

R

Register, das

Anke Lüdeling

lat. *regesta*, von *regerere* ‚eintragen‘

Register hat viele Lesarten. Hier geht es nicht um *Sündenregister*, *Ober-tonregister*, *Schnellfinderegister* oder *Transmissionsregister*, sondern um *Register* als funktionale Varietät in der Sprache. Die Grundidee ist einfach: Sprache kann zu unterschiedlichen Zwecken (Funktionen) und in verschiedenen Situationen verwendet werden und jede Funktion/Situation beeinflusst den Sprachgebrauch auf allen linguistischen Ebenen. Die meisten Unterschiede sind quantitativ – ein bestimmtes Register gebraucht ein Wort oder eine Wortartsequenz oder einen Satztyp häufiger als ein anderes Register. Registerstudien vergleichen also linguistische Kategorien in gesprochenen und geschriebenen Texten (Korpora) miteinander und versuchen dann, die ermittelten Unterschiede in einem Modell funktional zu erklären.

Es ist immer günstig, einen Begriff anhand eines klaren Beispiels zu erläutern. Am besten für eine Registeruntersuchung wäre daher ein Korpus aus Äußerungen mit einem sehr klar definierten Zweck in einer gut definierten Situation. Ich habe hierfür das in anderen Studien bisher zu Unrecht etwas vernachlässigte ‚KaTiKo‘-Korpus verwendet, auf dessen Metadaten wir weiter unten eingehen. Um die Charakteristika der in KaTiKo vertretenen Varietät herauszuarbeiten, vergleichen wir dieses Korpus mit fünf anderen Varietäten, die in dem gesprochenen ‚Berlin Map Task Corpus‘ (BeMaTaC), dem scripted-speech-Korpus ‚Parlamentsreden‘ und drei geschriebenen Korpora, nämlich dem Lernerkorpus ‚Falko‘, dem historischen Briefkorpus ‚Fürstinnenkorrespondenz‘ und dem historischen Kräuterkorpus ‚Ridges‘ vertreten sind. Alle Korpora sind annotiert mit Wortarten und Lemmata. Zunächst wollen wir die Wörter und dann die Wortarten betrachten.

Welche Wörter werden in den unterschiedlichen Registern verwendet? Tabelle 1 listet die häufigsten Lemmata in allen Korpora (*d* steht für alle definiten Artikel). Die zehn häufigsten Lemmata in KaTiKo (erste Spalte) kommen auch in der Liste der häufigsten Lemmata der anderen Korpora vor (fett markiert sind diejenigen Wörter, die in mindestens drei der anderen Listen auch auftauchen). Aber schon ab Position fünf finden sich bei KaTiKo Lemmata, die in den anderen Korpora nicht in den Top

Ten zu finden sind (hier kursiv markiert). Das KaTiKo unterscheidet sich also lexikalisch von den anderen Varietäten.

KaTiKo	Fürstinnenkorrespondenz	Ridges	Falko	Parlamentsreden	BeMaTaC
d	d	d	d	d	ja
in	ich	und	sein	und	ich
und	und	sein	und	sein	d
ein	euer	ein	ein	ein	Okay
<i>von</i>	sein	in	haben	in	sein
zu	Gnaden	mit	nicht	wir	ein
<i>deutsch</i>	zu	zu	zu	haben	haben
<i>Untersuchung</i>	haben	es	in	werden	rechts
<i>an</i>	es	werden	es	ich	also
<i>Beispiel</i>	Liebden	von	man	zu	zu

Tab. 1: Die jeweils zehn häufigsten Lemmata in den verschiedenen Korpora.

Der Eindruck, dass KaTiKo eine abweichende Varietät repräsentiert, verstärkt sich, wenn man häufige Bigramme und die häufigsten pränominalen Wörter betrachtet. KaTiKo verhält sich auch hier anders als die übrigen Korpora. Fast alle Bigramme (Lemmata) enthalten ein Nomen. Vergleicht man diese Bigramme mit denen aus der Fürstinnenkorrespondenz und aus BeMaTaC, stellt man fest, dass die Fürstinnen viel freundlicher sind und die BeMaTaC-ProbandInnen viel persönlicher und deutlich dynamischer.

KaTiKo: *die deutsch, Beispiel die, in die, an Beispiel, deutsch Sprache, Untersuchung anhand, in Althochdeutsch, Untersuchung an, die Verwendung, in Frühneuhochdeutsch*

Fürstinnen: *euer Gnade, euer Liebden, d lieb, lieb Gott, haben ich, dass ich, und ich, ich euer, wollen ich, ich es*

BeMaTaC: *ich haben, an d, um d, ich sein, so ein, d Motorrad, ich gehen, über d, auf d, alle klar*

Vergleichen wir als Nächstes die Wortartverteilungen. Auch hier ist KaTiKo ungewöhnlich: Es gibt, im Vergleich zu den anderen Korpora, erstaunlich viele Nomina und nur eine verschwindend geringe Anzahl von Verben.

Wir beginnen mit einem Blick auf die Verben, denn, wie schon Schiller 1805 an Goethe schrieb (Brief 998 im Briefwechsel Goethe/Schiller), „soll das Register nur eine empirische Aufzählung der Prädikate sein.“ In allen bisher in der Literatur beschriebenen Varietäten (Zeitungskorpora, Lernertexte, Webkorpora, Konversationen etc.) weisen die Verben eine

Register, das

Zipfsche Verteilung auf. Ein Beispiel dafür ist in Abb. 1 der Anfang der Verteilung der Verben in den Parlamentsreden. Die Verteilung der Verben im KaTiKo (Abb. 2) verblüfft: Es gibt nur ganze elf Verben, die alle jeweils einmal vorkommen. Für eine solche Verteilung gibt es bisher keine Theorie. Die Forschung steht vor einem Rätsel.

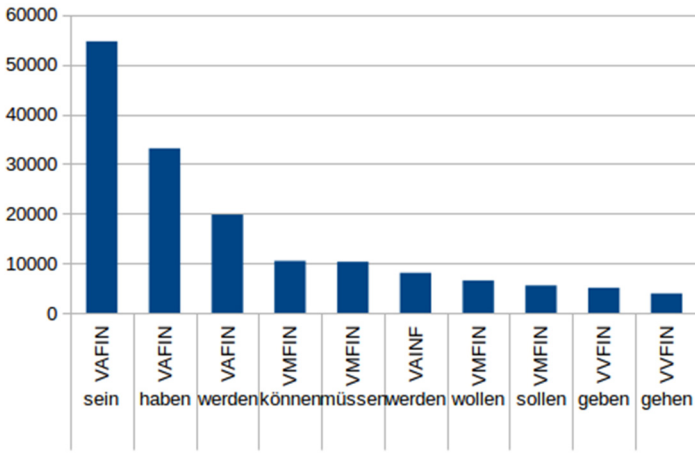


Abb. 1: Die häufigsten Verben im Parlamentsredenkorpus sind zipfisch verteilt.

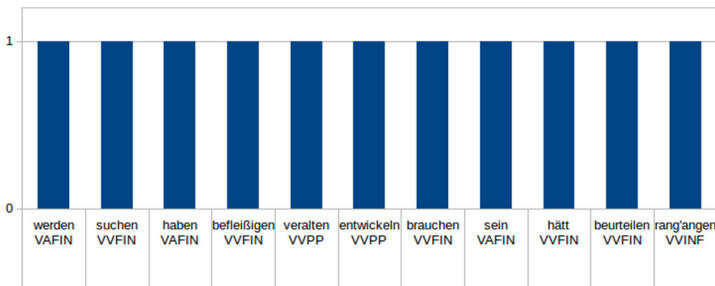


Abb. 2: Die Verben im KaTiKo sind ungewöhnlich verteilt.

Als nächstes betrachten wir die Adverbien: Die Listen der häufigsten Adverbien in den Vergleichskorpora sind einander nicht unähnlich. Aneinander gereiht lesen sie sich wie ein dadaistisches Gedicht:

*auch so aber wohl noch also nur sehr gar dann
auch noch so nur dann aber hier sehr jetzt schon
ja rechts also dann so jetzt links oben aber sozusagen da mal
auch so wohl noch ganz aber dann gar wieder gern
auch nur so aber mehr sehr immer noch dann viel*

Das KaTiKo unterscheidet sich auch in dieser Hinsicht von den anderen Korpora. Es enthält nur magere zwei Adverbien, die jeweils einmal vorkommen – diese würden bestenfalls einen Titel für das Adverbgedicht abgeben: *mal hier*.

Die Nomina, die in KaTiKo vorkommen, sind fast alle komplex und abstrakt; Konkreta sucht man fast vergeblich:

Untersuchung, Beispiel, Wörterbuch, Sprache, Verwendung, Jahrhundert, Vergleich, Analyse, Studie

Die KaTiKo-Varietät erscheint aufgrund der bisherigen Beobachtungen **unpersönlich, statisch, unpoetisch, nominal** und deutet insofern eine höchst literate Ausbaustufe an. Um welche merkwürdige und bisher offenbar nie beschriebene Varietät könnte es sich handeln?

Die Metadaten zu KaTiKo liefern hierfür eine Erklärung: KaTiKo ist das Korpus aus allen Titeln für Bachelor- und Masterarbeiten, die von Karin Donhauser betreut worden sind. Das Register ist durch folgende Situationsvariablen bestimmt:

Situation: Prüfung
Zweck: kurze Zusammenfassung eines komplexen Inhalts
Verhältnis zwischen Sprecher und Rezipientin: asymmetrisch, konzeptuell
und medial schriftlich

Wissenschaft beschreibt nicht bloß, sondern versucht auch Vorhersagen zu machen. Nicht anders verhält es sich im spezifischen Fall mit Registerstudien: Man versucht, von den vorliegenden Daten auf künftige zu extrapolieren, wie die Wissenschaftler in Abb. 3 illustrieren.

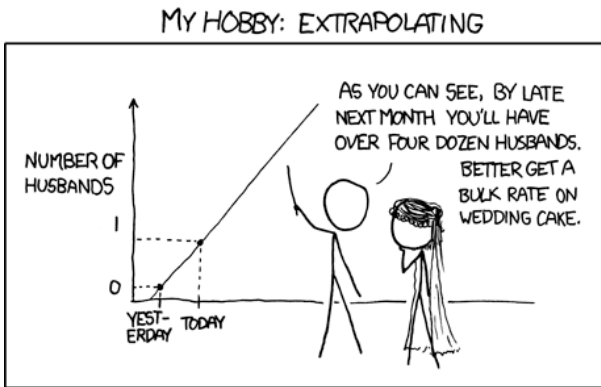


Abb. 3: Wissenschaftler bei der Arbeit (Quelle: xkcd.com/605).

Auf Basis der bisherigen in KaTiKo erfassten Titel werden die nächsten Arbeiten zu folgenden Themen geschrieben werden (für ‚VERB‘ können beliebige Verben eingesetzt werden, wobei lediglich darauf zu achten ist, dass sich niemals im Korpus ein Verb wiederholt):

*Ein Beispiel der Analyse der Verwendung des vorangestellten Wörterbuchs
Die Entwicklung von Schillers Partikelverben in Meister Eckhards Predigten
VERBT der Vergleich der teutschen Sprachgeschichte?
Ausgewählte am-Progressive im Althochdeutschen
Es VERBT hier mal historisch
Ein Jahrhundert von linken syntaktischen Entwicklungen*

Danksagung: Ich möchte mich bei Uta Kabelitz für die Extraktion des KaTiKo bedanken und bei Martin Klotz für die große Hilfe bei der Aufbereitung und Auswertung der Korpora.