

Strukturwandel der Urlaubswelt(en)

Framesemantische Clusteranalysen zur Erfassung thematischer Muster auf Ansichtskarten¹

Marcel Naef, Nicolas Wiedmer, Kyoko Sugisaki

1 Einleitung

Für die thematische Annotation von Texten haben sich in der Korpuslinguistik Ansätze etabliert, welche die Thema-Entfaltung vorwiegend nach informationsstrukturellen Gesichtspunkten angehen. Besonders prominent belegen dies die Prager Baumbank (vgl. Hajič 1998) und das Potsdamer Kommentarkorpus (vgl. Stede & Mamprin 2016). Das *Thema* wird dabei als semantische Kategorie gesehen, die auf der Ebene des Satzes angesiedelt ist und hauptsächlich von ihrer syntaktischen Position und ihrer diskursiven Salienz in Bezug auf die im vorangegangenen Satz erwähnten Entitäten bestimmt wird.

Für die thematische Annotation eines Ansichtskarten-Korpus scheint uns der informationsstrukturelle Ansatz allerdings ungeeignet zu sein, insofern *sprachliche* Lesbarkeits Hinweise (*sensu* Hausendorf et al. 2017: 69–105) bei dieser Textsorte im Vergleich zu *vertrauheitsbasierten* stark in den Hintergrund treten.² Deshalb gehen wir in unserer Analyse von einer thematischen Struktur im Sinne eines framesemantischen Rahmens aus, welcher für die Ansichtskarte die thematische Kohärenz auch dann gewährleistet, wenn sich wie im folgenden Beispiel auf sprachlicher Ebene (fast) keine typischen Themaentwicklungshinweise in den Mitteilungstexten finden.

-
- 1 Der Beitrag basiert auf einem ursprünglich auf Englisch erschienenen Artikel (Sugisaki et al. 2019). Die darin enthaltenen Beobachtungen wurden 2018 an den Workshops *Computational Methods in the Humanities (COMHUM)* an der Universität Lausanne und *Die Welt als Urlaubsort – Raumreferenz(en) und mehr auf der Ansichtskarte* an der Universität Zürich vorgestellt. Wir danken den Teilnehmer*innen für die wertvollen Vorschläge und Kommentare, die wesentlich zur Weiterentwicklung der hier vorgelegten Ideen beigetragen haben.
 - 2 Es handelt sich dabei um eine prototypische Eigenschaft sogenannter *kleiner Texte* (vgl. Hausendorf 2009: 16; und spezifisch für das Beispiel der Ansichtskarte Kato et al. 2021).

Abb. 1a und 1b: Beispiel für prototypische Themaentwicklung



Quelle: [anko] 150702 und [anko] 150701 (1984)

Transkript: Aus Arosa senden/wir Euch viele/herzl. Grüsse./Essen gut, Wetter gut,/alles O.K.

Anhand des Beispiels wird ersichtlich, dass die eingeführten Themen (*Essen, Wetter und Zustand/Befinden*) nicht weiter ausgeführt, sondern in willkürlich anmutender Abfolge parataktisch aneinandergereiht werden, wobei sich kein lexikalischer Zusammenhang zwischen den thematischen Feldern feststellen lässt. Dass ein*e Leser*in einen solchen Text nicht als unzusammenhängend empfindet, ist die Leistung eines semantischen Frames, wie wir im Folgenden weiter ausführen werden.

Wir verwenden den Begriff des *semantischen Frames* im Sinne von Busse (2012: 563), welcher den Frame als eine Struktur des Wissens definiert, in welcher der Kern des Frames (das Thema) mit Konstituenten des Wissens verbunden ist. Diese variieren je nach Kontext einer konkreten Situation und legen die Anschlussbedingungen für die Realisierung von zugehörigen thematischen Einheiten fest. Im Fall des »Urlaubsframes« (Diekmannshenke 2008: 90) wird der thematische Kern des Frames durch den Zustand des *im Urlaub Seins* gebildet. Die Konstituenten des Wissens, die Leerstellen des Frames, geben die Anschlussbedingungen vor, welche mögliche Realisierungen von thematischen Einheiten anlässlich des Schreibens einer Urlaubsansichtskarte erfüllen müssen, um als entsprechende Füllwerte des Frames fungieren zu können.

Im Rahmen unserer empirischen Analyse haben wir einen Katalog von Leerstellen des Urlaubsframes rekonstruktiv identifiziert, die beim Berichten über den Urlaub in Ansichtskarten potenziell gefüllt werden. Auf dieser Basis stellen wir im vorliegenden Beitrag einen neuen Ansatz für die thematische Annotation aus framesemantischer und textlinguistischer Perspektive vor und präsentieren eine datengeleitete Analyse zur Entwicklung und Veränderung des Urlaubsframes über den Zeitraum von 1950 bis 2010. Nachfolgend werden wir zuerst auf die thematische Struktur in den Mitteilungstexten eingehen, um dann unser Annotationsschema vorzustellen und den Annotationsprozess zu beschreiben (Kapitel 2). Schließlich werden wir die Ergebnisse unserer korpuslinguistischen Analyse des annotierten Teilkorpus präsentieren, welche wir durch Anwendung

des Clustering-Verfahrens HKA (Hauptkomponentenanalyse, Englisch: Principal Component Analysis) gewonnen haben (Kapitel 3).³

2 Thematische Annotation des Urlaubsframes

Im Berichten über den Urlaub zeigen sich in Ansichtskarten immer wieder gleiche thematische Muster. Diese Muster haben sich über die Zeit hinweg entwickelt und wurden stetig reproduziert (vgl. Hausendorf 2009: 13), um die Hauptfunktion von Ansichtskarten im Sinne der Kontaktherstellung (vgl. Hausendorf & Kesselheim 2008: 154–157; Brinker et al. 2014: 118–119) zu erfüllen. Mit dieser Funktion ist die Aufgabe verbunden, während des Urlaubs den sozialen Kontakt zu Familienmitgliedern und befreundeten Personen zu pflegen und aufrechtzuerhalten. Die Muster sind also durch eine Konventionalisierung dieser spezifischen Form der Urlaubskommunikation entstanden, also dadurch, dass die aus dem Urlaub schreibenden Personen immer wieder auf bewährte Lösungsroutinen für das kommunikative Problem der Pflege des sozialen Kontaktes zurückgegriffen haben. Soziale Veränderungen in der Gesellschaft (speziell im Urlaubs- und Ferienkontext des modernen Tourismus) haben zwar auch in der Kommunikation mittels Ansichtskarten einen Niederschlag gefunden (vgl. dazu z. B. den Beitrag von St. Wolff in diesem Band), dennoch lässt die Analyse unserer thematischen Annotation die Vermutung zu, dass die gefundenen Kategorien einen Themenkatalog abbilden, der über die Zeit konstant geblieben ist (vgl. Hausendorf 2008: 333). Das primäre Ziel unserer Annotation war demnach die Definition thematischer Kategorien, welche regelmäßig in den Mitteilungstexten von Ansichtskarten bedient werden, und in einem nächsten Schritt die Auswertung der Veränderungen dieser Kategorien über die Zeit vorzunehmen.

2.1 Entwicklung des Annotationsschemas

Für die Erstellung des Annotationsschemas haben wir nach Sichtung und manueller Analyse von Hunderten von Ansichtskarten aus dem Ansichtskartenkorpus [anko] die in den Mitteilungen immer wieder erwähnten thematischen Hauptkategorien bestimmt. Schließlich haben wir das so entstandene thematische Annotationsschema mit 14 studentischen Testteilnehmer*innen überprüft. Im Laufe dieser Überprüfung haben wir das Schema immer weiter verfeinert und erweitert. Schließlich haben wir das finale Annotationsschema (vgl. Tabelle 1) definiert, mit welchem eine umfangreiche und nachvollziehbare Annotation der thematischen Strukturen in Ansichtskartentexten erstellt

3 Die framesemantische Annotation des Teilkorpus und die im Folgenden beschriebene Entwicklung des thematischen Annotationsschemas entstanden im Rahmen des vom SNF und der DFG geförderten Projektes »Textsortenentwicklung zwischen Standardisierung und Variation: Das Beispiel der Ansichtskarte« unter Mitarbeit von Noah Bubenhofer, Heiko Hausendorf, Maaike Kellenberger, David Koch, Jan Langenhorst, Marcel Naef, Josephine Obert, Joachim Scharloth, Michaela Schnick, Kyoko Sugisaki und Nicolas Wiedmer. Neben der im vorliegenden Beitrag und in Sugisaki et al. (2019) vollzogenen HKA diente das framesemantisch annotierte Teilkorpus als Grundlage für Wiedmer (i. Vorb.) und den Beitrag von N. Wiedmer in diesem Band.

werden konnte. Die einzelnen Schritte wollen wir nun in den folgenden Abschnitten detailliert ausführen.

2.1.1 Identifizierung thematischer Kategorien

Um die Leerstellen des Urlaubsframes zu bestimmen, haben wir zunächst eine Reihe von Fragen definiert, welche durch die Mitteilungen immer wieder (in ähnlicher, wenn auch nicht immer gleicher Form) beantwortet werden (vgl. Ziem 2008: 94–95). Wir gehen davon aus, dass jeder Satz in einem Ansichtskartentext eine Antwort auf mindestens eine der Fragen darstellt, welche im Annotationsschema in Tabelle 1 aufgeführt sind.

Tab. 1: Annotationsschema zum Urlaubsframe

#	Kategorie	Frage	Beispiel für mögliche Antwort im Text
A	Extra-Diegetisches	Wie thematisiere ich die Welt außerhalb der Urlaubswelt?	Vielen Dank für Deinen lieben Brief!
B	Unterkunft	Wie/wo wohne ich hier?	Unser Campingplatz ist direkt am Meer.
C	Aktivität	Was mache ich hier?	Wir baden jeden Tag im See und wandern viel.
D	Essen und Trinken	Was/wo esse/trinke ich hier?	Riesenplatte mit Meeresfrüchten € 14!
E	Gefühle	Wie fühle ich mich hier?	Wir geniessen es unbeschreiblich hier!
F	Wahrnehmung	Was nehme ich hier mit meinen Sinnen wahr?	Überall blühen und duften die Lavendelfelder.
G	Vorkommnisse	Was ist mir hier Unvorhergesehenes passiert?	Schon am dritten Tag hatten wir eine Autopanne...
H	Wissen	Was weiß ich oder glaube ich zu wissen über diesen Ort?	Diese heilige Stadt ist über 5000 Jahre alt.
I	Ort	Wo bin ich hier?	Dieses Jahr sind wir alle zusammen auf Ibiza.
J	Kennenlernen neuer Leute	Mit wem verkehre ich hier?	Habe gestern einen netten Typen kennengelernt.
K	Hin- und Rückreise	Wie bin ich hierher/von hier weg gekommen?	Der Flug war schrecklich: Turbulenzen ohne Ende!
L	Urlaubsgrund	Was möchte ich mit dieser Art von Urlaub erreichen?	Zum Glück zeigt die Badekur endlich Wirkung!
M	Urlaubsart	Welche Art von Urlaub mache ich hier?	Liebe Grüsse aus unseren Wanderferien!
N	Wetter	Wie ist das Wetter hier?	Jetzt scheint die Sonne und es ist heiss (39°!).

Die thematischen Frame-Kategorien haben wir grundsätzlich in zwei Klassen aufgeteilt. Die eine Klasse beinhaltet diejenigen Themen, welche sich auf die Urlaubswelt beziehen (B-N in Tabelle 1), in der anderen Klasse sind diejenigen Themen angesiedelt, welche sich auf etwas beziehen, das sich außerhalb der Urlaubswelt befindet. Die thematischen Kategorien beziehen sich nicht auf einzelne Themenaspekte (wie Sonnenschein, Regen, Schneefall, Kälte oder Hitze), sondern stellen jeweils die thematische Hauptkategorie dar (wie *Wetter*), welche alle Einzelentitäten umfasst. Gerade die Hauptkategorie *Wetter* gehört zu den am häufigsten thematisierten Kategorien, und die einzelnen Varianten ihrer Realisierung zeigen eine Vielfalt an Bedeutungen auf, welche das Wetter für die Urlauber*innen hat. Daraus schließen wir, dass das Wetter eine zentrale Thematik für das Erleben und Beschreiben von Urlaub darstellt (Hausendorf 2019 beschäftigt sich speziell mit der Thematisierung des Wetters in Ansichtskarten und deren Bedeutung). Aufgrund der Häufigkeit der Thematisierung kann auch den Hauptkategorien *Essen und Trinken*, *Kennenlernen neuer Leute* sowie *Unterkunft* eine hohe Relevanz zugesprochen werden. Ebenfalls konnten wir beobachten, dass auch die Kategorie *Aktivität* einen festen Bestandteil des prototypischen Urlaubsberichts darstellt, in welchem die geplanten oder spontan ausgeführten Urlaubsaktivitäten thematisiert werden (vgl. den Beitrag von N. Wiedmer in diesem Band). Im Gegensatz zu geplanten Aktivitäten (wie Wandern, Schwimmen, Sightseeing etc.) bezieht sich die Kategorie *Vorkommnisse* auf unvorhergesehene Ereignisse wie Unfälle, Krankheit, Pannen oder Naturkatastrophen, welche oft einen einschneidenden Einfluss auf den Urlaub der schreibenden Personen haben. Die Thematisierung der Kategorie *Ort* erfolgt in der Regel am Anfang des Textes und umfasst oft die Beschreibung des Aufenthaltsorts – nicht selten mit positiv konnotierten Adjektiven, was gleichzeitig eine Thematisierung der Hauptkategorie *Gefühle* darstellt, welcher wir auch Bewertungen aller Art zugeordnet haben. Erklärende Kommentare zu den Urlaubsorten sowie stereotype Zuschreibungen von Eigenschaften der Länder und ihrer Bewohner*innen werden durch die Kategorie *Wissen* abgedeckt. Demgegenüber stehen Beobachtungen, welche auf der menschlichen Sinneswahrnehmung beruhen (wie Darstellungen von Farben, Gerüchen oder Landschaften) und zur Kategorie *Wahrnehmung* gehören. Ausgehend von den Mitteilungstexten lassen sich auch verschiedene Typen von Urlaub unterscheiden. Beschreibungen dieser Art erfasst die Kategorie *Urlaubsart*. Dazu gehören sämtliche Hinweise auf einen meist mit Aktivitäten im Zusammenhang stehenden Typ von Urlaub (wie z.B. Wanderurlaub, Klassenfahrt, Skiferien). Wenn hingegen Hinweise auftreten, die das mit dem Urlaub verbundene Ziel betreffen (im Sinne z.B. von Genesung in der Kur, Erholung von Strapazen, Erreichen von Fitness für einen Wettkampf), werden diese der Kategorie *Urlaubsgrund* zugeordnet). *Hin- und Rückreise* stellt eine besondere Kategorie im Annotationsschema dar, da sie die Thematisierungen des Übergangs von der Daheimwelt in die Urlaubswelt und zurück beinhaltet. Zwar ist die Hin- und Rückreise nicht alleine in der Urlaubswelt angesiedelt, sie stellt aber die Verbindung zwischen der Kategorie *Extra-Diegetisches* (Grüße, Danksagungen, Bitten während der Abwesenheit) und den Kategorien B-N dar.

Für die thematische Annotation haben wir die Einheit *Satz* gewählt.⁴ Im Gegensatz zu anderen möglichen Einheiten (auf Wort-, Phrasen- oder Absatzebene) sind Sätze ideale Einheiten für die Analyse der thematischen Struktur, da jeder Satz eine Proposition enthält, welche mindestens eine Frage im Annotationsschema beantwortet. Ein Satz kann dabei auch mehr als eine Proposition enthalten, bspw. aufgrund von Koordination (s.u. Beispiel (1)) oder aufgrund der Inhärenz semantischer Kategorien (s.u. Beispiel (2)). In solchen Fällen muss ein Satz mit mehr als einer Kategorie annotiert werden:

(1) Kategorien *Ort* und *Aktivität*:

Heute waren wir in Meissen, nachher kurz wieder Stadt, und gleich dann die Oper. ([anko] 60905) (1996)

(2) Kategorien *Ort* und *Aktivität*:

Gestern waren wir in der Nordsee baden. ([anko] 100327) (1973)

2.1.2 Überprüfung der thematischen Kategorien

Um die Robustheit und die Nachvollziehbarkeit unseres thematischen Annotationsschemas zu überprüfen, haben wir es 2016 an der Universität Zürich im Rahmen der Lehrveranstaltung »Text(sorten)linguistik am Beispiel der Ansichtskarte« mit 14 Germanistik-Student*innen getestet. Die Studierenden wurden nach einer 45-minütigen Einführung in das Annotationsschema dazu aufgefordert, jeweils 12 Karten individuell zu annotieren. Dazu erhielten sie eine Tabelle, in der jeder Satz in einer Zelle angezeigt wurde. Die Studierenden ordneten dann die Sätze den Kategorien des Annotationsschemas zu. Die Kategorie *Vorkommnisse* war zu diesem Zeitpunkt noch nicht Teil des Annotationsschemas, sondern wurde aufgrund dieses Tests nachträglich dem Schema hinzugefügt. Für den Fall, dass ein Satz keiner Kategorie zugeordnet werden könnte, wurde den Probanden das Annotations-Tag *X* zur Verfügung gestellt. Die durch die Studierenden erstellten Annotationen wurden schließlich mit einem Goldstandard verglichen, welcher durch vier Projektmitarbeiter*innen⁵ erstellt wurde.

Die statistischen Scores der studentischen Annotationen waren unter Berücksichtigung der kurzen Instruktionszeit sehr zufriedenstellend: Die *Precision* lag zwischen 83,89 % und 98,58% (Durchschnitt: 93,30 %) mit einem *Recall* zwischen 85,93 % und 97,20 % (Durchschnitt: 92,61 %). Allerdings gab es bei einigen Kategorien beträchtliche Unterschiede im Hinblick auf *Precision* und *Recall* (vgl. Tabelle 2). Als ausgewogenes Maß benutzten wir *Matthews Correlation Coefficient* (MCC) anstelle des *Cohens Kappa*, um den Divergenzen in den Häufigkeiten der einzelnen Kategorien Rechnung zu tragen (vgl. Powers 2012). Wie in Tabelle 2 ersichtlich wird, erwiesen sich die Kategorien *Unterkunft*, *Ort*, *Wissen*, *Urlaubsart*, *Urlaubsgrund*, *Gefühle* und *Wahrnehmung* als problematisch. Bei der Diskussion der Resultate mit den Studierenden zeigte sich, dass sich der tiefe Recall der Kategorien *Unterkunft* und *Ort* auf Missverständnisse in den Briefings zur

4 Als primärer Indikator für die Einheit *Satz* galten bei der Segmentierung die Satzzeichen *Punkt* (.), *Ausrufezeichen* (!) und *Fragezeichen* (?). Wurden die Satzzeichen weggelassen oder waren sie beispielsweise durch Verblässung der Schrift nicht mehr erkennbar, so erfolgte die Segmentierung aufgrund der sekundären Indikatoren (a) *neuer Absatz*, (b) *Initialmajuskel* nach potenziellem Satzende und potenziellem neuem Satzanfang.

5 Maaïke Kellenberger, David Koch, Marcel Naef und Nicolas Wiedmer.

Annotationssystematik zurückführen lässt. Beispielsweise ordneten die Studierenden die Kategorie *Ort* oftmals nicht zu, wenn der Urlaubsort nicht im Fließtext, sondern beispielsweise im Datum (»Paris, 07.08.1966«) erwähnt wurde. Basierend auf dieser Diskussion haben wir Annotationsrichtlinien mit Definitionen der Kategorien und beispielhaften Zweifelsfällen erstellt.

Tab. 2: Precision, Recall und MCC (in Prozent)

#	Kategorie	Precision	Recall	MCC
A	Extra-Diegetisches	88.61	93.23	89.59
B	Unterkunft	88.89	84.21	86.19
C	Aktivität	96.92	92.99	93.91
D	Essen und Trinken	88.89	94.12	91.17
E	Gefühle	81.03	97.92	87.90
F	Wahrnehmung	75.29	81.01	76.90
G	Vorkommnisse	–	–	–
H	Wissen	73.21	98.80	84.13
I	Ort	96.24	86.72	88.79
J	Kennenlernen neuer Leute	84.62	82.50	83.13
K	Hin- und Rückreise	92.86	100.00	96.30
L	Urlaubsgrund	73.33	78.57	75.68
M	Urlaubsart	38.89	50.00	43.53
N	Wetter	97.76	90.34	93.40

2.2 Manuelle Annotation

Basierend auf der im vorherigen Abschnitt beschriebenen Testannotation haben die beiden Annotator*innen⁶ B und C sodann eine satzbasierte mehrklassige thematische Annotation mit 14 Kategorien (vgl. Tabelle 1) von 1.120 Ansichtskarten durchgeführt. Die Karten wurden mittels einer Zufallsstichprobe aus unserem standarddeutschen Ansichtskartenkorpus [anko] (vgl. den Beitrag von K. Sugisaki et al. in diesem Band) ausgewählt, wobei von den 1950er Jahren bis zu den 2010er Jahren jeweils 160 Karten pro Jahrzehnt verwendet wurden. Durch dieses Sampling konnte sichergestellt werden, dass auch für die Jahrzehnte, für welche wir weniger Beispiele im Korpus haben, die gleiche Anzahl an Karten annotiert werden konnte.

Die manuelle Annotation erfolgte in drei Schritten. Zunächst haben die zwei Annotator*innen A und B jedem Satz eine oder mehrere Kategorien des Annotationsschemas zugeordnet. Zusätzlich zum Annotationsschema wurde ihnen ein Annotationsleit-

6 Annotator A: David Koch, Annotatorin B: Maaïke Kellenberger, Annotator C: Nicolas Wiedmer.

faden ausgehändigt, in welchem alle Kategorien erläutert und anhand von Beispielen verdeutlicht wurden. Jeder Ansichtskartentext bzw. Satz wurde von einer der beiden Annotator*innen annotiert, welche einen Kommentar notierte, wenn sie sich bei einer Zuordnung nicht sicher war. Nach einer ersten Phase wurden problematische Kategorien bzw. Zuordnungen von den Annotator*innen B und C diskutiert, wobei gemeinsam entschieden wurde, in welchem Fall welche Kategorie annotiert werden sollte. Diese beiden Schritte waren jedoch noch nicht ausreichend, um eine einheitliche und konsistente Annotation zu erhalten. Das Problem war, dass einige Kategorien nicht eindeutig von anderen unterschieden werden konnten, was dazu führte, dass die Annotator*innen A und B die Kategorien *Wahrnehmung*, *Gefühle* und *Wissen* nicht einheitlich annotiert hatten. Unser ursprünglicher Ansatz war, dass diejenige Kategorie gewählt werden sollte, deren Frage durch den Satz bestmöglich beantwortet wurde. Dies ließ jedoch Spielraum für individuelle Interpretationen. So wurde z. B. ein Satz wie »Der Strand ist herrlich zum wandern.« ([anko] 170217 (1998)) von Annotator A der Kategorie *Gefühle* (*Wie fühle ich mich hier?*) und von Annotatorin B der Kategorie *Wahrnehmung* (*Was nehme ich hier mit meinen Sinnen wahr?*) zugeordnet. Aus diesem Grund haben wir eine Reihe von lexikalischen Indikatoren bestimmt, deren Vorkommen die Kategorie *Gefühle* nahelegt (wie »genießen«, »schön«, »wunderbar«, »toll«, »mies« etc.), welche alle Gefühle in Form von Meinungen, Bewertungen sowie vom emotionalen Befinden der schreibenden Person ausdrücken. Solche lexikalischen Indikatoren wurden auch für die Kategorie *Wissen*, welche auch stereotypes Wissen im Sinne von Vorurteilen umfasst, festgelegt (mit Indikatoren wie »alle hier«, »typisch«, »man« etc.). Allerdings ist das Vorkommen eines lexikalischen Indikators kein eindeutiges Kriterium für die Zuordnung zu einer bestimmten Kategorie. Daher musste der Kontext jedes einzelnen Satzes im Gesamttext einer Karte berücksichtigt werden. Für die Kategorie *Wahrnehmung* wurde die Vorgabe gemacht, dass entsprechende Beschreibungen tatsächlich sinnlich wahrnehmbar (seh-, hör- und/oder riechbar) sein müssen. Nach dieser neuen Definition der Kategorienzuordnung haben die Annotator*innen B und C in einem dritten Schritt sämtliche Karten erneut einheitlich annotiert.

3 Datenanalyse

Basierend auf der manuellen Annotation der 1.120 Karten (49.261 Tokens und 6.713 Sätze) unterzogen wir den annotierten Text einer Hauptkomponentenanalyse (HKA). Die HKA ist ein Verfahren zur Dimensionsreduktion, das latente Dimensionen einer Textsammlung offenlegt, wobei die Kovarianz eliminiert wird und gleichzeitig ein Großteil der Varianz in den Daten erhalten bleibt (vgl. Moisl 2015). Sie wurde in der Korpuslinguistik als Ansatz zur Zuschreibung von Autorschaft oder Stilanalyse (z. B. Baayen et al. 1996) und als Faktoranalyse zur Untersuchung sprachlicher Register (z. B. Biber 1995) eingesetzt.

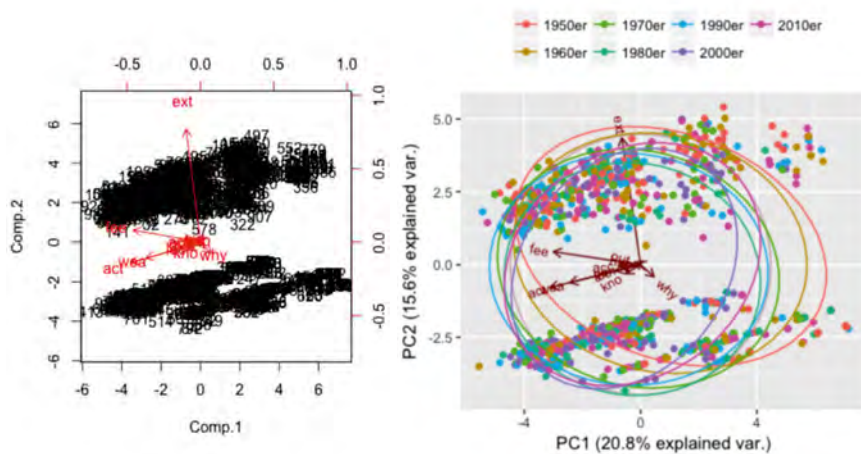
Um für zu wenig häufig vorkommende Kategorien ein besseres Resultat zu erhalten, wurden für die HKA die thematischen Kategorien *Urlaubgrund* und *Urlaubsort* zu einer neuen Kategorie *Weshalb* zusammengefasst; die Kategorien *Kennenlernen neuer Leute* und *Essen und Trinken* wurden zur Kategorie *Aktivität* aggregiert. Wir erhoben die Frequenz

der 10 verbleibenden Frames (ohne Kategorie X) in jedem Text und normalisierten die log-transformierte Zählung auf 1.000 Wörter.

Tab. 3: HKA-Ladung der Komponenten 1–4

Kategorie	Komp. 1	Komp. 2	Komp. 3	Komp. 4
Extra-Diegetisches	-0.123	0.962		
Unterkunft	-0.140			
Aktivität (Kennenlernen neuer Leute/Essen und Trinken)	-0.595	-0.173	0.307	0.482
Gefühle	-0.574			0.136
Wahrnehmung	-0.126			
Vorkommnisse				
Wissen			0.130	
Ort	-0.167		0.719	-0.644
Hin- und Rückreise				
Wetter	-0.465	-0.136	-0.587	-0.523
Weshalb (Urlaubsgrund/Urlaubsart)			0.116	0.221

Abb. 2a und 2b: Hauptkomponentenanalyse und historische Entwicklung der Komponenten 1 und 2



Die HKA identifizierte vier Hauptkomponenten, die 62,67 % der Varianz von 10 Variablen ausmachen.⁷ In Tabelle 3 findet sich eine Übersicht über die Ladungen. In den Abbildungen 2 und 3 sind alle Karten unseres Datensatzes und die Richtung der Variablen (vgl. für die Komponenten 1–2 Abb. 2a; vgl. für die Komponenten 3–4 Abb.) dargestellt. Im Folgenden besprechen wir diese vier Komponenten unter Berücksichtigung ihrer historischen Entwicklung (vgl. für Komponente 1–2 Abb. 2b; vgl. für Komponente 3–4 Abb. 3b) im Detail.

Komponente 1: Bei der ersten Komponente, die 20,78 % der Varianz ausmacht, deutet die HKA darauf hin, dass *Aktivität*, *Gefühle* und *Wetter* hochgradig korrelieren, d.h., dass sich viele Mitteilungen finden, in denen diese Kategorien zusammen auftreten. Den höchsten Wert weist die untenstehende *Karte A* als linker Pol der Komponente 1 (vgl. Abb. 2a) auf, den tiefsten Wert die *Karte B* als rechter Pol der Komponente 1. Die beiden Karten zeigen nicht nur in der Dimension der Komponente 1 die größte Differenz, sondern auch in den annotierten Karten in Bezug auf die Profilierung des Urlaubsframes insgesamt. Die folgende *Karte A* zeigt stellvertretend, dass der Wert von *Urlaubsart/-grund* (*why* in Abbildung 2a) hoch ist, während *Karte B* in Bezug auf die Kategorien mit einem negativen Wert (*Aktivität*, *Gefühle*, *Wetter*, *Ort*, *Unterkunft*, *Extra-Diagetisches*) in Tabelle 3 stark ausgeprägt ist:

Karte A (höchster Wert):

Ausflug 19.6.93 {Urlaubsart} ([anko] 110187) (1993)

Karte B (tiefster Wert):

Transkript: Zinal, 21.7.64./Unsere Lieben, nach den paar strengen Stunden der Expo – es war vom Sonntag mörderisch heiss es glich einem langsamen Selbstmord – sind wir nun doch noch für ein paar Tage hier im wirklich malerischen & sehr ruhigen Bergdörfchen Zinal (Walliser Hochthal) gelandet. {Wetter, Aktivität, Wahrnehmung, Gefühle, Wissen, Ort} Im Grande Hotel des Diablons (ganz alt, aber [unclear] mit [unclear] franz. Küche) sind wir in allen Teilen gut aufgehoben. {Ort, Wahrnehmung, Gefühle, Unterkunft, Essen und Trinken} Wir ruhen uns Beide gut aus, mein [unclear] Mann hat dies nach seiner Bruchoperation nötig und ich laufe auch schon lange zum [unclear] {Aktivität, Gefühle} Jetzt bricht eben ein Sturmwind mit Gewitter los, hoffentl. kommt kein Dauerregen. {Wetter} – Frau [NN] hat sich dann wieder fest gemeldet/Liebe Grüsse, Ihre H. [unclear] [NN] ([anko] 150111) (1964)

Wir interpretieren die Komponente als Etablierung prototypischer Themen auf Ansichtskarten. In Abbildung 2b lässt sich durch die tendenzielle Bündelung der älteren Karten (50er und 60er Jahre) auf der rechten Seite (Phänotyp Karte A) und die Platzierung neuerer Karten auf der linken Seite (Phänotyp Karte B) beobachten, dass sich

7 Die Kaiser-Meyer-Olkin(KMO)-Faktoranalyse ergab eine *Adequacy* von 0.6, was auf eine akzeptable *Adequacy* der Stichprobe schließen lässt. Die KMO-Werte liegen zwischen 0.49 und 0.67 und somit weit über dem Akzeptanzlevel (0.5). Der Bartlett-Test auf Sphärizität war signifikant ($X^2(55)=383.42$, $p < 0.001$). Alle vier Hauptkomponenten hatten einen Eigenwert über 5.69, was weit über dem Akzeptanzlevel liegt (1).

auf Ansichtskarten eine Bewegung von eher verstreuten und kargen hin zu prototypischen Themenrepertoires (*Aktivität, Gefühle, Wetter, Ort, Unterkunft, Extra-Diegetisches*) nachzeichnen lässt. Außerdem zeigt sich in der Analyse, dass Urlauber*innen vor dem Aufkommen des Massentourismus in den 1960er Jahren öfter über den Grund des Urlaubs geschrieben haben.

Komponente 2: Bei der zweiten Komponente (15,61 % der Varianz) sind unsere Karten deutlich in zwei Cluster gegliedert, hauptsächlich in Abhängigkeit des Vorkommens der Kategorie *Extra-Diegetisches*. *Aktivität* und *Wetter* korrelieren negativ mit *Extra-Diegetisches*. *Karte A* mit dem höchsten Wert beinhaltet die Kategorien *Aktivität* (aber auch *Urlabsart*) (»Wanderferien«) und *Wetter* (»sonnigen und warmen«), während auf *Karte B* mit dem tiefsten Wert fast ausschließlich die adressierte Person bzw. die Welt außerhalb des Urlaubs, also die Kategorie *Extra-Diegetisches*, zum Thema wird. *Karte A* und *Karte B* stellen die Dimensionen der Komponente 2, also die Adressat*innenorientierung und die Urlaubsaktivität mit Wetterbezug, als antonymische Pole exemplarisch gegenüber:

Karte A (höchster Wert):

Aus unseren bisher sehr sonnigen und warmen Wanderferien im Berner Oberland senden wir Euch herzliche Grüsse {Wetter, Aktivität, Urlaubsart, Ort}

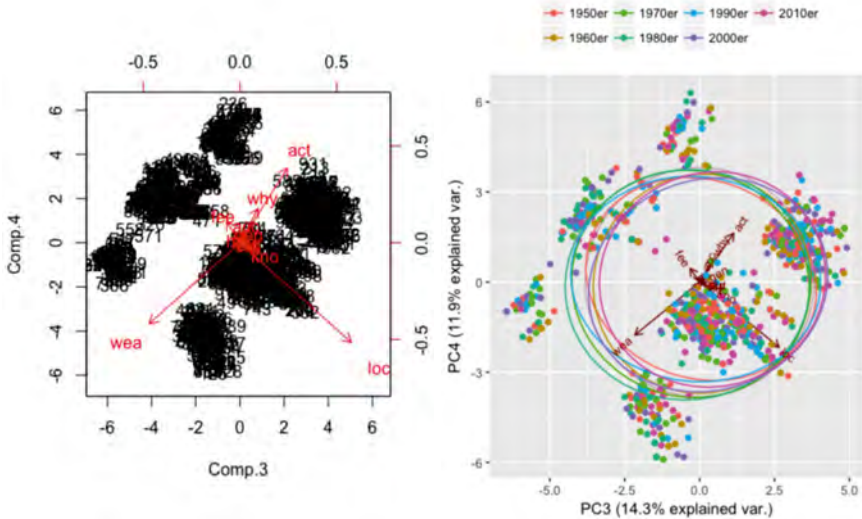
Gret + Ralph [NN] Sandra + Katja ([anko] 100203) (1973)

Karte B (tiefster Wert):

Transkript: Meine liebe große Dame! Besten Dank für Ihre beiden Karten & guten Wünsche zum Geburtstag, worüber ich mich sehr freue. {Extra-Diegetisches, Gefühle} Sind Sie immer noch in Kreuzingen? {Extra-Diegetisches} Ich glaubte Sie wären ab 4.II. wieder im Clubhaus. {Extra-Diegetisches} – Wie geht es Ihnen – hatten Sie mit der Kur Erfolg? {Extra-Diegetisches} Ich komme Ende dieser oder Anfang nächsten Monats wieder zurück & hoffe dann auf ein Wiedersehen. {Extra-Diegetisches} Bis dahin noch weiter hin recht gute Wünsche & Liebe Grüsse auch an Herrn Sohn von Ihrer kleinen Dame. {Extra-Diegetisches} ([anko] 300233) (1957)

Wir interpretieren die Komponente als Ausdruck des Wandels im Dominanzverhältnis der Textfunktionen von Ansichtskarten: Abbildung 2b zeigt in Hinblick auf Komponente 2, dass sich die Mehrheit der Karten in den 50er und 60er Jahren eher oben befindet, während sich die Ansichtskarten seit den 70er Jahren allmählich nach unten verschieben. Die Anzahl der Karten mit Korrespondenz-Charakter nach Phänotyp *Karte B*, in denen die Adressat*innen-orientierte Thematisierung der Welt außerhalb des Urlaubs (*Extra-Diegetisches*) im Vordergrund steht, nimmt über die Jahrzehnte konstant ab, und es etablieren sich Karten, die sich des »klassischen« Themenrepertoires zur Darstellung der Urlaubswelt bedienen. Es scheint, dass die dominante Funktion der Ansichtskarte sich allmählich verschoben hat von einer kontaktorientierten »Ich denke an Dich«-Schreibweise hin zu einem (selbst)darstellungsorientierten Urlaubsbericht, also von einer Art »Ich-Du«-Botschaft zu einer Art »Ich-Es« (= das Im-Urlaub-Sein)-Botschaft (vgl. dazu und zur funktionalen Ausprägung von Ansichtskarten im Allgemeinen Kato et al. 2021).

Abb. 3a und 3b: Hauptkomponentenanalyse und historische Entwicklung der Komponenten 3 und 4



Komponente 3: Bei der dritten Komponente (14,32 % der Varianz) deutet die HKA darauf hin, dass *Ort* und *Aktivität* (und bis zu einem gewissen Grad auch *Wissen* und *Urlaubsgrund/-art*) miteinander korrelieren. Gleichzeitig lässt sich eine negative Korrelation mit *Wetter* feststellen. *Karte A* mit dem höchsten Wert handelt hauptsächlich von *Ort* (»Silvanien«, »Nordgriechenland« etc.) und *Aktivität* (»durchgewandelt«), aber auch *Wissen* (»ein Relikt aus der Eiszeit, an der mazedonisch-albanischen Grenze«) und *Urlaubsart* (»auf einer Reise von Silvanien durch Nordgriechenland, Mazedonien, Bulgarien«), während es bei *Karte B* mit dem tiefsten Wert hauptsächlich um das *Wetter* (»Leider spielt das Wetter nicht mit. Sehr kalt und Regen«) geht:

Karte A (höchster Wert):

Transkript: 17.05.07 Lieber Coni/Bernhard + ich befinden uns auf einer Reise von Silvanien durch Nordgriechenland, Mazedonien, Bulgarien. {Urlaubsgrund, Aktivität, Ort} Wir haben die Vikoschlucht durchwandert und am Ochridsee die Eichenwälder an den Hängen. {Ort, Aktivität} Gestern waren wir auf dem Ohrid-See, ein Relikt aus der Eiszeit, an der mazedonisch-albanischen Grenze. - {Ort, Wissen, Aktivität} Ganz herzliche Dank für den feinen Alpkäse u. die Güegi. {Extra-Diegetisches} Beides war über die Ostertage bei dem vielen Besuch hoch willkommen. {Extra-Diegetisches} Mit Gruss/Bernhard/Barbara ([anko] 160869) (2007)

Karte B (tiefster Wert):

Transkript: Von den viel zu kurzen, aber zu [unclear] Ferien-Tage ganz herzliche Grüsse. {Gefühle} Leider spielt das Wetter nicht mit. {Gefühle, Wetter} Sehr kalt und Regen. {Wetter}/A. [NN] ([anko] 110001) (1990)

Wir interpretieren die Komponente als Wandel soziokultureller Aspekte auf Ansichtskarten: In Abbildung 3b ist ersichtlich, dass Urlauber*innen im Zeitraum der 70er, 80er und 90er Jahre ausführlich über das Wetter berichteten (PC 3 auf der rechten Seite in Abb. 3b). In den 2000er und 2010er Jahren wird ausgiebig über *Ort, Aktivität, Wissen und Urlaubsgrund/-art* berichtet (PC 3 auf der linken Seite in Abb. 3b). In der extensiven Thematisierung des Wetters im Zeitraum von 1970 bis Ende 1990 lässt sich nachzeichnen, dass das Wetter und die Art und Weise, darüber zu schreiben, offenbar einen wichtigen Baustein zur Konstruktion der massentouristischen Urlaubswelt bildete, anhand dessen sich nicht zuletzt Vorstellungen von gelungenem und misslungenem Urlaub ausgedrückt haben (vgl. dazu Hausendorf 2019). Seit der Jahrtausendwende tendieren Urlauber*innen zu einer Orientierung an *Aktivität* in Bezugnahme von *Ort, Wissen und Urlaubsgrund/-art* in der Darstellung der Urlaubswelt. Individualität und Originalität einer Reise und Urlaubserfahrungen gewinnen demnach für die Selbstdarstellung der Urlauber*innen in einer leistungsorientierten Gesellschaft an Gewicht (vgl. auch Bubenhofer et al. 2014).

Komponente 4: Bei der vierten Komponente (11,95 % der Varianz) deutet die HKA auf einen hohen Korrelationsgrad zwischen *Ort* und *Wetter* hin. Die beiden Kategorien korrelieren dabei negativ mit *Aktivität* (und leicht mit *Urlaubsgrund/-art* und *Gefühle*). Den höchsten Wert hat ein mit *Aktivität* und *Urlaubsart* sowie *Gefühlen* (»von den schönen, aber sehr sportlichen Skiferien«) geschmückter Urlaubsgruß *Karte A*, während der tiefste Wert in einer mit *Ort* und *Wetter* kombinierten Grußformel (»Grüsse aus dem sonnigen Spanien«) bei *Karte B* festgestellt wurde. Diese Muster ließen sich jeweils auf den 10 Karten finden, auf denen diese Komponente am stärksten ausgeprägt ist. Komponente 4 zeigt, dass der Gruß typischerweise mit der Thematisierung des *Ortes* und *Wetters* oder der *Aktivität* kombiniert wird:

Karte A (höchster Wert):

Transkript: Lieber Pius, /von den schönen, aber sehr sportlichen Skiferien die herzlichsten Purzelbaumgrüsse {Aktivität, Gefühle, Urlaubsart}/[unclear]/Chlaus/Berthe [NN] Dieter [anko] 110025 (1987)

Karte B (tiefster Wert):

Transkript: Lieben Dank für Fredis Karte und herzlichen Grüsse aus dem sonnigen Spanien, wo ich zwei Ferienwochen verbringe, {Ort, Extra-Diegetisches, Wetter}/Eure Hanni ([anko] 100297) (197X)

Wir interpretieren diese Komponente als Sprachmusterwandel in der Texthandlung (vgl. Hausendorf et al. 2017: 242–244) des Grüssens: In den 70er und 80er Jahren war eine mit *Wetter* und *Ort* kombinierte Grußformel offenbar üblicher als in anderen Jahrzehnten (vgl. Abbildung 3b: unten). Sie bündelte gewissermassen die stereotypen Erwartungen an einen Urlaubs(kurz)bericht im Zeitalter des Massentourismus.

4 Schlussbemerkungen

In unserer Studie haben wir gezeigt, dass die HKA ein geeignetes Verfahren ist, um den Wandel von thematischen Textsortenmustern in Ansichtskarten zu modellieren. Dabei zeigte sich auch, dass unser Annotationsschema eine adäquate Grundlage für datengeleitete Analysen bietet. Die HKA weist auf vier zentrale Aspekte hin, die für den Textsortenwandel entscheidend sind: Erstens kam es zu einer allmählichen Standardisierung in Hinblick auf die Entwicklung des Themas innerhalb eines Urlaubsframes und zu einer Entwicklung hin zu prototypischen Themen-Slots innerhalb des Urlaubsframes (wie *Aktivität*, *Wetter* und *Gefühle*). Zweitens zeigte die HKA, dass die dominante Funktion von Ansichtskarten sich allmählich von der Kontaktfunktion nach dem Typus *Korrespondenzkarte* hin zur Darstellungsfunktion im Sinne einer (selbst-)darstellenden Beschreibung der Urlaubswelt nach dem heute dominanten Typus der *Urlaubsansichtskarte* verschoben hat. Drittens konnten wir zeigen, dass sich über die HKA auch Hinweise auf soziokulturelle Verschiebungen nachzeichnen lassen, deren genaue Interpretation allerdings weitere Forschung erfordert. Viertens deckte das Verfahren auch einen Wandel sprachlicher Muster innerhalb der für Ansichtskarten konstitutiven Texthandlung des Grüßens auf, der Rückschlüsse auf eine besondere Erwartungshaltung gegenüber den Kategorien *Wetter* und *Ort* im Zeitalter des Massentourismus zulässt. Die framesemantische Hauptkomponentenanalyse hat also gezeigt, dass sich auf Ansichtskarten für den Zeitraum von 1950 bis 2010 ein stabiler Themenkatalog nachzeichnen lässt. Dank dieser Stabilität lassen sich die komponentialen Veränderungen in Verteilung und Gewichtung der thematischen Struktur als Indikator für den Wandel von Versprachlichungsmustern zur textsortenspezifischen Konstruktion von Urlaubswelten auf Ansichtskarten interpretieren.

Literatur

- Baayen, Herald/van Halteren, Hans/Tweedie, Fiona (1996): Outside the cave of shadows. Using syntactic annotation to enhance authorship attribution. *Literary and Linguistic Computing* 11 (3), 121–132.
- Biber, Douglas (1995): Dimensions of register variation. A cross-linguistic comparison. Cambridge: University Press.
- Brinker, Klaus/Cölfen, Hermann/Pappert, Steffen: (2014): Linguistische Textanalyse. Eine Einführung in Grundbegriffe und Methoden. 8., neu bearb. und erw. Aufl. Berlin: Erich Schmidt.
- Bubenhofer, Noah/Scheurer, Patricia (2014): Warum man in die Berge geht. Das kommunikative Muster »Begründen« in alpinistischen Texten. In: Stefan Hauser/Ulla Kleinberger/Kersten Sven Roth (Hg.): Musterwandel – Sortenwandel. Aktuelle Tendenzen der diachronen Text(sorten)linguistik. Bern: Peter Lang (= Sprache in Kommunikation und Medien 3), 245–274.
- Busse, Dietrich (2012): Frame-Semantik. Ein Kompendium. Berlin/Boston: De Gruyter.
- Diekmannshenke, Hajo (2008): Text-Bild-Kommunikation am Beispiel der Postkarte. In: Inge Pohl (Hg.): Semantik und Pragmatik – Schnittstellen. Frankfurt am Main: Lang

- (Sprache – System und Tätigkeit 59), 85–107. Hajič, Jan (1998): Building a Syntactically Annotated Corpus. The Prague Dependency Treebank. In: Eva Hajičová (Hg.): Issues of Valency and Meaning. Studies in Honour of Jarmila Panevová. Prag: Karolinum, 106–132.
- Hausendorf, Heiko (2008): Zwischen Linguistik und Literaturwissenschaft: Textualität revisited. Mit Illustrationen aus der Welt der Urlaubsansichtskarte. Zeitschrift für germanistische Linguistik 36 (3): 319–342.
- Hausendorf, Heiko (2009): Kleine Texte. Über Randerscheinungen von Textualität. In: Germanistik in der Schweiz. Online-Zeitschrift der Schweizer Akademischen Gesellschaft für Germanistik 6. Online unter: https://www.sagg-zeitschrift.unibe.ch/6_09/hausendorf.html
- Hausendorf, Heiko (2019): Das Ferienwetter auf der Ansichtskarte. Ein Wetterbericht im Schnittpunkt von Textlinguistik, Medienlinguistik und Korpuspragmatik. In: Juliane Schröter/Susanne Tienken/Yvonne Ilg/Joachim Scharloth/Noah Bubenhofer (Hg.): Linguistische Kulturanalyse. Berlin/Boston: De Gruyter (= Germanistische Linguistik 314), 293–321.
- Hausendorf, Heiko/Kesselheim, Wolfgang (2008): Textlinguistik fürs Examen. Göttingen: Vandenhoeck & Ruprecht.
- Hausendorf, Heiko/Kesselheim, Wolfgang/Kato, Hiloko/Breitholz, Martina (2017): Textkommunikation. Ein textlinguistischer Neuanatz zur Theorie und Empirie der Kommunikation mit und durch Schrift. Berlin/Boston: De Gruyter.
- Kato, Hiloko/Naef, Marcel/Sugisaki, Kyoko/Wiedmer, Nicolas (2021): »Wie ihr seht, hat es hier nicht mehr Platz zum Schreiben« – Eine text- und korpuslinguistische Untersuchung zur Rolle der Lesbarkeitsquellen kleiner Texte. In: Stefan Pappert/Kersten Sven Roth (Hg.): Kleine Texte. Bern: Peter Lang, 71–116.
- Moisl, Hermann (2015): Cluster analysis for corpus linguistics. Berlin/Boston: De Gruyter.
- Powers, David M. W. (2012): The problem with kappa. In: Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL), 345–355.
- Stede, Manfred/Mamprin, Sara (2016): Information structure in the Potsdam Commentary Corpus: Topics. In: Proceeding of the 9th International Conference on Language Resources and Evaluation (LREC), 1718–1723.
- Sugisaki, Kyoko/Wiedmer, Nicolas/Naef, Marcel/Hausendorf, Heiko (2019): Tracing changes in thematic structure of holiday picture postcards from 1950s to 2010s. In: Proceedings of the Workshop on Computational Methods in the Humanities 2018 (COMHUM), 67–75.
- Ziem, Alexander (2008): Frame-Semantik und Diskursanalyse – Skizze einer kognitions-wissenschaftlich inspirierten Methode zur Analyse gesellschaftlichen Wissens. In: Ingo Warnke/Jürgen Spitzmüller (Hg.): Methoden der Diskurslinguistik. Berlin/Boston: De Gruyter, 89–117.

