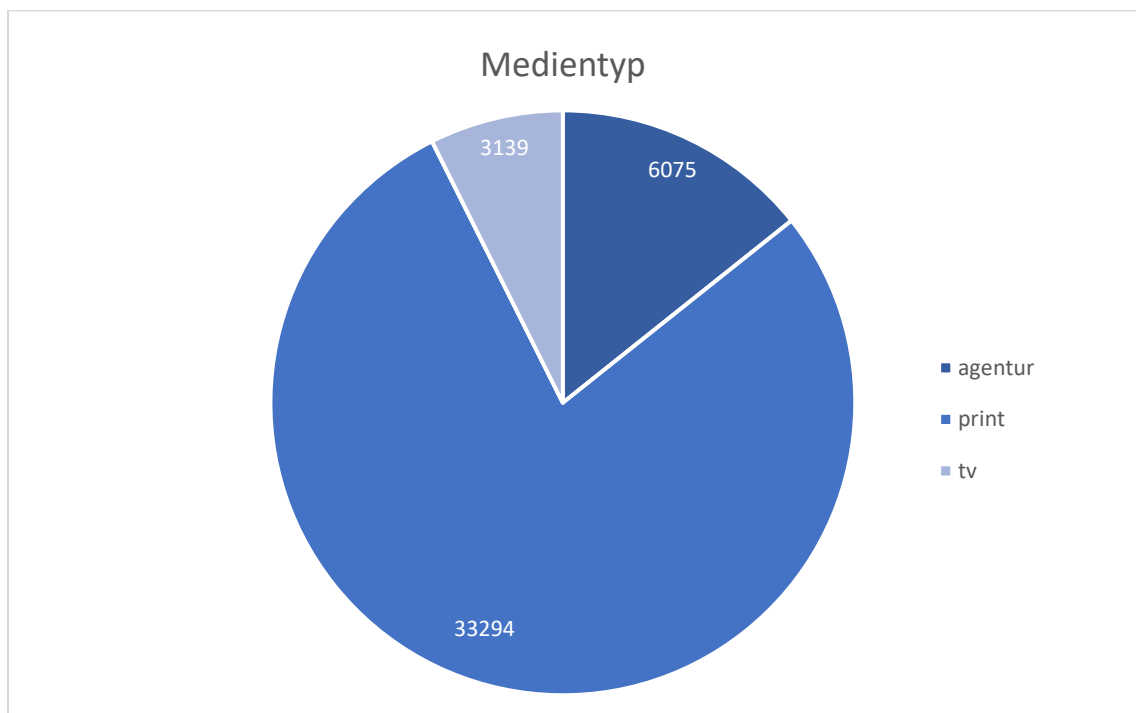


Abbildung 21: Aufteilung der *gehören*-Passiv-Belege auf die drei Medientypen **agentur**, **print** und **tv** (absolut)

So ist es nicht überraschend, dass ein Großteil (über 78 Prozent) der absoluten Belege auf den Medientyp **print** entfallen. Dies stellt jedoch ein verzerrtes Bild dar.

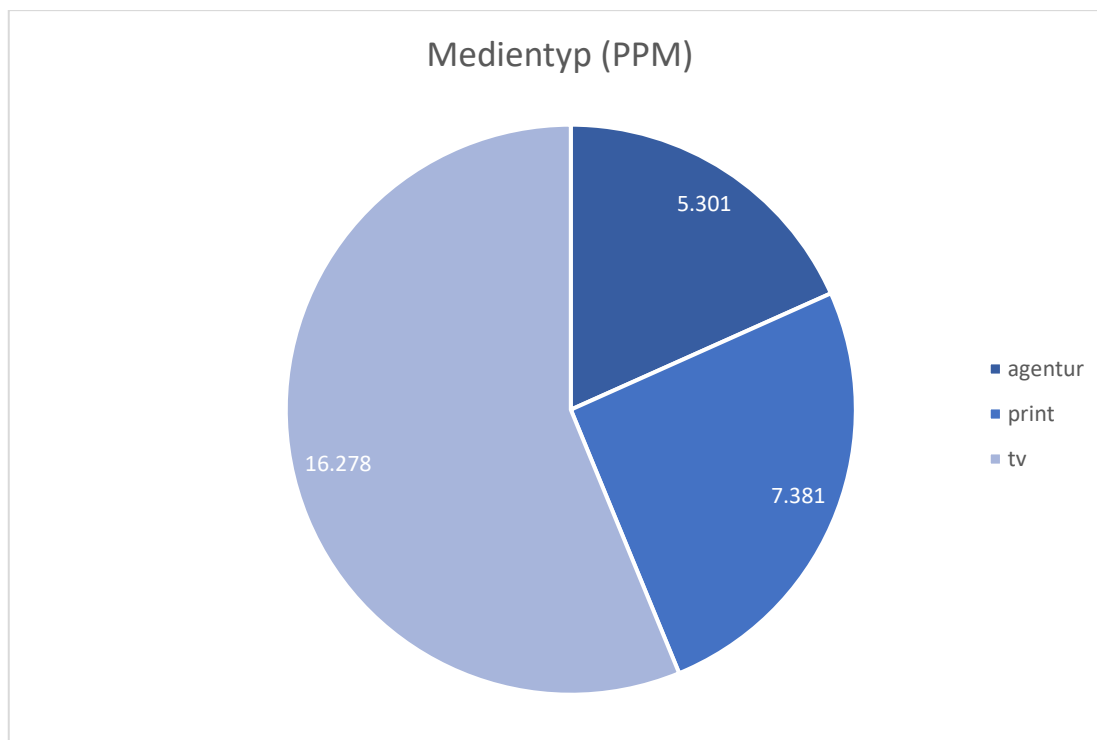


Abbildung 22: Aufteilung der gehören-Passiv-Belege auf die drei Medientypen **agentur**, **print** und **tv** (normalisiert)

Am deutlichsten sieht man die Verhältnisse, wenn man den Blick auf die normalisierten Treffer wirft: Während bei **agentur** gehören-Passiv-Belege am wenigsten vorkommen (ca. 5,3 PPM) und sich diese Anzahl für **print** nur etwas erhöht (7,4 PPM), sind bei **tv** deutlich mehr Belege vorhanden (16,3 PPM).

Medientyp	gehören-Passiv	Frequenz (PPM)
agentur	6175	5,301397256
print	33294	7,38060079
tv	3139	16,27804253
Gesamtergebnis	42608	7,26028691

Tabelle 5: absolute und normalisierte Belegzahlen zum gehören-Passiv nach Medientyp (**agentur**, **print**, **tv**)

Die Frequenz der gesamten Belege im Vergleich zu der Anzahl der Gesamttoken beläuft sich auf 7,26 PPM, also ähnlich den normalisierten Zahlen für den Medientyp **print**.

Interessant ist an dieser Stelle auch, dass die beiden Medien für **agentur** relativ weiter auseinander gehen: Während der OTS mit 10,6 PPM Belegen über dem Durchschnitt liegt, beläuft sich die Belegdichte bei APA-Meldungen digital auf 2,97 PPM.

Wie in Kapitel 3.1 beschrieben, werden im amc die Medien in unterschiedliche Regionen aufgeteilt, wobei zwei der Kategorien sich auf Gesamtösterreich beziehen, **agesamt** und **spezifisch**.

Regionen	Anzahl <i>gehören</i> -Passiv	Frequenz (PPM)
agesamt	21587	6,513
amitte	3104	7,038
aost	5559	6,985
asuedost	3585	8,119
awest	4569	8,698
spezifisch	4104	11,715

Tabelle 6: absolute und normalisierte Belegzahlen des *gehören*-Passivs nach Region (mitsamt **agesamt** und **spezifisch**)

Die absoluten Zahlen zu den Regionen vermitteln ein verzerrtes Bild, da für **agesamt** zwar mit Abstand am meisten Belege versammelt sind, aber auch im amc-Korpus am meisten Token in dieser Kategorie vorhanden sind.

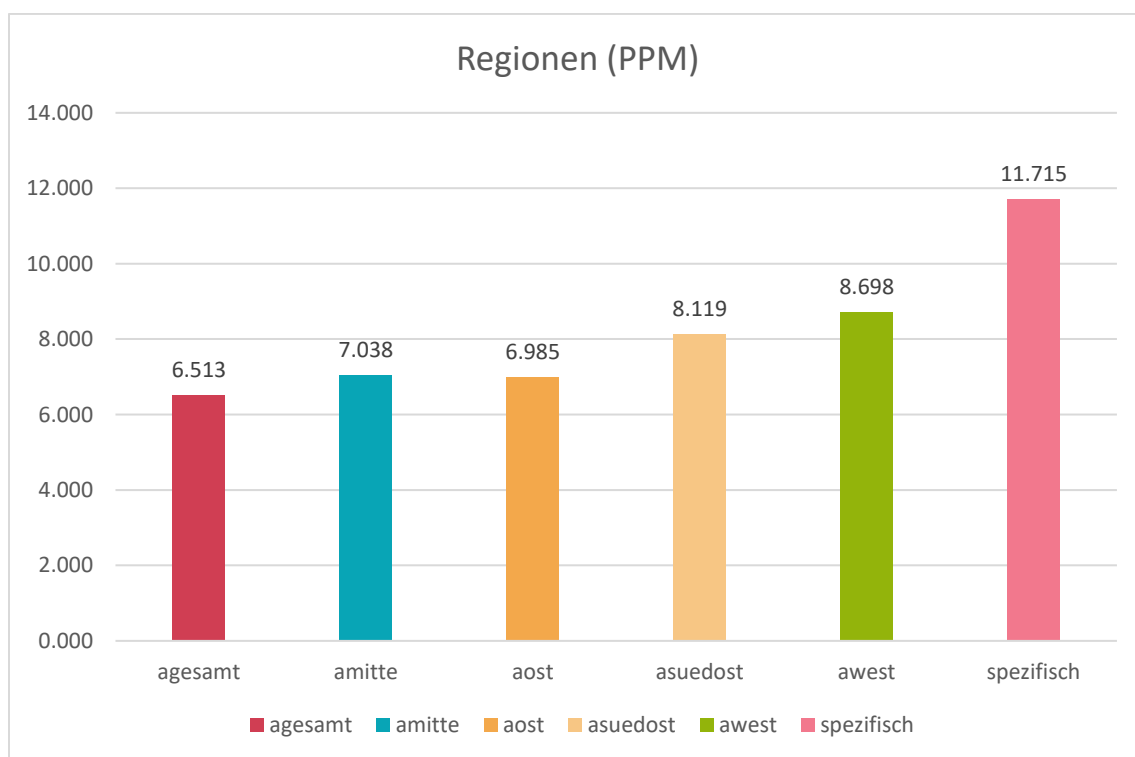


Abbildung 23: Belegzahlen des *gehören*-Passivs nach Region (mitsamt **agesamt** und **spezifisch**) (normalisiert)

Bei den normalisierten Zahlen zeigt sich genau das Gegenteil. Die Kategorie **spezifisch** weist die meisten *gehören*-Passiv-Belege in PPM auf.

Die einzelnen Regionen haben, wie in Kapitel 3.1 zum amc dargelegt, nur vier bis fünf Medien unter sich versammelt, die allesamt im **print**-Medientyp verortet sind.

In **agesamt** und **spezifisch** sind weitaus mehr Medien versammelt, wobei erstere weitaus mehr Tokens im Gesamtkorpus ausmachen als zweite.

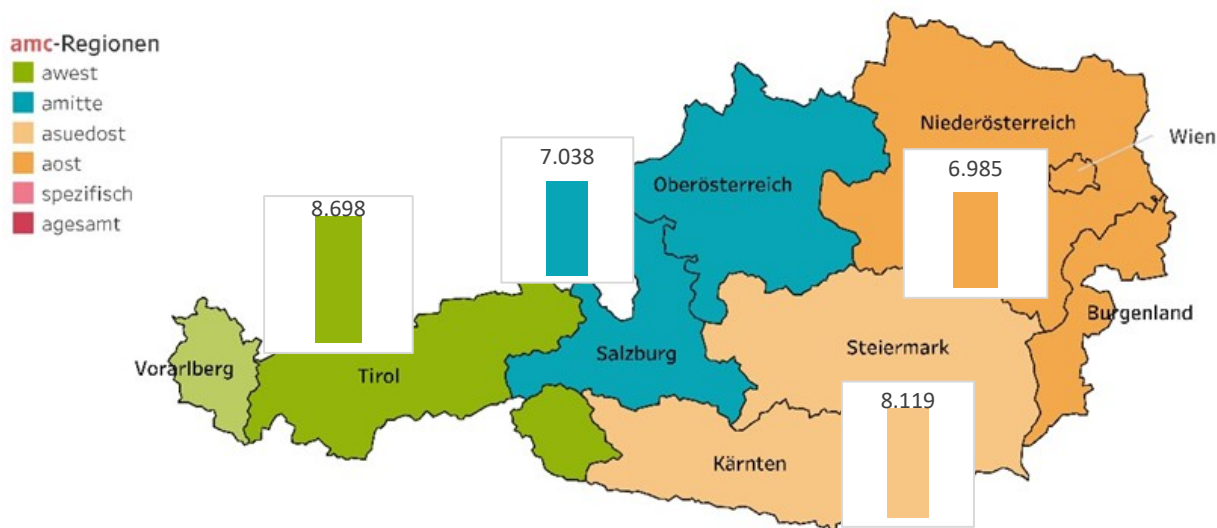


Abbildung 24: gehören-Passiv-Belege nach Regionen (normalisiert), dargestellt auf der Landkarte CC-BY Theresa Ziegler

Es fällt ein leicht steigender Trend von Ost- nach West-Österreich auf, bzw. ein Unterschied zwischen Mittel- und Südbairisch, jedoch sind die Differenzen eher gering und eine Aussage über das „Warum“ wäre bloße Spekulation.

4.2 Morphologie des gehören-Verbs

Dieses Unterkapitel beschäftigt sich mit der morphologischen Gestalt des *gehören*-Verbs, das wie ein Auxiliar verwendet wird, um die Passiv-Konstruktion zu bilden.

In der hier vorliegenden Datensammlung kommen diese fünf Verbformen vor, wobei mit Abstand die meisten Belege auf *gehört* entfallen, etwas über die Hälfte (50,73 Prozent), gefolgt von *gehören* mit 34,05 Prozent, was circa einem Drittel der Gesamtanzahl entspricht. Die übrigen 15,22 Prozent entfallen auf die Formen *gehöre*, *gehörte*, *gehörten*.

Gehörst wird von den Parsern nicht erkannt: Das Problem hierbei könnte die kombinierte Suche via CQL-Abfrage sein, sodass dies nicht am *gehörst* liegt (dieses scheint morphologisch eindeutig), sondern an den VVPPs, mit denen die Konstruktion gebildet wird, oder es besteht die Möglichkeit, dass diese aufgrund syntaktischer Ungewöhnlichkeiten nicht in das strenge Schema passen, und deshalb – zumindest in der hier gewählten Abfrageweise – nicht in den Belegen auftauchen.

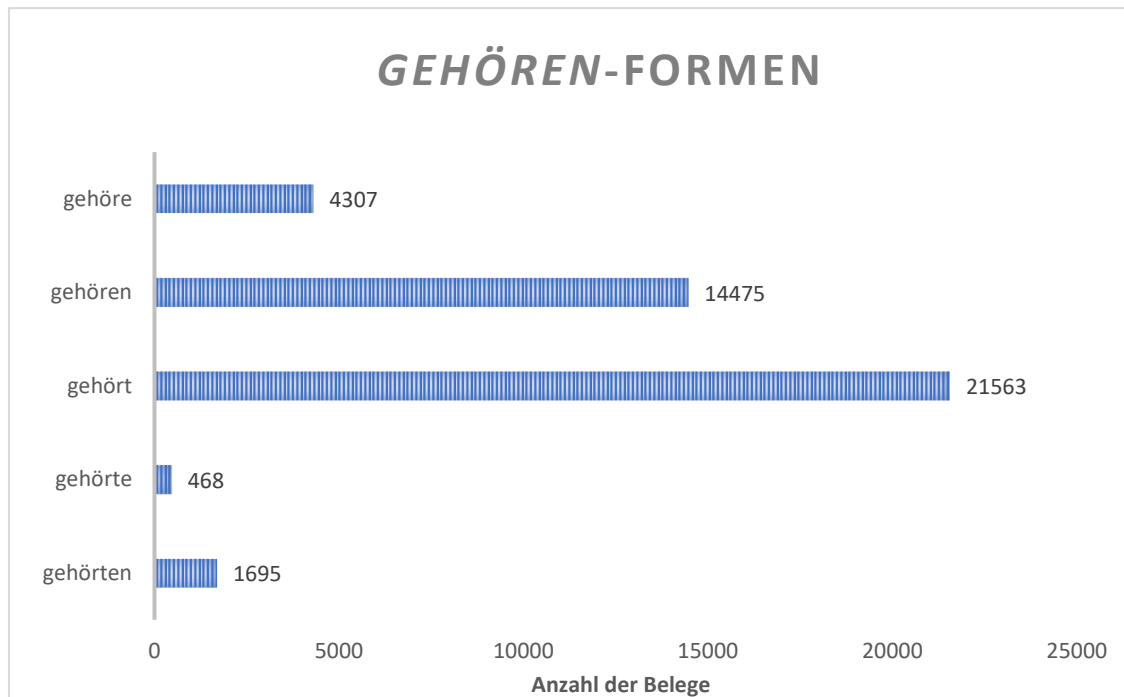


Abbildung 25: Die Anzahl der gehören-Formen im Korpus (absolut)

Ein Problem bei der Interpretation dieser Ergebnisse ist die Formgleichheit im Paradigma:

- *gehört*: *er/sie/es gehört* (3P Sgl Indikativ Präsens), *gehört* (Perfektpartizip), *ihr gehört* (2P Pl Indikativ Präsens)
- *gehöre*: *ich gehöre* (1P Sgl Indikativ Präsens), *er/sie/es gehöre* (3P Sgl Konjunktiv Präsens)
- *gehörte*: *er/sie/es gehörte* (3P Sgl Konjunktiv Präteritum)
- *gehören*: *gehören* (Indikativ), *wir gehören* (1P Pl Indikativ Präsens), *sie gehören* (3P Pl Indikativ Präsens)
- *gehörten*: *wir gehörten* (1P Pl Konjunktiv Präteritum), *sie gehörten* (3P Pl Konjunktiv Präteritum)

Die infiniten Verbformen werden hier vollständigshalber angeführt, vor allem der Indikativ *gehören* wäre sehr ungewöhnlich. Das Auftauchen von *gehören* als Perfektpartizip ist theoretisch möglich, wenn beispielsweise Plusquamperfekt-Formen gebildet werden, beispielsweise *Das hätte schon viel früher gemacht gehört* (siehe Kapitel 2.3; Stathi 2010, 332).

Ein Blick in die qualitativen Daten kann hier Aufschlüsse geben, welche der Formen die jeweils dominierenden sind:

Verb- bestimmung	<i>gehöre</i>	<i>gehören</i>	<i>gehört</i>	<i>gehörte</i>	<i>gehörten</i>	Gesamt- ergebnis
Indikativ		169	226			395
Präsens		169	226			395
3P Pl		169				169
3P Sgl			226			226
Konjunktiv	46			5	26	77
Präsens	46					46
3P Sgl	46					46
Präteritum				5	26	31
3P Pl					26	26
3P Sgl				5		5
Gesamtergebnis	46	169	226	5	26	472

Tabelle 7: Formen von *gehören* morphologisch aufgeschlüsselt, qualitatives Sample (absolut)

In der hier vorliegenden Tabelle ist es für die jeweilige Form eindeutig, welche Verbbestimmung ihr zuteilwird. Die Belege für *gehört* entfallen auf *er/sie/es gehört* (3P Sgl Indikativ Präsens) und für *gehören* auf *sie gehören* (3P Pl Indikativ Präsens). *Gehöre* ist ausschließlich für 3P Sgl Konjunktiv Präsens vorgesehen (keine 1P Indikativ Formen), *gehörte* und *gehörten* jeweils für 3P Konjunktiv Präteritum, einmal im Singular, einmal im Plural.

Es wäre jedoch ein Trugschluss anzunehmen, dass andere mögliche Formen nicht existieren, sie stellen jedoch eindeutig Ausnahmen dar. Einige Beispiele sind entweder im Zuge der Filterung der Daten oder bei Stichproben aufgefallen, beispielsweise 1. oder 2. Person Plural Indikativ Präsens.

(K40762) „Sollten wir dann noch den Aufstieg vergeigen, gehören wir erschlagen.“
OBERRUND_2012_05 (Unterstreichung C.M.)

(K45328) „I tritt euch die Tür ein, ihr gehört's alle weggeräumt, I stich euch ab!“
BVZ_2022_0609 (Unterstreichung, C.M.)

Auffällig an diesen Beispielen ist, dass sie allesamt in der direkten Rede stehen sowie weitere Besonderheiten mündlicher und substandardsprachlicher Sprechweisen vorkommen. Das Zitat aus der Burgenländischen Volkszeitung zeigt beispielsweise eindeutige Marker von schriftlich festgehaltenem Dialekt, wie *i* statt *ich* und das klitisierte Personalpronomen *'s* am finiten Prädikatteil; in K40762 wird *vergeigen* als Verb im Nebensatz verwendet.

Da vor allem im quantitativen Korpus nur das KWIC extrahiert und aufgrund dieses Umstandes nicht auf weitere Prädikatsteile kontrolliert werden kann, bzw. die Einschätzung der Parser, ob es sich beim *Gehört* um ein VVPP handelt, aufgrund der verkürzten Sätze nicht in der Filterung beachtet wurde, lässt sich nur schwer einschätzen, wie häufig analytische

Bildungen mit *gehören* ebenfalls im Partizip in der Konstruktion tatsächlich vorkommen.⁵⁰
Es finden sich allerdings vereinzelt Beispiele wie folgendes:

(K1635) Eigentlich hätte Mickelson dafür disqualifiziert gehört, meinten einige Spieler.
APA_2018_0618

Hier wird eine Konjunktiv-II-Form im Plusquamperfekt gebildet, *gehören* selbst ist in der rechten Satzklammer als Perfektpartizip vorhanden. Diese Unterscheidung kann im quantitativen Korpus aufgrund der verkürzten Beispiele nicht getroffen werden, weshalb nur anekdotisch und eher zufälligerweise derartige Belege gefunden werden konnten.

Während im qualitativen Korpus alle *gehört*-Formen auf dritte Person Singular entfallen (siehe Tabelle 7), kann dies nicht für das quantitative ausgeschlossen werden – wie eben auch das Vorkommen anderer Personen/Anzahl-Formen.

Im Allgemeinen ist jedoch die Konstruktion mit 3. Person am häufigsten – unabhängig von Singular und Plural – gemeinsam mit Indikativ Präsens, auch wenn es eine nicht unbeachtliche Anzahl an Konjunktiv-Belegen in den Daten gibt, sowohl in Präsens- als auch Präteritum-Formen. Auf diesen Umstand wird im Kapitel zur Diskussion noch ein weiterer Blick geworfen.

4.3 Semantik der VVPPs

Dieses Unterkapitel beschäftigt sich mit dem anderen Verbteil des Prädikats, nämlich dem Perfektpartizip, das aufgrund seiner Annotation im Korpus auch als VVPP (VollVerb Perfektpartizip) bezeichnet wird.

In Kapitel 4.1 wurden im Zuge der Datenreinigung ungewöhnliche Formen des Verbs besprochen, die von Parsern nicht erkannt werden, manche davon wurden hinzugefügt, wenn es sich dabei tatsächlich um *gehören*-Passiv-Belege gehandelt hat. Teilweise scheint es sich dabei um potenzielle mündliche ad-hoc Bildungen zu handeln, die einen sprechsprachlichen Duktus aufweisen (*entparteipolisiert* (K20740), *atterseeisiert* (K2521)), bei anderen handelt es sich eventuell um nicht- bzw. sub-standardsprachliche Wörter wie beispielsweise *zerdepert* (K12275) (vgl. Online-Duden: *zerdeppern*). Diese Verben stellen Ausnahmen dar, sie kommen teilweise nur einmal im Korpus vor.

⁵⁰ Auch im qualitativen Sample, das zum Zweck der genaueren Betrachtung vollständiger Sätze und „Ausnahmen“ gedacht war, kommen diese nicht vor.

Im Gegenzug dazu wird nun ein Blick auf die am häufigsten vorkommenden Verben im qualitativen und quantitativen Teil der Untersuchung geworfen:

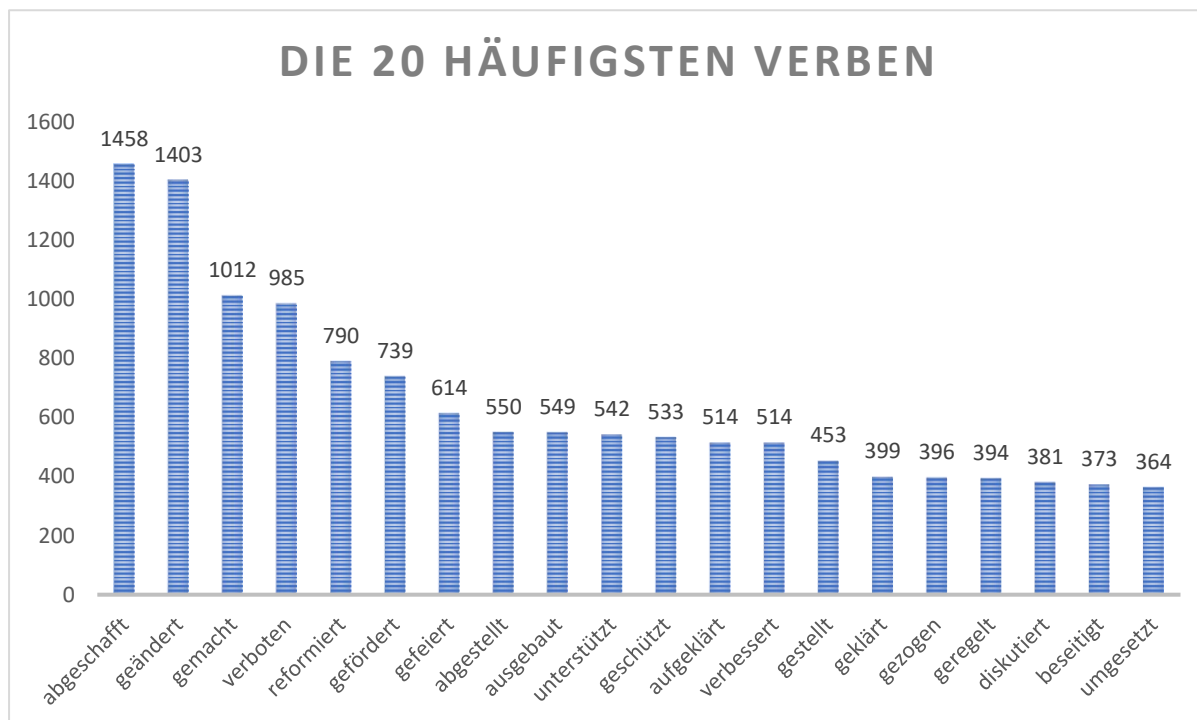


Abbildung 26: Die 20 häufigsten Verben im quantitativen Korpus (absolut)

Auf die 20 häufigsten Verben entfallen 12963 Belege, dies entspricht ungefähr 30,49 Prozent der Gesamtbelege, also stellen diese fast ein Drittel der VVPPs im Korpus dar. Insgesamt gibt es 2198 verschiedene Verben im Korpus, die jeweils mit einem Mittelwert von 19,34 – also gerundet 19 – Belegen vorkommen.

Im qualitativen Sample gibt es 472 valide Belege, die 239 verschiedene Perfektpartizipia aufweisen.

Die 20 häufigsten Verben im qualitativen Sample sind beinahe, aber nicht ganz deckungsgleich mit denen im quantitativen Subkorpus:

quantitativ	qualitativ
abgeschafft	abgeschafft
abgestellt	abgestellt
aufgeklärt	angepasst
ausgebaut	aufgeklärt
beseitigt	ausgebaut
diskutiert	ausgetauscht
geändert	geändert
gefeiert	gefeiert
gefördert	gefördert
geklärt	geklärt
gemacht	gelöst
geregelt	gemacht
geschützt	geregelt
gestellt	gezogen
gezogen	überarbeitet
reformiert	überprüft
umgesetzt	unterstützt
unterstützt	verändert
verbessert	verbessert
verboten	verboten

Tabelle 8: die 20 häufigsten Verben in beiden Teilen der Untersuchung, kontrastiert

Die grün markierten Verben sind jene, die übereinstimmen, insgesamt 14, sowie jeweils weitere sechs Verben, die sich sehr wohl im Rest des Korpus finden. Dies ist der Fall, da es sich um ein randomisiert ausgewähltes Sample handelt und die Verhältnisse der einzelnen VVPPs nicht kontrolliert wurden. Deshalb existiert hier ein Unterschied zwischen den beiden Teilen der Untersuchung, die grundlegend auf denselben Daten basieren.

Bei Betrachtung des qualitativen Samples wurde ein Beispiel für jedes der 20 am häufigsten vorkommenden Verben gewählt, auch um der gewünschten Kontextualisierung der quantitativen Datensammlung anhand vollständiger Sätze Rechnung zu tragen:

VVPP	Anzahl	Beispielsatz
verboten	18	S487 Sehr lange Tiertransporte oder viel zu enge Ställe gehören sofort verboten. TT_2022_1005
gemacht	16	S344 Bei ihm (dem Angeklagten, Anm. d. Red.) gehört doch ein psychiatrisches Gutachten gemacht. OOEN_2019_12110
geändert	15	S225 „Insgesamt besteht noch dringend Diskussionsbedarf, denn es gehört noch einiges geändert, bevor dieses Gesetz in Kraft treten darf.“ OTS_2017_0223
abgestellt	11	S83 Machtmissbrauch und Korruption gehörten abgestellt. OEREICHE_2013_0823
geregelt	10	S192 Ich glaube zum Beispiel nicht, dass der Jugendschutz länderspezifisch geregelt gehört, weil sich Jugendliche über Grenzen hinweg amüsieren. VN_2013_1015

abgeschafft	10	S328 Ist das Sitzenbleiben noch zeitgemäß oder gehört es abgeschafft? OOEN_2013_1011
verbessert	8	S314 Der öffentliche Nahverkehr und die Zusammenarbeit mit Tirol gehörten verbessert, sagen sie. TT_2018_1011
ausgetauscht	7	S419 Meiner Meinung nach gehört die gesamte Regierung sofort ausgetauscht. NOEN_2014_0901
gezogen	7	S165 Geisterfahrer gehören rasch aus dem Verkehr gezogen, nicht nur auf der Straße. KRONE_2019_0317
gefeiert	6	S415 Klar ist nämlich: Cupsiege sind schön und gehören gefeiert. NOEN_2013_0930
gefördert	6	S69 „Doch das musische Talent dieses jungen Mannes gehörte aus meiner Sicht gefördert.“ NVT_2012_1207
geklärt	6	S483 „Diese Dinge gehören vor diesem Schöffensenat geklärt.“ TT_2022_0510
überprüft	6	S417 Alle zwei Jahre gehört jeder Feuerlöscher überprüft, damit er im Fall des Falles seinen Job tun kann: Den Brand zu löschen, bevor alles verloren ist. NOEN_2014_0402
unterstützt	6	S172 „In Krisen ist Solidarität wichtig. Schutzbedürftige gehören unterstützt.“ KRONE_2020_0315
aufgeklärt	5	S222 Die Korruptionsvorwürfe innerhalb von EULEX gehören ebenfalls restlos aufgeklärt, (...). OTS_2015_0224
gelöst	5	S326 Dieser für die betroffene Bevölkerung unzumutbare Zustand gehört schleunigst gelöst. OOEN_2011_0906
überarbeitet	5	S76 Sie wollen von Kepplinger, der auch für Veterinär-Agenden zuständig ist, wissen, ob das Hundehaltungsgesetz überarbeitet gehöre. OEREICHE_2010_1004
verändert	5	S54 In der Digitalisierung gehört das Breitbandregime dringend verändert. KURIER_2017_1224
angepasst	4	S471 Beispielsweise gehöre die Ausgleichszulage an die Armutsschwelle angepasst. KRONE_2022_0817
ausgebaut	4	S369 Die Suizid-Prävention, gerade im Jugendbereich, gehört massiv ausgebaut (...). MWVOLL_2020_0927

Tabelle 9: Beispiele für die häufigsten Verben aus dem qualitativen Sample

(S487) Sehr lange Tiertransporte oder viel zu enge Ställe gehören sofort verboten.
TT_2022_1005

(S369) Die Suizid-Prävention, gerade im Jugendbereich, gehört massiv ausgebaut (...).
MWVOLL_2020_0927

Die jeweiligen Kontexte der beiden Sätze machen die Schwierigkeit der Sentimentbeurteilung klar: *verboten* ist tendenziell negativ, jedoch sollen durch die Verbote die Zustände der Tierhaltung verbessert werden. *Ausbauen* ist eher als neutral oder gar positiv konnotiert, das, was jedoch ausgebaut werden soll, ist „Suizid-Prävention“. Der Kontext beschreibt also in beiden Fällen eine gewünschte Verbesserung der jeweiligen Situationen, die durch die Gesamtbedeutung der Sätze ausgedrückt wird.

Die Bedeutung für das Patiens hängt also von der Semantik des VVPPs, aber auch des gesamten Satzes ab, deren Sentiment nicht immer genau zu bestimmen ist.⁵¹ Wie positiv oder negativ muss eine Bedeutung sein, damit das Verb allein nicht mehr neutral ist? Welche Unterschiede lassen sich im Vergleich der Verben allein und der Gesamtaussage beobachten?

Bei einer weiteren Betrachtung der am häufigsten vorkommenden Verben und einem Versuch, diese nach Sentimenten einzuteilen, fällt auf, dass ohne weiteren Kontext in vielen der Fälle nicht eindeutig gesagt werden kann, ob das Verb negativ, neutral oder positiv ist. Die Einschätzungen versuchen, die inhärente Konnotation in die eine oder andere Richtung zu annotieren; für ein Großteil der Verben kommt es jedoch darauf an, in welchem Kontext das Verb eingebettet ist.

Quantitativ (K)	Qualitativ (S)
abgeschafft	abgeschafft
abgestellt	abgestellt
aufgeklärt	angepasst
ausgebaut	aufgeklärt
beseitigt	ausgebaut
diskutiert	ausgetauscht
geändert	geändert
gefeiert	gefeiert
gefördert	gefördert
geklärt	geklärt
gemacht	gelöst
geregelt	gemacht
geschützt	geregelt
gestellt	gezogen
gezogen	überarbeitet
reformiert	überprüft
umgesetzt	unterstützt
unterstützt	verändert
verbessert	verbessert
verboten	verboten

Tabelle 10: Sentimente der 20 häufigsten Verben im quantitativen Korpus und qualitativen Sample: positiv (grün), neutral (grau), negativ (orange)

⁵¹ Eine tiefergehende Untersuchung in Richtung Sentimentanalyse könnte hier aufschlussreich sein, jedoch waren einige Tests, die diesbezüglich aufgestellt wurden, nicht ertragreich, da manche Kontexte in der Concordance nicht vollständig aufgeführt sind und Teilsätze nur bedingte Aussagekraft haben, bzw. das getestete Package textblob nicht für alle einzelnen Verben Werte annotiert hat, und tiefergreifende Methoden angewandt werden müssten.

Ausnahmen sind hier Sätze mit VVPPs, die eindeutig positiv sind oder eine sehr stark negative, fast schon pejorative Bedeutung haben, da dadurch diese Bewertung eindeutiger ist. Die Diskussion dieser Ergebnisse und der Bedeutung der Konstruktion wird in Kapitel 5.2 *Semantik des gehören-Passivs* eingehender besprochen.

4.4 Transitivity der Verben

Grundsätzlich wird davon ausgegangen, dass vor allem transitive Verben wie bei einem *werden*-Passiv vorkommen; eine Ausnahme sind laut Forschungsliteratur unpersönliche Formen, bei denen meist ein eigentlich transitives Verb intransitiv verwendet wird (vgl. Szatmári 2002: 182).

Von den 472 Belegen belaufen sich insgesamt vier auf subjektlose Passive:

(S115) „Mit denen gehört abgefahren, frei nach Hannes Amesbauer!“ KLEINE_2018_0706

(S195) „Ein paar Mal gehört uns in den Arsch getreten, weil sonst geht nichts weiter.“ VN_2020_0124

(S294) Hier gehöre angesetzt, „damit auch die Sicherheit im Sektor erhöht wird“, fasst Schelling zusammen. STANDARD_2016_0312

(S403) „In diesem Verband stinkt es mehr als in den von ihm abzuholenden Müllsäcken, hier gehört ordentlich aufgeräumt, aber nicht durch die Raumpflegerin, sondern durch den Staatsanwalt.“ NOEN_2011_0208

S403 stellt eine direkte Entsprechung zu dem Beleg dar, der bei Szatmári (2002: 182) als Beispiel für das unpersönliche Passiv verwendet wird. *Aufräumen* kann sowohl mit als auch ohne Akkusativ-Komplement verwendet werden und stellt in zweiterem Fall oft die Bedeutung ‚eine Wohnung, ein Zimmer in Ordnung bringen‘ dar (vgl. Grammis: *aufräumen*). Auch *ansetzen* scheint sowohl transitiv als auch intransitiv verwendbar zu sein.

Jemandem in den Arsch treten und *mit jemandem abfahren* fallen jedoch aus diesem Muster heraus, wobei es sich hierbei um – mehr oder weniger – gefestigte Formulierungen ohne Akkusativobjekt handelt, die hier gemeinsam mit dem *gehören*-Passiv zusammengebracht werden und aus diesem Grund subjektlos sind.

Diese Beispiele wurden auch in direkter Rede geäußert, weshalb es auch möglich ist, dass es sich um nicht standardsprachliche Ausdrücke handelt. Mehr zu diesem Umstand in Kapitel 4.7 *Aspekt der Mündlichkeit* und 5.3 *Weitere Aspekte der Konstruktion*.

4.5 Subjekte

In diesem Kapitel wird zuerst auf die formale Gestalt des Subjekts eingegangen, die die Realisierung des Subjekts in den Belegen des qualitativen Samples betrachtet.

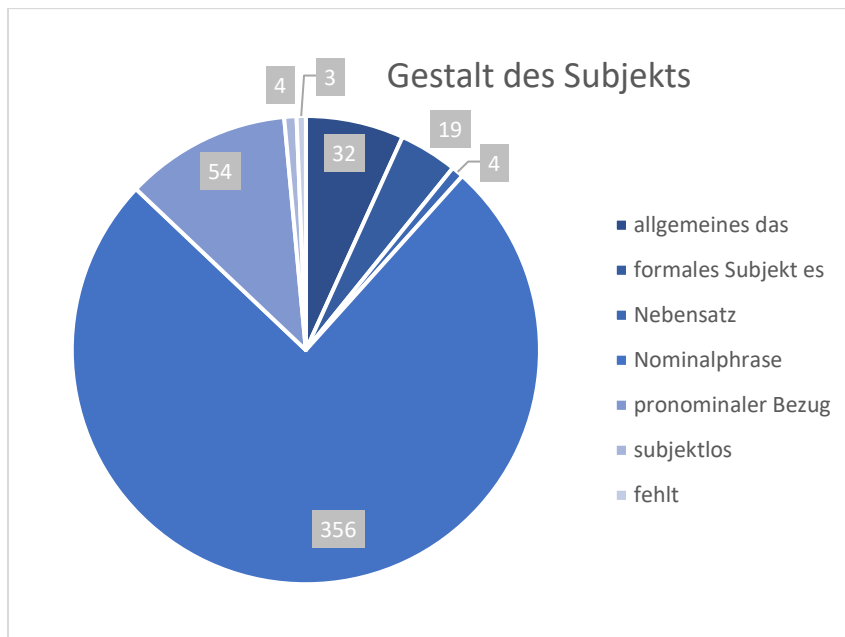


Abbildung 27: Gestalt des Subjekts im qualitativen Sample

Die meisten Belege entfallen auf Nominalphrasen, die entweder ein Substantiv oder ein Pronomen als Kopf besitzen, nämlich 356 von ihnen. Das entspricht 75,42 Prozent der Belege des qualitativen Samples.

Zudem gibt es 54 weitere Belege mit verweisenden Pronomen, die zumeist einen pronominalen Bezug auf Nominalphrasen herstellen. Subjektlose Beispielsätze kommen fast gar nicht vor ($n=4$; Belege siehe Kapitel 4.3), aber häufiger wird ein *das* ($n=32$) oder *es* ($n=19$) als Subjekt verwendet, das sich auf den gesamten Sachverhalt bezieht, über den im Kontext gesprochen wird, anstatt sich auf eine konkrete Nominalphrase oder einen Nebensatz zu beziehen.

Das formale Subjekt *es* kommt dabei teilweise mit einem weiteren Subjekt vor, das das Patiens darstellt (S468), teilweise ist es das alleinige Subjekt des Satzes (S45).

(S468) „Ich hoffe, dass die Gruppe so beieinanderbleiben kann. Es gehören sicher ein paar Spieler ausgetauscht – das ist normal in dem Sport“, sagt Schiechl (...) KLEINE_2022_0306

(S45) „Es gehört mit einem Masterplan durchstrukturiert, wie die Umsetzung zu erfolgen hat.“ KURIER_20140929

Zudem wird vier Mal auch ein Nebensatz als Subjekt verwendet, ohne ein formales *es* hinzuzufügen.

(S152) Wer aber glaubt, sich hinter einer angeblichen Anonymität alles erlauben zu können, gehört ganz streng bestraft. KRONE_2016_0626

(S310) Dass ausgerechnet die Staatsfirmen Verbund und ÖBB noch immer auch Kernenergie verkaufen, gehört sofort abgestellt. TT_2014_0911

(S381) „Für uns gehört klar geregelt, wer die Kosten für welche Dinge trägt und auch, wer wofür haftet.“ BVZ_2016_0803

(S490) Faktum ist, dass zuerst einmal unzweifelhaft geklärt gehört, ob und in welchem Ausmaß tatsächlich Finanzdoping der Bullen-Boliden stattgefunden hat. WZ_2022_1012

Nicht nur die Struktur der Subjekte wurde grob annotiert, sondern auch die Semantik, was sich primär auf die Einteilung in **belebt** und **unbelebt** bezieht, ähnlich wie sie auch bei Szatmári (2002: 182) vorgenommen wurde.

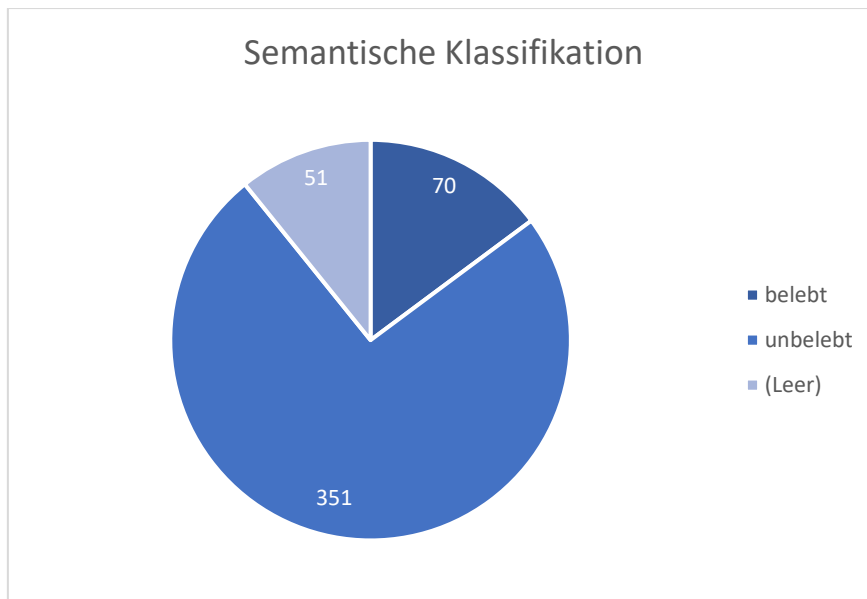


Abbildung 28: Einteilung der Subjekte in **belebt** und **unbelebt**

Die 51 nicht-annotierten Belege entfallen auf **subjektlos**, die drei fehlenden Belege sowie das formale *das* und *es*, da diese in dieser Hinsicht auf nicht genauer bestimmbare Kontexte bezogen sind und diese Unterscheidung daher nicht eindeutig getroffen werden kann.

Der Großteil der Belege ist beläuft sich auf unbelebte Subjekte (n=351; 74,4% der Gesamtbelege), also Abstrakta und Konkreta, Sachverhalte und Dinge, mit denen etwas gemacht gehört.

Die letzte Kategorie ist **belebt**, also menschliche und tierische Subjekte, mit 70 Belegen, wobei 69 davon auf Personen und Personengruppen entfallen.

Kategorisierung	Anzahl gehören-Passiv
Menschlich	69
Tierisch	1
Unbelebt/ohne Subjekt	401
Gesamtergebnis	472

Tabelle 11: Unterteilung der belebten Subjekte in **tierisch** und **menschlich**, qualitatives Sample

Dazu gehören namentliche Nennungen wie *Trump* oder *die Grüne Politikerin Roth*, Personengruppen, die nach bestimmten sozio-demographischen Gesichtspunkten eingeteilt werden können, wie *Familien*, *Jugendliche* oder *Frauen*, Berufsgruppen und Branchenbezeichnungen wie *Lehrer*, *Politiker* und *Unternehmer* (ohne die entsprechenden Entgenderungen), sowie Bezeichnungen für marginalisierte Gruppen und Minderheiten (teilweise auch mit einer entsprechenden Wortwahl) oder scheinbare ad-hoc Bildungen wie *die Gruppe Menschen*, *der sitzenden Art Tanzmuffel*. Eine Auflistung aller Subjekte, die im qualitativen Sample vorkommen, findet sich in Anhang C (Tabelle C.4).

Die Überlegung wurde angestellt, für die menschlichen Subjekte Beneficiens und Malificiens zu annotieren; ebenso wie die Beurteilung des Sentiments der Verben lässt sich das vor allem in den betont positiven und stark negativen Bedeutungen eindeutig bestimmen, ansonsten obliegt dies teilweise der spezifischen – und auch subjektiven – Interpretation der Kontexte und der VVPPs (siehe auch Kapitel 4.3 *Semantik der VVPPs*). Aus diesem Grund wurde darauf verzichtet.

Es lässt sich aber festhalten, dass besonders bei den menschlichen Subjekten auch jene Fälle vorkommen, die stark negative bis hin zu pejorisierende Bedeutungen aufweisen. Zur weiteren Diskussion wird auf Kapitel 5.2 zur Semantik der Konstruktion verwiesen.

Das Beleg, bei dem tierisch annotiert wurde, zeigt, welche Sätze ebenfalls mit dem *gehören*-Passiv gebildet werden können und wie direkte Reden teilweise in den Text integriert werden.

(S044) Honigbienen seien „wichtige Dienstleister (Bestäubung) und Produzenten (Honig), die unterstützt gehören“. Kurier_2014_0619

Es handelt sich hierbei um Gleichsetzungsnominative, die *Honigbienen* als *Dienstleister und Produzenten* auszeichnen, es wird dann im Nebensatz mit *die* darauf Bezug genommen, wodurch eine Kette aus Verweisen entsteht, worum es sich bei dem Subjekt der *gehören*-Konstruktion handelt. Aus diesem Grund wurde es als **tierisch** annotiert, auch weil sich die Frage stellt, wie *Dienstleister und Produzenten* in diesem Fall zu annotieren wäre, wenn **menschlich** als Klassifikation wegfällt.

Grundsätzlich sei festgestellt, dass ein Großteil der Subjekte unbelebte Nominalphrasen ausmacht, ein kleinerer Teil entfällt auf grammatikalisch-formale Subjekte, die beispielsweise auf den größeren Kontext, in dem die Formulierung eingebettet ist, verweisen, sowie auf Personen und Personengruppen, mit denen ‚etwas gemacht gehört‘.

4.6 Präpositionalphrasen

Bei den Präpositionalphrasen handelt es sich um zwei verschiedene Aspekte, die untersucht wurden, die primär nichts miteinander zu tun haben, beide jedoch konkret bei Stathi (2010, 327–329; 338) angesprochen wurden.

4.6.1 Vorhandensein einer lokativen Präpositionalphrase

Um die lokativen Präpositionalphrasen (PP) genauer betrachten zu können, wurden zunächst alle in Frage kommenden Phrasen aus dem qualitativen Korpus extrahiert – jene PPs, die eine für lokale Attribution typische Präposition aufweisen – und in einem zweiten Schritt versucht, zu bestimmen, ob es sich tatsächlich um eine lokative Präpositionalphrase handelt und inwieweit man sie in statisch oder direktiv einteilen kann.

In manchen Belegen ist die Unterscheidung zwischen einer lokativen Präpositionalphrase, einer fakultativen Ergänzung oder einer anders gearteten Präpositionalphrase nicht eindeutig zu treffen, da die einzelnen Ausdrücke teilweise stark konventionalisiert sind. Ein erster Blick auf die Belege zeigt die Problematik der Annotation:

Lokative Präpositionalphrase	Anzahl <i>gehören-Passiv</i>
am Müllhaufen der politischen Inkompetenz abgelegt	1
auf das Abstellgleis verfrachtet	1
auf den Asphalt geheftet	1
auf den rechten Weg gebracht	1
auf die Parallelstraße verwiesen	1
aufs Schild gehoben	1
aus dem Verkehr gezogen	6
aus der Welt verbannt	1
aus ihren Sesseln geschleudert	1
in Angriff genommen	1
in den Arsch getreten	1
in den Hintergrund gerückt	1
in die Pflicht genommen	1
in die Realität zurück geholt	1
in Verwahrung genommen	1
ins Visier genommen	1
vor Gericht gestellt	1
Gesamtergebnis	22

Tabelle 12: lokative Präpositionalphrasen im qualitativen Sample

Auf das Abstellgleis (gestellt) ist eines der Beispiele dafür, bei denen im Gegensatz zum ursprünglichen lokativen Kontext, der bei Stathi (2010: 327–328) angeführt wird, die Phrase im übertragenen Sinne gemeint ist; aus diesem Grund wird auch zusätzlich eine

metaphorische Verwendung annotiert. Von den 22 Belegen mit lokativer Präpositionalphrase sind acht metaphorisch.

Lokative Präpositionalphrasen, auch mit der Modifikation, dass sie im übertragenen Sinne gemeint sein können, kommen also insgesamt nicht sehr häufig vor. Aufgrund des prozentuellen Anteils (4,7 Prozent) lässt sich grob abschätzen, wie viele der betrachteten Belege im quantitativen Teil der Untersuchung eine derartige Präpositionalphrase aufweisen, nämlich ungefähr 1981.

Aus dem Verkehr (gezogen) ist die einzige Präpositionalphrase, die öfter als einmal in dem Sample vorkommt; es handelt sich hierbei um eine gefestigte Formulierung.

(S448) Grob faktenresistente Ärzte gehören sofort aus dem Verkehr gezogen.
PROFIL_2015_0309

(S55) Geisterfahrer gehören rasch aus dem Verkehr gezogen. KURIER_2019_0317

Manche der betreffenden Beispielen sind metaphorisch gemeint (S448), manche referieren tatsächlich auf Verkehrsgeschehen (S55).

Art der Präposition	Anzahl gehören-Passiv
direktiv	21
statisch	1
Keine Präpositionalphrase	450
Gesamtergebnis	472

Tabelle 13: direkte oder statische Lesart der lokativen Präpositionen

Von den 22 Präpositionalphrasen sind alle bis auf eine Ausnahme (S480 *am Müllhaufen der politischen Inkompetenz abgelegt*) direktiv.

4.6.2 Vorhandensein eines Agens-Anschluss

Von den 472 Belegen im qualitativen Sample sind nur vier Sätze, die einen Agensanschluss aufweisen, von denen wird eines mit *durch* und die anderen drei mit *von* gebildet.

Agensanschluss	Anzahl gehören-Passiv
durch die Versicherten	1
von einem niedergelassenen Arzt	1
von uns	2
Kein Agensanschluss	468
Gesamtergebnis	472

Tabelle 14: Agensanschluss in gehören-Passiv-Belegen

Dies bedeutet, dass gerundet 0,85 Prozent der Belege einen Agensanschluss aufweisen, was in Bezug auf das Korpus eine ungefähre Anzahl von 360 *gehören*-Passiven ergibt, in denen der Agens in einer Präpositionalphrase explizit erwähnt wird. Er stellt also eine Ausnahme dar.

4.7 Aspekt der Mündlichkeit

Bei der Untersuchung des Aspekts der Mündlichkeit bestand am Anfang die Herausforderung darin, wie dieser Umstand in dem schriftlichen Korpus überhaupt zu betrachten ist. Durch unterschiedliche Zugänge, die sich aus den Daten und den spezifisch durchgeführten Annotationen im qualitativen Sample ableiten, soll ein Blick darauf geworfen werden.

Einer der Punkte sind die Ressorts, in denen die *gehören*-Passive auftauchen, mithilfe derer ursprünglich auch die Belege im quantitativen Korpus kontextualisiert werden sollen; es stellte sich jedoch heraus, dass diese Ergebnisse weniger aufschlussreich sind als erhofft.

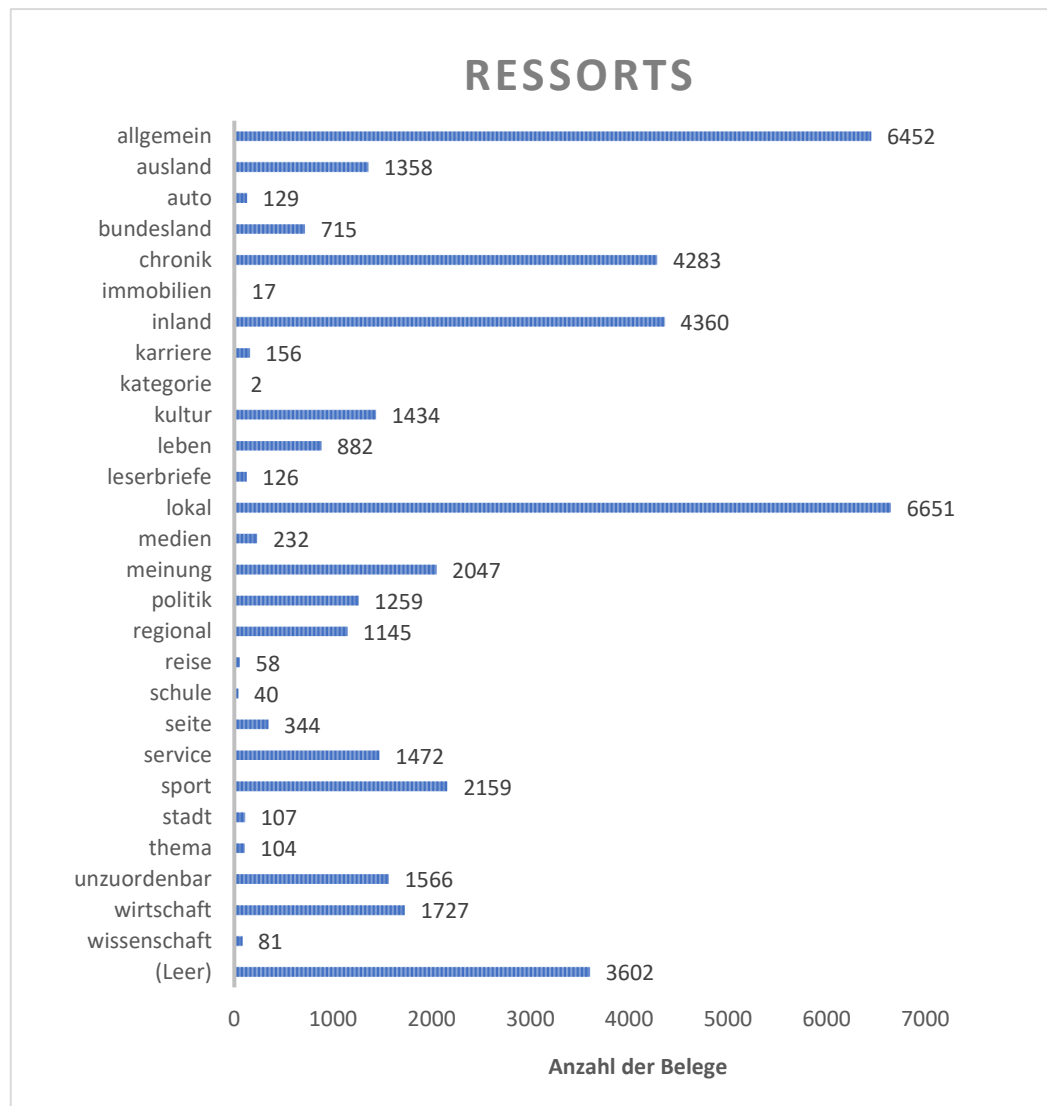


Abbildung 29: Die Ressortzuteilungen im quantitativen Korpus (nur erste Zuteilung) (absolut)

Die fünf häufigsten Primärzuteilungen sind in absteigender Reihenfolge abgebildet: **lokal** (n=6651), **allgemein** (n=6452), **inland** (n=4360), **chronik** (n=4283) und keine Zuteilung („Leer“, n= 3602). Dies sind allesamt Ressorts, die eher allgemein gehalten sind.

Die Ressorts **meinung** (n=2047) und **leserbriefe** (n=126) sind jedenfalls vertreten, stehen allerdings bezüglich ihrer Anzahl nicht explizit heraus.

Ein eindeutiger Zusammenhang zwischen bestimmten Ressortzuteilungen und dem *gehören*-Passiv kann aus den Daten nicht erschlossen werden; es ist jedoch festzuhalten, dass – aus der Benennung zu schließen – betont meinungsbasierte Ressorts nicht dominieren.

Anhand der erstgereihten Ressorts allein die tatsächliche Zuteilung festzumachen, stellt sich für das quantitative Korpus als herausfordernd dar, jedoch geben auch die genaueren Einteilungen im qualitativen Sample nur bedingt weitere Aufschlüsse, unter anderem aufgrund der bereits genannten fragwürdigen Zuordnung und der detaillierten Einteilung, die wiederum einen Vergleich erschwert.

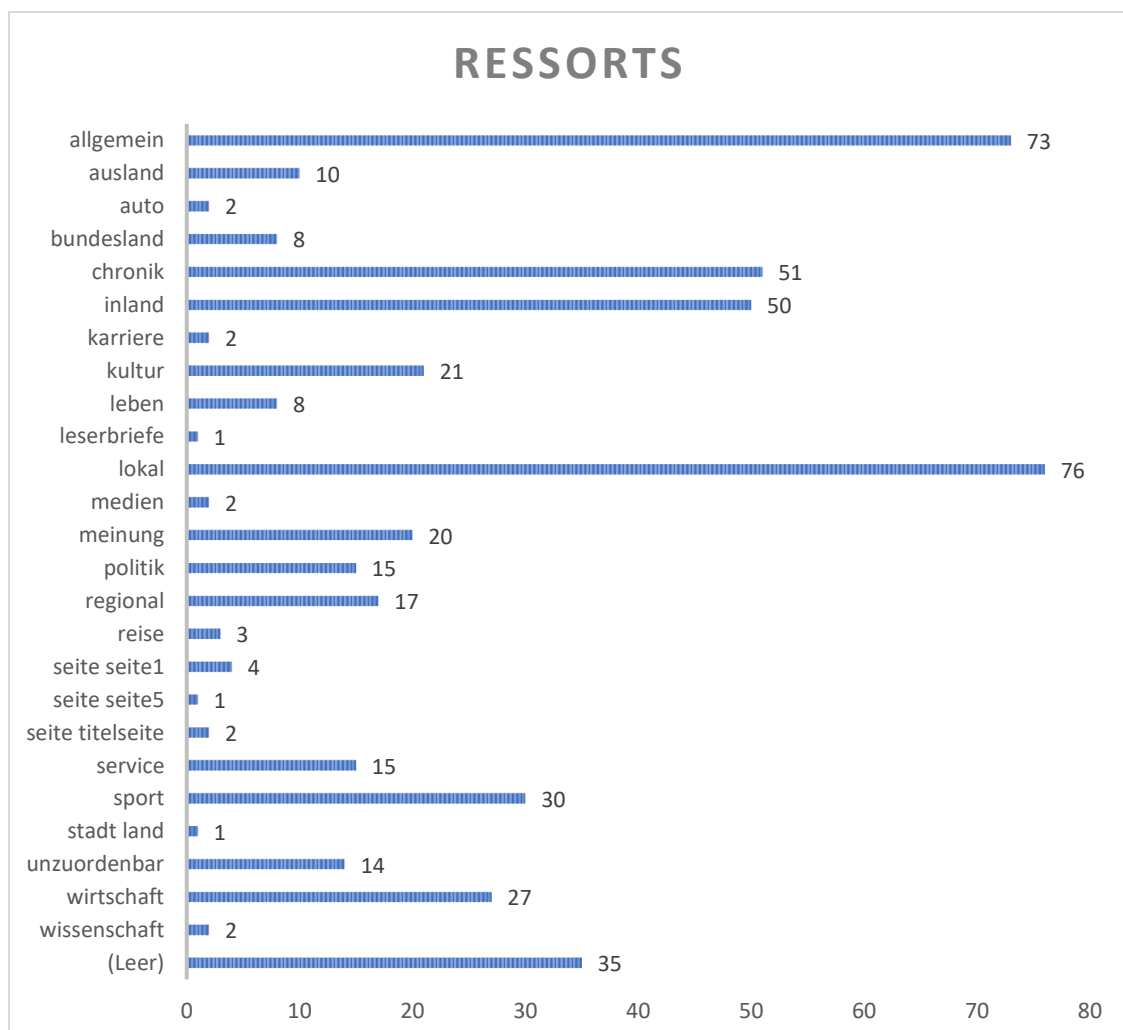


Abbildung 30: Zuteilung zu den Ressorts im qualitativen Sample (nach Gruppierung der Erstreichung) (absolut)

Anhand des Ressorts **seite** wird ein Problem dieser Zuteilung deutlich, nämlich, dass hier teilweise pro Zeitung sehr unterschiedliche Themen unter einem potenziell übergeordneten Ressort verstanden werden, da Titelseite, Seite 1 und Seite 5 zwar alle zu **seite** gezählt werden, aber sich mit hoher Sicherheit auf unterschiedliche Konzepte beziehen.

Genauso können diese erstgereihten Kategorien bei anderen als zweit-, dritt-, viert- oder fünftgereichte Zuordnung auftauchen. Hier wurde als Beispiel die Zuteilung zum Ressort **meinung** gewählt.

meinung	20
meinung	9
meinung leserbriefe	4
meinung leserbriefe service	4
meinung politik	2
meinung unzuordenbar	1

Es lässt sich dabei nicht differenzieren, was nun den Unterschied zwischen **meinung leserbriefe** und **meinung leserbriefe service** genau ausmacht. Diese Umstände sowie der Fakt, dass die Ressorts zur besseren Lesbarkeit für das amc gruppiert werden, und daher unklar ist, was die einzelnen Bezeichnungen nun genau bedeuten, also die genauen Ressort-Titel von Zeitung zu Zeitung differieren, machen die Betrachtung dieser Metainformation nicht so ergiebig wie erhofft; für die konzeptionelle Mündlichkeit sogar unbrauchbar.

Aus diesem Grund werden diesbezüglich größtenteils die annotierten Informationen aus dem qualitativen Korpus herangezogen. In diesem Zusammenhang wurden direkte und indirekte Rede, mit und ohne Anführungszeichen annotiert, sowie die Informationen zu Indikativ und Konjunktiv ein weiteres Mal betrachtet, um diesbezüglich Regularitäten herauszufinden.

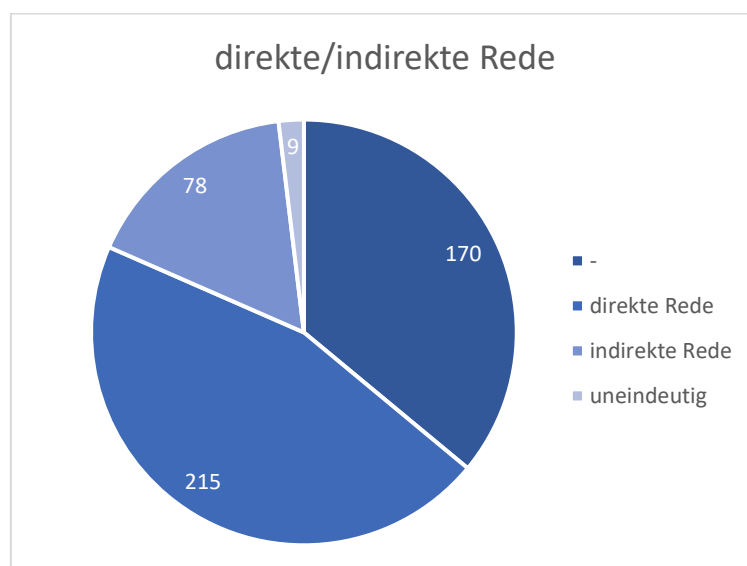


Abbildung 31: Klassifizierung der Redeweise im qualitativen Sample

Der Großteil der Belege entfällt auf direkte oder indirekte Rede, ein Umstand, der trotz der hypothetisierten Mündlichkeit der Konstruktion unerwartet hoch ausfällt, denn die Anzahl beläuft sich auf 62,1 Prozent, also beinahe zwei Drittel, der Belege im qualitativen Korpus. Die Mehrheit der Belege entfällt dabei auf direkte Rede (n=215), 78 Belege wurden als indirekte Rede annotiert, und 170 Belege fallen in keine der beiden Kategorien und gelten als nicht weiter differenzierte Aussagesätze.

Uneindeutig (n=9) beschreibt jene Belege, bei denen sich aus dem Kontext nicht unterscheiden lässt, ob es sich um direkte oder indirekte Rede oder eine nicht weiter differenzierte schriftliche Formulierung der Konstruktion handelt.

In Zusammenhang damit wurde ein Blick auf die Verbindung der Redeweise mit Modus des Verbes geworfen:

Modus	-	direkte Rede	indirekte Rede	uneindeutig	Gesamtergebnis
Indikativ	161	213	14	8	396
Konjunktiv	9	2	64	1	76
Gesamtergebnis	170	215	78	9	472

Tabelle 15: Modus des Verbes und Klassifikation der Redeart im qualitativen Korpus

Insgesamt gab es bei der Betrachtung des qualitativen Samples Schwierigkeiten bei der Beurteilung, ob es sich um eine indirekte Rede oder eine Verwendung des Konjunktivs aus anderen Gründen handelt. Der Kontext gibt in den meisten Fällen ausreichend Aufschluss, man könnte jedoch dafür oder dagegen argumentieren, ob es ein einfacher Aussagesatz im Konjunktiv ist oder derart gekennzeichnete indirekte Rede.

Generell wurden die Belege, die im Indikativ gehalten wurden und keine Anführungszeichen als Marker für direkte Rede hatten, als einfacher Fließtext annotiert, Belege, die einen Konjunktiv aufweisen, hingegen tendenziell als indirekte Rede eingeschätzt.

Klassifikation	Modus		
Mit Anführungszeichen	Indikativ	Konjunktiv	Gesamtergebnis
-	161	9	170
nein	161	9	170
direkte Rede	213	2	215
ja	167	1	168
nein	46	1	47
indirekte Rede	14	64	78
nein	14	64	78
uneindeutig	8	1	9
nein	8	1	9
Gesamtergebnis	396	76	472

Tabelle 16: Klassifikation der Redeart, Verwendung von Anführungszeichen und Modus des Verbes im qualitativen Sample

Die Unterscheidung zwischen mit und ohne Anführungszeichen betrifft im Endeffekt nur die direkte Rede, da die übrigen Kategorien keine Anführungszeichen enthalten. Es gibt aber auch Fälle von direkter Rede, in denen andere Markierungen für die Mündlichkeit gewählt wurden, beispielweise in Interviews.

Zudem wurden alle Fälle der Transkriptionen im Medientyp **tv** als direkte Rede gewertet, von denen kommen n=28 Fälle im qualitativen Sample vor; bei diesen ist von mündlicher Äußerung auszugehen, die schriftlich für das Korpus festgehalten wurde.

Dies gilt auch für das quantitative Korpus, in dem 3139 Belege auf **tv** entfallen. Hier sollte auch noch einmal darauf hingewiesen werden, dass in normalisierten Zahlen die Kategorie **tv** mit Abstand die höchste Dichte an *gehören*-Passiven aufweist (siehe Kapitel 4.1).

In manchen Belegen finden sich unterschiedliche Diskursmarker wie „meiner Meinung nach“, die das Gesagte platzieren. Sie setzen also den Satz mit *gehören*-Passiv entweder in einen meinungsbetonten Kontext oder machen deutlich, dass es von jemandem direkt geäußert wurde; hierbei ist jedoch ohne weitere Hinweise aus dem gesamten Satz nicht zu unterscheiden, ob es sich um direkte mündlich geäußerte Rede handelt

(S419) T. B. (anonymisiert, C.M.), Kellnerin, Berndorf: Meiner Meinung nach gehört die gesamte Regierung sofort ausgetauscht. NOEN_2014_0901

Dieses Beispiel stammt augenscheinlich aus einem Leserbrief und zeigt, wie der Ausdruck des *gehören*-Passivs konzeptionell mündlich konnotiert ist, schriftlich geäußert und explizit in einen meinungsbetonten Kontext gebracht wird, der sich auf die sprechende Person bezieht, obwohl die Passivität tendenziell – auch implizit über die Agensdezentrierung – den*die Sprechende*n hintenanstellt.

5. Diskussion der Ergebnisse

In diesem Kapitel werden die Ergebnisse der Untersuchung unter den Gesichtspunkten der Forschungsliteratur aus dem Theorieteil betrachtet, vor allem in Bezug auf die grammatikalisierte Form des *gehören*-Passivs im Feld der Passivkonstruktionen im Allgemeinen.

Es wird zuerst auf die Grammatik, dann auf die Semantik eingegangen und im Anschluss werden weitere Aspekte der *gehören*-Passiv-Konstruktion beleuchtet, die über die grammatisch-semantische Gestaltung hinausgehen, im Besonderen die konzeptionelle Mündlichkeit und die damit zusammenhängende untersuchte Varietät. Zuletzt wird versucht, zusammenfassend den Status der Grammatikalisierung abzuschätzen.

5.1 Grammatik des gehören-Passivs

In diesem Kapitel wird die Grammatik von *gehören* betrachtet, vor allem im Paradigma der Formen sowie weitere grammat(ikal)ische Aspekte in der syntagmatischen Umgebung.

Gemeinsam mit den Belegen, die bei Stathi (2010: 331) im Paradigma vorkommen, sind für alle Personen und Anzahl zumindest vereinzelte Beispiele attestiert, da in der hier vorliegenden Untersuchung die in der Tabelle fehlende erste Person Plural nachgewiesen werden kann (siehe Beleg (K40762), Kapitel 4.2 *Morphologie des gehören-Verbs*).

Insgesamt sind allerdings andere Formen als die dritte Person Singular und Plural sowohl im Indikativ als auch Konjunktiv in der klaren Minderheit. Auch wenn man die einzelnen Formen im qualitativen Sample disambiguiert, entfällt der Großteil auf Formen der dritten Person Singular und Plural (siehe Kapitel 4.2).

Konjunktiv II Plusquamperfektformen sind bei Stathi (2010: 332) und in dieser Untersuchung attestiert. Das Vorkommen ist möglicherweise durch Konjunktiv II mitbedingt, der in einer periphrastischen Form ausgedrückt wird anstatt synthetisch, *Er gehörte disqualifiziert* im Vergleich zu *Er hätte disqualifiziert gehört* (siehe Beleg (K1635)). Hier muss eben angemerkt werden, dass die Beispiele in der direkten oder indirekten Rede vorkommen.

Die hohe Anzahl an Konjunktiven lässt sich im Zusammenhang mit dem Aspekt der Mündlichkeit erklären, nämlich dass es sich bei den Daten sehr häufig auch um indirekte Reden handelt, die unter anderem über die Konjunktivform markiert werden (siehe Kapitel 4.7 *Aspekt der Mündlichkeit*).

Andere Tempora sind in den betrachteten Daten nicht aufzufinden, auch wenn ihre Existenz nicht ausgeschlossen werden kann, da nur ein kleines Sample diesbezüglich genauer untersucht wurde.

Wenn man nun auf die Definition von Auxiliarverben blickt, die in Kapitel 2.2 angegeben wurde (siehe Amrhein 1995: 86–87, nach Szatmári 2002: 183), sieht man, dass das *gehören*-Passiv, wie es hier in der Untersuchung auftritt, grundsätzlich auf diese Weise verwendet wird. *Gehören* als Auxiliar ist „Träger der Verbflexion“, der andere Verbteil ist ein Partizip, das „als Träger des verbsemantischen Kerns im komplexen Prädikat fungiert“ (Amrhein 1995: 86–87, nach Szatmári 2002: 183). Der erste Punkt scheint erfüllt zu sein, aber wie auch angemerkt wird, geht dabei die Funktion über die grundlegende grammatische hinaus und es wird zudem ein modaler Aspekt hineingebracht. Diese wird in zwei der folgenden Kapitel zur Semantik (5.2) und Grammatikalisierung (5.4) der Konstruktion noch weiter besprochen.

So wie es für das Passiv typisch ist, wird eben der Agens eliminiert. Die Agensdezentrierung geht so weit, dass nicht nur das Patiens in das Fokus gestellt wird, sondern dass auch

der elidierte Agens nur selten angeschlossen wird. Agensanschlüsse in Form von Präpositionalphrasen kommen vor, sie sind jedoch Ausnahmen, sowohl in Stathi (2010: 338) Untersuchung, in der zwei Fälle vorkommen, als auch in der Betrachtung hier, mit vier Beispielen im qualitativen Korpus (siehe Kapitel 4.6.2).

Die andere Präpositionalphrase ist die lokative, die gemeinsam in der Konstruktion mit *gehören* nach wie vor, wenn auch nicht häufig vorkommt, meistens in direkter Verwendung. Es sind dabei Verben bzw. Ausdrücke, die die PP verlangen; im Beispiel von *aus dem Verkehr ziehen* handelt es sich sogar um eine gefestigte Wendung (siehe Kapitel 4.6.1).

Dies ist kein Gegenbeweis für den möglichen Grammatikalisierungspfad, der bei Stathi (2010: 329) angenommen wird, jedoch ist dieser für die Gestaltung der Konstruktion im Gegenwartsdeutschen nur von geringer Relevanz. Die Steuerungsfaktoren, ob eine lokative Präpositionalphrase verwendet wird oder nicht, scheinen nicht explizit an das *gehören*-Passiv gebunden zu sein; es ist fraglich, ob sie es jemals waren. Dieser Umstand wird in Kapitel 5.4 zur Grammatikalisierung ein weiteres Mal aufgegriffen.

Die Transitivität der Verben ist wie erwartet ähnlich dem *werden*-Passiv (siehe Kapitel 2.3). Es werden vorwiegend transitive Verben verwendet, die unterschiedliche Valenz-Wertigkeiten haben, einige wenige werden eben auch intransitiv verwendet (siehe Kapitel 4.4 *Transitivität der Verben*).

Subjektlose Belege kommen dadurch ebenfalls vor (siehe Kapitel 4.5); das wird in der Forschungsliteratur als unpersönliches Passiv bezeichnet (siehe Szatmári (2002: 182), Vogel (2006)).

Die vorkommenden Subjekte sind großteils Nominalphrasen oder pronominale Verweise auf wiederum Nominalphrasen; es finden sich im qualitativen Sample auch objektfähige Nebensätze sowie formale *es*-Subjekte und ein demonstratives *das*, die sich von den pronominalen Verweisen unterscheiden und auch deshalb so annotiert werden, um hier genauere Unterscheidungen treffen zu können (siehe Kapitel 4.5 *Subjekte*).

Dies ist nur ein Teil der Diskussion, weitere Aspekte werden in Bezug auf die Grammatikalisierung besprochen, da diese im Zusammenspiel von Grammatik und Semantik zu betrachten sind.

5.2 Semantik des *gehören*-Passivs

Die Bedeutung der Konstruktion wird in der Forschungsliteratur teilweise sehr stark mit der Semantik der VVPPs in Verbindung gebracht (siehe Stathi 2009; 2010); aus diesem Grund

wird ein weiterer Blick auf die explizit erwähnten Perfektpartizipia bei Stathi (2010: 335) und Szatmári (2002: 180) geworfen:

Verben	Stathi (2010: 335)	Szatmári (2002: 180)	Qualitatives Sample	Quantitatives Korpus
abgenommen	-	2	0	18
abgeschafft	88	1	10	1458
abgeschoben	6	0	1	162
abgestellt	-	1	11	550
abgewählt	6	0	2	55
aufgetaut	-	1	0	0
ausgemistet	-	1	0	44
behandelt	-	1	3	239
beseitigt	6	0	3	373
bestraft	23	1	4	350
eingesperrt	8	0	2	128
entsorgt	-	1	2	191
erschossen	10	0	1	45
geändert	7	0	15	1403
gefeiert	-	1	6	614
gemacht	-	1	16	1012
gestoppt	-	1	1	178
getrennt	-	1	2	146
legalisiert	-	1	0	3
reformiert	-	1	4	790
überprüft	-	1	6	230
verboten	55	0	18	985
verfilmt	-	1	0	6
verhängt	-	1	0	16

Tabelle 17: Übersicht über die VVPPs aus der Forschungsliteratur

Die beiden Verben *abschaffen* und *bestrafen* sind die einzige Überschneidung in den beiden Artikel zum *gehören*-Passiv, diese beiden Verben kommen auch in der hier vorliegenden Untersuchung vor.

Nicht alle Verben, die in der Forschungsliteratur angeführt werden, finden sich im qualitativen Sample, jedoch tauchen sie – mit Ausnahme von *aufgetaut* – im quantitativen Korpus auf (siehe Tabelle 17).

Weitere Verben, die häufig im Korpus vorkommen, sind: *gefördert*, *ausgebaut*, *unterstützt*, *geschützt*, *aufgeklärt*, *verbessert*, *gestellt*, *geklärt*, *gezogen*, *geregelt*, *diskutiert* und *umgesetzt* (siehe Kapitel 4.3 *Semantik der VVPPs*).

Die Semantik der am häufigsten vorkommenden Verben scheint auf den ersten Blick neutraler oder gar positiver als bisher angenommen, besonders die Verwendung von *feiern*, *fördern* und *verbessern*.

Jedoch ist aufgrund der Implikation, dass der Zustand, in dem sich das Patiens befindet, nicht gewünscht ist und geändert gehört, die Bedeutung teilweise negativer als andere potenzielle Formulierungen mit demselben Perfektpartizip. Man vergleiche:

Die Schüler*innen gehören gefördert.

Die Schüler*innen müssen/sollen gefördert werden.

Die Schüler*innen werden gefördert.

Diese Gegenüberstellung demonstriert den Bedeutungsunterschied, der durch *gehören* und seine Semantik innerhalb der Konstruktion verursacht wird, eben die deontische Modalität, die gleichsam dafür sorgt, dass der Satz weniger positiv konnotiert wirkt, auch wenn die Schüler*innen in beiden Fällen als Beneficiens annotiert werden können. In dem einen Fall kommt ihnen bereits die Förderung zu, im anderen bleibt es eine Forderung im allgemeinen Interesse.

An dieser Stelle ist anzumerken, dass dies eine nicht eindeutig quantifizierbare Bewertung ist. Es wäre bei zukünftigen Untersuchungen einerseits wünschenswert, Sentimentanalysen zu machen, die eine Einschätzung darüber abgeben, ob es sich tendenziell um positive, neutrale oder negative Konnotationen handelt, andererseits eventuell eine Befragung durchzuführen, um dies durch Sprachnutzer*innen quantifizieren zu lassen.

Explizit positive Verben wie *feiern*, *ehren* und *würdigen* scheinen bei Stathi (2009; 2010) im Korpus nicht auf; eine überwältigende Menge der Belege – mehr als zwei Drittel – sind explizit oder implizit negativ konnotiert. Eine Erklärung dafür erschließt sich bei der Betrachtung des hier vorliegenden Korpus nicht auf den ersten Blick, genauso wenig wie für die dominierende Anzahl explizit stark negativer Perfektpartizipia.

In der Diskussion rund um die Semantik der Perfektpartizipia bzw. der Konstruktion im Allgemeinen wird die semantische Klassifikation nach Stathi (2010: 337) noch einmal aufgegriffen:

A. *gehören* + [Participle II [PUBLIC AFFAIRS, PUBLIC OPINION]]

B. *gehören* + [Participle II [UNPLEASANT ACTION ON PATIENT (PUNISHMENT)]]

Der zweite Punkt, der hier angeführt wird, also Verben mit einer ‚unpleasant action‘ im Sinne einer Bestrafung für den Patiens, als Unterordnung zu Punkt 1, soll nun im Folgenden diskutiert werden.

Aus diesem Grund wurde die Aufzählung bei Stathi (2010: 335) in eine Liste überführt und deren Belegzahlen aus dem quantitativen Korpus jeweils dazugetragen.

VVPPs bei Stathi (2010: 335)	Anzahl im quant. Korpus
abgelöst	71
abgerissen	20
abgeschafft	1458
abgeschoben	162
abgewählt	55
aufgehängt	106
aufgelöst	163
ausgemerzt	28
ausgemustert	6
ausgepeitscht	2
ausgerottet	11
ausgetauscht	213
ausgewechselt	28
ausgewiesen	32
aussortiert	27
bedroht	2
beschränkt	30
beseitigt	373
bestraft	350
bombardiert	-
diffamiert	-
disqualifiziert	10
eingeschränkt	52
eingesperrt	128
eliminiert	23
entlassen	38
erschossen	45
geschlachtet	11
geschlagen	12
geschlossen	264
gesprengt	6
gesteinigt	7
hingerichtet	3
kontrolliert	275
kritisiert	10
rausgeschmissen	2
rausgeworfen	-
überwacht	60
umgebracht	20
verbannt	103

verboten	985
vergast	17
verhaftet	13
vernichtet	28
verprügelt	4
weggejagt	-
weggesperrt	5
Gesamtanzahl	5258

Tabelle 18: VVPPs bei Stathi (2010: 335) und im quantitativen Korpus

Dabei wurden nach Angabe in Stathi (2009: 294) die Ergebnisse der Abfrage *gehören* + participle II miteinbezogen, das sind 738 Belege. Die Summe der häufigsten Belege, die auch in Tabelle 17 angeführt werden, beläuft sich bei Stathi (2010: 335) auf 209. Wenn man alle angeführten VVPPs zur negativen Bedeutung für das Patiens im quantitativen Korpus betrachtet, versammeln diese 5258 von 42508 Belegen; das entspricht 12,37%, wobei auf *verboten* und *abgeschafft* beinahe die Hälfte (n=2443) entfallen und vier Verben – *bombardiert*, *diffamiert*, *rausgeworfen* und *weggejagt* – gar nicht vorkommen (siehe Tabelle 18).

Es wird keine Gesamtanzahl der unterschiedlichen VVPPs angegeben; die Verben in Tabelle 18 sind 47 an der Zahl. Im Vergleich dazu sind im quantitativen Korpus 2198 unterschiedliche Verba vorhanden. Dadurch wird ein Abgleich erschwert, da die Untersuchung einer so großen Anzahl an Verben den Rahmen sprengen würde. Die Betrachtung der häufigsten Verben in den beiden Teiluntersuchungen hat aber gezeigt, dass diese allein eine Mehrheit an negativen Bedeutungen nicht widerspiegeln (siehe Kapitel 4.3, Tabelle 10).

Zu den häufigsten Verben in Stathis (2010: 335) Untersuchung gehört auch *erschießen* (VVPP *erschossen*), sie selbst kommt jedoch nicht auf konkrete Belege zu diesem Perfektpartizip zu sprechen. Ein Beispiel für diese explizit stark negative Bedeutung – im Besonderen auf menschliche Patiens-Subjekte – findet sich im qualitativen Sample der hier vorliegenden Untersuchung:

(S181) Warum sollte man Ariel Muzicant angreifen? Würde er sagen, alle Araber in Österreich gehören erschossen, dann ja. Das sagt er aber nicht, sondern er sagt, die sollen friedlich hier miteinander leben. WZ_2010_0102 (Unterstreichung, C.M.)

Die Formulierung mit *erschossen* zeigt, wie das *gehören*-Passiv dazu verwendet werden kann, menschenfeindliche Aussagen zu formulieren. In dem hier vorliegenden Beispiel wird es durch den angeschlossenen *würde*-Satz relativiert, ebenso wie durch den hypothetischen Kontext, in den es darüber hinaus eingebettet wird. Auch wenn die Konstruktion im Indikativ steht, ist sie als indirekte Rede zu lesen, die jedoch nicht tatsächlich getätigt wurde –

zumindest nicht von der Person, von der hier die Rede ist, nämlich Muzicant. Sie wurde aber von der schreibenden Person zu Papier gebracht und zeigt eindrucksvoll, wie stark negative Formulierungen mit dem *gehören*-Passiv gebildet werden können.⁵²

Aus der Kombination von derart pejorativ gemeinten VVPPs und menschlichen Patiens-Subjekten, die teilweise auf Minderheiten Bezug nehmen, ergeben sich die besonders negativen Bedeutungen, die mit der Konstruktion des *gehören*-Passivs ausgedrückt werden können, gegen die die explizit positiven verblassen.

Wenn man die hier bezüglich Semantik angesprochenen Ergebnisse und Aussagen Revue passieren lässt, fallen ein paar Punkte auf, die in der bisherigen Untersuchung des *gehören*-Passivs angeschnitten wurden, aber Fragen offenlassen: Wo ist die Grenze zwischen bloßer öffentlicher Angelegenheit und einer mit einer explizit unangenehmen Bedeutung für den Patiens im Sinne einer Bestrafung und warum ist dies nicht so einfach einzuschätzen? Wie verhält es sich mit „negativen Bedeutungen“ für Patiens, die eigentlich keine ‚Public Affair‘ waren? Ob es sich nun um Bestrafungen, ‚destruction‘ oder Kritik handelt, die als ‚unpleasant action‘ zusammengruppiert werden (vgl. Stathi 2010: 337), scheint ebenfalls keine triviale Frage zu sein.

Aussagen wie „Die Regierung gehört abgesetzt“ (wie in Beleg S419) sind nicht zweifelsfrei mit Aussagen, die als Drohungen gegen Menschenleben zu verstehen sind (wie in Beleg S181, mit VVPP *erschossen*), in eine Kategorie zusammenzuwerfen, noch dazu, wenn weitere sich – zumindest nicht in einem vergleichbaren Sinn – nicht auf den von Stathi (2010: 337) allgemein gesehenen Kontext von öffentlichen Themen und Meinungen beziehen, der laut ihr für alle *gehören*-Passiv-Belege gilt.

Es macht einen bedeutenden Unterschied, ob in öffentlichen Diskussionen eben beispielsweise nach einer Absetzung der Regierung gerufen wird (wie es in Österreich in den letzten Jahren aufgrund politischer Skandale durchaus öfter vorgekommen ist) oder ob diskriminierende und fremdenfeindliche Aussagen getätigt bzw. (re)produziert werden, die aufgrund von allgemein gegen die Menschenwürde verstoßenden Bedeutungen zu ahnden sind oder sich gar Gedankengut anschließen, das in Verbindung zum nationalsozialistischen Regime und dessen Verbrechen steht. Diese sind deshalb stark zu verurteilen und durch verschiedene

⁵² Eines der Anliegen in dieser Arbeit war, möglichst nicht die menschenfeindlichen Aussagen zu reproduzieren und gleichzeitig dennoch die Belege der Untersuchung in ihrer Gesamtheit abzubilden. Dieses – wenn auch hypothetisch formulierte – Beispiel stellt eine gewisse Ausnahme dar, die zur Darstellung der drastischen Formulierungen dient, die eben teilweise durch das *gehören*-Passiv – auch in einem öffentlich publizierten Kontext – getätigt werden.

Gesetze in Österreich, aber auch in Deutschland juristisch relevant⁵³ – und gelangen gerade aus diesem Grund oft erst durch Berichterstattung in den Nachrichten an die Öffentlichkeit.

Der Äußerungskontext scheint nämlich potenziell ursprünglich ein nicht-öffentlichkeits-wirksamer gewesen zu sein, vor allem bei derart negativen Verwendungen, auch wenn hiermit darauf hingewiesen werden soll, dass derartige Belege sich mindestens auf zweistellige Zahlen beliefern (siehe Tabelle 18), wonach dieses Thema keinen „Einzelfall“ darstellt und damit Äußerungen in diese Richtung belegt werden können.

Es ist – nicht nur aus diesem Grund – schwierig, bei einer semantischen Analyse der Konstruktion 100 Prozent der Belege in dem allgemein gefassten Bereich der „Public Affairs und Public Opinion“ zu verorten, noch dazu wenn auch bei Stathi (2010: 336–337) aus diesem herausfallende Beispiele genannt werden. Und wo dabei die Grenze zwischen bloßer öffentlicher Angelegenheit und einer mit einer explizit unangenehmen Bedeutung für das Patiens verläuft, lässt sich für viele der vorkommenden Verben nicht eindeutig bestimmen, unter anderem, weil der Grad der ‚unpleasantness‘ stark fluktuiert.

Im Speziellen, wenn von generellen Verwendungskonventionen die Rede ist (vgl. Stathi 2010: 337), ist es demnach zudem zu kurz gegriffen, diese auf den semantischen Kern der VVPPs zu beziehen und nach diesem die Gesamtkonstruktion einzuteilen. Vor allem der Bezug auf die Perfektpartizipia allein (auch graphisch über die Setzung der Klammern) ist hier nicht günstig gewählt, da erst die Passivierung mit *gehören* dafür sorgt, dass die Belege als „matters of public opinion“ oder eben „unpleasant actions on the patient [Hervorhebung C.M.]“ (Stathi 2010: 337) gelesen werden können, und es sich dabei nicht zwingend um bereits bestehende Bedeutungen der Partizipia handelt, die mit *gehören* bloß zusammengeführt werden.

Es ist also ein Unterschied, ob die Bedeutung des Partizips selbst oder in der gesamten Konstruktion oder in dem Kontext des gesamten Satzes gesehen wird; eine Unterscheidung, die Stathi (2009; 2010) bei der Betrachtung der Verben nicht vornimmt.

Beschränkt ist ein Verb, das bei Stathi (2010: 335) eben als typisch für die öffentlichkeits-wirksame negativ-bestrafende Bedeutung für das Patiens angeführt wird und zu dem auch im qualitativen Sample folgendes Beispiel zu finden ist:

(S472) Auch die Handyzeiten gehören beschränkt. KRONE_2022_1018

⁵³ An dieser Stelle sei auf das Verbotsgesetz hingewiesen, dass seit 1947 besteht (vgl. Verbotsgesetz). In einem Bericht des OSCE, in dem ein Überblick über die Gesetzeslage gegeben wird, werden eben auch bloße Aussagen, die in der Verbindung mit der Ideologie stehen, als möglicher Straftatbestand gesehen (vgl. OSCE (2004)).

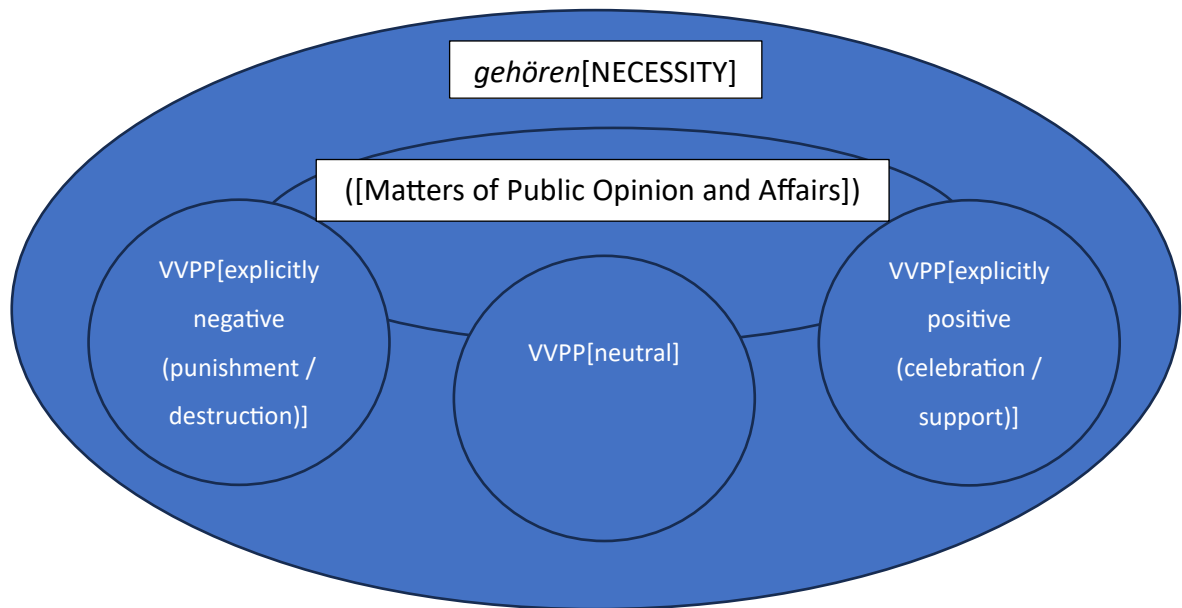
Die Bedeutung von *punishment* ‚Bestrafung‘ ist für unbelebte Subjekte, auf die, wie in Kapitel 4.5 gezeigt, ein Großteil der Belege im qualitativen Sample entfällt (n=351), schwerer nachzuvollziehen, da die Konnotationen eher auf Strafe in einem rechtlichen oder Ächtung in einem öffentlichen Kontext verweisen und in dieser Lesart eher mit menschlichen Subjekten in Verbindung gebracht werden. Ähnlich verhält es sich für die weitere mögliche Bedeutung im Sinne der *destruction* ‚Zerstörung‘ (vgl. Stathi 2010: 337); es handelt sich um eine Beschränkung, die sich in diese Gruppierung nur schwer einfügen lässt. Dies ist nur ein Beispiel, aber diese Tendenz lässt sich generell für unbelebte Subjekte festhalten, dass sich die explizite Negativität nicht derart stark widerspiegelt.

Der Ansatz der (deontischen) Modalität, die grundsätzlich über die Verwendung von *gehören* in der Konstruktion mit teilweise sehr unterschiedlichen Verben im Perfektpartizip ausgedrückt wird, scheint der treffendere Weg der Beschreibung zu sein, der dann in weiterer Folge nach Gesichtspunkten der VVPPs eingeteilt werden kann, also eine semantische Primärklassifikation der Konstruktion von *gehören* ausgehend. Man könnte eine sekundäre Differenzierung der „semantischen Kerne“, den bedeutungstragenden Verbpartizipia, anschließen; die Frage stellt sich hierbei, ob dies zwingend notwendig ist, um die Semantik der Konstruktion zu definieren bzw. nach welchen Kriterien man die Verben einteilen soll. Hier wurde eine Dreiteilung nach dem Sentiment der Verben vorgenommen:

- gehören[NECESSITY] + participle II of a verb (most often transitive)
- optional: matters of public opinion and affairs
 - VVPP[explicitly negative (punishment / destruction)]
 - VVPP[neutral]
 - VVPP[explicitly positive (celebration / support)]

Der explizit negativen Bedeutung, die angelehnt an Stathi (2009; 2010) als *punishment* oder *destruction* bezeichnet wird, werden auf der positiven Seite *celebration* und *support* gegenübergestellt, wobei nur ein Teil der vorkommenden Verben diesen beiden Gruppen eindeutig zugeordnet werden kann, wie etwa *eingesperrt* und *vernichtet* sowie *gefeiert* und *gefördert* (siehe Tabelle 9 in Kapitel 4.2 und 18).

Der Punkt der ‚Matters of Public Opinions and Affairs‘ kann theoretisch auf alle Partizipia zutreffen, muss aber nicht vorhanden sein. Dies spielt ebenfalls in den Aspekt hinein, dass diese Bedeutung teilweise durch den Kontext der Äußerung und Auffindung in einem öffentlichen Medium beeinflusst sein kann (siehe Kapitel 5.3 für weitere Diskussionen dieses Punktes).



Gemeinsam mit diesem Versuch einer semantischen Klassifikation lässt sich auch eine Lesart der Konstruktion formulieren:

Ich mache es nicht (oder ich bin nicht in der Lage, es zu machen), aber es ist (aufgrund von öffentlichem Interesse | meiner Ansicht nach) notwendig, dass der vorgeschlagene Sachverhalt durchgeführt wird.

Der entscheidende Punkt ist eben dieses „Notwendig“, das durch *gehören* in diesem Passiv ausgedrückt wird, der modale Aspekt, die deontische Modalität, die auch in der Forschungsliteratur fokussiert wird, bei Stathi (2010: 332) als Begründung der grammatikalisch vorkommenden Verbformen angeführt und in Bezug auf die Semantik hintenangestellt wird.

Es liegt ein aktueller Zustand vor, der in einen gewünschten Zustand überführt werden soll; in welcher Form das passiert, wird durch die VVPPs ausgedrückt. Die Hervorhebung eines Mangels am Ist-Zustand sorgt für die implizite Negativität, während die Durchführung der Aktion, die durch das Perfektpartizip Einzug in die Konstruktion erhält, als notwendig und/oder wünschenswert dargestellt wird.

Der Ausdruck dieses Sachverhalts über die Passivität anstatt über eine Agens-zentrierte Konstruktion, obwohl es sich um eine – wie positiv oder negativ auch immer geartete – Meinungsäußerung handelt, ist der Punkt der meiner Ansicht nach in der Forschungsliteratur nicht ausreichend herausgearbeitet wird.⁵⁴

⁵⁴ Eine tiefergehende Untersuchung zur Pragmatik und, wie diese Konstruktion genau funktioniert, könnte weitere Aufschlüsse geben. Eine Vermutung ist, dass es sich hier eventuell um ein Hedging-Phänomen handeln könnte, im Sinne der linguistic Politeness (zu den Konzepten siehe Kaltenböck/Mihatsch/Schneider 2010; Ginsburg et al. 2016; Culpeper/Haugh/Kádár 2017).

Der*die Sprechende nimmt sich durch die Passivität zurück, die deontische Modalität sorgt jedoch für die Lesart im Sinne von „das sollte/müsste gemacht werden“. Die Eliminierung des Agens und die Bedeutung der Konstruktion, als etwas „allgemeiner Gültiges“, das nicht explizit nur den*die Sprechende*n betrifft, scheint hier mit hineinzuspielen.

Wer spricht, wird in manchen Beispielen dennoch eindeutig, nämlich jene, in denen beispielsweise Diskursmarker den *gehören*-Passiv-Satz als Meinung der Sprachproduzent*innen ausdrücken:

(S419) T. B. (anonymisiert, C.M.), Kellnerin, Berndorf: Meiner Meinung nach gehört die gesamte Regierung sofort ausgetauscht. NOEN_2014_0901

Ebenso wird dies implizit durch die Integration der*s Sprechenden selbst erreicht, wie es beispielsweise bei den Belegen zur 1. Person Plural der Fall ist.

(K40762) „Sollten wir dann noch den Aufstieg vergeigen, gehören wir erschlagen.“ OBERRUND_2012_0509

In diesem Beispiel wird zum Ausdruck gebracht, dass der*die Sprechende der Meinung ist, dass das mit dem *gehören*-Passiv vorgeschlagene Verhalten das Angemessene ist, und stellt sich gleichzeitig an die Stelle des Patiens. Er*sie wäre dabei selbst leidtragend, wenn auch nicht die ausführende Macht, was nach wie vor der Logik der hier angeführten Lesart entspricht.

Dies weicht die Verhüllung des*r Sprechenden auf, die bei der Konstruktion vorliegt und macht deutlich, dass diese Rolle mit dem Patiens und dem Agens übereinstimmen kann; in den meisten Fällen ist sie in der Meinungsäußerung gesondert davon zu betrachten.

Dass *gehören*-Passive tendenziell ein negatives Sentiment aufweisen, kann nicht von der Hand gewiesen werden. Dies liegt jedoch nicht an der Semantik der VVPPs allein bzw. nicht primär an ihnen, wie die neutralen und sogar positiven Belege der Konstruktion belegen. Eine Bestätigung, dass es sich bei zwei Drittel der *gehören*-Passiv-Belege um öffentlichkeitswirksame negativ konnotierte Bestrafungen handelt, kann hier nicht erbracht werden. Das vorliegende Korpus zeigt ein differenzierteres Bild, auch wenn die konkrete Semantik und vor allem das Sentiment der einzelnen Partizipia und den Konstruktionen mit ihnen (aber nicht durch sie) nicht genau quantifiziert werden können. Diese Schlussfolgerungen beziehen sich sowohl auf die Verhältnisse der vorliegenden Belegzahlen als auch auf die Vorannahme, dass pejorisierende Verben wie *erschießen* und *vergasen* nicht mit (stark) negativen Ausdrücken wie *wegjagen* und *verbieten* oder neutraleren Bedeutungen wie *kritisieren* oder *austauschen* in eine Kategorie zusammengruppiert werden können.

Die Bedeutung der gesamten Konstruktion an der Semantik der VVPPs festzumachen, auch in der grafischen Darstellung durch die Klammern des semantischen Musters (siehe Stathi 2010: 337), und anhand dessen eine dominierende Form der Konstruktion im Sinne der Konstruktionsgrammatik festzuhalten, ist diskutierten Punkten nach weder zielführend noch notwendig.

Szatmári (2002) mit ihrem Kleinstkorpus aus 18 Belegen kommt zu diesem Schluss, dass das *gehören*-Passiv ausdrückt, „wie man in einer bestimmten Situation eine bestimmte Angelegenheit erledigen soll (sollte) bzw. muss (müsste).“ (Szatmári 2002: 182). Dieser Anhaltspunkt gemeinsam mit der deontischen Modalität liegt auch dem Versuch der hier vorliegenden semantischen Klassifikation zugrunde.

Stathi (2010: 336) formuliert eine ähnliche Aussage in Bezug auf *gehören*-Passive, die aus den bei ihr präsentierten Schemata herausfallen: „But still, by the use of *gehören* the speaker claims that the action described is the only one that is appropriate.“

Diese wird mit einem Hinweis auf die Fußnote 21 versehen, dass diese Bedeutung auch eine der ersten belegten ist (vgl. Stathi 2010: 336), nämlich in: „Zwischen dem Gamsarchen und der Hochscharte gehören die Aichenköpfe aufgeführt [...]“ von 1874 (Stathi 2010: 328).

5.3 Weitere Aspekte der Konstruktion

Diese betreffen vor allem die Diskussion rund um die Standard- und Schriftsprachlichkeit der Konstruktion und dadurch auch des Korpus. Ein nicht unbeträchtlicher Teil der Belege ist von Mündlichkeit geprägt, was dazu führt, dass die betrachtete Varietät relativiert werden muss. Die Ausführungen hier spiel(t)en auch teilweise in der Diskussion rund um die Semantik der Konstruktion (Kapitel 5.2) eine Rolle; hier werden sie jedoch mit Fokus auf die Standard(schrift)sprache betrachtet.

Die normalisierten Belegzahlen für den Medientyp sind bei **tv** am höchsten, also tatsächlich mündlich produzierte und transkribierte Fälle. Genauso sind alle TV-Sender bei den einzelnen Medien in den normalisierten Zahlen unter den Top 5 (siehe Kapitel 4.1 *Quantitative Verhältnisse der gehören-Passive*).

Die Prämisse, dass es sich hierbei zweifellos um österreichische Standardschriftsprache handelt, muss im weiteren Verlauf ebenfalls modifiziert werden, weil ein guter Teil der Belege von ehemals medialer und/oder konzeptioneller Mündlichkeit (siehe Kapitel 3.3 zur Annotation) geprägt sind.

Bei vielen Fällen handelt es sich um direkte oder indirekte Rede (direkt: 215, indirekt 78, von 472 Belegen im qualitativen Sample, Kapitel 4.7), die ebenfalls meistens, aber nicht

immer für die Öffentlichkeit bestimmt waren; soweit dies aus den Kontexten erkennbar ist (siehe Kapitel 5.2 zur Semantik der Konstruktion).

Die Form der Sätze ist grundsätzlich orthographisch standardnah normiert, teilweise finden sich jedoch auch dialektal geprägte Belege (schriftliches Festhalten von Dialektaussagen – siehe Kapitel 4.2; (K45328) „I tritt euch die Tür ein, ihr gehört's alle weggeräumt, I stich euch ab!“ BVZ_2022_0609 (Unterstreichung, C.M.)).

Die genaue Identität der Sprecher*innen bleibt dabei bis zu einem gewissen Grad unklar, auch wenn bei Durchsicht der Samples sich Gruppen wie Politiker*innen und Expert*innen, die zu ihrem Gebiet befragt werden, herauskristallisieren:

(S91) Bezirksvize Franz Lerch bietet jetzt an, sich die betroffenen Stellen anzusehen. „Notwendige Sperrflächen gehören vernünftig abgesichert, aber der ohnehin knappe Parkraum darf nicht sinnlos vernichtet werden“, so der VP-Politiker. Heute_2010_0820 (Unterstreichung, C.M.)

(S471) „Angesichts der Teuerung ist es jedoch notwendig, die Lücken zu schließen“, betont Expertin Marie Hasdenteufel. Beispielsweise gehöre die Ausgleichszulage an die Armutschwelle angepasst. Krone 2022_0817 (Unterstreichung, C.M.)

Dies sind Sprecher*innengruppen, die beispielsweise auch bei Koppensteiner (2023, 387–388) im Zuge der edukativ-sprachbezogenen oder formell-berufsbezogenen Zugänge zu typisch standardsprachlichen Domänen auftauchen, jedoch kommt er zu dem Schluss, dass die Beurteilungen durch Lai*innen (abseits der Nachrichtensprecher*innen) zu heterogen sind, um eindeutig klassifizierte Kategorien von Standardsprecher*innen anzuführen.

In manchen der Belege – wie den hier angegebenen – ist es eindeutig, wer die Aussage tätigt und welchen Berufsstand diese Person hat bzw. in welcher Rolle sie spricht. Dies gilt jedoch nicht für alle bzw. würde es mehr Kontext oder Recherche brauchen, um das eindeutig zu lösen; eine Leistung, die hier in dieser Arbeit nicht erbracht werden kann, weshalb auf die genauere Betrachtung verzichtet wurde.

Manche Aussagen sind auch hypothetisch (wie zum Beispiel der Beleg in Bezug auf Ariel Muzicant, siehe Kapitel 5.2), manche waren vermutlich nie für den Kontext der öffentlichen Äußerung bestimmt, beispielsweise in Drohungen (vergleiche (K45328); siehe Kapitel 4.2). Eine Gemeinsamkeit dieser Beispiele ist die stark negative Bedeutung der angeblich geäußerten *gehören*-Passiv-Phrase, die eben bei letzterem auch für Veröffentlichung in einem Nachrichtenkontext gesorgt haben.

Für einen Teil der Belege ist es also weniger Standardschriftsprache als Gebrauchsstandard, der mündlich geäußert wird und in schriftlicher Form vorliegt. Der schriftsprachliche

Aspekt kommt also in manchen Belegen als Transkription oder anderweitige Niederschrift zu tragen. Unklar ist dabei auch der Faktor, inwiefern von den Schreiber*innen der Zeitungstexte in die Gestalt der als direkte Zitate dargestellten *gehören*-Passiv-Belege eingegriffen wurde, inwieweit indirekte Reden Wiedergaben oder Paraphrasierungen von Äußerungen sind oder ob sie durch die (Nicht-)Akzeptanz durch Modellschreiber*innen vorkommen.

Bei den Belegen, die nicht eindeutig als (ehemals) medial mündlich konnotiert gekennzeichnet sind, finden sich ebenfalls Beispiele, die von Lai*innen (im Sinne der Professionen im Journalismus) geschrieben wurden und als Leserbriefe oder andere Formen der direkten Meinungskundgebung (es finden sich Beispiele mit „Name, via E-Mail“) im Korpus auftauchen.

(S333) ..., dann hat die Behörde in meinen Augen auch ihre Existenzberechtigung endgültig verloren und gehört aufgelöst! Mag. W. K. [anonymisiert, C.M], per E-Mail. OOEN_2015_0307

Auch wurden die Transkripte der TV-Sender nicht ausgeschlossen, da diese ja durch Metadaten problemlos von den übrigen Belegen unterschieden werden können; bei diesen Daten handelt es sich sogar um dezidiert medial mündlich geäußerte Belege, die im Nachtrag schriftlich festgehalten wurden.

Nachrichtensprecher*innen werden ebenfalls laut Ammon (1995: 79; 2017: 11–12) zur Produktion von Modelltexten (auf mündlicher Ebene) herangezogen; die Frage, die sich hierbei meiner Ansicht nach stellt, vor allem auch nach eingehender Beschäftigung mit den Belegen, ist die Form ihres (persönlichen) Einflusses, da das Format der Nachrichten ein Vorlesen nahelegt und die vorgelesenen Texte voraussichtlich von einem Team aus Schreiber*innen produziert wurden (und daher hier auch wieder ein schriftsprachlicher Bezug vorliegt; siehe auch Koppensteiner (2023: 379) zu den diesbezüglichen Ausführungen der schriftsprachlichen Grundprägung). Der Kontext der öffentlichen standardsprachlichen Äußerung scheint dennoch gegeben und bei Koppensteiner (2023: 390) wird eben diese Sprecher*innengruppe als der Prototyp der österreichischen Standardsprachlichkeit angesehen.

Nicht alle Belege aus dem TV, die im Korpus vorliegen, stammen aber von Nachrichtensprecher*innen (wie im Beispiel S373), tendenziell eher von anderen Personen (S360 und S368), beispielsweise in Interviews:

(S373) Jansch Christian (ORF) Die Impfaufklärung gehört also ausgebaut. Zum Stoppen der vierten Welle kommt das allerdings zu spät. (Im Gespräch mit Wolfgang Mückstein) MWVOLL_2021_0901

(S360) Anton Hannes (Burgenland Tourismus) Es läuft vieles sehr gut, aber es gehört einfach vieles noch zusammen geschlossen (sic) (...). MWVOLL_2017_0331

(S368) Meini-Reisinger Beate (NEOS - Das Neue Österreich) Ich würde mal sagen, das gehört zentral entschieden, was besser zentral entschieden wird (...). MWVOLL_2020_0803

Also zeichnet sich auch hier eine ähnliche Situation ab, dass hier ebenfalls die bereits angesprochenen Personengruppen sowie „die allgemeine Bevölkerung“ an der direkten Produktion beteiligt sind, wie es zumindest beim Medientyp **print** der Fall ist; nur **agentur** scheint von diesem Prinzip ausgenommen.

Vor allem durch direkte mündliche Äußerungen, die standardorthografisch transkribiert wurden – soweit dies aus den vorliegenden Belegen beurteilt werden kann –, finden sich im Allgemeinen Marker sprechsprachlicher Ausformungen, die in der geschriebenen Form markiert sind (oder teilweise auch als normativ inkorrekt bewertet werden könn(t)en).

(S195) „Ein paar Mal gehört uns in den Arsch getreten, weil sonst geht nichts weiter.“ VN_2020_0124

(S354) Für uns macht das auf jeden Fall Sinn. Jugendliche gehören getrennt von den Erwachsenen. MWVOLL_2014_0428⁵⁵

In S195 findet sich außer dem umgangssprachlichen Ausdruck *in den Arsch treten* Verb-Zweit-Stellung um Nebensatz. In S354 wird die Präpositionalphrase, die aufgrund der Verbklammer des Deutschen zwischen den Teilen des Prädikats stehen sollte, ins Nachfeld gerückt, eine informelle, potenziell Substandard- und/oder sprechsprachliche Ausdrucksweise. Dieses Beispiel stammt auch aus den Transkriptionen des ORF.

Diese Aspekte rund um das Zusammenspiel von Gesagtem und Geschriebenem machen es schwierig, eine genaue Einschätzung der untersuchten Varietät abzugeben sowie die (konzeptionelle) Mündlichkeit der Konstruktion einzuschätzen,

Möglicherweise handelt es sich beim *gehören*-Passiv um eine eigentlich non- bzw. substandardsprachliche Konstruktion, die aufgrund direkter Wiedergaben von Gesagtem Einzug in ein präsupponiert standardsprachliches Korpus erhalten hat. Ihr Gebrauch scheint aber in dem vorliegenden Untersuchungskontext etabliert: Ein Beispiel mit Anführungszeichen ohne anderweitig gekennzeichnete direkte Rede ist im qualitativen Sample nicht zu finden, was dafür spräche, dass sie nicht markiert ist.

Diese Ergebnisse haben Bedeutung für die Varietät, die im amc reproduziert bzw. die hier durch die Betrachtung des *gehören*-Passivs untersucht wird. Einerseits bedeutete es nämlich beim Titel dieser Arbeit *Schrift* bei *Standardschriftsprache* in Klammer zu setzen, da das

⁵⁵ Diese Formulierung zeigt möglicherweise den bei Stathi (2010, 328–329) angesprochenen Übergang zwischen lokativen und metaphorisch verwendeten Präpositionalphrasen, die eventuell dazu beigetragen haben, dass *gehören* heute als Passiv verwendet werden kann.

nicht eindeutig für alle Belege gewährleistet werden kann, auch wenn die betrachteten Belege in schriftlicher Form vorliegen, andererseits ist das eine Relativierung, die bei der Betrachtung des gesamten Korpus im Allgemeinen und dem hier spezifisch gefilterten Subkorpus bedacht werden muss.

5.4 Grammatikalisierung des *gehören*-Passivs

All diese Punkte fließen in die Überlegungen mit ein, in welcher Form das *gehören*-Passiv nun grammatikalisiert ist, und welche Schlüsse aus der durchgeführten Untersuchung vor dem Hintergrund der Forschungsliteratur gezogen werden können.

Anknüpfend an den letzten Gedanken im Kapitel zur Grammatik der Konstruktion und den Darlegungen im Kapitel zur Semantik stellt sich in Bezug auf das Dargelegte zum Verb *gehören* allgemein die Frage, inwiefern dieser Aspekt der Notwendigkeit, des Erforderlichseins, der Modalität bereits in der Bedeutung konventionalisiert ist, bevor es als Passiv verwendet wurde.

Die ursprüngliche Bedeutung des Lexems bezieht sich laut Kluge/Seebold (1999: 307) auf Mitgliedschaft (*zu jemandem gehören*) bzw. auf Eigentum (*jemandem gehören*) (vgl. Grammis: *gehören*. in Bezug auf Dinge), in dieser wird das Verb auch heute noch verwendet (siehe Szatmári 2002: 187; Online-Duden: *gehören*):

Das Buch gehört mir.

Er gehört zu uns.

Bereits im 14. Jahrhundert, wie Stathi (2009: 292–293; 2010: 326–328) in ihrer diachronen Untersuchung gezeigt hat, gestützt durch das DWB der Brüder Grimm, verschiebt sich die Bedeutung von *gehören*, vor allem in Kombination mit lokativen Präpositionalphrasen, die allerdings in dieser Form als Komplement von *gehören* zu sehen sind. Die Umschreibungen der Belege bzw. die Semantik der Konstruktion zeigen bereits die Modalität, die aufgrund pragmatischer Interferenzen lesbar wird: Zugehörigkeit zu einer Gesetzesinstitution trägt eine Verpflichtung in sich; in diesen Kontexten wird *gehören* auch verwendet.

Diese Semantik ist ebenfalls nach wie vor erhalten, jedoch ist der Kontext von Administration und Gesetz nicht zwingend notwendig – wie sich auch bereits in Beispielen des 16. Jahrhunderts zeigt.

Die Bedeutung der Verpflichtung, Notwendigkeit oder Erforderlichkeit – wie auch immer man *obligation* und *necessity* übersetzen möchte – scheint sich schon vor der Verwendung

als passivähnliche Konstruktion entwickelt und sich dann auf weitere syntaktische Konstruktionen ausgeweitet zu haben.⁵⁶

Es lässt sich nämlich argumentieren, dass das *gehören*-Passiv nicht die einzige Konstruktion ist, die mit in ähnlicher Semantik gebildet wird, wenn man beispielsweise auf *gehören* mit Qualitativ (siehe Lasch 2018)⁵⁷ oder das *gehören*-Reflexiv blickt. Eine Verwandtschaft, wie Szatmári (2002: 186) sie bei letzterem sieht, lässt sich nicht abstreiten; sie mutmaßt über eine mögliche parallele bzw. verwandte Entwicklung.

Qualitativ: Diese gehören kalt zu geben. (Lasch 2018: 16)

Reflexiv: Das gehört sich nicht.

Passiv: Das gehört behoben.

Auffällig ist bei diesen Beispielen eine ähnliche Struktur der allgemeinen Formulierungen, welche auch beim Passiv häufig zu finden ist (Kap. 4.5 *Subjekte*).

Wenn man auf die Definition von Grammatikalisierung nach Hopper/Traugott (2012) blickt, fällt vor allem der letzte Teil des Satzes auf: „(...) continue to develop new grammatical functions.“ (Hopper/Traugott 2012: XV) Eventuell hat sich *gehören* nach seinem Bedeutungsshift auf Notwendigkeit und Obligation von diesem Zustand aus auf weitere syntaktische Kontexte und Konstruktionen ausgeweitet und als Modalitätsmarker herausgebildet. Eine endgültige Antwort kann hier nicht gegeben werden; weitere Untersuchungen, die eben auch andere Konstruktionen mit *gehören* genauer betrachten, könnten hier weitere Aufschlüsse geben.

Die syntaktische Umgebung, in der nach Stathi (2010: 328–329) das *gehören*-Passiv entstanden ist, ist dieselbe, in der der Shift zur obligatorischen Bedeutung passierte, nämlich mit lokativer Präpositionalphrase. Eben diese Präpositionalphrasen wurden auch hier in der Untersuchung genauer betrachtet und es zeigt sich, dass viele davon tendenziell eher

⁵⁶ Die Bedeutung der öffentlichen Äußerung wird dabei von Stathi (2009: 298–299) auch als Aufgreifen der Diskurstradition gesehen, nämlich dass durch sie der Kontext der Gesetzes- und Administrations erhalten ist, in dem der Bedeutungsshift auftrat – von Mitgliedschaft zu obligatorischer Mitgliedschaft.

Man könnte für die ‚Public Matters‘ auch argumentieren, dass genauso gut eine offensichtliche Verbindung zwischen ihnen und Berichterstattung in öffentlichen Zeitungen und Zeitschriften besteht, die bei Stathi (2010: 332–333) als Beispiele angeführt werden und auch hier als Korpusgrundlage dienen.

⁵⁷ Lasch (2018) betrachtet *gehören* mit Qualitativ in Bezug auf nonagentivische Konstruktionen und geht dabei auch auf das *gehören*-Passiv im Vergleich an, stützt sich dabei auf die Forschungsliteratur von Eroms (2000) und Stathi (2010), die auch hier verwendet wird. Seine Ausführungen gehen auf unterschiedliche Arten von Konstruktionen ein und bringen weitere theoretische Modelle mit, die den Rahmen hier gesprengt hätten (u.a. konstruktionsgrammatische). In kontrastierenden Untersuchungen zwischen unterschiedlichen Formulierungen mit *gehören* könnten diese wiederum aufgegriffen werden.

metaphorisch gemeint oder konventionalisiert sind, sowie dass ohne dem lokativen Komplement die Wertigkeit des Verbes nicht erfüllt und der gesamte Satz ungrammatisch wäre.

Die Kinder gehören ins Bett.

Die Kinder gehören ins Bett gelegt.

*Die Kinder gehören gelegt.

Das erste Beispiel zeigt die Konstruktion nur mit lokativer Präpositionalphrase, das zweite dann die Erweiterung um das Perfektpartizip, wobei – wie im Kontrast zum dritten Satz gezeigt wird – eine syntaktische Reanalyse passierte und das lokative Komplement zum VVPP gehört und als solches auch nicht weggelassen werden kann.

(S130) Schwarze Schafe gehören aus dem Verkehr gezogen. KRONE_2011_0708

Bei diesem Beispiel aus dem qualitativen Sample kann die Präpositionalphrase nicht eliminiert werden, da sie ein Komplement des Verbes ist und in dem Sinne eine feste Wendung darstellt. *Legen, stellen, ziehen, ...* andere typische Verben, die üblicherweise mit lokativen Präpositionalphrasen assoziiert werden, könnten durchaus der Auslöser für die syntaktische Reanalyse sein, die die Bildung von *gehören* als Passiv mit Perfektpartizip ermöglicht hat. Dies ist auf Basis der Darlegungen bei Stathi (2009; 2010) allein jedoch nicht zweifelsfrei abzuklären.

Es lässt sich keine eindeutig abgegrenzte Klasse an Verben erkennen, die mit dem *gehören*-Passiv vorkommen wie es für *give*- und *get*-Passive beispielsweise gesehen wird (dreiwertige Transferverben), außer dass es sich zumeist um transitive Verben handelt (jedoch ist auch dies nicht absolut restriktiv, wie einige intransitive Verwendungen sowie Zweifelsfälle zeigen (siehe Kapitel 4.4 *Transitivität der Verben*)).

Auch bezüglich ihrer Semantik und der Gruppierung danach lassen sich bestenfalls allgemeine Tendenzen erkennen. Die verwendeten Verben sind zu heterogen, um hier klare Grenzen zu setzen, nicht einmal das Sentiment lässt sich für sie eindeutig bestimmen, weil diese teilweise stark kontextabhängig sind.

Die Beurteilung von Persistenz- und Dekategorialisierungserscheinungen gestaltet sich aus diesen Gründen schwierig; einerseits, weil die Ergebnisse sehr heterogen sind, sowohl grammatikalisch als auch semantisch, und andererseits, weil es kaum kontrastive Untersuchungen gibt.

Das *gehören*-Passiv scheint vor allem in einer relativ fixierten Form vorzukommen bzw. die Tendenz zu einer bestimmten Anordnung der Konstruktion zu haben:

[Nominalphrase mit einem Substantiv als Kopf, zumeist unbelebt]
+
[gehören in 3P Sgl oder Pl Indikativ Präsens]
+
[Perfektpartizip mit tendenziell neutraler/negativer Bedeutung]

Aber wie die Heterogenität in der syntaktischen und semantischen Gestaltung sowie die Verwendung in unterschiedlichsten Kontexten zeigt, funktioniert es auch außerhalb davon.

Die unterschiedlichen Aspekte der Grammatikalisierung, wie sie auch in Kapitel 2.1 besprochen wurden, sollen zum Abschluss noch einmal zusammengefasst werden:

Der Verlust an phonologischer Substanz kann in einem schriftlichen Korpus nicht einfach festgestellt werden; Stathi (2010: 333) und Szatmári (2002: 184) beziehen sich beide auf den Satzakzent, wobei erstere abspricht, dass dieser emphatisch auf *gehören* gelegt werden kann, und zweiteire vermerkt, dass es nur bei „expressiver Ausdrucksweise“ die Betonung trägt.

Gehören kann gemeinsam mit dem Perfektpartizip am Weg der Auxiliarisierung gesehen werden, auch wenn manche der als typisch angesehenen Schritte der Grammatikalisierung nicht eindeutig nachvollzogen werden können, wie der Verlust an syntaktischer Freiheit oder ein fehlerhaftes Paradigma (siehe Stathi 2010: 331–332).

Die Agenstilgung bzw. die Rückstellung des Agens, während der Patiens in die Subjektposition gerückt wird, ist einer der stärksten Indikatoren, dass es sich hier um eine passivische Konstruktion handelt; diese teilt sich das *gehören*-Passiv mit den anderen Passiven (vgl. Stathi 2010: 338). Der Agens wird auch nur selten durch einen Anschluss kenntlich gemacht; davon finden sich zwei Belege (von 738) bei Stathi (2010: 338) und vier (von 472 im qualitativen Sample) in der hier vorliegenden Untersuchung.

Im Zuge der Grammatikalisierung wird häufig der Verlust an semantischer Substanz angenommen – auch als *semantic bleaching* (bspw. Wegener 1998: 37, 43, nach Szatmári 2002: 183–184) – bezeichnet. Hier ist eher ein semantischer Shift passiert bzw. passierte einer, bevor sich überhaupt die Passivform mit *gehören* entwickelte. Die Bedeutung entfernte sich von Mitgliedschaft, Zugehörigkeit und Besitz und zeigt Notwendigkeit bzw. Erforderlichsein an (vgl. Stathi 2010: 327), sodass das Verb nun beim *gehören*-Passiv sowohl als Hilfsverb als auch als Modalitäts-Marker fungiert. Diese Funktion bleibt als „Restsemantik“ erhalten.

Mit diesem Marker kann ausgedrückt werden, dass die Durchführung des diskutierten Sachverhalts notwendig oder wünschenswert ist, also ‚gemacht werden muss oder soll‘, was unter anderem anhand der Umschreibung des Passivs auf diese Weise deutlich wird. Dies ist

der entscheidende Punkt, der in der Forschungsliteratur teilweise herausgearbeitet wird (siehe Diskussion zur Semantik in Kapitel 5.2).

Die grundlegende Gestaltung der Konstruktion zeigt damit einen Gewinn an pragmatischer metakommunikativer Stärke, Verkürzung des modalen Aspekts innerhalb der Passivkonstruktion (vgl. Szatmári 2002: 188), eventuell aufgrund von Sprachökonomie. Diese ist vermutlich relevanter für Sprechsprachliches oder Substandard als für Schriftsprache, aber dennoch hat es Einzug in das Zeitungskorpus gefunden, das eher als standardsprachlich gesehen wird.

Ob es sich beim *gehören*-Passiv um eine Konstruktion mit bestimmter Bedeutung im Sinne der Konstruktionsgrammatik handelt, ist fraglich (vgl. Stathi 2009: 298; 2010: 338), aber wie Eroms (2000: 406) kurz und knapp bereits festhielt: „Immerhin ist damit eine sehr kompakte, syntaktisch geregelte Konstruktion ermöglicht.“, eben durch die Modalität, was über *gehören* ausgedrückt wird.

6. Konklusion

Das *gehören*-Passiv ist eine spezielle Form des Passivs, die sich durch die Semantik auszeichnet, die sie ausdrückt; ein Umstand, der dazu geführt hat, dass die Passivität der Konstruktion diskutiert wurde.

Die verhältnismäßig wenig vorhandene Forschungsliteratur stellte eine Herausforderung dar, vor allem auch weil die Anzahl der Belege, die dazu behandelt wurden, nur einen kleinen Einblick geben. Aus diesem Grund wurde für die hier vorliegende Arbeit eine datenbasierte Untersuchung vorgenommen, eine Herangehensweise, die einen korpuslinguistischen Zugriff beinhaltet. Hier fiel die Wahl auf das Austrian Media Corpus (amc) (Kapitel 3.1).

Bei der Sammlung und Bearbeitung der Daten traten viele Probleme und Herausforderungen auf, die daher rührten, dass eine derartige Datensammlung eines analytischen Passiv-Prädikats (oder anderer ähnlicher Konstruktionen) im amc bislang nicht durchgeführt wurde und deshalb auf keine Erfahrungswerte zurückgegriffen werden konnte, während das Urheberrecht des Korpus es gleichzeitig erschwerte, auf andere Methoden außerhalb des Interfaces und den darin implementierten Möglichkeiten auszuweichen (Kapitel 3.2).

Es wurden unterschiedliche Methoden getestet, um innerhalb der Rahmenbedingungen möglichst viele Belege filtern zu können. Ein nicht-unbeträchtlicher Teil der Arbeit floss in diesen Schritt der Masterarbeit hinein und das Ergebnis ist nach wie vor verbesserungswürdig, jedoch ist es das, das mit den mir zur Verfügung stehenden Mitteln möglich war und

dennoch eine ansehnliche Anzahl an Belegen liefert, nämlich knapp 41.000 Belege, von denen ungefähr 1.500 false Positives sind (siehe Kapitel 4.1).

Es ist vor allem keine Aussage über das exakte Verhältnis von true und false Negatives möglich, die extrapolierten Hochrechnungen suggerieren, dass zwischen 2010 und 2022 rund 59.000 *gehören*-Passive im Korpus zu finden wären, von denen also ungefähr 18.000 nicht extrahiert werden konnten (false Negatives).

Unter Betrachtung unterschiedlicher Kombinationen fiel die Entscheidung für CQL-Abfrage mit spaCy plus eigener spaCy Filter (Kapitel 3.2.2). Diese hatte möglicherweise zur Folge, dass die Ergebnisse durch die doppelte Filterung mit demselben Parser (ungewollt) vereinheitlicht wurden und deshalb auch im qualitativen Sample eine bestimmte, aus der Kombination der Parser und Filter geformte Gestalt der Konstruktion überwiegt.

Aufgrund der überwältigenden Mehrheit der Belege kann gesagt werden, dass dieses Resultat nicht allein durch diesen Umstand bewirkt wird; eine Vermutung ist nur, dass aus diesem Grund die – teilweise explizit gewünschten – Ausnahmen gegenüber der dominierenden Form auch im qualitativen Kontext nicht auftauchen.

Die detailliertere Betrachtung der Konstruktion wurde durch das Hinzuziehen eines qualitativen Samples ergänzt, um weitere Aspekte untersuchen zu können, die bei der Menge des quantitativen Subkorpus nicht möglich sind – aus urheberrechtlichen, programmietechnischen oder zeitlichen Gründen.

Die dabei untersuchten Aspekte betreffen entweder Morphologie oder Semantik, Teilbereiche der Konstruktion wie *gehören* oder das VVPP und des Weiteren den Kontext, in dem die Passivform auftaucht (Kapitel 4 zur Betrachtung der Ergebnisse).

Zusammenfassend lässt sich festhalten, dass das *gehören*-Passiv eine gefestigte Konstruktion zu sein scheint – die Passivbildung mit *gehören* stellt per se keine ungewöhnliche Ausnahme dar. In den beiden Artikeln, die sich eingehender mit ihm befassen, wird es auch tatsächlich als Passivform gedeutet; diesem Umstand kann und will auch hier nicht widersprochen werden. Der Agens wird dezentriert, der Patiens als Subjekt fokussiert; *gehören* wird als Auxiliar verwendet und bildet gemeinsam mit einem Perfektpartizip die kompakte Konstruktion.

Die am häufigsten anzutreffende grammatische Form von *gehören* ist dabei dritte Person Singular oder Plural. Andere Formen kommen im qualitativen Sample nicht vor, das heißt jedoch nicht, dass gar keine anderen Personen usw. im gesamten Korpus vorkommen (Kapitel 4.2). Weitere Untersuchungen der Konstruktion sollen sich vor allem auch auf diese Fälle fokussieren.

Die Semantik der Verben ist tendenziell neutral bis negativ zu werten, auch wenn es sowohl Beispiele für betont positive Konnotationen gibt – beispielweise *Das gehört gefeiert*. – als auch für sehr stark negative bis hin zur Pejorisation (Kapitel 4.3 und 5.2). Die Subjekte, die hierbei das Patiens darstellen, sind zu einem großen Teil Nominalphrasen, aber nicht nur. Unbelebte Substantiva überwiegen, es werden aber auch belebte – zum Großteil menschliche – Subjekte verwendet (Kapitel 4.5).

Im Gegensatz zu anderen Auxiliaren, die im Zuge der Grammatikalisierung die Eigensemantik verlieren, behält *gehören* etwas davon bei. Diese ist jedoch von der Ursprungsbedeutung der Zugehörigkeit zu einem Ausdruck der Obligation geworden, ein semantischer Shift, der schon vor der Verwendung als Passiv passiert ist. Das Verb hat also diese (deontische) Modalität augenscheinlich nicht nur in der Passivkonstruktion erhalten, sie ist jedoch der Grund, warum sie funktioniert.

Es wurde die Hypothese aufgestellt, dass die Konstruktion einen möglichen Ursprung im Gesprochenen hat und eben über diese Einzug in ein standardsprachlich bzw. schriftliches Korpus, wie es das amc ist, erhält. Die vorliegenden Belege stützen die These, dass die Konstruktion eine starke Verbindung zur Sprechsprachlichkeit hat, was unter anderem dadurch belegt werden kann, dass ein beträchtlicher Teil der Bildungen in einer mündlichen Situation geäußert wurde und durch schriftliche Fixierung in das amc Einzug hielten: Interviews, Transkripte von Radio- und Nachrichtensendungen, direkte (und indirekte) Reden von konkreten Personen in Zeitungsartikeln und dergleichen (siehe Kapitel 5.3).

Ebenso wie die Schriftsprachlichkeit muss auch die Standardsprachlichkeit der Konstruktion relativiert werden, aufgrund der Fälle von dialektal bzw. substandardlich geprägten Ausformungen. Die Form der Mündlichkeit hat also weitreichendere Bedeutung für die untersuchte Varietät als am Anfang angenommen, weshalb auch der Titel der vorliegenden Arbeit zumindest relativiert werden musste (siehe Kapitel 5.3).

Dennoch ist der Kontext des Zeitungskorpus nicht außer Acht zu lassen, da das Passiv in beträchtlicher Beleganzahl angeführt wird, nicht nur in zitierter Mündlichkeit, von Laiensprecher*innen und -schreiber*innen, aber auch von Journalist*innen und Berufssprecher*innen.

In diesem Spannungsfeld befindet sich die *gehören*-Konstruktion, die also auch in der schriftlichen Fixierung Aspekte der Mündlichkeit zeigt, jedoch von Sprecher*innen unmarkiert verwendet wird und dadurch als in dieser Form akzeptiert und etabliert gelten kann.

Es wäre ein Einfaches, aufgrund der Situation in der Forschungsliteratur eine gestreamlinete Konstruktion mithilfe der Belege zu präsentieren, die Untersuchung zeichnet jedoch

ein differenzierteres Bild. Der „klassische“ Satz mit einem *gehören*-Passiv mag zwar folgende Form annehmen: *Das Gesetz gehört geändert*. [Nominalphrase mit einem Substantiv als Kopf, zumeist unbelebt] + [*gehören* in 3P Sgl oder Pl Indikativ Präsens] + [Perfektpartizip mit tendenziell neutraler/negativer Bedeutung]. Es finden sich aber auch genug Belege, die aus dieser Blaupause herausfallen.

Die Untersuchung zeigt also eine Tendenz in der Verwendung, im Besonderen in Bezug auf die Grammatik, aber das *gehören*-Passiv wird produktiv genutzt, um Unterschiedliches zum Ausdruck zu bringen. Die Modalität, die an die verbleibende Eigensemantik von *gehören* gebunden ist, sorgt dafür, dass ein komplexerer Sachverhalt kurz und knapp ausgedrückt werden kann, unabhängig davon, welches Sentiment die gesamte Aussage erhält.

Zum Abschluss soll ein Ausblick gegeben werden, denn weitere Untersuchungen zum *gehören*-Passiv sind wünschenswert, sowohl im amc als auch in weiteren Korpora.

Aufgrund des hier gewählten – trotz Einsprüchen, hauptsächlich – standardschriftsprachlichen Korpus wären weitere Untersuchungen der Konstruktion in einem sprechsprachlichen oder Dialektkorpus interessant.

In zukünftigen Untersuchungen wäre eine Erhebung alternativer Formen wünschenswert. Hierzu braucht es NLP-Skripte, die mithilfe von Word Embeddings diese synonymen gebrauchten Konstruktionen heraussuchen und einen Vergleich ermöglichen. So könnten andere Passivkonstruktionen mit einem modalen Aspekt kontrastiv betrachtet werden. Dies scheint sinnvoller als zu testen, ob sich alle *gehören*-Passive in *werden*-Passive mit *sollen* oder *müssen* umformen lassen, da die Antwort auf diese Frage keine Aufschlüsse über die parallele Verwendung der beiden Konstruktionen geben würde.

Die hier gesammelten und annotierten Daten könnten für die Arbeit mit dem amc genutzt werden, um mithilfe weiterer digitaler Methoden tiefergehendere Einblicke zu gewinnen, wie eben Machine Learning Ansätze, die bereits in Fußnoten und den Kapiteln zur Methode erwähnt wurden, sowie Trainieren eines eigenen Parsers für das *gehören*-Passiv mit den annotierten Belegen.

Zusätzlich könnten Vergleiche mit anderen Konstruktionen von *gehören* angestellt werden, besonders in Hinblick auf die vermutete, über das Passiv hinausgehende Etablierung von *gehören* als Marker von (deontischer) Modalität. In diesem Kontext wäre die Betrachtung in einem historischen Korpus, ähnlich wie es bei Stathi (2009; 2010) gemacht wurde, von Interesse, aber auch im hier verwendeten amc.

Der Wunsch nach Zugriff auf mehr Daten aus dem amc besteht, auch um Probleme bei der Filterung zu minimieren. Zudem befindet sich ein neues Analyse Tool CorpSum in der Entwicklung befindet, das ein neues Interface bereitstellt, welches genauere Betrachtungen und Visualisierungen zu Metadaten ermöglicht.

Eine mögliche Weiterentwicklung wäre ebenfalls die Untersuchung weiterer Passivformen auf ähnliche Weise, wie beispielsweise das Rezipienten-Passiv.

7. Literaturverzeichnis

- Ágel, Vilmos (1996): Was gibt's Neues übers Passiv? Funktion, Typen, Bildung. In: Deutschunterricht für Ungarn 11 / II (1996), 76–87.
- Ágel, Vilmos/Hennig, Mathilde (2006): Theorie des Nähe- und Distanzsprechens. In: Ágel, Vilmos/Hennig, Mathilde (Hg.): Grammatik aus Nähe und Distanz. Theorie und Praxis am Beispiel von Nähertexten 1650-2000. Tübingen: Niemeyer, 3–31.
- Ágel, Vilmos/Hennig, Mathilde (2007): Überlegungen zur Theorie und Praxis des Nähe- und Distanzsprechens. In: Ágel, Vilmos/Hennig, Mathilde (Hg.): Zugänge zur Grammatik der gesprochenen Sprache. Tübingen: Niemeyer, 179–214.
- Altinok, Duygu (2021): Mastering spaCy. Birmingham/Mumbai: Packt Publishing.
- Ammon, Ulrich (1995): Die deutsche Sprache in Deutschland, Österreich und der Schweiz: Das Problem der nationalen Varietäten. Berlin/New York: De Gruyter.
- Ammon, Ulrich (2017): Grundbegriffe für plurizentrische Sprachen und ihre Operationalisierung – mit Bezug auf die deutsche Sprache und das „Variantenwörterbuch des Deutschen“: In: Sieburg, Heinz/Solms, Hans-Joachim (Hg.): Das Deutsche als plurizentrische Sprache: Ansprüche - Ergebnisse - Perspektiven, Zeitschrift für deutsche Philologie: Berlin: Erich Schmidt Verlag, 5–22.
- Ammon, Ulrich/Bickel, Hans/Lenz, Alexandra Nicole (Hg.) (2016): Variantenwörterbuch des Deutschen: Die Standardsprache in Österreich, der Schweiz, Deutschland, Liechtenstein, Luxemburg, Ostbelgien und Südtirol sowie Rumänien, Namibia und Mennonitensiedlungen. Berlin/New York: De Gruyter.
- Amrhein, Jürgen (1995). Valenztheoretische Aspekte der Auxiliarisierung. Die Grammatikalisierung des *werden*-Passivs und der Agensanschluß mit *von*. In: Jahrbuch der ungarischen Germanistik 1995, 85–108.
- Askedal, John Ole (2005): Grammatikalisierung und Persistenz im deutschen „Rezipienten-Passiv“ mit *bekommen* / *kriegen* / *erhalten*: In: Leuschner, Torsten/Mortelmans, Tanja/De Groot, Sarah (Hg.): Grammatikalisierung im Deutschen. Berlin/New York: De Gruyter, 211–228.
- Auer, Peter (2014): Enregistering pluricentric German: In: Soares Da Silva, Augusto (Hg.): Pluricentricity. Berlin/New York: De Gruyter, 19–48.
- Bybee, Joan / Perkins, Revere / Pagliuca, William (1994): The Evolution of Grammar. Tense, Aspect and Modality in the Languages of the World. Chicago [u.a.]: University of Chicago Press.
- Culpeper, Jonathan / Haugh, Michael / Kádár, Dániel Z. (Hg.) (2017): The Palgrave Handbook of Linguistic (Im)politeness. London: Palgrave Macmillan UK.
- Dalianis, Hercules (2018): Evaluation Metrics and Evaluation: In: Clinical Text Mining. Cham: Springer International Publishing, 45–53.
- De Cillia, Rudolf / Ransmayr, Jutta (2019): Österreichisches Deutsch macht Schule. Bildung und Deutschunterricht im Spannungsfeld von sprachlicher Variation und Norm. Wien/Köln/Weimar: Böhlau.
- Diewald, Gabriele (1999): Die Modalverben im Deutschen: Grammatikalisierung und Polyfunktionalität. Berlin/New York: De Gruyter.
- Dorn, Amelie / Höll, Jan / Ziegler, Theresa / Koppensteiner, Wolfgang / Pirker, Hannes (2023): Die österreichische Presselandschaft digital: Das Austrian Media Corpus (amc) und sein Potenzial für die Linguistik.: In: Kupietz, Marc / Schmidt, Thomas (Hg.): Neue Entwicklungen in der Korpuslandschaft der Germanistik: Beiträge zur IDS-Methodenmesse 2022, Korpuslinguistik und Interdisziplinäre Perspektiven auf Sprache: Paper präsentiert bei der Institut für deutsche Sprache. Tübingen: Narr Francke, 43–55.
- Eisenberg, Peter (2020): Der Satz. 5., aktualisierte und überarbeitete Auflage. Grundriss der deutschen Grammatik. Berlin: J.B. Metzler.
- Elspaß, Stephan (2005): Sprachgeschichte von unten: Untersuchungen zum geschriebenen Alltagsdeutsch im 19. Jahrhundert: Reihe germanistische Linguistik. Tübingen: Niemeyer.
- Elspaß, Stephan / Dürscheid, Christa / Ziegler, Arne (2017): Zur grammatischen Pluriarealität der deutschen Gebrauchsstandards – oder: Über die Grenzen des Plurizentritätsbegriffs.: In:

- Zeitschrift für deutsche Philologie (2017. Thematisches Sonderheft: „Das Deutsche als plurizentrische Sprache. Ansprüche – Ergebnisse – Perspektiven“), 69–91.
- Engel, Ulrich (2009): Deutsche Grammatik. Neubearbeitung, 2., Aufl. München: Iudicium.
- Eroms, Hans-Werner (2000): Syntax der deutschen Sprache. Berlin/New York: De Gruyter.
- Gaeta, Livio (2018): Im Passiv sprechen in den Alpen. In: Sprachwissenschaft (43 (2)), 159–185.
- Ginsburg, Shiphra / Van Der Vleuten, Cees / Eva, Kevin W. / Lingard, Lorelei (2016): Hedging to save face: a linguistic analysis of written comments on in-training evaluation reports. In: Advances in Health Sciences Education 21 (1), 175–188.
- Glauninger, Manfred (2015b): Die Nationalvarietät „österreichisches Deutsch“ als kakanisches Paradoxon. In: Cwanek-Florek, Ewa / Nöbauer, Irmgard (Hg.): Sprachliche Wechselbeziehungen in der Habsburgermonarchie. Wien: Polnische Akademie der Wissenschaften (Symposien und Seminare am Wissenschaftlichen Zentrum der Polnischen Akademie der Wissenschaften in Wien 13), 11–18.
- Glauninger, Manfred Michael (2013): Deutsch im 21. Jahrhundert: ‚pluri‘-, ‚supra‘- oder ‚postnational‘? In: Sava, Doris/Lăzărescu, Ioan/Scheuringer, Hermann (Hg.): Im Dienste des Wortes: lexikologische und lexikografische Streifzüge ; Festschrift für Ioan Lăzărescu, Forschungen zur deutschen Sprache in Mittel-, Ost- und Südosteuropa. Passau: Stutz, 123–132.
- Glauninger, Manfred Michael (2015a): (Standard-)Deutsch in Österreich im Kontext des gesamtdeutschen Sprachraums. Perspektiven einer funktional dimensionierten Sprachvariationstheorie: In: Lenz, Alexandra N. / Glauninger, Manfred Michael (Hg.): Standarddeutsch im 21. Jahrhundert: theoretische und empirische Ansätze mit einem Fokus auf Österreich, Wiener Arbeiten zur Linguistik: Göttingen: V & R unipress [u.a.], 11–58.
- Goldberg, A. E. (1995). Constructions. A construction grammar approach to argument structure. Chicago: The University of Chicago Press.
- Heine, Bernd (1993): Auxiliaries. Cognitive Forces and Grammaticalization. New York: OUP.
- Heine, Bernd / Kuteva, Tania (2002): World lexicon of grammaticalization. Cambridge: Cambridge Univ. Press.
- Heinold, Simone Beatrice (2015): Tempus, Modus und Aspekt im Deutschen. Ein Studienbuch. Tübingen: Narr Verlag.
- Helbig, Gerhard (1989). Das Passiv - und kein Ende. Deutsch als Fremdsprache 26.4, 215–221.
- Hoey, Michael (2005): Lexical priming: a new theory of words and language. London: Routledge.
- Höll, Jan (2020): Wissensstrukturen im Kontext der deutschen Rechtschreibreform. Framebasierte Analyse eines medialen Diskurses in Österreich. Masterarbeit, Universität Wien.
- Hopper, Paul J. (1990): „Principles of Grammaticization: Towards a Diachronic Typology.“ In: Winfred P. Lehmann (Hg.), Language Typology 1987. Systematic Balance in Language. Papers from the Linguistic Typology Symposium, Berkeley, 1-3 December 1987. (CILT 67). Amsterdam/Philadelphia: Benjamins, 157–170.
- Kaltenböck, Gunther / Mihatsch, Wiltrud / Schneider, Stefan (Hg.) (2010): New approaches to hedging. Bingley: Emerald.
- Kant, Immanuel (1787/2003): Kritik der reinen Vernunft. Hg. von Ingeborg Heidemann. Stuttgart: Reclam.
- Kluge, Friedrich / Seebold, Elmar (1999): Etymologisches Wörterbuch der deutschen Sprache. 23. Aufl. bearb. von Elmar Seebold. Berlin New York: De Gruyter.
- Koch, Peter / Österreicher, Wulf (1985): Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld der Sprachtheorie und Sprachgeschichte. In: Romanistisches Jahrbuch 36, 15–43.
- Koch, Peter / Österreicher, Wulf (1990): Gesprochene Sprache in der Romania: Französisch, Italienisch, Spanisch. Tübingen. Niemeyer.
- Koch, Peter / Österreicher, Wulf (1994): Schriftlichkeit und Sprache. In: Günther, Hartmut / Ludwig, Otto (Hg.): Schrift und Schriftlichkeit. Ein interdisziplinäres Handbuch internationaler Forschung. 1. Halbband. Berlin/New York, 587–604.
- Koppensteiner, Wolfgang (2015): Das österreichische Deutsch im plurizentrischen Kontext. Eine korpuslinguistische Untersuchung der österreichischen Presse im Zeitraum von 1986 – 2013. Diplomarbeit, Universität Wien.

- Koppensteiner, Wolfgang (2023): Standard(s) in Österreich: Attitudinal-perzeptive Analysen. Dissertation, Universität Wien.
- Koppensteiner, Wolfgang / Lenz, Alexandra N. (2017): Theoretische und methodische Herausforderungen einer perzeptiv-attitudinalen Standardsprachforschung. Perspektiven aus und auf Österreich. In: Sieburg, Heinz / Solms, Hans-Werner (Hg.), *Das Deutsche als plurizentrische Sprache. Ansprüche – Ergebnisse – Perspektiven*. Sonderheft Zeitschrift für deutsche Philologie 136, 43–68.
- Koppensteiner, Wolfgang / Lenz, Alexandra N. (2021): Standard(s) aus der Perspektive von „Nicht-LinguistInnen“ in Österreich: In: Hoffmeister, Toke / Hundt, Markus / Nath, Saskia (Hg.): *Laien, Wissen, Sprache*. Berlin/New York: De Gruyter, 391–416.
- Korecky-Kröll, Katharina (2020): Morphological Dynamics of German Adjective Gradation in Rural Regions of Austria: In: *Zeitschrift für Dialektologie und Linguistik* 87 (1), 25–65.
- Lasch, Alexander (2018): „Diese gehören kalt zu geben.“ Die Konstruktion „gehören“ mit Qualitativ. In: *Sprachwissenschaft* (43 (2)), 159–185.
- Lenz, Alexandra N. (2018): GIVE- und GET-Passive im Deutschen und Luxemburgischen. Von der Affinität transferentieller Verben zur Passiv-Auxiliarisierung: In: *Sprachwissenschaft* (43 (2)), 187–219.
- Lenz, Alexandra N. / Dorn, Amelie / Ziegler, Theresa (2021): Lexik aus areal-horizontaler und vertikal-sozialer Perspektive – Erhebungsmethoden zur inter- und intraindividuellen Variation. In: *Sprachwissenschaft* 46/4, 387–431.
- Louw, B. (1993). “Irony in the text or sincerity in the writer? The diagnostic potential of semantic prosodies”, in M. Baker, G. Francis and E. Tognini-Bonelli (Hg.), *Text and technology: In honour of John Sinclair*. Philadelphia: John Benjamins, 157–176.
- Louw, Bill (1993): Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies: In: *Text and Technology: In Honour of John Sinclair*. Philadelphia: Benjamins Pub. Co., 157–176.
- Malik, Preeti / Mittal, Varsha / Nautiyal, Lata / Ram, Mangey (2022): NLP techniques, tools, and algorithms for data science: In: Sharma, Abhinav/Jain, Arpit/Kumar Arya, Ashwini/Ram, Mangey (Hg.): *Artificial Intelligence for Signal Processing and Wireless Communication*. Berlin/New York: De Gruyter, 123–148.
- Moore, Colette (2006): The use of videlicet in Early Modern slander depositions: A case of genre-specific grammaticalization: In: *Journal of Historical Pragmatics* 7 (2), 245–263.
- Moosmüller, Sylvia (1991): *Hochsprache und Dialekt in Österreich: soziophonologische Untersuchungen zu ihrer Abgrenzung in Wien, Graz, Salzburg und Innsbruck*. Sprachwissenschaftliche Reihe. Wien/Köln/Weimar: Böhlau.
- Muhr, Rudolf / Schrod, Richard / Wiesinger, Peter (Hg.) (1995): *Österreichisches Deutsch. Linguistische, sozialpsychologische und sprachpolitische Aspekte einer nationalen Variante des Deutschen*. Wien: Holder-Pichler-Tempsky.
- Nuyts, Jan (2006): “Modality: Overview and Linguistic Issues”. In: Frawley, Williams (Hg.): *The Expression of Modality*. Berlin/New York: De Gruyter, 1–26.
- Palmer, Frank R. (2001): *Mood and Modality*. 2. Ausg. Cambridge [u.a.]: Cambridge University Press.
- Ransmayr, Jutta / Mörth, Karlheinz / Ďurčo, Matej (2017): AMC (Austrian Media Corpus) - Korpusbasierte Forschungen zum österreichischen Deutsch: In: Resch, Claudia / Dressler, Wolfgang U. (Hg.): *Digitale Methoden der Korpusforschung in Österreich, Sitzungsberichte / Österreichische Akademie der Wissenschaften, Philosophisch-Historische Klasse*. Wien: Verlag der Österreichischen Akademie der Wissenschaften, 27–38.
- Reis, Marga (1976): Zum grammatischen Status der Hilfsverben. In: *Beiträge zur Geschichte der deutschen Sprache und Literatur (PBB)* 1976 (98).
- Scheuringer, Hermann (1996): Das Deutsche als pluriareale Sprache: Ein Beitrag gegen staatlich begrenzte Horizonte in der Diskussion um die deutsche Sprache in Österreich. In: *Die Unterrichtspraxis / Teaching German* 29/2, 147–153.
- Schiller, Anne / Teufel, Simone / Thielen, Christine / Stöckert, Christine (1999): Guidelines für das Tagging deutscher Textcorpora mit STTS (Kleines und großes Tagset). <https://www.sfs.uni-tuebingen.de/resources/stts-1999.pdf> [Zuletzt zugegriffen 26.2.2024].

- Schumann, Michael (2010): Modalität. In: Hentschel, Elke (Hg.): Deutsche Grammatik. Berlin: De Gruyter, 180–183.
- Söll, Ludwig (1974, 1985): Gesprochenes und geschriebenes Französisch. Berlin.
- Soukup, Barbara (2009): Dialect use as interaction strategy: a sociolinguistic study of contextualization, speech perception, and language attitudes in Austria. Austrian studies in English. Wien: Braumüller.
- Srinivasa-Desikan, Bhargav (2018): Natural language processing and computational linguistics: a practical guide to text analysis with Python, Gensim, spaCy, and Keras. Expert Insight. Birmingham/Mumbai: Packt Publishing.
- Stathi, Katerina (2009): The Origin of Semantic Change in Discourse Tradition: A Case Study. In: Mosegaard Hansen, Maj-Britt/Visconti, Jacqueline (Hg.): Current Trends in Diachronic Semantics and Pragmatics: Bingley, England: Emerald, 291–302.
- Stathi, Katerina (2010): Is German *gehören* an auxiliary? The grammaticalization of the construction *gehören* + participle II. In: Stathi, Katerina / Gehweiler, Elke / König, Ekkehard (Hg.): Grammaticalization: current views and issues. Studies in language companion series (SLCS). Amsterdam/Philadelphia: Benjamins Pub. Co, 323–342.
- Stein, Dieter (2001): Distanz und Nähe in interaktiver Internetkommunikation. In: Beißwenger, Michael (Hg.): Chat-Kommunikation. Sprache, Interaktion, Sozialität & Identität in synchroner computervermittelter Kommunikation. Perspektiven auf ein interdisziplinäres Forschungsfeld. Stuttgart: Ibidem, 339–347.
- Szatmári, Petra (2002): *Das gehört nicht vom Tisch gewischt...* Überlegungen zu einem modalen Passiv und dessen Einordnung ins Passiv-Feld. In: Jezikoslovlje (3.1–2), 171–192.
- Szczepaniak, Renata (2011): Grammatikalisierung im Deutschen: eine Einführung. 2. Aufl. Tübingen: Narr Verlag.
- Thaler, Verena (2007): Mündlichkeit, Schriftlichkeit, Synchronizität. Eine Analyse alter und neuer Konzepte zur Klassifizierung neuer Kommunikationsformen. In: ZFGL 35 (1–2), 146–181.
- Traugott, Elisabeth C. / Hopper Paul J. (2012): Grammaticalization. 2. Aufl. Cambridge [u.a.]: Cambridge Univ. Press.
- Traugott, Elisabeth C. (1988) Pragmatic strengthening and grammaticalization. In: Berkeley Linguistics Society. Proceedings of the Fourteenth Annual Meeting, February 13–15, 1988. General Session and Parasession on Grammaticalization, 406–416.
- Traugott, Elisabeth C. / König, Ekkehard (1991): The semantics-pragmatics of grammaticalization revisited. In: Traugott, Elisabeth C. / Heine, Bernd (Hg.): Approaches to Grammaticalization. Amsterdam: John Benjamins. (Typological Studies in Language 19), 189–218.
- Vogel, Petra Maria (2006): Das unpersönliche Passiv: eine funktionale Untersuchung unter besonderer Berücksichtigung des Deutschen und seiner historischen Entwicklung. Studia linguistica Germanica. Berlin/New York: De Gruyter.
- Vogel, Petra Maria (2010): Passiv. In: Hentschel, Elke (Hg.): Deutsche Grammatik. Berlin: De Gruyter, 224–230.
- Wegener, Heide (1998): Zur Grammatikalisierung von Modalpartikeln. In: Barz, Irmhild/Öhlschläger, Günther (Hg.): Zwischen Grammatik und Lexikon. Berlin/New York: De Gruyter, 37–56.
- Whitsitt, Sam (2005): A critique of the concept of semantic prosody. In: International Journal of Corpus Linguistics 10 (3), 283–305.
- Wiesinger, Peter (2014): Das österreichische Deutsch in Gegenwart und Geschichte. 3. Aufl. Wien/Münster: Lit.
- Wöllstein-Leisten, Angelika/Eisenberg, Peter/Dudenredaktion (Bibliographisches Institut) (Hg.) (2016): Duden, die Grammatik: unentbehrlich für richtiges Deutsch. 9. Auflage. Berlin: Dudenverlag.
- Wright, Georg H. von (1951): An essay in Modal Logic. Amsterdam: North Holland.
- Ziegler, Theresa (2021): Über Geschenk-s-körbe und Schokolade-Ø-torten (zu runden Geburtstagen). In: Wiener Linguistische Gazette (WLG) 89 (2021), 55–83.
- Zifonun, Gisela / Hoffmann, Ludger / Strecker, Bruno (2011): Grammatik der deutschen Sprache. Berlin: De Gruyter.

7.1 Verwendete Links

- amc: Das amc in Zahlen. <https://amc.acdh.oeaw.ac.at/dokumentation/amc-in-zahlen/> [Zuletzt zugegriffen 15.11.2023]
- amc: Das amc! <https://amc.acdh.oeaw.ac.at/about-amc/> [Zuletzt zugegriffen 15.11.2023]
- amc: Korpusinhalt-Attribute. doc.ressort2: <https://amc.acdh.oeaw.ac.at/dokumentation/korpusinhalt-attribute/#doc.ressort2> [Zuletzt zugegriffen 15.11.2023]
- amc: Korpusinhalt-Attribute. <https://amc.acdh.oeaw.ac.at/dokumentation/korpusinhalt-attribute/> [Zuletzt zugegriffen 15.11.2023]
- amc: Normalisierung. <https://amc.acdh.oeaw.ac.at/normalisierung/> [Zuletzt zugegriffen 15.11.2023]
- amc: Publikationen. <https://amc.acdh.oeaw.ac.at/publikationen/> [Zuletzt zugegriffen 15.11.2023]
- BBAW: DWB. <https://www.bbaw.de/forschung/deutsches-woerterbuch-von-jacob-grimm-und-wilhelm-grimm> [Zuletzt zugegriffen 15.11.2023]
- Grammis: *aufräumen*. <https://grammis.ids-mannheim.de/verbs/view/400265/1> [Zuletzt zugegriffen 28.2.2024]
- Grammis: *gehören*. <https://grammis.ids-mannheim.de/verbvalenz/400655> [Zuletzt zugegriffen 15.11.2023]
- Online-Duden: *gehören*. <https://www.duden.de/rechtschreibung/gehen> [Zuletzt zugegriffen 15.11.2023]
- Online-Duden: *politisieren*. <https://www.duden.de/rechtschreibung/politisieren> [Zuletzt zugegriffen 15.11.2023]
- OSCE (2004): Überblick über die österreichische Gesetzeslage im Bereich "Nationalsozialistische Wiederbetätigung und Rassismus". <https://www.osce.org/files/f/documents/e/2/31049.pdf> [Zuletzt zugegriffen 26.2.2024]
- Presse (2020): ORF gehört „gehört gehört“ <https://www.diepresse.com/5754844/orf-gehört-gehört-gehört-gehört> [Zuletzt zugegriffen 20.2.2024]
- Radioszene (2012): SWR 1 Baden-Württemberg – Reichweite mit Melodie. <https://www.radioszene.de/36600/swr-1-baden-wuerttemberg-reichweite-mit-melodie.html> [Zuletzt zugegriffen 23.7.2023]
- Spacy: Models. <https://spacy.io/usage/models> [Zuletzt zugegriffen 16.3.2024]
- Standard (2020): „Ö1 gehört gehört“. Radiosender darf Slogan unbegrenzt verwenden. <https://www.derstandard.at/story/2000117600042/oe1-gehört-gehört-radiosender-darf-slogan-unbegrenzt-verwenden> [Zuletzt zugegriffen 20.2.2024]
- Verbotsgesetz = Bundesrecht konsolidiert: Gesamte Rechtsvorschrift für Verbotsgesetz 1947, Fassung vom 26.2.2024. <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10000207> [Zuletzt zugegriffen 26.2.2024]
- Wörterbuchnetz: DWB. <https://woerterbuchnetz.de/?sigle=DWB&lemid=G05071> [Zuletzt zugegriffen 15.11.2023]
- amc: amc_4.2: Medienliste. <https://amc.acdh.oeaw.ac.at/dokumentation/medienliste/> [Zuletzt zugegriffen 15.11.2023]
- amc: Austria Media Corpus (amc), Version 4.2, zugänglich über [\[http://hdl.handle.net/21.11115/0000-000C-4F08-4\]](http://hdl.handle.net/21.11115/0000-000C-4F08-4), abgerufen am 15.12.2023.
- amc: Versionsarchiv 4.2. https://amc.acdh.oeaw.ac.at/dokumentation/dokumentation-versionsarchiv/#amc_4.2 [Zuletzt zugegriffen 15.11.2023]
- Spacy. <https://spacy.io/> [Zuletzt zugegriffen 16.3.2024]

7.2 Verwendete Packages

- Package pandas <https://pandas.pydata.org/docs/> [Zuletzt zugegriffen 15.11.2023]
- Package re <https://docs.python.org/3/library/re.html> [Zuletzt zugegriffen 15.11.2023]
- Model de_dep_news_trf-3.5.0 https://github.com/explosion/spacy-models/tree/de_dep_news_trf-3.5.0 [Zuletzt zugegriffen 15.11.2023]
- Package spaCy 3.7.2 <https://pypi.org/project/spacy/> [Zuletzt zugegriffen 15.11.2023]
- Python 3.11.3 <https://www.python.org/downloads/release/python-3113/> [Zuletzt zugegriffen 15.11.2023]

8. Abbildungsverzeichnis

Abbildung 1: Die nationalen Zentren der deutschen Sprache (VWB 2016: XLII)	12
Abbildung 2: Soziales Kräftefeld einer Standardvarietät (Ammon 1995: 80; 2017: 11)	14
Abbildung 3: Anpassung der Theorie Hoppers (1990) als hierarchisches Baumdiagramm (Askedal 2005: 215)	18
Abbildung 4: Stufen der Grammatikalisierung (Szczenepaniak 2011: 37)	19
Abbildung 5: Passiv-Mikrofeld 1 [-Modalfaktor] (Szatmári 2002: 175)	21
Abbildung 6: Passiv-Mikrofeld [+Modalfaktor] im Passiv-Feld (Grobgliederung) (Szatmári 2002: 178).....	33
Abbildung 7: <i>gehören</i> -Formen im Korpus (Stathi 2010: 332)	35
Abbildung 8: Verben als Perfektpartizip in der Konstruktion (Stathi 2010: 335)	36
Abbildung 9: Liste der im Kleinkorpus aufscheinenden Vollverben (Szatmári 2002: 180)	37
Abbildung 10: Semantische Merkmale der Patiens-Lexeme des Kleinkorpus (Szatmári 2002: 182).....	40
Abbildung 11: Einordnung des <i>gehören</i> -Passivs ins Passiv-Mikrofeld 2 [+Modalfaktor] (Szatmári 2002: 189)	42
Abbildung 12: Grafiken zum amc (Dorn et al. 2023: 45); CC-BY Theresa Ziegler	44
Abbildung 13: Zuteilung der einzelnen Medien zu den Regionskategorien (vgl. amc: Korpusinhalt-Attribute #doc.region)	45
Abbildung 14: Abfrage des <i>gehören</i> -Verbs über die NoSketch Engine des amc	47
Abbildung 15: Confusion matrix (Dalianis 2018: 47).....	51
Abbildung 16: Berechnung der Precision- und Recall-Werte (vgl. Dalianis 2018, 47)	51
Abbildung 17: Anzahl der <i>gehören</i> -Passiv-Belege nach Jahren (absolut)	68
Abbildung 18: Anzahl der <i>gehören</i> -Passiv-Belege nach Jahren (normalisiert).....	68
Abbildung 19: Anzahl der <i>gehören</i> -Passiv-Belege in den einzelnen Medien des Korpus (absolut).....	70
Abbildung 20: Anzahl der <i>gehören</i> -Passiv-Belege in den einzelnen Medien (normalisiert, PPM).....	72
Abbildung 21: Aufteilung der <i>gehören</i> -Passiv-Belege auf die drei Medientypen agentur , print und tv (absolut).....	73
Abbildung 22: Aufteilung der <i>gehören</i> -Passiv-Belege auf die drei Medientypen agentur , print und tv (normalisiert)	74

Abbildung 23: Belegzahlen des <i>gehören</i> -Passivs nach Region (mitsamt agesamt und spezifisch) (normalisiert)	75
Abbildung 24: <i>gehören</i> -Passiv-Belege nach Regionen (normalisiert), dargestellt auf der Landkarte CC-BY Theresa Ziegler	76
Abbildung 25: Die Anzahl der <i>gehören</i> -Formen im Korpus (absolut)	77
Abbildung 26: Die 20 häufigsten Verben im quantitativen Korpus (absolut)	80
Abbildung 27: Gestalt des Subjekts im qualitativen Sample	85
Abbildung 28: Einteilung der Subjekte in belebt und unbelebt	86
Abbildung 29: Die Ressortzuteilungen im quantitativen Korpus (nur erste Zuteilung) (absolut)	90
Abbildung 30: Zuteilung zu den Ressorts im qualitativen Sample (nach Gruppierung der Erstreichung) (absolut)	91
Abbildung 31: Klassifizierung der Redeweise im qualitativen Sample	92

9. Tabellenverzeichnis

Tabelle 1: Accuracy, Precision, Recall und F1-Score der unterschiedlich bearbeiteten Daten	58
Tabelle 2: Absolute Zahlen der Extraktionsvarianten	58
Tabelle 3: Entwicklung der Belegzahlen der quantitativen Datensammlung	65
Tabelle 4: Anzahl der <i>gehören</i> -Passiv-Belege nach Jahr seit 2010 (absolut und normalisiert)	67
Tabelle 5: absolute und normalisierte Belegzahlen zum <i>gehören</i> -Passiv nach Medientyp (agentur , print , tv)	74
Tabelle 6: absolute und normalisierte Belegzahlen des <i>gehören</i> -Passivs nach Region (mitsamt agesamt und spezifisch)	75
Tabelle 7: Formen von <i>gehören</i> morphologisch aufgeschlüsselt, qualitatives Sample (absolut)	78
Tabelle 8: die 20 häufigsten Verben in beiden Teilen der Untersuchung, kontrastiert	81
Tabelle 9: Beispiele für die häufigsten Verben aus dem qualitativen Sample	82
Tabelle 10: Sentimente der 20 häufigsten Verben im quantitativen Korpus und qualitativen Sample: positiv (grün), neutral (grau), negativ (orange)	83
Tabelle 11: Unterteilung der belebten Subjekte in tierisch und menschlich , qualitatives Sample	86

Tabelle 12: lokative Präpositionalphrasen im qualitativen Sample	88
Tabelle 13: direktive oder statische Lesart der lokativen Präpositionen	89
Tabelle 14: Agensanschluss in <i>gehören</i> -Passiv-Belegen	89
Tabelle 15: Modus des Verbes und Klassifikation der Redeart im qualitativen Korpus	93
Tabelle 16: Klassifikation der Redeart, Verwendung von Anführungszeichen und Modus des Verbes im qualitativen Sample	93
Tabelle 17: Übersicht über die VVPPs aus der Forschungsliteratur.....	97
Tabelle 18: VVPPs bei Stathi (2010: 335) und im quantitativen Korpus.....	100

10. Anhang

Anhang A: Github Repository

https://github.com/claudiamattes98/gehoeren_passiv

Siehe auch 7.2 für die Quellenangabe der dazu verwendeten Packages

Anhang B: Medien-/Normalisierungstabelle

Medien, Kürzel, Erscheinungsjahr, Berechnungen zur Normalisierung

Kürzel	Medien	Erscheinung	Gesamttoke	Anzahl gehören	Frequency (ppm)
ACADEMIA	Academia	2010-2022	1203049	14	11,637
APA	APA-Meldungen digital	2010-2022	808625584	2401	2,969
ARBEITW	Arbeit und Wirtschaft	2010-2022	2894712	32	11,055
ATVVOLL	ATV (Fernsehsender)	2010-2020	2764395	43	15,555
AUGUSTIN	Augustin	2010-2022	7837702	66	8,421
BAUERNZT	Bauernzeitung	2010-2022	26656177	228	8,553
BVZ	Burgenländische Volksze	2010-2022	114256910	840	7,352
DATUM	Datum	2010-2022	4183858	27	6,453
DIEWIR	Die Wirtschaft	2010-2022	2057690	11	5,346
ECHO	Echo	2010-2022	6161286	82	13,309
EMEDIA	E-Media	2010-2022	6127868	20	3,264
FALTER	Falter	2010-2022	39427630	466	11,819
FORMATDB	Format	2010-2022	16057872	153	9,528
FURCHE	Die Furche	2010-2022	19456683	152	7,812
GEWINN	Gewinn	2010-2022	9452545	62	6,559
GRAZER	Der Grazer	2010-2022	9830921	178	18,106
HEUTE	Heute	2010-2022	39533500	336	8,499
HOR	Horizont	2010-2022	13004727	76	5,844
IM	Industriemagazin	2010-2022	3816587	16	4,192
KLEINE	Kleine Zeitung	2010-2022	395821518	3157	7,976
KONSUM	Der Konsument	2010-2022	4509928	30	6,652
KRONE	Kronen Zeitung	2010-2022	566921812	4738	8,357
KTNMONAT	Neuer Kärntner Monat	2010-2013	1418954	4	2,819
KTZ	Kärntner Tageszeitung	2010-2014	33578615	244	7,267
KURIER	Kurier	2010-2022	424615493	2052	4,833
KW	Kärntner Wirtschaft	2010-2022	5766638	58	10,058
MEDIANET	Medianet	2010-2022	30667362	122	3,978
MWVOLL	ORF (Radio und TV)	2010-2022	186314769	2993	16,064
NEWS	News	2010-2022	24173854	261	10,797
NOEN	NÖ Nachrichten	2010-2022	602616808	3917	6,500
NVB	Neues Volksblatt	2010-2022	84205457	509	6,045
NVT	Neue Vorarlberger Tages	2010-2022	86579077	538	6,214
OBERRUND	Oberländer Rundschau	2010-2022	44701889	389	8,702
OEREICHE	Österreich	2010-2022	159390863	1192	7,478
OOEN	OÖ Nachrichten	2010-2022	285940722	1773	6,201
OTS	Original Text Service	2010-2022	356161679	3774	10,596
PRESSE	Die Presse	2010-2022	227059850	1489	6,558
PRO7VOLL	Pro 7 (Fernsehsender)	2010-2020	699018	19	27,181
PROFIL	Profil	2010-2022	31251158	288	9,216
PULSVOLL	Puls 4 (Fernsehsender)	2010-2020	2055487	60	29,190
SAT1VOLL	Sat1 (Fernsehsender)	2010-2020	1002785	24	23,933
SBGW	Salzburger Woche	2010-2022	53251435	759	14,253
SN	Salzburger Nachrichten	2010-2022	190121229	1822	9,583
SOLI	Solidarität	2010-2022	544095	5	9,190
SPORTZTG	Sportzeitung	2010-2022	8713975	29	3,328
STANDARD	Der Standard	2010-2022	198494782	2058	10,368
STMONAT	Steirer Monat	2010-2012	891854	2	2,243
SVZ	Salzburger Volkszeitung	2010-2014	17631357	63	3,573
TREND	Trend	2010-2022	11924278	70	5,870
TT	Tiroler Tageszeitung	2010-2022	195086645	2229	11,426
TKOMP	TT Kompakt	2010-2022	24898771	169	6,787
TMEDIA	TV-Media	2010-2022	28628135	73	2,550
VN	Vorarlberger Nachrichten	2010-2022	174055984	1244	7,147
WIBLATT	Wirtschaftsblatt	2010-2016	36184385	303	8,374
WIENER	Wiener	2010-2022	4459027	42	9,419
WIENERIN	Wienerin	2010-2022	6361728	41	6,445
WOMAN	Woman	2010-2022	13898968	156	11,224
WZ	Wiener Zeitung	2010-2022	214688519	709	3,302

Anhang C: qualitatives Sample

Tabelle C.1: Form des Subjekts

Form des Subjekts	Anzahl <i>gehören-Passiv</i>
allgemeines <i>das</i>	32
formales Subjekt <i>es</i>	19
Nebensatz	4
Nominalphrase	356
pronominaler Bezug	54
subjektlos	4
fehlt	3
Gesamtergebnis	472

Tabelle C.2: Semantische Klassifikation der Subjekte

Semantische Klassifikation	Anzahl <i>gehören-Passiv</i>
belebt	70
unbelebt	351
Keine Klassifikation	51
Gesamtergebnis	472

Tabelle C.3: Klassifizierung der Redeart

Klassifizierung	Anzahl <i>gehören-Passiv</i>
-	170
direkte Rede	215
indirekte Rede	78
uneindeutig	9
Gesamtergebnis	472

Tabelle C.4: menschliche Subjekte im qualitativen Sample

Subjektphrase	Bezug auf	belebt/unbelebt	menschlich/tierisch	Form des Subjekts
die Banker	-	belebt	Menschlich	Nominalphrase
Ferienarbeiter	-	belebt	Menschlich	Nominalphrase
die Unteren	-	belebt	Menschlich	Nominalphrase
sie	beide	belebt	Menschlich	pronominaler Bezug
Steuerhinterzieher		belebt	Menschlich	Nominalphrase
beide		belebt	Menschlich	Nominalphrase
die Honigbienen		belebt	Tierisch	pronominaler Bezug
rechte Populisten		belebt	Menschlich	Nominalphrase
Pfuscher		belebt	Menschlich	Nominalphrase
Geisterfahrer		belebt	Menschlich	Nominalphrase
Eltern, die den Bildungserfolg ihrer Kinder gefährden		belebt	Menschlich	Nominalphrase
Unternehmer		belebt	Menschlich	Nominalphrase
sie	Hillary Clinton	belebt	Menschlich	pronominaler Bezug
Herr Michalis		belebt	Menschlich	Nominalphrase
Unternehmer		belebt	Menschlich	Nominalphrase
sie	Krisenpflegeeltern	belebt	Menschlich	pronominaler Bezug
dieses Nationalteam		belebt	Menschlich	Nominalphrase
die Bürger		belebt	Menschlich	Nominalphrase
diese Leute		belebt	Menschlich	Nominalphrase
die	Steirer	belebt	Menschlich	pronominaler Bezug
Frau und Mutter		belebt	Menschlich	Nominalphrase
alle türkisen Ministerinnen		belebt	Menschlich	Nominalphrase
schwarze Schafe		belebt	Menschlich	Nominalphrase
die ganze Familie		belebt	Menschlich	Nominalphrase
die	Volkvertreter	belebt	Menschlich	pronominaler Bezug
Die derzeitige Politikaste samt EU-Politikern		belebt	Menschlich	Nominalphrase
Psychopathen		belebt	Menschlich	Nominalphrase
er	der Franzose	belebt	Menschlich	pronominaler Bezug
sie	unsere Politiker	belebt	Menschlich	pronominaler Bezug
ihnen	Kriegsflüchtlinge	belebt	Menschlich	pronominaler Bezug
die FPÖ		belebt	Menschlich	Nominalphrase
jeder		belebt	Menschlich	Nominalphrase
Wer aber glaubt, sich hinter einer angeblichen Anonym		belebt	Menschlich	Nebensatz
die anderen Leute		belebt	Menschlich	Nominalphrase
Die dafür verantwortlichen Politiker der letzten Jahre		belebt	Menschlich	Nominalphrase
Geisterfahrer		belebt	Menschlich	Nominalphrase
sie alle	Zuspätkommer, Schwätzer,	belebt	Menschlich	Nominalphrase
Schutzbedürfte		belebt	Menschlich	Nominalphrase
er	#unklar	belebt	Menschlich	pronominaler Bezug
alle Araber in Österreich		belebt	Menschlich	Nominalphrase
die	Kids	belebt	Menschlich	pronominaler Bezug
diese	Familien	belebt	Menschlich	pronominaler Bezug
kriminelle Ausländer		belebt	Menschlich	Nominalphrase
SPÖ und ÖVP		belebt	Menschlich	Nominalphrase
Jugendliche		belebt	Menschlich	Nominalphrase
Frauen		belebt	Menschlich	Nominalphrase
Pilnacek		belebt	Menschlich	Nominalphrase
die Salzburger		belebt	Menschlich	Nominalphrase
Mödlhammer		belebt	Menschlich	Nominalphrase
Chefs, die ihre männlichen Mitarbeiter nicht in Karenz		belebt	Menschlich	Nominalphrase
Erziehungsberechtigte wie diese		belebt	Menschlich	Nominalphrase
Trump		belebt	Menschlich	Nominalphrase
Alle, die beteiligt waren,		belebt	Menschlich	Nominalphrase
die politisch Verantwortlichen		belebt	Menschlich	Nominalphrase
der	#unklar	belebt	Menschlich	pronominaler Bezug
Solche Menschen, der meist sitzenden Gattung Tanzm		belebt	Menschlich	Nominalphrase
Lehrer		belebt	Menschlich	Nominalphrase
Jugendliche		belebt	Menschlich	Nominalphrase
die	Menschen	belebt	Menschlich	pronominaler Bezug
Leistungsträger		belebt	Menschlich	Nominalphrase
behinderte Menschen		belebt	Menschlich	Nominalphrase
die Juden		belebt	Menschlich	Nominalphrase
die Grüne Politikerin Roth		belebt	Menschlich	Nominalphrase
er	Louay	belebt	Menschlich	Nominalphrase
schwarze Schafe		belebt	Menschlich	Nominalphrase
die Bezirksschulräte		belebt	Menschlich	Nominalphrase
diese Menschen		belebt	Menschlich	Nominalphrase
alle		belebt	Menschlich	Nominalphrase
grob faktenresistente Ärzte		belebt	Menschlich	Nominalphrase
es/ein paar Spieler		belebt	Menschlich	formales Subjekt es

Tabelle C.5: Ressortzuteilung im qualitativen Sample

Ressort	-	direkte Rede	indirekte Rede	uneindeutig	Gesamtergebnis
allgemein	36	25	9	1	71
ausland	3	6	1		10
auto	1		1		2
bundesland	5	3			8
chronik	9	24	14	1	48
inland	8	26	14	2	50
karriere	2				2
kultur	13	4	2	1	20
leben	3	3	1		7
leserbriefe	1				1
lokal	16	48	8	1	73
medien	2				2
meinung	18		1		19
politik	5	7	3		15
regional	3	10	2	2	17
reise	1	1	1		3
seite	4		3		7
service	13		2		15
sport	8	17	4		29
stadt land	1				1
unzuordenbar	1	11	2		14
wirtschaft	5	10	8		23
wissenschaft	1	1			2
(Leer)	11	19	2	1	33
Gesamtergebnis	170	215	78	9	472

Anhang D: quantitatives Korpus

Tabelle D.1: Anzahl der *gehören*-Passiv-Belege in den einzelnen Medien

Medien	Anzahl <i>gehören</i> -Passiv
ACADEMIA	14
APA	2401
ARBEITW	32
ATVVOLL	43
AUGUSTIN	66
BAUERNZT	228
BVZ	840
DATUM	27
DIEWIR	11
ECHO	82
EMEDIA	20
FALTER	466
FORMATDB	153
FURCHE	152
GEWINN	62
GRAZER	178
HEUTE	336
HOR	76
IM	16
KLEINE	3157
KONSUM	30
KRONE	4738
KTNMONAT	4
KTZ	244
KURIER	2052
KW	58
MEDIANET	122
MWVOLL	2993
NEWS	261
NOEN	3917
NVB	509
NVT	538
OBERRUND	389
OEREICHE	980
OESTERREICH	212
OOEN	1773
OTS	3674
PRESSE	1489
PRO7VOLL	19
PROFIL	288

PULSVOLL	60
SAT1VOLL	24
SBGW	759
SN	1822
SOLI	5
SPORTZTG	29
STANDARD	2058
STMONAT	2
SVZ	63
TREND	70
TT	2229
TTKOMP	169
TMEDIA	73
VN	1244
WIBLATT	303
WIENER	42
WIENERIN	41
WOMAN	156
WZ	709
Gesamtergebnis	42508

Tabelle D.2: Anzahl *gehören*-Passiv-Belege pro *gehören*-Form

Form	Anzahl <i>gehören</i> -Passiv
gehörten	1695
gehörte	468
gehört	21563
gehören	14475
gehöre	4307
Gesamtergebnis	42508

Tabelle D.3: Die zwanzig häufigsten Verben

VVPP	Anzahl
abgeschafft	1458
geändert	1403
gemacht	1012
verboten	985
reformiert	790
gefördert	739
gefeiert	614
abgestellt	550
ausgebaut	549
unterstützt	542

geschützt	533
aufgeklärt	514
verbessert	514
gestellt	453
geklärt	399
gezogen	396
geregelt	394
diskutiert	381
beseitigt	373
umgesetzt	364
Gesamtanzahl	12963

Tabelle D.4: Anzahl *gehören*-Passiv-Belege in den einzelnen Ressorts (Erstreichung)

Ressort	Anzahl <i>gehören</i>-Passiv
allgemein	6452
ausland	1358
auto	129
bundesland	715
chronik	4283
immobilien	17
inland	4360
karriere	156
kategorie	2
kultur	1434
leben	882
leserbriefe	126
lokal	6651
medien	232
meinung	2047
politik	1259
regional	1145
reise	58
schule	40
seite	344
service	1472
sport	2159
stadt	107
thema	104
unzuordenbar	1566
wirtschaft	1727
wissenschaft	81
(Leer)	3602
Gesamtergebnis	42508

11. Abstracts

Die Grammatikalisierung von Passiv-Konstruktionen wird in der Forschungsliteratur immer wieder aufgegriffen, nicht nur die üblichen Fälle von *werden* und *sein*, sondern auch weitere Verben wie *geben* und *bekommen*, die ebenfalls den Prozess der Auxiliarisierung durchlaufen haben.

Eines dieser Verben ist *gehören*. Es hat bislang nur wenig Aufmerksamkeit bekommen, bis auf einige Ausnahmen, größtenteils in Form von kurzen Erwähnungen in Werken zu Passiven, aber auch Artikel von Szatmári (2002) und Stathi (2009; 2010).

Mit einem datenzentrierten Zugang, der das Austrian Media Corpus (amc) heranzieht, wird die Konstruktion in einem Textkorpus aus unterschiedlichen Medien untersucht, welches aus Daten von österreichischen Zeitungen und Nachrichtendiensten für sprachwissenschaftliche Zwecke zusammengestellt wurde.

Der Fokus der Untersuchung liegt auf den grammatischen Formen von *gehören* und der Bedeutung sowohl der Perfektpartizipia als auch der gesamten Konstruktion.

Neben einer ausführlichen Methodenreflexion, die die Extraktion aus dem amc, die Weiterbearbeitung mithilfe eines NLP-Skripts und die dabei auftretenden Herausforderungen beschreibt, werden unterschiedliche Aspekte des *gehören*-Passivs beleuchtet, um die Grammatikalisierung einzuschätzen und ihre deontische Modalität in der syntaktisch-kompakten Packung zu untersuchen.

The grammaticalization of different passive-forms gets discussed periodically, not only the default cases of *werden* and *sein*, but more verbs that went through auxiliarization like *geben* and *bekommen*.

One of them is *gehören* which didn't get much attention in the past, except for short mentions and the articles of Szatmári (2002) and Stathi (2009; 2010).

Thus a data-driven approach using the Austrian Media Corpus (amc) was chosen to shed a light on the construction in a text corpus of different media, collected from austrian news outlets for linguistic purposes.

The focus therein lies on the grammatic forms of *gehören* and the semantic meaning of the participle II and the whole construction.

Besides a reflection of the method, how the *gehören*-passive was extracted from the amc, processed further with an NLP-script and the challenges that arose during that process, different aspects of the construction are discussed to survey its grammaticalization and analyse its deontic modality in the syntactically compact package.

Key-Words: Grammaticalization, Auxiliarization, Passive, NLP, Corpuslinguistics, written Standard Language