

Kollokationen in der Fachsprache des Marketings: Entwicklung und Anwendung eines Extraktionsmodells

Guadalupe Ruiz Yepes

Abstract The aim of this article is to explore collocations in the language of marketing. The method used relies on the description of scientific papers in the field of marketing published in specialised journals on the basis of a computer-held text archive or corpus in three languages (German, English and Spanish). In the first chapters different definitions for collocations are presented and the methods which have been applied for the extraction of collocations from corpora up to now are described and analysed. On this basis a new model for the extraction of collocations from corpora is presented and applied to the corpus. The results of the extraction of collocation candidates for German, English and Spanish are compared with each other and final conclusions are drawn. The results and findings of this piece of research can be applied to improve the performance in the fields of Translator Training and Second Language Acquisition.

Keywords collocations, languages for special purposes, marketing, corpus linguistics, phraseology of marketing

1 Einführung

Während Muttersprachlern Kollokationen wie *Geld abheben*, *Konto eröffnen* oder *Antrag stellen* bekannt sind, stellen sie für diejenigen, die es wagen, einen Text in einer Fremdsprache zu schreiben, eine häufige Fehlerquelle dar. Die Tatsache, dass *wir uns* im Deutschen *die Zähne putzen*, während *wir uns* im Spanischen *die Zähne waschen* und im Englischen *die Zähne bürsten*, ist sehr schwer mit semantischen Regeln zu erklären. Die semantische Verträglichkeit zweier Wörter gewährleistet also nicht ihre häufige Verbindung im Gebrauch der Sprache. Dies gilt nicht nur für die Allgemeinsprache, sondern insbesondere auch für die Fachsprachen. Die wohlbekannte Erkenntnis, dass Fachsprachen nicht nur aus Terminologie bestehen, sondern sich auch durch syntaktische und textstrukturelle Merkmale charakterisieren, ist für diese Arbeit besonders relevant, da die Bildung von Kollokationen fachsprachenspezifisch sein kann. Diese Erkenntnis öffnet auch die Türen der Forschung im Bereich der fachsprachlichen Phraseologie, die im Vergleich zu der gemeinsprachlichen Phraseologie noch in den Kinderschuhen steckt. Picht (1987) hat als Erster im Rahmen der Terminologieforschung Interesse für die fachsprachliche Phraseologie gezeigt und auf die Relevanz des gemeinsprachlichen Verbs beim Eintreten in eine Fachwendung hingewiesen (vgl. Picht 1988, 1993). Besonders in der Übersetzerausbildung ist die Problematik der „scheinbar harmlosen gemeinsprachlichen Verben“ wie *ziehen* (*einen Wechsel ziehen*, *die Wurzel ziehen*) nicht erkannt und wird somit nicht behandelt (vgl. Picht 1988: 187 f.). Dasselbe gilt für die Fremdsprachenausbildung. Hauptziel dieser Arbeit ist es, die Kollokationen aus einem multilingualen Korpus (Deutsch, Spanisch, Englisch) aus Marketingtexten¹ zu extrahieren und miteinander zu vergleichen. Es handelt sich vorwiegend um Kollokationen, die als solche noch nicht anerkannt sind, weil sie

¹ Unter „Marketingtexte“ verstehen wir keine Werbetexte, sondern wissenschaftliche Artikel über das Thema „Marketing“.

in einer noch wenig untersuchten Fachsprache vorkommen: der Fachsprache des Marketings. Um dieses Ziel zu erreichen, werden Kollokationen definiert und von anderen Wortverbindungen abgegrenzt und die bisher angewendeten Methoden der Extraktion von Kollokationen werden beschrieben. Mit der Kombination dieser und anderer Methoden wird ein Modell zur Extraktion von Kollokationen aus Korpora entwickelt und anschließend angewendet. Abschließend werden die Ergebnisse analysiert und Schlussfolgerungen gezogen. Die Ergebnisse dieser Studie können sowohl im Bereich der Übersetzerausbildung wie auch im Bereich der Fremdsprachendidaktik im Rahmen eines Wirtschaftsstudiums angewendet werden.

2 Die Marketingsprache als Fachsprache

Der Terminus *Fachsprache* ist laut Fluck (1996: 11) in der Fachsprachenforschung noch nicht „endgültig definiert“, da sich die meisten Linguisten über dessen Definition nicht einig sind. Dennoch vertreten die meisten von ihnen die Meinung, dass die Fachsprache eine Sonderform der Gemeinsprache ist, sich aus der Gemeinsprache entwickelt hat und sich der grammatischen Mittel der Gemeinsprache bedient (vgl. Arntz/Picht/Schmitz 2014). In der Fachsprachenforschung koexistieren zahlreiche Versuche, die Fachsprache zu gliedern, um für jeden Teilbereich der Fachsprachen eine Typologie aufzustellen. Obwohl sich die Linguisten in diesem Bereich auch nicht einig sind, haben sich zwei Arten von Gliederung durchgesetzt: die horizontale und die vertikale Gliederung (vgl. Hoffmann 1984: 58 ff.). Was die Wirtschaftssprache angeht, gab es auch mehrere Versuche, diese zu klassifizieren. Hundts Typologie (1995), die zu den meistzitierten gehört, unterscheidet zwischen verschiedenen kommunikativen Bezugswelten, in denen über Wirtschaft gesprochen wird: Alltag, Institutionen und Theorie/Wissenschaft (vgl. Hundt 1995: 8). Innerhalb der Bezugswelt Theorie/Wissenschaft bilden die Fachsprachen der Volks- und Betriebswirtschaftslehre den Kern der Wirtschaftswissenschaftssprache, wobei die Fachsprache des Marketings der Fachsprache der Betriebswirtschaftslehre untergeordnet ist.²

Mefferts Definition von Marketing verdeutlicht seine zentrale Philosophie: „In der klassischen Interpretation bedeutet Marketing die Planung, Koordination und Kontrolle aller auf die aktuellen und potenziellen Märkte ausgerichteten Unternehmensaktivitäten. Durch eine dauerhafte Befriedigung der Kundenbedürfnisse sollen die Unternehmensziele verwirklicht werden.“ (Meffert et al. 2015: 12) Der Wirtschaftswissenschaftszweig der Marketingforschung ist länderübergreifend und somit stark dem Einfluss der englischen Sprache ausgesetzt. Viele Begriffe und die damit verbundenen Termini oder Bezeichnungen wurden aus dem Englischen entlehnt, weil sowohl englische als auch amerikanische Theoretiker Vorreiter in dieser Disziplin waren und nicht nur die deutsche Fachsprache des Marketings, sondern auch die des Spanischen stark beeinflusst haben (vgl. Feix 1982: 84). Die Fachsprache des Marketings ist bisher kaum untersucht worden. Die Anzahl der dieser Fachsprache gewidmeten Arbeiten ist entsprechend überschaubar. Lediglich im Bereich der Terminologie konnten aufschlussreiche Untersuchungen aufgefunden werden. Besonders von Interesse ist die Klassifikation der Terminologie des Marketings nach deren Zugehörigkeit zu verschiedenen Fachsprachen. Als Subdisziplin der Wirtschaft weist das Marketing eine „sogenannte wirtschaftliche Basisterminologie“ auf, die in allen Subdisziplinen der Wirtschaft beobachtet werden kann (*Außenwirtschaft, Finanzierung usw.*). Zudem kommen aufgrund des wissenschaftlichen Charakters vie-

² Laut Bongard (2000: 48) ist das Marketing eine Subdisziplin der Betriebswirtschaftslehre.

ler Textsorten im Bereich Marketing (Experimentalstudien, Fallberichte usw.) Termini aus der Fachsprache der Statistik und Mathematik vor (*Varianzanalyse, Korrelationskoeffizient, usw.*). Die wichtigste Gruppe bilden jedoch die Termini aus der Wirtschaftspsychologie, die für die Fachsprache des Marketings unabdingbar sind (*Konsumentenverhalten, Kundenzufriedenheit, usw.*). Diese oft emotional gefärbten Wörter dienen der Beschreibung des Kundenverhaltens und seiner Konsumbedürfnisse (vgl. Konovalova/Ruiz Yepes 2016: 106).

3 Kollokationen: Definitionen

Für das Phänomen *Kollokation* existieren zahlreiche Definitionen. Seit dieser Begriff 1957 von Firth eingeführt wurde, haben sich Forscher aus den verschiedensten Forschungsbereichen damit beschäftigt. Dementsprechend finden sich Definitionen von *Kollokation* aus der Phraseologie, aus der Computerlinguistik, aus der Korpuslinguistik usw. Jede Disziplin – und innerhalb dieser jeder Ansatz – weist eine Definition von *Kollokation* auf, die ihre Bedürfnisse deckt. Daraus folgt, dass der Begriff sehr schwammig ist und keine genauen Grenzen aufweist. Trotz dieser Unklarheit können zwei Traditionen unterschieden werden: Die eine folgt Firths empirischen Postulaten im Rahmen des britischen Kontextualismus, die andere hat ihren Ursprung in der deutsch-französischen Lexikographie (Hausmann 1984). Anders als im britischen Kontextualismus, bei dem die Häufigkeit des Miteinandervorkommens zweier Wörter im Mittelpunkt steht, legt Hausmann den Schwerpunkt auf die Bedeutungsbeziehung dieser Wörter.

Die Kollokation besteht aus einer Basis, die semantisch autonom und somit ko-kreativ ist, und einem Kollokator, der zu dieser Basis affin oder kollokativ ist. (Hausmann 1984: 401)

Hausmanns (1989: 1010) Kollokationstypologie unterscheidet sechs verschiedene Kombinationsmöglichkeiten von Kollokatoren und Basen:

- Verb + Substantiv (Objekt) bzw. Substantiv (Objekt) + Verb (z. B. *ein Formular ausfüllen, einen Antrag stellen* usw.)
- Adjektiv + Substantiv (*schütteres Haar, starker Kaffee*)
- Substantiv (Subjekt) + Verb (*Glocken läuten*)
- Substantiv + (Präposition) + Substantiv (*eine Tafel Schokolade, eine Schachtel Zigaretten*)
- Adverb + Adjektiv (*bitter enttäuscht, voll genervt, hoch begabt*)
- Verb + Adverb bzw. Adverb + Verb (*scharf kritisieren*)

Für die Zwecke dieser Arbeit werden sowohl Firths empirische Tradition wie auch Hausmanns phraseologischer Ansatz berücksichtigt. Empirische Kollokationen, wie sie vom britischen Kontextualismus definiert werden, können zwar im Rahmen der Korpuslinguistik besser erforscht werden, weil quantitative Merkmale beim Extrahieren von Kollokationen aus dem Korpus im Mittelpunkt stehen. Dafür vereinfachen die sprachwissenschaftlichen Kategorien der lexikalischen Kollokation die Analyse und Klassifizierung dieses Phänomens. Laut Bartsch muss eine gute Definition von *Kollokation* beide Aspekte berücksichtigen:

The aim of devising a definition of collocations is twofold. In the first instance, it serves as the basis for the computer-aided extraction of collocation candidates, i. e. potential collocations, from the corpus. In a second step, a definition of collocations based on qualitative

linguistic criteria must be devised to enable the systematic description of the structural and functional properties of collocations and the distinction between different types of collocational structures. (Bartsch 2004: 68)

Dementsprechend werden in diesem Artikel verschiedene Definitionen von *Kollokation* verwendet. In der Phase der Extrahierung von Kollokationen aus dem Korpus ist die Definition von Evert (2009: 1214, in Anlehnung an Firth) angemessen: „a combination of two words that exhibit a tendency to occur near each other in natural language“. Mit „word pair“ meint Evert allerdings nicht Funktionswörter, sondern nur bedeutungstragende Wörter. Dementsprechend fällt für den Fall *kick the bucket* der Artikel *the* weg, sodass nur vom Wortpaar (*kick, bucket*) gesprochen wird. Dagegen wird in einer fortgeschrittenen Phase dieser Arbeit, zum Beispiel in der kontrastiven Analyse der Kollokationen in den Sprachen Deutsch, Spanisch und Englisch, mit der schon oben erwähnten Beschreibung von Hausmann gearbeitet (1984: 401).

4 Abgrenzung des Begriffs *Kollokation* von anderen Wortverbindungen

In der Phraseologie werden Kollokationen an der Nahtstelle zwischen freien Wortverbindungen und Idiomen angesiedelt. Ob Kollokationen als eine selbstständige Kategorie oder als eine Subkategorie von Phraseologismen verstanden werden, hängt davon ab, ob wir es mit einer „engen“ oder einer „weiteren“ Phraseologismen-Auffassung zu tun haben (vgl. Reder 2006: 44). Allerdings hängt es auch davon ab, ob die Idiomatizität oder die Festigkeit bzw. Stabilität als konstituierendes Merkmal von Kollokationen genommen wird. Da Kollokationen nichtidiomatischer Natur sind, bilden sie bei einer engen Phraseologismen-Auffassung eine selbstständige Kategorie zwischen den freien Wortverbindungen und den Phraseologismen. So klassifiziert Burger (2003: 36) in seiner eher engen Auffassung von Phraseologismen das gesamte phraseologische Material in referentielle, strukturelle und kommunikative Phraseologismen. Referentielle Phraseologismen nominativer Natur unterteilt er nach dem Grad der Idiomatizität in Idiome, Teil-Idiome und Kollokationen. Für eine weitere Phraseologismus-Auffassung wiederum gilt lediglich die Festigkeit als konstituierendes Merkmal. Demnach existieren nur phraseologische oder nichtphraseologische Wortverbindungen und die „Kollokationen als nichtidiomatische, jedoch feste Wortverbindungen bilden in diesem Modell eine Subkategorie der Phraseologismen“ (Reder 2006: 44). In diesem Zusammenhang werden sie auch schwache oder nichtidiomatische Phraseologismen genannt.

Unabhängig davon, für welche Auffassung man sich entscheidet, bleibt die Abgrenzungsproblematik ungelöst, denn die Kategorien lassen sich nicht immer leicht voneinander trennen. Kollokationen als lexikalische Einheiten bilden sowohl in Bezug auf die Idiomatizität wie auch in Bezug auf die Festigkeit eine Kategorie mit unscharfen Grenzen. Laut Reder (2006: 48) kann „die Zuordnung der Mehrworteinheiten zu freien Wortverbindungen oder Kollokationen sowie Idiomen [...] umstritten sein“.

Wie schon erwähnt, müssen für die Extraktion von Kollokationen aus Korpora neben quantitativen auch qualitative Methoden eingesetzt werden. Für diese Arbeit bedeutet dies nicht nur das Einsetzen von Patterns oder Mustern, sondern auch das Festlegen, ob und inwiefern wir es bei einer Wortkookkurrenz auch wirklich mit einer Kollokation zu tun haben. Dafür müssen die Idiomatizität und die Festigkeit einer Wortkookkurrenz nicht nur geprüft werden, sondern es müssen auch nachvollziehbare Kriterien für die Identifikation von Kollokationen herausgearbeitet werden, die auf diesen beiden Aspekten beruhen.

Im Online-Duden-Wörterbuch lautet die Definition/Bedeutung von *Idiom*: „eigentümliche Wortprägung, Wortverbindung oder syntaktische Fügung, deren Gesamtbedeutung sich nicht aus den Einzelbedeutungen der Wörter ableiten lässt (z. B. Angsthase = sehr ängstlicher Mensch)“. Kollokationen sind genau aus diesem Grund keine Idiome. Das erste identifizierende Kriterium für eine Kollokation könnte dementsprechend folgenderweise lauten: „Wortverbindung, deren Gesamtbedeutung sich aus den Einzelbedeutungen der Wörter ableiten lässt“. Die Festigkeit oder Stabilität wiederum kann als formale, lexikalische, syntaktische, psycholinguistische, pragmatische und semantische Festigkeit vorkommen. Besonders bedeutend für die Kollokationen ist die pragmatische Stabilität, bei der es sich um eine von der Rekurrenz des Miteinandervorkommens abhängige Stabilität (vgl. Caro Cedillo 2004: 41) handelt. Dies trägt zu einem weiteren Unterschied zwischen Idiomen und Kollokationen bei, denn Idiome weisen eine Art „innere“ syntaktische oder semantische Stabilität auf, Kollokationen dagegen eher eine „äußere“ Stabilität, die mit der Festlegung des Gebrauchs zu tun hat. Das zweite identifizierende Kriterium lautet für uns: „Polylexikalische Standards im Bereich der Kollokationen sind vor allem an Festlegungen des Gebrauchs gebunden und nicht unbedingt an bestimmte semantische und/oder syntaktische Restriktionen, wie es bei phraseologischen Ausdrücken mehr oder weniger der Fall ist.“ (Rothkegel 1994: 517)

Andererseits kann nur von einer Kollokation die Rede sein, wenn mehrere Optionen für den Platz des Kollokators zur Verfügung stehen. Nicht-Muttersprachler werden oft aufgrund der falschen Kollokationen als solche identifiziert. Wer nicht *to brush my teeth* sondern *to wash my teeth* oder Ähnliches sagt, wird von Muttersprachlern sofort als Englischlernender identifiziert. Obwohl wir mehrere Verben mit ähnlicher Bedeutung haben (*to clean, to wash*), die den Platz von *to brush* einnehmen könnten, kann hier nur *to brush* verwendet werden. Alles andere würde dem Muttersprachler sofort auffallen. Das dritte identifizierende Kriterium lautet dementsprechend: Eine Kollokation kann nur als solche betrachtet werden, wenn dem Sprecher mehrere Kollokatoren zur Verfügung stehen, die zwar semantisch passend für das Zusammenscheinen mit einer Basis sind, aber nur einer dieser Kollokatoren im Gebrauch bevorzugt wird.

Diese drei Kriterien sind bedauerlicherweise nicht alle gleich überprüfbar. Während das erste Kriterium recht konkret und gut überprüfbar ist, lassen sich Kriterien zwei und drei nur durch Recherchen in Korpora und dann auch nur teilweise nachweisen.

5 Extraktion von Kollokationen aus Korpora

So wie es verschiedene Definitionen von *Kollokation* gibt, gibt es auch verschiedene Methoden der Extraktion von Kollokationen aus Korpora. Bisher sind grundsätzlich drei Methoden verwendet worden: auf Basis eines Kookkurrenzalküls, auf Basis von Kollokationsmustern oder durch hybride Methoden. Die erste Methode nimmt das Kollokationskonzept von Firth als Ausgangspunkt und legt den Schwerpunkt auf die Häufigkeit des Miteinandervorkommens von Basis und Kollokator. Die zweite Methode nimmt Hausmanns binäre Kollokationsauffassung als Ausgangspunkt und legt den Schwerpunkt auf die Wortarten von Basis und Kollokator. Die dritte Methode ist eine Kombination aus den ersten beiden.

5.1 Auf Basis eines Kookkurrenzalküls

Wenn man sich dafür entschieden hat, die Kollokationen auf Basis eines Kookkurrenzalküls aus dem Korpus zu extrahieren, ist es unabdingbar, sich zunächst für einen Kookkurrenztyp zu entscheiden. Evert (2009: 1215) unterscheidet zwischen Oberflächenkookkurrenz („surface cooccurrence“), Textkookkurrenz („textual cooccurrence“) und syntaktischer Kookkurrenz („syntactic cooccurrence“). Im Rahmen dieser Studie ist es sinnvoll, die syntaktische Kookkurrenz als Ausgangspunkt zu nehmen, denn sie befasst sich mit der Kookkurrenz von Wörtern, die in einer bestimmten syntaktischen Beziehung zueinander stehen. Die bloße Kookkurrenz zweier Wörter reicht jedoch nicht aus, um eine starke Anziehung zweier Wörter zu beweisen. Mithilfe der Assoziationsmaße kann festgestellt werden, ob zwei Wörter nur durch Zufall nebeneinander erscheinen oder ob von einer Kollokation die Rede ist. Mit diesen Maßen kann auch unterschieden werden, ob es sich bei der Kookkurrenz zweier Wörter um „starke“ oder „schwache“ Kollokationen handelt. Die statistischen Assoziationsmaße vervollständigen somit die Definition der „empirischen Kollokation“ (Evert 2009: 1215).

Je nach Forschungsinteressen und Forschungszwecken wurden verschiedene Assoziationsmaße entwickelt. Für die Korpuslinguisten oder Forscher bedeutet dies, dass sie sich für bestimmte Assoziationsmaße entscheiden müssen. Um Kollokationen von bloßen Kookkurrenzen zu unterscheiden, werden Listen erstellt, in denen Wortpaare nach Erscheinungshäufigkeit im Korpus eingeordnet sind. Die häufigen Wortpaare werden als „starke“ Kollokationen und die nicht so häufigen Wortpaare als „schwache Kollokationen“ bezeichnet. Laut Evert (2009: 1236) gibt es kein „perfektes“ Assoziationsmaß. Aus diesem Grund sollte man für eine Studie mehrere verwenden, um verschiedene Ergebnisse zu erzielen und diese untereinander vergleichen zu können. Assoziationsmaße werden in zwei Gruppen unterteilt, bei denen die Fragestellung grundsätzlich entgegengesetzte Ziele verfolgt: „effect-size measures“ und „significance measures“. Laut Evert (2009: 1228) stellt sich der Forscher bei den „effect-size measures“ die Frage: „How much does observed cooccurrence frequency exceed expected frequency?“ und bei den „significance measures“: „How unlikely is the null hypothesis that the words are independent?“ Innerhalb der ersten Gruppe haben sich folgende Maße als nützlich erwiesen: *mutual information (MI)*, *dice*, *odds*, *ratio*, und innerhalb der zweiten Gruppe sind folgende Maße bekannt: *z-score*, *t-score*, *simple II*, *chi-squared*, *log-likelihood*. Evert schlägt vor, sich für Assoziationsmaße aus beiden Gruppen zu entscheiden, weil die Ergebnisse sehr verschieden ausfallen und man sich ein genaues Bild der Kookkurrenz zweier Wörter machen kann. Er empfiehlt *MI*, *t-score* und *simple II* (*log-likelihood* basiert auf *simple II*). Kilgariff und Rundell (2002: 810) empfehlen, die Funktionswörter eines Korpus herauszufiltern, bevor man ein Assoziationsmaß anwendet, weil sonst die Anzahl von „falschen“ Kollokationen (*noise*) im Endergebnis zu hoch wäre. Sie schlagen auch vor, das Korpus vorher zu lemmatisieren, weil so alle Bigramme desselben Lemmas zusammen erscheinen (vgl. Kilgariff/Rundell 2002: 811).

5.2 Auf Basis von Kollokationsmustern (Patterns)

Heid (1994) unterscheidet sechs verschiedene Kollokationsmuster, die auf Hausmanns binärer Kollokationsauffassung basieren: Nomen + Nomen, Adjektiv + Nomen, Verb + Nomen (als Objekt), Verb + Nomen (als Subjekt), Adverb + Verb und Adverb + Adjektiv. In späteren Studien nennen Weller und Heid (2010: 3195) die extrahierten Kollokationsmuster „collocation candidates“. Als Ausgangspunkt nehmen sie die Definition von *Kollokation* von Bartsch (2004:

76): „[collocations are] lexically and/or pragmatically constrained recurrent cooccurrences of at least two lexical items which are in a direct syntactic relation with each other“. Zum Extrahieren von Kollokationen aus Korpora müssen diese mindestens mit POS-Tags annotiert sein, einen höheren *recall* erreicht man jedoch, wenn die Korpora auch noch syntaktisch analysiert (*parsed*) werden. Die Ergebnisse werden dann in einem relationalen Datenbanksystem gespeichert. Abschließend erfolgt die Aussortierung durch die Korpuslinguisten mithilfe von Interfaces wie LexiView (Evert et al. 2005).

5.3 Hybride Methoden

Pamies und Pazos (2005: 317–329) vergleichen die verschiedenen Assoziationsmaße untereinander und kommen zu der Schlussfolgerung, dass alleine mathematische Methoden nicht hinreichend sind und stattdessen die Anwendung von „hybriden“ Methoden vorzuziehen ist. Damit sind Methoden gemeint, die eine Kombination aus einer Kookkurrenzberechnung und linguistischen Kriterien in Form von Kollokationsmustern anwenden.

Nun stellt sich eine Frage: In welcher Reihenfolge sollen diese Filtermethoden angewendet werden? Dazu gibt es verschiedene Ansätze:

1. Zunächst werden statistische Verfahren verwendet, und um das Filtern zu verfeinern, werden zusätzliche Kollokationsmuster oder Patterns auf die erzielten Ergebnisse angewendet (vgl. Smadja 1993).
2. Es werden zunächst Kollokationsmuster eingesetzt und im zweiten Schritt die statistischen Methoden auf die erzielten Ergebnisse angewendet (vgl. unter anderem Heid 1998, Krenn 2000, Evert 2005).

Auf der Suche nach einer Software, die die Extraktion von Kollokationen ermöglicht, hat sich ConcGram 1.0 als besonders nützlich erwiesen. Diese Software wurde als eine „inklusive“ Suchmaschine für phraseologische Einheiten entwickelt und funktioniert auf Basis eines Kookkurrenz kalküls. Es bleibt also den Wissenschaftlern überlassen zu entscheiden, welche dieser Kookkurrenzen bedeutende Wortkombinationen darstellen und welche nur zufällige Wortkookkurrenzen sind. Das Programm erstellt nach Eingabe eines Befehls eine Liste von „unique words“, die als Grundlage für das Anzeigen von „Concgrams“ dienen, wobei Concgrams sowohl aneinandergrenzende wie nicht aneinandergrenzende Wortkookkurrenzen sind, die in einer beliebigen Reihenfolge im Korpus erscheinen können. Sobald die Concgrams-Listen erstellt sind, können statistische Verfahren in Gang gesetzt werden, die eine Reduzierung der Listen ermöglichen und klare Hinweise auf bedeutungslose Wortkookkurrenzen geben, die sicher ignoriert werden können. Bei den angewendeten statistischen Verfahren handelt es sich um die *t-score*- und *MI*-Tests, deren Formeln ausführlich von Barnbrook (1996: 88–106) erklärt werden. Nun erfährt man bei der computergestützten Korpusanalyse, dass eine maschinell erstellte Häufigkeitsliste – trotz Anwendung der Assoziationsmaße – nicht immer direkt verwertbar ist, sondern einer zusätzlichen Selektion durch den Analysierenden bedarf. Aus diesen Gründen ist für diese Arbeit der erste von den oben genannten Ansätzen am geeignetsten. Das heißt, als Erstes werden statistische Verfahren verwendet, und um das Filtern zu verfeinern, werden zweitens auf die erzielten Ergebnissen zusätzlich manuell die im Kapitel 3 aufgeführten Kollokationsmuster von Hausmann (1989) angewendet.

6 Korpora

Für viele Forschungsfragen können bestehende, im Idealfall frei verfügbare Korpora verwendet werden. Doch da es sich in der Regel um Korpora handelt, die hauptsächlich aus Presstexten bestehen, ist es sinnvoll, selbst ein Korpus oder mehrere Korpora zusammenzustellen, die den eigenen Bedürfnissen möglichst gut entsprechen: „It is a truism that there is no such thing as a ‚good‘ or a ‚bad‘ corpus, because how a corpus is designed depends on what kind of corpus it is and how it is going to be used.“ (Hunston 2009: 155) Diese Aussage von Hunston gehört zu den meistzitierten Sätzen in der Korpuslinguistik, denn sie besagt, dass die Zusammenstellung eines Korpus vom Forschungsvorhaben der Forscher abhängt, und verleiht diesen somit absolute Freiheit beim Konzipieren der Korpora.

In dieser Arbeit wird ein vergleichbares Korpus zur Analyse gezogen. Es handelt sich um drei Korpora in drei Sprachen (Deutsch, Spanisch und Englisch), die miteinander vergleichbar sind, weil sie derselben Textsorte angehören. Alle drei bestehen aus in wissenschaftlichen Fachzeitschriften veröffentlichten Artikeln im Bereich Marketing.

Im Falle des englischen Korpus handelt es sich um alle Artikel, die im Zeitraum 2000–2013 in der Fachzeitschrift „The Journal of Marketing Management“ veröffentlicht wurden. Das englische Korpus besteht aus 1.131.744 Wörtern. Im deutschen Korpus sind alle Artikel enthalten, die im Zeitraum 2000–2012 in der Fachzeitschrift „Der Markt – International Journal of Marketing“ erschienen sind. Das deutsche Korpus besteht aus 903.430 Wörtern. Im Falle des spanischen Korpus handelt es sich um alle im Zeitraum 2004–2014 in der Fachzeitschrift „Revista española de Investigación de Marketing“ erschienenen Artikel. Das spanische Korpus besteht aus 1.279.954 Wörtern.

Das Anreichern von Korpora mit linguistischen Informationen geschieht durch das Annotieren. Die bekanntesten und meist verwendeten Formen der Annotation sind das Part-of-Speech-Tagging, bei der die Wortart der einzelnen Token bestimmt wird, die Lemmatisierung, bei der die Grundformen der Token bestimmt werden, und das Parsing, bei dem die Gruppen von Token syntaktisch kategorisiert werden. Es können jedoch auch andere Metainformationen hinzugefügt werden, z. B. zur Semantik von einzelnen Token. Viele Forscher bevorzugen annotierte Korpora, weil die Abfragen konkretisiert werden können. Man kann zum Beispiel gezielt nach allen Verb-Nomen-Konstruktionen oder Adverb-Verb-Konstruktionen usw. suchen. Da für uns jedoch das häufige Erscheinen von bestimmten Kookkurrenzen im Mittelpunkt steht und die Software Concgram 1.0 auch nur mit Textdateien ohne Annotationen arbeitet, sind unsere Korpora nicht annotiert. Die Kollokationsmuster oder Patterns, die den genannten Wortart-Konstruktionen entsprechen, werden manuell angewendet.

Die Extraktion der Kollokationen wird nur im englischen Korpus vorgenommen. Mit dem deutschen und spanischen Korpus gehen wir semasiologisch vor. Ausgehend von den extrahierten Kollokationen in englischer Sprache – und deren Begrifflichkeit – werden die Entsprechungen auf Deutsch und Spanisch mittels der Programm-Funktion „concordances“ mit dem Format „keyword in context“ gesucht, um sie miteinander zu vergleichen. Die englische Sprache wurde somit als Basis gewählt, weil sie die deutsche und die spanische Sprache des Marketings geprägt hat.

7 Entwicklung eines Extraktionsmodells

Wie in Kapitel 4 angedeutet, bilden Kollokationen als lexikalische Einheiten eine Kategorie mit sehr unscharfen Grenzen und die Zuordnung der Mehrworteinheiten zu freien Wortverbindungen oder Kollokationen kann sehr umstritten sein (vgl. Reder 2006: 48). Bei der Analyse einer Wortkookkurrenz geht es in dieser Arbeit hauptsächlich darum festzustellen, ob es sich um eine Kollokation handelt. In den folgenden Seiten wird zum Herausfiltern von Kollokationen aus dem englischen Korpus ein Modell angewendet, das aus folgenden Phasen besteht:

1. Vorarbeit mit ConcGram 1.0 oder einer anderen Software

Nachdem die drei Korpora für die drei Sprachen zusammengestellt wurden, indem die kleinen Dateien in eine große Datei fusioniert wurden, müssen ConcGram-Listen (Listen mit häufigen Wortkookkurrenzen oder Kollokationskandidaten) erstellt werden. Dieser Vorgang besteht aus zwei unumgänglichen Schritten: An erster Stelle müssen *unique words*³-Listen erstellt werden. Dazu muss unter dem Menü-Punkt „Statistics“ die „Unique Words“-Funktion gewählt werden. Mithilfe dieser Liste führt das Programm in einem zweiten Schritt eine Recherche im gesamten Korpus durch und erstellt eine 2-Wörter-Concgram-Liste⁴. Somit haben wir eine erste Liste mit Kollokationskandidaten, die jedoch mit der Anwendung von Assoziationsmaßen verfeinert werden soll.

2. Anwendung der Assoziationsmaße: *t*-score und mutual information (MI)

Sobald die Concgrams-Listen erstellt sind, können statistische Verfahren in Gang gesetzt werden, die eine Reduzierung der Listen ermöglichen und klare Hinweise auf bedeutungslose Wortkookkurrenzen geben. Wie von Evert (2009: 1228) empfohlen, werden in dieser Phase die Assoziationsmaße *t*-score und *mutual information* (MI) angewendet. Somit haben wir ein Assoziationsmaß der Gruppe „effect-size measures“ und ein Assoziationsmaß der Gruppe „significance measures“. Sowohl die einen wie die anderen basieren auf einer sehr allgemeinen Beschreibung von statistischen Verfahren von Stubbs (1995: 7): „Many statistical calculations compare (a) how often something is actually observed and (b) how often it might be expected merely by chance.“ Concgram 1.0 ermöglicht es, entweder eine *t*-score- oder eine MI-„Grenze“ oder -„Sperre“ anzuwenden oder beide gleichzeitig. Wenn die Entscheidung getroffen ist und die ConcGrams- und *unique word*-Listen ausgewählt wurden, erfolgt die Erstellung der Werte automatisch. Möchte man die grammatischen Wörter in einer Kollokation bei der Analyse berücksichtigen, empfiehlt es sich, *t*-score anzuwenden. Wenn es nicht der Fall ist, sollte man MI verwenden, weil bei diesem Test nur die lexikalischen Wörter berücksichtigt werden.

3. Manuelle Anwendung von Kollokationsmustern oder Patterns

Dieses Modell ist für die Extraktion von Kollokationen aus nicht annotierten Korpora konzipiert. Aus diesem Grund müssen die Kollokationsmuster bestehend aus Basis und Kollokator nach Hausmanns Typologie (1989) manuell angewendet werden. Die Typologie berücksichtigt sechs mögliche Konstrukte: Verb + Substantiv (Objekt) bzw. Substantiv (Objekt) + Verb; Adjektiv + Substantiv; Substantiv (Subjekt) + Verb; Substantiv + (Präposition) + Substantiv; Adverb + Adjektiv; Verb + Adverb bzw. Adverb + Verb.

³ „Unique Word“-Listen sind keine *hapax legomena*-Listen, sondern Types-Listen.

⁴ Wir arbeiten nur mit ConcGram-Listen mit Kollokationskandidaten bestehend aus zwei Wörtern, weil die Formeln der Assoziationsmaße *t*-score und MI nur für solche Kollokationskandidaten zugeschnitten sind.

4. *Ausschluss der Konkurrenten des Kollokators auf semantischer Ebene*

Wie schon im Kapitel 4 angesprochen, kann nur von einer Kollokation die Rede sein, wenn mehrere Optionen für den Platz des Kollokators zur Verfügung stehen. In dieser vierten Phase des Extraktionsmodells wird das dritte identifizierende Kriterium angewendet. Dieses lautet: Eine Kollokation kann nur als solche betrachtet werden, wenn dem Sprecher mehrere Kollokatoren zur Verfügung stehen, die zwar semantisch passend für das Zusammensehen mit einer Basis sind, von denen aber nur einer im Gebrauch bevorzugt wird. Diese vierte Phase des Extraktionsmodells besteht also darin, Synonyme des Kollokators im Korpus zu suchen und zu prüfen, ob sie auch in Verbindung mit derselben Basis vorkommen.⁵

5. *Prüfung der Wortkookkurrenz auf ihre Idiomatizität und Festigkeit*

In dieser Phase geht es vor allem darum, die Idiomatizität und die Festigkeit der Kollokationen zu prüfen, indem die Kollokationen den zwei ersten im Kapitel 4 vorgestellten Kriterien unterzogen werden. Es handelt sich also darum festzustellen, dass es sich bei den Kollokationen nicht um Idiome handelt, so dass sie das erste identifizierende Kriterium erfüllen: „Wortverbindung, deren Gesamtbedeutung sich aus den Einzelbedeutungen der Wörter ableiten lässt“. Es geht aber auch darum, dass die Kollokationen eine pragmatische Stabilität aufweisen und somit das zweite identifizierende Kriterium erfüllen: „Polylexikalische Standards im Bereich der Kollokationen sind vor allem an Festlegungen des Gebrauchs gebunden und nicht unbedingt an bestimmte semantische und/oder syntaktische Restriktionen, wie es bei phraseologischen Ausdrücken mehr oder weniger der Fall ist.“ (Rothkegel 1994: 517)

8 Die Analyse der Wortkookkurrenzen: *satisfaction* und *loyalty*

Hauptziel dieses Kapitels ist es, die Nutzbarkeit des in Kapitel 7 vorgeschlagenen Modells darzustellen. Dazu werden seine Phasen systematisch durchgegangen. Ab Phase 4 erfolgt zudem eine vergleichende Analyse für die drei Sprachen des Korpus, um einen möglichen Anwendungsbereich des Modells zu erläutern.

Phase 1: Vorarbeit mit ConcGram 1.0 oder einer anderen Software

Da die Software ConcGram 1.0 nur aus einem Korpus automatisch Listen mit Kollokationskandidaten erstellen kann, müssen an erster Stelle alle Dateien, die das Korpus bilden sollen, fusioniert werden. Dazu verfügt die Software über eine spezifische Funktion unter „Tools Menu“, die alle ausgewählten kleinen Dateien in eine neue große Textdatei kopiert. So wurden alle Artikel derselben Sprache in einem Korpus fusioniert, bevor sie der Analyse unterzogen wurden. Wie schon im Kapitel 7 erklärt, werden mithilfe der Software Listen von ConcGrams erstellt, die wiederum auf Listen von *unique words* basieren.

Phase 2: Anwendung der Assoziationsmaße: t-score und mutual information (MI)

Die Wörter *customer* und *consumer* – Kunde und Konsument auf Deutsch und *cliente* und *consumidor* auf Spanisch – sind im Korpus sehr gut vertreten. Sie sind in den Häufigkeitslisten sehr hoch platziert und wurden aufgrund dieser hohen Platzierung als Untersuchungsge-

⁵ Uns ist bewusst, dass nicht alle in einem Synonymlexikon aufgeführten „Synonyme“ in allen Kontexten wirklich als Synonym zu bezeichnen sind und dass es auch Wörter gibt, die zwar nicht als solche aufgeführt sind, aber in einem bestimmten Kontext synonym verwendet werden können. Diese Problematik der Definition von *Synonym* wird hier jedoch nicht näher behandelt.

genstand in dieser Analysephase ausgewählt. Im englischen Korpus befindet sich das Wort *consumer* auf Platz 31 der allgemeinen Häufigkeitsliste. Wenn aber die grammatischen Wörter weggelassen (wie *the* auf Platz 1 oder *of* auf Platz 2) und nur die Substantive berücksichtigt werden, befindet sich *consumer* nach *marketing*, *research*, *brand* und *journal* auf Platz 5. Ähnliches gilt für das Wort *customer* und für die oben erwähnten spanischen und deutschen Entsprechungen (*cliente/consumidor* und *Kunde/Konsument*). Besonders häufig treten sie mit den Wörtern *satisfaction/loyalty*, *satisfacción/lealtad* und *Zufriedenheit/Bindung* auf, die zur im Kapitel 1 dieser Arbeit erwähnten Wirtschaftspsychologie gehören. Bei einer näheren Beobachtung dieser Wortverbindungen sind interessante Ergebnisse ans Licht gekommen. Aber an erster Stelle muss überprüft werden, ob wir es hier wirklich mit Kollokationen zu tun haben. Die Ausgangswörter dieser Untersuchung sind *customer* und *consumer*, weil sie die Basis der Kollokationen bilden. Bei den Wörtern *satisfaction* und *loyalty* handelt es sich um die Kollokatoren.

Da das Korpus nicht annotiert ist, konnten die Funktionswörter nicht herausgefiltert werden und die Anzahl von „falschen“ Kollokationen (*noise*) im Endergebnis ist sehr hoch. Aus diesen Grund zeigt Tabelle 1 die Kookkurrenzen *customer satisfaction* und *customer loyalty* nach *MI*-Wert geordnet und nicht nach Häufigkeit, da sich in unmittelbarer Nähe sehr viele „falsche Kollokationen“ befinden.

Tab. 1: Häufige Kookkurrenzen von *customer*, nach *MI*-Wert geordnet

Single origin	Co-occ Word	Instances	%	t-score	MI-Value
<i>customer</i>	<i>dysfunction</i>	5	0,001771	2,193754	5,723688
<i>customer</i>	<i>satisfaction</i>	168	0,059520	12,778880	5,135948
<i>customer</i>	<i>centrality</i>	7	0,002480	2,746655	5,112253
<i>customer</i>	<i>delight</i>	6	0,002126	2,376527	5,069184
<i>customer</i>	<i>retention</i>	41	0,014526	6,209386	5,046594
<i>customer</i>	<i>disposition</i>	21	0,007440	4,424282	4,855478
<i>customer</i>	<i>complaints</i>	18	0,006377	4,316985	4,849219
<i>customer</i>	<i>engagement</i>	155	0,054914	12,094509	4,846715
<i>customer</i>	<i>loyalty</i>	128	0,045349	10,901133	4,777269
<i>customer</i>	<i>competitor</i>	7	0,002486	2,542439	4,678600

Da die Wörter *satisfaction*, *loyalty*, *customer* und *consumer* schon am Anfang der Studie aufgrund ihrer Häufigkeit im Korpus aufgefallen waren, wurden Kookkurrenzen dieser Wörter gesucht und es stellte sich heraus, dass auch sie vergleichsweise sehr häufig sind.

Als Kontrast zu den in Tabelle 1 vorgestellten Beispielen und um den Umfang des Korpus besser darzustellen, zeigt Tabelle 2 die häufigsten Kookkurrenzen. Die Kookkurrenz *et al* erscheint 2075 Mal im Korpus und steht somit an erster Stelle mit einem *t-score* von 45,074707. Sowohl diese wie auch die Kookkurrenzen *Journal Marketing* und *Journal Research* schulden ihr Dasein der Tatsache, dass die Aufsätze samt Bibliographie im Korpus eingespeist wurden.

Tab. 2: Häufigste Kookkurrenzen im englischen Korpus

Single origin	Co-occ Word	Instances	%	t-score	MI-Value
<i>et</i>	<i>al</i>	2075	0,735146	45,074707	6,511850
<i>Journal</i>	<i>Marketing</i>	1616	0,572528	36,989315	3,295906
<i>such</i>	<i>us</i>	1283	0,454550	33,110470	3,620115
<i>be</i>	<i>can</i>	1004	0,355704	29,674728	3,977694
<i>Journal</i>	<i>Research</i>	914	0,323818	26,661165	3,035505
<i>been</i>	<i>has</i>	872	0,308938	28,846928	5,333758

Tabelle 3 zeigt die Kookkurrenzen *consumer satisfaction* und *consumer loyalty* und acht andere Beispiele, die nach *MI*-Wert in unmittelbarer Nähe erscheinen.

Tab. 3: Häufige Kookkurrenzen von *consumer*, nach *MI*-Wert geordnet

Single origin	Co-occ Word	Instances	%	t-score	MI-Value
<i>consumer</i>	<i>global</i>	50	0,017714	6,486381	3,446542
<i>consumer</i>	<i>cross</i>	6	0,002126	2,218779	3,408322
<i>consumer</i>	<i>interpretive</i>	6	0,002126	2,218779	3,408322
<i>consumer</i>	<i>theft</i>	4	0,001417	1,811625	3,408322
<i>consumer</i>	<i>perceptions</i>	69	0,024446	7,514796	3,391003
<i>consumer</i>	<i>interpreting</i>	5	0,001771	2,018438	3,361017
<i>consumer</i>	<i>culture</i>	84	0,029760	8,267654	3,352179
<i>consumer</i>	<i>loyalty</i>	7	0,002480	2,384688	3,341208
<i>consumer</i>	<i>satisfaction</i>	72	0,025509	7,684147	3,312745
<i>consumer</i>	<i>demographics</i>	9	0,003189	2,696507	3,305229

Aus diesen Tabellen können schon die ersten Rückschlüsse gezogen werden. Während die Kookkurrenzen von *satisfaction* und *loyalty* mit *customer* beide einen *t-score*-Wert von über 10 aufweisen (12,78 und 10,90 jeweils), sind ihre *t-score*-Werte in Verbindung mit *consumer* viel niedriger (7,68 und 2,38 jeweils). Hinsichtlich soziolinguistischer und pragmatischer Faktoren könnte dies darauf hindeuten, dass die Wörter mit der Wurzel *consum-* auf Englisch eine negative semantische Prosodie oder Konnotation haben.

Phase 3: Manuelle Anwendung von Kollokationsmustern oder Patterns

Wie schon erwähnt, konnten die Funktionswörter im Korpus nicht herausgefiltert werden, weil das Korpus nicht annotiert ist. Aus diesem Grund mussten die Kollokationsmuster nach Hausmanns Typologie (1984) manuell angewendet werden. Der Schwerpunkt wurde dabei auf die Kollokationen mit dem Muster „Substantiv + Substantiv“ gelegt. Dass letztlich die Kookkurrenzen der Wörter *satisfaction*, *loyalty*, *customer* und *consumer* gewählt wurden, hat mehr mit der Intuition der Autorin zu tun, die das Gefühl hatte, dass die Kookkurrenzen dieser Wörter interessante sprachliche Tendenzen zeigen könnten. Die Intuition kann also aus einer solchen Studie nie ganz ausgeschlossen werden. Wie Stubbs (1995: 13) argumentiert: „The statistics are produced mechanically by the software and can be calculated for any new corpus, but their interpretation clearly involves subjective judgements.“

Phase 2 und 3 des Modells werden nur für das englische Korpus verwendet. Mit dem deutschen und spanischen Korpus wird (wie in Kapitel 6 erwähnt) von einer Bezeichnung auf Englisch ausgegangen und nach ihren Entsprechungen auf Deutsch und Spanisch gesucht.

Phase 4: Ausschluss der Konkurrenten des Kollokators auf semantischer Ebene

Wie in der Einführung dieses Kapitels erwähnt, erfolgt ab Phase 4 eine vergleichende Analyse für die drei Sprachen des Korpus. Dazu werden die Synonyme der Kollokatoren in den drei Sprachen im Korpus gesucht und geprüft, ob sie auch in Verbindung mit derselben Basis vorkommen.

Für das englische Wort *loyalty* wurden im Online-Wörterbuch “Collins Cobuild Synonyms Dictionary” folgende Synonyme gefunden: „faithfulness, commitment, devotion, allegiance, reliability, fidelity, homage, patriotism, obedience, constancy, dependability, trustworthiness, steadfastness, troth (archaic), fealty, staunchness, trueness, trustiness, true-heartedness“.⁶ Während die Wortkombination *consumer loyalty* 7 Mal und die Wortkombination *customer loyalty* 128 Mal erscheint, kommen die Synonyme von *loyalty* (mit Ausnahme von *customer commitment* 2 Mal) kein einziges Mal in Verbindung mit *consumer/customer* vor. Für das englische Wort *satisfaction* wurde ein ähnliches Phänomen festgestellt. Im “Collins Cobuild Synonyms Dictionary” wurden folgende Synonyme gefunden: „fulfilment, pleasure, achievement, joy, triumph, happiness, relish, glee, gratification, pride, complacency, contentment, content, comfort, ease, pleasure, well-being, happiness, enjoyment, peace of mind, gratification, satiety, repletion“. Während die Wortkombination *consumer satisfaction* 72 Mal und *customer satisfaction* 168 Mal im Korpus vorkommen, erscheinen die Synonyme von *satisfaction* kein einziges Mal mit den Bezeichnungen *consumer/customer*.

Tab. 4: Kollokationskandidaten der Wörter *customer* und *consumer*

Sprache	Kollokator	Basis	Kollokation	H.	Struktur
Englisch	<i>loyalty</i>	<i>customer</i>	<i>customer loyalty</i>	128	Subs. + Subs.
		<i>consumer</i>	<i>consumer loyalty</i>	0	Subs. + Subs.
	<i>commitment</i>	<i>customer</i>	<i>customer commitment</i>	2	Subs. + Subs.
		<i>consumer</i>	<i>consumer commitment</i>	0	Subs. + Subs.
	<i>satisfaction</i>	<i>customer</i>	<i>customer satisfaction</i>	168	Subs. + Subs.
		<i>consumer</i>	<i>consumer satisfaction</i>	72	Subs. + Subs.

Für das deutsche Wort *Loyalität* stehen folgende Synonyme im Online-Duden-Synonymwörterbuch: „Ergebenheit, Solidarität, Gerechtigkeit, Anständigkeit, Redlichkeit, Zuverlässigkeit, Anhänglichkeit, Geradlinigkeit, Standhaftigkeit, Unwandelbarkeit, Pflichtgefühl, Verschwiegenheit, Konstanz, Treue“. Mit den Wörtern *Kunde* oder *Konsument* erscheinen jedoch nur *Kundenloyalität* (19 Mal), *Loyalität der Kunden* (1 Mal) und *Loyalität der Konsumenten* (1 Mal). *Treu* erscheint nur als Attribut *treue Kunden* (4 Mal), es erscheint aber kein Mal *Kundentreue*. Für das deutsche Wort *Bindung* stehen folgende Synonyme im Online-Duden-Synonymwörterbuch: „Beziehung, [innere] Verbundenheit, Liaison, Verbindung, Verhältnis, Zusammenhalt; (bildungssprachlich) Konnex“. Die Bezeichnung *Kundenbindung* erscheint im

⁶ Selbstverständlich können viele dieser „Synonyme“ aufgrund des Kontexts von vornherein ausgeschlossen werden. Dies gilt auch für die deutschen und spanischen Beispiele. Die Ausarbeitung der Frage, welche Synonyme berücksichtigt werden sollten und welche nicht, würde jedoch eine neue Forschungslinie öffnen und würde den Rahmen dieser Arbeit sprengen.

deutschen Korpus 286 Mal, was sehr überraschend war, weil die wörtliche Entsprechung auf Englisch nur zwei Mal auftaucht.

Tab. 5: Kollokationskandidaten der Wörter *Kunde* und *Konsument*

Sprache	Kollokator	Basis	Kompositum/ Possessivmarkierung	H.	Struktur
Deutsch	<i>Loyalität</i>	<i>Kunde</i>	<i>Kundenloyalität</i>	19	Determinativkompositum
			<i>Loyalität des Kunden</i>	0	Genitiv
		<i>Konsument</i>	<i>Konsumentenloyalität</i>	0	Determinativkompositum
			<i>Loyalität des Konsumenten</i>	0	Genitiv
	<i>Bindung</i>	<i>Kunde</i>	<i>Kundenbindung</i>	286	Determinativkompositum
			<i>Bindung des Kunden</i>	4	Genitiv
		<i>Konsument</i>	<i>Konsumentenbindung</i>	0	Determinativkompositum
			<i>Bindung des Konsumenten</i>	1	Genitiv
	<i>Zufriedenheit</i>	<i>Kunde</i>	<i>Kundenzufriedenheit</i>	197	Determinativkompositum
			<i>Zufriedenheit des Kunden</i>	5	Genitiv
		<i>Konsument</i>	<i>Konsumentenzufriedenheit</i>	1	Determinativkompositum
			<i>Zufriedenheit des Konsumenten</i>	0	Genitiv

Für das deutsche Wort *Zufriedenheit* stehen folgende Synonyme im Online-Duden-Synonymwörterbuch: „Erfüllung, Sättigung, Befriedigung, Beglückung, Behagen, Freude, Frohsinn, Glückseligkeit, Augenweide, Fröhlichkeit, Glück, Lust, Wonne, Ausgeglichenheit, Ausgewogenheit, Einhelligkeit, Einklang, Eintracht, Friede, Gleichgesinntheit, Gleichgewicht, Übereinstimmung, Wohlklang, Freundschaft, Frieden, Brüderlichkeit, Einigkeit, Harmonie“ und noch viele mehr, die hier nicht erwähnt werden sollen. *Kundenzufriedenheit* erscheint 197 Mal, *Zufriedenheit des Kunden* 5 Mal und *Konsumentenzufriedenheit* 1 Mal. Die Synonyme wurden kein einziges Mal in Verbindung mit den Wörtern *Konsumenten* und *Kunden* im Korpus gefunden.

Im Falle der spanischen Sprache finden wir im „Diccionario de sinónimos y antónimos Espasa-Calpe“ (2005) für *lealtad* folgende Synonyme: „fidelidad, nobleza, franqueza, amistad, adhesión, cumplimento, devoción, honradez, observancia“. Während im spanischen Korpus die Kombinationen *lealtad del consumidor* 31 Mal und *lealtad del cliente* 17 Mal gefunden wurde, haben wir die Wörter *consumidor* und *cliente*, bis auf *fidelidad del cliente* (1 Mal) und *fidelidad de los clientes* (2 Mal), kein einziges Mal gefunden. Für das spanische Wort *satisfacción* wurde ein ähnliches Phänomen festgestellt. Im „Diccionario de sinónimos y antóni-

mos Espasa-Calpe” (2005) stehen folgende Synonyme: „regocijo, complacencia, gozo, agrado, contento, placer, pago, indemnización, reparación, recompensa, solución, devolución, resarcimiento“. Nach einer erweiterten Suche konnten jedoch diese Synonyme nicht in Verbindung mit *consumidor/cliente* gefunden werden, aber sehr wohl in Verbindung mit *satisfacción* (*satisfacción del cliente* 79 Mal und *satisfacción del consumidor* 61 Mal). Für die spanische Sprache wurden auch die Entsprechungen im Plural gefunden (*satisfacción de los consumidores* 3 Mal und *satisfacción de los clientes* 15 Mal sowie *lealtad de los consumidores* 3 Mal und *lealtad de los clientes* 12 Mal).

Tab. 6: Kollokationskandidaten der Wörter *cliente* und *consumidor*

Sprache	Kollokator	Basis	Kollokation	H.	Struktur	
Spanisch	<i>lealtad</i>	<i>cliente</i>	<i>lealtad del cliente</i>	17	Subs. + de + Art. (Pl.) + Subs. (Pl.)	
			<i>lealtad de los clientes</i>	12		
		<i>consumidor</i>	<i>cliente</i>	<i>lealtad del consumidor</i>	31	Subs. + de + Art. (Pl.) + Subs. (Pl.)
				<i>lealtad de los consumidores</i>	3	
	<i>compromiso</i>	<i>cliente</i>	<i>compromiso del cliente</i>	1	Subs. + de + Art. (Pl.) + Subs. (Pl.)	
			<i>compromiso de los clientes</i>	1		
		<i>consumidor</i>	<i>cliente</i>	<i>compromiso del consumidor</i>	4	Subs. + de + Art. (Pl.) + Subs. (Pl.)
				<i>compromiso de los consumidores</i>	1	
	<i>satisfacción</i>	<i>cliente</i>	<i>satisfacción del cliente</i>	79	Subs. + de + Art. (Pl.) + Subs. (Pl.)	
			<i>satisfacción de los clientes</i>	15		
		<i>consumidor</i>	<i>cliente</i>	<i>satisfacción del consumidor</i>	61	Subs. + de + Art. (Pl.) + Subs. (Pl.)
				<i>satisfacción de los consumidores</i>	15	

Diese Korpusrecherche zeigt, dass die deutsche Entsprechung für *customer loyalty* auf Deutsch *Kundenbindung* ist. Sowohl für den Übersetzer als auch für den Fremdsprachenlerner bedeutet so ein Fall eine regelrechte Herausforderung, denn nach gesundem Menschenverstand würde man *customer loyalty* mit *Kundenloyalität* übersetzen und nicht mit *Kundenbindung*, weil dies semantisch eher der englischen Bezeichnung *customer commitment* entspricht.

Sowohl für Englisch als auch für Deutsch gilt, dass die Bezeichnungen *customer/Kunden* gegenüber den Bezeichnungen *consumer/Konsument* bevorzugt werden, was höchstwahrscheinlich mit den semantischen Prosodien und einer negativen Konnotation der Wörter zusammenhängt, wie auch im Fall *Konsumgesellschaft*. Anders als auf Deutsch und Englisch wird im spanischen Korpus die Bezeichnung *consumidor* genauso oft wie die Bezeichnung *cliente* verwendet, was darauf hinweist, dass sich die semantischen Prosodien dieser Wörter auf Spanisch auf einer Ebene befinden.

Ähnlich wie auf Englisch wurden wenige Beispiele mit der Variante *compromiso* (*commitment* auf Englisch) gefunden. Für den Übersetzer bedeutet dies, dass die spanische Entspre-

chung für das Wort *Kundenbindung* nicht wie erwartet *compromiso de los clientes/del cliente* ist, sondern *lealtad del cliente/de los clientes* oder *lealtad del consumidor/de los consumidores*.

Auf syntaktischer Ebene erschließt sich aus dieser vergleichenden Analyse, dass die Kollokationen in einer bestimmten Sprache nicht immer als solche in anderen Sprachen zu finden sind. Verschiedene Sprachen bedienen sich also verschiedener Typen von Wortverbindungen, um dieselben Inhalte zu vermitteln. Während im untersuchten englischen Korpus Kollokationen des Typs „Substantiv + Substantiv“ ihre Entsprechungen im deutschen Korpus in Form von Komposita oder Possessivmarkierungen finden, erscheinen sie im spanischen Korpus als Präpositionalphrasen.

Phase 5: Prüfung der Wortkookkurrenz auf ihre Idiomatizität und Festigkeit

In Bezug auf die Idiomatizität der Wortverbindungen *consumer loyalty*, *customer loyalty*, *consumer satisfaction* und *customer satisfaction* gilt die Definition „Wortverbindung, deren Gesamtbedeutung sich aus den Einzelbedeutungen der Wörter ableiten lässt“. Sie sind somit keine Idiome, aber potenzielle Kollokationskandidaten.

Die Festigkeit oder Stabilität ist in diesem Fall durch den Gebrauch etabliert, der sich wiederum aus den Werten der Assoziationsmaße deuten lässt. Die Tatsache, dass die Kookkurrenzen von *satisfaction* und *loyalty* mit *customer* beide einen *t-score*-Wert von über 10 aufweisen, aber ihre *t-score*-Werte in Verbindung mit *consumer* viel niedriger sind, kann so verstanden werden, dass *satisfaction* und *loyalty* nur in Verbindung mit *customer* Kollokationen bilden, aber nicht in Verbindung mit *consumer*.

9 Schlussfolgerungen

Die Auseinandersetzung mit der Frage, wann bei einer Wortkookkurrenz von einer Kollokation die Rede sein kann, und die Analyse verschiedener Wortkookkurrenzen hat zu verschiedenen Schlussfolgerungen geführt. Es kann, zum Beispiel, nur von einer Kollokation gesprochen werden, wenn man mehrere Optionen zur Auswahl hat. Somit könnte man sagen, dass eine Kollokation nur eine solche ist, wenn dem Sprecher mehrere Wörter zur Verfügung stehen, die zwar semantisch passend für das Zusammenscheinen mit einem anderem Wort sind, aber nur eine dieser Möglichkeiten tatsächlich verwendet wird. Ein weiterer wichtiger Schritt besteht darin, die Wortkookkurrenz auf ihre Idiomatizität und Festigkeit zu prüfen. Wenn die Wortkookkurrenz kein Idiom ist, aber eine pragmatische Festigkeit aufweist und an Festlegungen des Gebrauchs gebunden ist, ist die zweite Bedingung auch erfüllt. Diese zwei Schritte werden an Wortkookkurrenzen, die von der Software nach Anwendung von Assoziationsmaßen und Kollokationsmustern vorgeschlagen wurden, durchgeführt. Werden all diese Bedingungen erfüllt, kann von einer Kollokation die Rede sein. Ist eine dieser Bedingungen nicht erfüllt, muss von einer „schwachen Kollokation“ die Rede sein.

Die Vergleiche der drei Sprachen Englisch, Deutsch und Spanisch hat zur Schlussfolgerung geführt, dass die Kollokationen in einer Sprache meistens durch andere Formen und Typen von Wortverbindungen zum Ausdruck gebracht werden. In den hier analysierten Fällen finden englische Kollokationen ihre Entsprechungen auf Deutsch in Form von Komposita oder Possessivmarkierungen und auf Spanisch in Form von Präpositionalphrasen – sei es mit oder ohne Artikel, im Plural oder Singular.

Unter den zahlreichen Anwendungsbereichen dieser Studie sind die Übersetzerausbildung und die Fremdsprachendidaktik besonders hervorzuheben. Für Übersetzer, besonders

wenn sie in die Fremdsprache übersetzen, stellen Kollokationen eine häufige Fehlerquelle dar, umso mehr angesichts „falscher Freunde“. Für den Fall *customer loyalty* wäre die nächstliegende Übersetzung ins Deutsche *Kundenloyalität*. Doch nach einer ausführlichen Recherche im Korpus haben wir festgestellt, dass die Bezeichnung *Kundenbindung* bevorzugt wird. Diese Erkenntnisse sind für Übersetzer von großer Bedeutung, besonders wenn sie in die Fremdsprache übersetzen. Die durchgeführte Korpusrecherche hat auch gezeigt, dass, während auf Spanisch die Bezeichnungen *consumidor* und *cliente* in gleicher Weise verwendet werden, auf Deutsch und Englisch *customer* bzw. *Kunde* bevorzugt werden. Dies hat mit soziolinguistischen und pragmatischen Faktoren zu tun – die Wörter mit der Wurzel *Konsum-* haben auf Englisch und Deutsch eine negative semantische Prosodie oder Konnotation, die auf Spanisch nicht zu finden ist. Für den Übersetzer hat dies auch eine klare Konsequenz: *consumidor* sollte nicht als *Konsument* oder *consumer* übersetzt werden.

Für die Fremdsprachendidaktik sind diese Erkenntnisse gleichermaßen wertvoll. Fremdsprachenlehrer haben eine Grundlage, um den Studierenden zu erklären, dass im Gebrauch einer Sprache nicht nur grammatikalische Aspekte wichtig sind, sondern dass es auch Gebrauchsformen gibt, die weder mit grammatikalischen noch mit semantischen Regeln erklärt werden können. Aspekte, die mit dem Gebrauch einer Sprache zu tun haben, pragmatischer Natur sind und nur in Verbindung mit der Kultur und den Werten einer Gesellschaft gelehrt und gelernt werden können.

Literatur

- Arntz, Reiner/Picht, Heribert/Schmitz, Klaus-D. (2014): *Einführung in die Terminologiearbeit*. 7. überarb. Aufl. Hildesheim: Olms.
- Barnbrook, Geoff (1996): *Language and Computers: A Practical Introduction to the Computer Analysis of Language*. Edinburgh: Edinburgh University Press.
- Bartsch, Sabine (2004): *Structural and Functional Properties of Collocations in English. A Corpus Study of Lexical and Pragmatic Constraints on Lexical Co-occurrence*. Tübingen: Narr.
- Bongard, Joachim (2000): *Werbewirkungsforschung. Grundlagen – Probleme – Ansätze*. Münster: LIT Verlag.
- Burger, Harald (2003): *Phraseologie: Eine Einführung am Beispiel des Deutschen*. 2. überarb. Aufl. Berlin: Schmidt.
- Caro Cedillo, Ana (2004): *Fachsprachliche Kollokationen. Ein übersetzungsorientiertes Datenbankmodell Deutsch-Spanisch*. Tübingen: Narr.
- Evert, Stefan (2005): *The Statistics of Word Cooccurrences: Word Pairs and Collocations*. Dissertation, Institut für maschinelle Sprachverarbeitung, Universität Stuttgart.
- Evert, Stefan (2009): "Corpora and collocations." *Corpus Linguistics. An International Handbook*. Volume 2. (Handbooks of Linguistics and Communication Science/HSK 29.2). Ed. Anke Lüdeling/Merja Kytö. Berlin: Mouton de Gruyter. 1212–1248.
- Feix, Nereu (1982): „Die Terminologie der Wirtschaftswissenschaften (Volkswirtschaft) im Sprachvergleich: Spanisch und Deutsch.“ *Fachsprachenforschung und -lehre: Schwerpunkt Spanisch*. Hrsg. José Rodríguez Richart. Tübingen: Narr. 81–87.
- Firth, John Rupert (1957): "A Synopsis of Linguistic Theory 1930–55." *Studies in Linguistic Analysis*. Oxford: The Philological Society. 1–32.
- Fluck, Hans-R. (1996): *Fachsprachen: Einführung und Bibliographie*. 5. Aufl. München: Francke.
- Häcki Buhofer, Annelies (1999): „Phraseologismen im Spracherwerb.“ *Wortbildung und Phraseologie*. Hrsg. Rainer Wimmer/Franz Josef Berens. Tübingen: Narr. 209–232.

- Hausmann, Franz Josef (1984): „Wortschatzlernen ist Kollokationslernen. Zum Lehren und Lernen französischer Wortverbindungen.“ *Praxis des neusprachlichen Unterrichts* 31: 395–406.
- Hausmann, Franz Josef (1989): „Praktische Einführung in den Gebrauch des Student’s Dictionary of Collocations.“ *Student’s Dictionary of Collocations*. Hrsg. Evelyn Benson/Morton Benson/Robert Ilson. Berlin: Cornelsen. iv–xviii.
- Heid, Ulrich (1994): „On Ways Words Work Together: Topics in Lexical combinatorics.“ *Euralex 1994: Proceedings. Papers submitted to the 6th EURALEX International Congress on Lexicography in Amsterdam, The Netherlands*. Hrsg. Willy Martin et al. Amsterdam: Euralex. 226–257.
- Heid, Ulrich (1998): „Towards a corpus-based dictionary of German noun-verb collocations.“ *Proceedings of the 8th Euralex International Congress*. Hrsg. Thierry Fontenelle et al. Liège: University of Liège. 301–312.
- Hoffmann, Lothar (1985): *Kommunikationsmittel Fachsprache. Eine Einführung*. Erschienen im Forum für Fachsprachen-Forschung. Tübingen: Narr.
- Hundt, Markus (1995): *Modellbildung in der Wirtschaftssprache: zur Geschichte der Institutionen- und Theorie-fachsprachen der Wirtschaft*. Tübingen: Niemeyer.
- Hunston, Susan (2009): „Collection Strategies and Design Decisions.“ *Corpus Linguistics. An International Handbook*. Volume 2. (Handbooks of Linguistics and Communication Science/HSK 29.2). Ed. Anke Lüdeling/Merja Kytö. Berlin: Mouton de Gruyter. 154–168.
- Kilgarriff, Adam/Rundell, Michael (2002): „Lexical Profiling Software and its Lexicographic Applications – a Case Study.“ *Proceedings of the Tenth EURALEX International Congress*. Hrsg. Anna Braasch/Claus Povlsen. Copenhagen, Denmark, August 13–17, 2002. Copenhagen: Københavns Universitet. 807–818.
- Konovalova, Anastasia/Ruiz Yepes, Guadalupe (2016): „Die Sprache des Marketings und ihre Übersetzung: Morphologische und semantische Aspekte der Terminologie.“ *MonTI: Monografías de Traducción e Interpretación*: 95–124.
- Krenn, Brigitte (2000): *The Usual Suspects: Data-Oriented Models for the Identification and Representation of Lexical Collocations*. PhD Thesis, DFKI & Universität des Saarlandes, Saarbrücken.
- Meffert, Heribert/Burmann, Christoph/Kirchgeorg, Manfred (2015): *Marketing. Grundlagen marktorientierter Unternehmensführung. Konzepte – Instrumente – Praxisbeispiele*. Heidelberg: Springer.
- Mel’čuk, Igor (1974): *Towards a Theory of Meaning. Text Linguistics Models*. Moskau: Nauka.
- Pamies, Antonio/Pazos, José Manuel (2005): „Extracción automática de colocaciones y modismos.“ *La creatividad del lenguaje: colocaciones idiomáticas y fraseología*. Hrsg. Juan de Dios Luque Durán/Antonio Pamies Bertrán. Granada: Granada Lingvística. 317–329.
- Picht, Heribert (1987): „Fachsprachliche Phraseologie. Die terminologische Funktion von Verben.“ *Terminology and Knowledge Engineering*. Hrsg. Hans Czap/Christian Galinski. Frankfurt: IN-DEKS. 21–34.
- Picht, Heribert (1988): „Fachsprachliche Phraseologie.“ *Textlinguistik und Fachsprache: Akten des Internationalen übersetzungswissenschaftlichen AILA-Symposiums, Hildesheim, 13.–16. April 1987*. (Studien zu Sprache und Technik 1). Hrsg. Reiner Arntz. Hildesheim et al.: Olms. 187–196.
- Picht, Heribert (1993): „Fachsprachliche Phraseologie.“ *Ausgewählte Texte zur Terminologie*. Hrsg. Christer Lauren/Heribert Picht. Wien: Term-Net/International Network for Terminology. 439–456.
- Reder, Anna (2006): *Kollokationen in der Wortschatzarbeit*. Wien: Praesens.
- Rothkegel, Anneli (1994): „Kollokationsbildung und Textbildung.“ *EUOPHRAS 92 Tendenzen der Phraseologieforschung*. Hrsg. Barbara Sandig. Bochum: Brockmeyer. 499–523.
- Smadja, Frank (1993): „Retrieving collocations from text: Xtract.“ *Computational Linguistics* 19.1, 143–177.
- Stubbs, Michael (1995): „Collocations and semantic profiles: on the cause of the trouble with quantitative Studies.“ *Functions of Language*. Vol. 2:1. Amsterdam: Benjamins. 1–33.
- Weller, Marion/Heid, Ulrich (2010): „Corpus-derived data on German multiword expressions for lexicography.“ *Proceedings of the Euralex International Congress 2010*. Leeuwarden: Euralex [CDROM].

Online-Quellen:

www.collinsdictionary.com/dictionary/english-thesaurus

www.duden.de

www.duden.de/hilfe/synonyme

www.rae.es

www.wordreference.com/sinonimos

Prof. Dr. Guadalupe Ruiz Yepes

Hochschule Heilbronn

Fakultät für International Business

Angewandte Spanische Studien für Wirtschaft und Tourismus

Tel.: 07131 504 6765

guadalupe.ruiz-yepes@hs-heilbronn.de