**HA Sprachtypologie bei Prof. Cysouw**

Huan Wie, MA Linguistik: Kognition und Kommunikation

Dieses Repository ist eine empirische Arbeit im Rahmen des Seminars „Einführung in die Sprachtypologie“ bei Prof. Dr. Michael Cysouw an der Universität Marburg.

**Inhalten dieses Repository**

* Daten von Übersetzungsbogen (Ordner Umfrage)
* Sortieren die gehobenen Daten (Ordner: sources)
* Analysieren und Interpretieren die Daten in R (Ordner: Sprachtypologie)

**Ideen und Hintergrund**

Um den Zusammenhang zwischen der Wortfolge im Satz und Nominalphrase in der Sprache herauszufinden, habe ich einen Fragebogen mit 10-deutschen Sätzen an Informanten mit verschiedene Muttersprachen gegeben. Von Januar bis Ende Februar 2018 habe ich insgesamt 26 Fragebogen von Informanten zurückbekommen.

Die insgesamt erhobene 26 Sprachen sind: Deutsch, Englisch, Niederländisch, Französisch, Italienisch, Griechisch, Portugiesisch, Polnisch, Russisch, Ukrainisch, Rumänisch, Armenisch, Türkisch, Tigrinja, Hindi, Vietnamesisch, Indonesisch, Kroatisch, Georgisch, Berber, Persisch und Spanisch von Mexiko.

Für mein Analysieren sind nur folgende Satzteilen relevant:

* 1. Ich koche Suppe.
* 3. Musik von Mozart
* 4. diese fünf großen Häuser [DET + Num + Adj-Umfang + N]
* 5. einen schönen großen roten Ball [ ART + Adj + Adj-Umfang + Farbe + N]
* 9. kleine Hunde [Adj + N]
* 10. das feuchte Holz [DET + Adj + N]

Mit Satz 1 möchte ich die Wortfolgen im Satz von der Sprache festzustellen. Leider gibt es viele Sprachen, wenn der Subjekt im Satz „ich“ ist, fällt den Subjekt weg und wird mit der Endung von Verb markiert. Daher habe ich die Wortfolgen von der Sprache immer auf alle gefragte zehn Sätzen berücksichtigt. „Musik von Mozart“ sind zwei Nomen, daher analysiere ich nur die Rheinfolge von „Musik“ und „Mozart“.

# Zusammenhang zwischen den Wortfolgen und den Wortstellungen in Nominalphrasen von Sprachen der Welt

**1. Einleitung**

Insgesamt gibt es ca. 6500 bis 7000 Sprachen auf der Welt. Etwa die Hälfte der Erdbevölkerung sprechen eine der 10 meistgesprochenen Sprachen. Die sind Mandarin-Chinesisch, Englisch, Spanisch, Hindi, Arabisch, Portugiesisch, Bengali, Russisch, Japanisch und Deutsch.

Verschiedene Sprachen haben verschiedene Wortfolgenmöglichkeiten im Satz, zum Beispiel ist die Wortfolge auf Deutsch Subjekt-Verb-Objekt, wie im Satz *„Ich koche Suppe.“*, aber auf Japanisch ist die Wortfolge Subjekt-Objekt-Verb, wie im Satz:

*Watashi wa su-pu wo tsukuru*

Ich Suppe kochen

Verschiedene Sprachen haben auch verschiedene Wortstellungsmöglichkeiten in Nominalphrasen, zum Beispiel die Wortstellung von Adjektiven und Nomen in Nominalphrasen. In Deutschen ist es Adjektiv-Nomen, wie „das feuchte Holz“, aber in Französischen ist es Nomen-Adjektiv:

Le bois humide

das Holz feuchte

Jedoch kann man nicht sagen, dass die Sprachen auf der Welt keine Gemeinsamkeiten haben. Denn alle Sprachen haben Wörter und Sätze, Wörter werden aus elementaren bedeutungslosen Lauten oder Gesten gebildet, alle Sprachen haben komplexe morphologische und/oder syntaktische Strukturen, alle Sprachen haben Frageausdrücke und Verneinungsausdrücke usw.

Es lassen sich oft universale Behauptungen nur aufstellen, wenn eine weitere Bedingung erfüllt ist: Implikationsuniversalien. (vgl. W. Croft 2009)

„Wenn eine Sprache L die Eigenschaft A hat, dann hat sie auch die Eigenschaft B“

Beispielsweise „Wenn eine Sprache ein [f] Laut hat, dann hat sie auch ein [s] Laut; Wenn in einer Sprache das Objekt dem Verb vorangeht *„Der Hund die Katze jagt“*, dann steht auch der Besitzer vor dem Besitz *„des Hundes Hütte“*.

Nach dieser Theorie wird in dieser empirischen Arbeit den Zusammenhangen zwischen den Wortfolgen im Satz und den Wortstellungen in Nominalphrasen von den erhobenen 26 Sprachen anschauen, ob es nach dieser Theorie auch Ausnahme gibt oder die Zusammenhangen nach Statistik signifikant sind. Diese Arbeit wird mit Statistik „R“ geschrieben und alle Daten werden in „R“ eingelesen und ausgewertet. Außerdem werden Karten für die Übersichten von verschiedene Themen der erhobenen Sprachen dargestellt.

**2. Eigene empirische Arbeit**

Um den Zusammenhangen zwischen der Wortfolge im Satz und den Wortstellungen in Nominalphrasen in der Sprache herauszufinden, wurde einen Fragebogen mit 10-deutschen Sätzen an verschiedene Muttersprachler gegeben. Von Dezember 2017 bis Ende Februar 2018 habe ich insgesamt 26 Fragebogen von Informanten zurückbekommen. Mit diesen 26 Sprachen werden die Implikationsuniversalien genau angeschaut.

**2.1 Überlegungen zur Wahl der Informanten und Sprachen**

Die Überlegungen zur Wahl der Personen war, dass wir nur die Muttersprachler suchen. Denn die L2-Leaner sind meistens unsicher und brauchen Lehrbücher oder Wörterbuch, um den Übersetzungsbögen ins anderen Sprachen zu übersetzen. Aber Muttersprachler können meistens die passenden Ausdrücke in ihrer Muttersprache besser finden.

Natürlich ist es beim Erheben optimal, wenn man die Sprachen von vielen Sprachfamilien als Quellen bekommen kann, am besten auch ein paar isolierte Sprachen oder die Sprachen mit wenige Sprecher dabei könnten. Die insgesamt erhobene 26 Sprachen sind: Deutsch, Englisch, Niederländisch, Französisch, Italienisch, Griechisch, Portugiesisch, Polnisch, Russisch, Ukrainisch, Rumänisch, Armenisch, Türkisch, Tigrinja, Hindi, Vietnamesisch, Indonesisch, Kroatisch, Georgisch, Berber, Persisch und Spanisch von Mexiko.

Obwohl es nur 26 Sprachen sind, aber diese erhobenen Sprachen sind gut auf den Sprachfamilien geteilt. Es gibt nicht nur Indoeuropäische Sprachen wie Deutsch, Englisch, Niederländisch, Französisch, Spanisch, Italienisch, Griechisch, Portugiesisch, Polnisch, Russisch, Ukrainisch, Rumänisch, Hindi, Kroatisch und Persisch, sondern auch die isolierten Sprachen wie Koreanisch, Japanisch und Chinesisch. Außerdem sind die Altaische Sprachen wie Türkisch, Afroasiatische Sprachen wie Tigrinja und Berber, Sinotibetische Sprachen wie Kantonesisch, Austroasiatische Sprache wie Vietnamesisch, Austronesisch wie Indonesisch und Südkaukasische Sprachen wie Georgisch (vgl. Wals.info).

Im Folgenden werden die Daten von den erhobenen Sprachen in „R" eingelesen und schließlich wird eine Karte für eine Übersicht der erhobenen Sprachen mit diesen Daten dargestellt.

Die Analyse von den erhobenen Sprachen ist als „Umfragedaten.csv“ unter Ordner „sources“ gespeichert. Hier wird es mit „read.csv()“ in R eingelesen und als „Daten“ genannt. Mit der Funktion „summary()“ kann man die Strukturen von der Datei anschauen.

Im Bild 1 werden alle erhobenen Sprachen nach ihrer Räumlichkeit auf der Weltkarte dargestellt. Mit diesem Bild sieht man, dass die erhobene Sprachen gut auf der ganzen Welt geteilt sind. Die meisten Sprachen sind von Europa, aber asiatische und afrikanische Sprachen sind auch dabei.

In Folgenden Bild 2 werden diese erhobenen Sprachen nach der Sprachfamilie dargestellt. Mit dieser Darstellung kann man einen besseren Überblick über die zusammengehörige Sprachfamilie. Die erhobenen Sprachen sind gut auf den Sprachfamilien geteilt. Es gibt 16 Indoeuropäische Sprachen und zwei Afroasiatische Sprachen. Außerdem sind die Sprachfamilie Südkaukasische Sprachen, Altaische Sprachen, Austroasiatische Sprachen, Austronesische Sprache, Sinotibetische Sprachen, Chinesisch, Japanisch und Koreanisch auch dabei. Darüber hinaus kann man sagen, dass die erhobenen 26 Sprachen sehr vielfältig sind.

Um die Struktur von den Sprachen mit Farben darzustellen, habe ich zuerst mit der Funktion „table()“ die Struktur von der Sprachfamilie angeschaut. Sodass ich die richtige Reihfolge von den Sprachen mit passenden Farben einordnen konnte. Die Auswahl von Farben für die Karte ist schwierig. Eine Seite sollten die näh zusammengehörige Sprachfamilie eine ähnliche Farbe bekommen, andere Seite sollten die Ortspunkte von Sprachen gut auffällig sind.

**2.2 Übersetzungssätzen**

Dieser Übersetzungsbogen wurde von mir und zwei weiteren Studentinnen zusammengestellt.

Mit diesem Übersetzungsbogen möchten wir herausfinden, wie bestimmte Sätze in verschiedene Sprachen aufgebaut sind, wie Superlativ und Modalpartikeln. Daher enthält der Übersetzungsbogen folgende Zehn Sätze:

1. Ich koche Suppe.

2. Der Bäcker backt ja das Brot am schnellsten.

3. Die Schüler hören am liebsten Musik von Mozart.

4. Die Touristen sehen diese fünf großen Häuser.

5. Die Kinder werfen doch einen schönen großen roten Ball.

6. Was machst du denn?

7. Du hast eben nicht genug gelernt!

8. Tim wirft den Ball am weitesten.

9. Kleine Hunde bellen aber am lautesten.

10. Das feuchte Holz verbrennt am schlechtesten.

Ich interessiere mich über den Zusammenhang zwischen den Wortfolgen und den Wortstellungen in Nominalphrasen der Sprachen. Für mein Analysieren sind nur folgende Satzteilen relevant:

- 1. Ich koche Suppe. [Subjekt + Verb + Objekt]

- 3. Musik von Mozart [ Nomen + Nomen]

- 4. diese fünf großen Häuser [DET + Num + Adj-Umfang + N]

- 5. einen schönen großen roten Ball [ ART + Adj + Adj-Umfang + Farbe + N]

- 9. kleine Hunde [Adj + N]

- 10. das feuchte Holz [DET + Adj + N]

Mit Satz 1 möchte ich die Wortfolgen im Satz von der Sprache festzustellen. Leider gibt es viele Sprachen, wenn der Subjekt im Satz "ich" ist, fällt den Subjekt weg und wird mit der Endung von Verb markiert. Daher habe ich die Wortfolgen von der Sprache immer auf alle gefragte zehn Sätzen berücksichtigt. Mit Phrase „Musik von Mozart" wird die Genitiv-Nomen-Reihenfolge analysiert, damit wird die Reihenfolge für Genitiv-Nomen oder Nomen-Genitiv in der Sprache festgestellt. Mit Sätzen 4, 5, 9, 10 werden die Reihenfolgen von verschiedenen Elementen in der Nominalphrase der erhobenen Sprachen festgestellt.

**2.3 Methode und Ablauf**

Die meisten Informanten sind meine Arbeitskollegen, Freunde oder Kommilitonen. Durch persönliche Gespräche, E-Mails und Nachtrichten von Facebook habe ich sie als Informanten gewonnen. Die Übersetzungsbogen sind entweder in ausgedrückten Papierformen an Informanten persönlich gegeben, oder als Word- und PDF-Datei an Informanten per E-Mail erreichtet. Die Informanten haben durchschnittlich ca. eine Woche gebraucht, bis sie uns die Übersetzungen zurückgeben. Viele Sprachen, wie Kroatisch, Armenisch und Rumänisch, hatten die Informanten weiter ihre Freunde für uns gefragt. Zusammenfassende kann man sagen, dass Enkodierung der Informanten nicht sehr schwierig war. Denn sie sind sehr motiviert, ihre Muttersprache für andre Person zu erklären und die Unterschieden zwischen ihre Muttersprache und Deutsch zu erkennen.

Die Umfragebogen „Umfrage Studie Sprachtypologie“ ist unten Ordner „Umfrage“ zu finden. Außerdem sind 22 erhobene Umfragebogen auch in demselben Ordner zu finden.

**3. Ergebnisse der Umfrage**

Die Ergebnisse der Umfragesätze sind zuerst nach Sprache analysiert. Die Wortfolge in Sätzen und die Wortstellung im Nominalphrasen sind nach jeder Sprache sehr genau berücksichtigt. Die Ergebnisse sind als Excel-Datei „Umfrage Analyse“ unter Ordner sources zu finden.

Für die Analyse in R wird die Excel-Datei „Daten\_Umfrage“ genutzt. In dieser Analyse sind Sprache nach Land, Gattung und Familie zuerst geachtet. Außerdem sind Longitude und Latitude von der Sprache nach Wals.info für das Erstellen der Karte geachtet. Die Analyse von Daten sind außer Wortfolge unter sehr viele kleine Elementpaare geteilt, wie „Adjektiv + Nomen“, „Musik + Mozart“, „Artikel + Nomen“, „Farbe + Nomen“, „Numerale + Nomen“, „Artikel + Farbe“, „Artikel + Numerale“, „Artikel + Adjektiv“, „Farbe + Adjektiv“ und „Adjektiv + Adjektiv-Umfrang“.

Im Folgenden werden nicht nur die Verteilung der Wortfolge von Subjekt, Objekt und Verb der erhobenen Sprachen in Weltkarte dargestellt, sondern auch die Wortstellungen in Nominalphrasen und die Reihenfolge von verschiedenen Arten der Adjektive in Nominalphrase der erhobenen Sprachen.

**3.1 Die Wortfolge von Subjekt, Objekt und Verb der erhobenen Sprachen**

Für die Wortfolgen von Subjekt, Objekt und Verb werden nicht nur nach dem Satz 1 „Ich koche Suppe“ festgestellt, sondern auch alle weitere neun Sätze von Umfragebogen.

Mit der Funktion „table()“ können wir die Verteilung der Wortfolg von der erhobenen Sprachen sehen. In diesen Sprachen gibt es insgesamt sechs Sprache, die Subjekt-Objekt-Verb Wortfolge haben, und insgesamt 20 Sprache, die Subjekt-Verb-Objekt Wortfolge haben.

**3.2 Die Wortstellungen in Nominalphrasen der erhobenen Sprachen**

**3.2.1 Die Reihenfolge von Adjektiv und Nomen in Nominalphrase**

**3.2.2 Die Reihenfolge von Artikel und Nomen in Nominalphrase**

**3.2.3 Die Reihenfolge von Farbe und Nomen in Nominalphrase "roter Ball"**

**3.2.4 Die Reihenfolge von Numerale und Nomen in Nominalphrase "fünf Häuser"**

**3.2.5 Die Wortfolge von Genitiv und Nomen in Nominalphrase "Musik von Mozart"**

**3.3 Die Reihenfolge von verschiedenen Arten der Adjektive in Nominalphrase**

**3.3.1 Die Reihenfolge von Artikel und Farbe in Nominalphrase "einen roten Ball"**

**3.3.2 Die Reihenfolge von Artikel und Numerale in Nominalphrase "diese fünf Häuser"**

**3.3.3 Die Reihenfolge von Artikel und Adjektiv in Nominalphrase "das feuchte Holz"**

**3.3.4 Die Reihenfolge von Farbe und Adjektiv in Nominalphrase "schönen roten Ball"**

**3.3.5 Die Reihenfolge von Adjektiv und Umfangsadjektiv in Nominalphrase "schönen großen Ball"**

**4. Zusammenhang zwischen den Wortfolgen und den Wortstellungen in Nominalphrasen der erhobenen Sprachen**

**4.1 Zusammenhang zwischen den Wortfolgen und die Wortstellung des Adjektivs und Nomen**

**4.2 Zusammenhang zwischen den Wortfolgen und den Wortstellungen der Artikel und Nomen**

**4.3 Zusammenhang zwischen den Wortfolgen und den Wortstellungen der Farbe und Nomen**

**4.4 Zusammenhang zwischen den Wortfolgen und den Wortstellungen der Numerale und Nomen**

**Korrelation**:

Sprachen mit Genitiv-Nomen-Reihenfolge haben meistens Objekt- Verb-Reihenfolge, und Sprachen mit Nomen-Genitiv-Reihenfolge haben meistens Verb-Objekt-Reihenfolge

|  |  |  |
| --- | --- | --- |
| page34image1962239216page34image1962211344 | page34image1961445152  Genitiv-Nomen  page34image1961450304 | Nomen-Genitiv |
| Objekt-V erb | z.B. Türkisch  434 | (z.B. Persisch)  30 |
| page34image2022755024  V erb-Objekt | page34image2022775984  (z.B. Schwedisch)  113 | z.B. Französisch  352 |

**5. Allgemeine Diskussion und Ausblick**

Kritik:

In vielen Sprachen sind die Bedeutung von jedem Wort nicht klar markiert oder geschrieben, wie Persisch, Koreanisch. Man kann nur mit Wörterbuch und Google Übersetzer die Bedeutung vom einzelnen Wort raussuchen, aber meistens ist man unsicher. Daher kann man nur die Art von Wort feststellen, nicht die genaue Bedeutung und Feinheit.

Satz 1: Ich koche Suppe.

In Vielen Sprachen wird das Subjekt nicht eingegeben, wenn es "ich" ist. "Ich" als Subjekt wird in der Endung von Verb makiert.

**Literatur- und Quellenverzeichnis**

- Koordinate für die erhobenen Sprachen: http://wals.info/languoid (aufgerufen am 22.02.2018)

- Croft W. Typology and universals[M]. Cambridge University Press, 2002.

- M. Haspelmath: *Sprachen der Welt*. Max-Planck-Institut für evolutionäre Anthropologie (1999-2005) <http://home.uni-leipzig.de/muellerg/su/haspelmath.pdf> (aufgerufen am 14.03.2018)