|  |
| --- |
| **DeNpKongru – Ein Erkenner zur Bestimmung der nominalen Kongruenz**  **in deutschen Nominalphrase aus Lernerkorpora** |
| **Christopher Michael Chandler**  Straße  Postleitzahl, Stadt  christohper.chandler@rub.de |

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Eu mi bibendum neque egestas congue quisque egestas diam in. Ante in nibh mauris cursus mattis. Risus commodo viverra maecenas accumsan lacus vel facilisis. Sagittis id consectetur purus ut. Convallis tellus id interdum velit laoreet id. Eget felis eget nunc lobortis mattis. Tempus iaculis urna id volutpat lacus laoreet non. Imperdiet massa tincidunt nunc pulvinar sapien et. Sed ullamcorper morbi tincidunt ornare massa eget egestas. Vitae turpis massa sed elementum. Turpis egestas sed tempus urna et pharetra pharetra massa. Venenatis tellus in metus vulputate eu scelerisque. Venenatis urna cursus eget nunc scelerisque viverra. Sed ullamcorper morbi tincidunt ornare massa eget egestas. Vitae turpis massa sed elementum. Turpis egestas sed tempus urna et pharetra pharetra massa. Venenatis tellus in metus vulputate eu scelerisque. Venenatis urna cursus eget nunc scelerisque viverra.

Einleitung

Kongruenz als eine grammatikalische Notwendigkeit im Sinne von Satzkorrektheit ist in vielen Sprachen vertreten. Das gilt beispielsweise bei europäischen Sprachen (Englisch, Deutsch, Englisch, Isländisch , Polnisch, Russisch, Spanisch), semitischen Sprachen (Arabisch, Hebräisch), uto-akztekischen Sprachen (Hopi) und auch bei isolierten Sprachen (Baskisch) (Corbett, 2003; Hofherr, 2020). Es lässt sich also sagen, dass Kongruenz ein sprachübergreifendes Phänomen ist.

Kongruenz im Allgemeinen bezeichnet das regelhafte linguistische Phänomen, dass bestimmte linguistische Elemente, wie z.B. Verben, Nomen, etc. mit anderen linguistischen Elementen, wie z.B. Pronomen, Adjektiven, etc. innerhalb einer linguistischen Äußerung miteinander übereinstimmen müssen (Corbett, 2003; Hofherr, 2020). Um das näher definieren zu können, kann man also in diesem Fall dann von morphologischer Kongruenz sprechen.

Da Kongruenz ein so stark präsentes Phänomen ist, ist es von Bedeutung festzulegen, wie man sie am besten beschreiben und untersuchen kann, denn Kongruenz bereitet Nicht-Muttersprachlern Schwierigkeiten. Bei Sprachsystemen z.B. Semantik-, oder Syntaxerkennung ist Kongruenz wichtig, denn diese ist oft bedeutungsentscheidend.

Kongruenz kann verschiedene Ausprägungen annehmen und wie diese Ausprägungen ausfallen hängt direkt von der jeweiligen Sprache ab. Deswegen lässt sich nicht verallgemeinert sagen, wie genau die Kongruenz auszusehen hat und welche Formen es geben soll, dennoch gibt es ein paar Kongruenzarten, die hier aufgeführt werden, um diese zu veranschaulichen.

Um die Kongruenz und deren Merkmale klarer zu untersuchen wird als Objektsprache das Deutsche herangezogen.

Bei beispielsweise Subjekt-Verb-Kongruenz müssen Subjekt und Verb bezüglich Numerus und Kasus miteinander übereinstimmen. Genus spielt hierbei keine Rolle.

|  |  |
| --- | --- |
| Kongruenz | Nicht-Kongruenz |
| Ich gehe | Ich gehst\* |
| 1 SG.NOM gehen.pres1 SG | 2 SG.NOM  gehen.pres1 SG |

Tabelle

Diese Form von Kongruenz ist in einigen Sprachen (Französisch, Spanisch, Russisch) vorhanden (Hofherr, 2020). Somit dürfte es auch den meisten Sprechern dieser Sprachen vertraut vorkommen.

Eine andere Variante der Kongruenz ist die Kongruenz bei Nomen, wo Numerus, Genus und Kasus in einer linguistischen Phrase alle eine Rolle spielen. Diese wird infolge als nominale Kongruenz bezeichnet. Bei nominaler Kongruenz spielt der Kopf der Nominalphrase eine entscheidende Rolle, denn alle betroffenen linguistischen Elemente müssen mit dem morphologischen Kopf übereinstimmen (Hofherr, 2020).

|  |  |
| --- | --- |
| Kongruenz | Nicht-Kongruenz |
| Der kleine Hund | die Kleine Hund\* |
| 1 SG.NOM bestimmt.mask.sing Artikel + mask.sing.nomen + mask.sing.nominative adjective | 1 SG.NOM bestimmt.fem.sing Artiekl + fem.sing.nomen + fem.sing.nominative adjective |

Tabelle

Wenn ein Element in einer Nominalphrase fehlt oder nicht richtig eingesetzt wurde, ist die gesamte Äußerung falsch. Im Vergleich zu anderen Sprachen wie z.B. Englisch, Italienisch, Französisch, weil so ein Kasussystem, gilt die deutsche Sprache als bei Nicht-Muttersprachlern bzw. Lernenden als Schwierig (Spinner & Juffs, 2008).

Um die genauen Schwierigkeiten in Bezug auf nominale Kongruenz im Deutschen bestimmen zu können, werden Text von Anglophonen und Frankophonen aus dem Lernerkorpus Merlin (Wisniewski et al., 2018) extrahiert. Gewählt wurden diese beiden Gruppen unter der Annahme, dass Frankophonen am Anfang eher weniger Schwierigkeiten bei nominaler Kongruenz als Anglophone haben, da Kongruenz, auch wenn weniger komplex, im Französischen stärker präsent ist als im Englischen. Es wird jedoch erwartet, dass diese Differenz mit steigendem Niveau abnimmt.

Die Ergebnisse sollen Auskunft darüber liefern, inwieweit man Strategien oder Ansätze entwickeln kann, um Lernenden effektivere Lernmethoden bezüglich des Erwerbs der nominalen Kongruenz im Deutschen anbieten zu können.

Verwandte Literatur

Die Kernthematik bezieht sich hier auf morphologische Kongruenz, insbesondere Kongruenz bei Nominalphrasen. In folgenden Kapiteln wird ein regelbasiertes Vorgehen aufgestellt, jedoch gibt es schon LLTSM-basierte Ansätze

Nominale Kongruenz

Die Kongruenzart, die untersucht werden soll, ist die nominale Kongruenz. Das Ziel ist automatisch bestimmen zu können, ob eine Nominalphrase korrekt bzw. kongruent oder inkorrekt bzw. nicht-kongruent.

Eine Nominalphrase gilt als korrekt, wenn alle seine Elemente miteinander kongruieren bzw. übereinstimmen. Gleichfalls gilt eine Kongruenz nicht, wenn mindestens ein Element inkorrekt ist.

Termini

Um die Kongruenz genauer untersuchen und beschreiben zu können, müssen bestimmte Termini festgelegt werden:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Domäne | |  |  |  |  |
|  | Ziel |  | Kontroller |  |  |  |
|  | Das blaue |  | Haus |  |  | Bedingung |
|  |  |  |  |  |  |  |
|  | Merkmal:  Numerus, Genus, Kasus  Wert: Singular, Neutrum, Nominativ | | |  |  |  |

Tabelle Umriss der Termini (Corbett, 2003, p. 2)

Der Kontroller bestimmt die Kongruenz und das Ziel muss mit den Eigenschaften des Kontrollers übereinstimmen. Die Domäne ist die syntaktische Umgebung, worin die Nominalphrase sich befindet. Merkmale sind die Eigenschaften der Kongruenz und es gibt außerdomän-Bedinungen, die Auswirkungen auf die Kongrunez haben (Corbett, 2003; Hofherr, 2020).

Intention des Sprechers

Obwohl Zielhypothesen der einzeln Texte gemacht wurde, in dem korrigierte Version festgelegt wurde (Boyd et al., 2014), wird die Intention des Sprechers bei der Bestimmung der Nominalphrasenarten nicht berücksichtigt. Der Hauptgrund dafür ist, dass diese sich nicht so leicht automatisch bestimmen lässt und es könnte mehrere Möglichkeit als richtig gelten:

|  |  |
| --- | --- |
| Satzarten | Sätze |
| Lernersatz | Ich sehe die Hund\* |
| Zielhypothese 1 | Ich sehe den Hund |
| Zielhypothese 2 | Ich sehe die Hunde |

Tabelle

Wenn man beispielsweise annimmt, dass die Nominalphrase Ich sehe die Hund\* falsch ist, hat man zwei Möglichkeiten, den Satz zu korrigieren. Ohne Weiteres ist die Intention des Lerners unbekannt und somit wird sie hier nicht beachtet.

Der Satz wird als solcher behandelt.

Vorgehensweise

In diesem Projekt ging es hauptsächlich darum, die Kongruenz der einzelnen Nominalphrasen zu bestimmen, die aus bestimmten Lernerkorpustexten extrahiert wurden.

Nominalphrasen bestimmen zu können setzt voraus, dass der Eingangstext schon entsprechend tokenisiert und nach Wortklasse getaggt wurde. Diese Aufgabe wurde im Vorfeld von dem ParZu-Tagger übernommen, sodass die Texte im Rahmen des Projekts nicht geparst werden müssen.

Folgende Wortklassen nach dem STTS Schema sind für die Nominalphrasen relevant: NOUN, PREP, ART und ADJA. Ausgeschlossen wurden alle anderen Wortklassen.

Nach dem erfolgreichen Parsen wurden die daraus stammenden Conll-Dateien nach geeigneten Nominalphrasen untersucht. Geeignete bedeutet, dass mindestens eine Wortart in Äußerung unbestimmter Länge mindestens

Diese wurden extrahiert und in einer gesonderten Datei gespeichert.

Ein Nominalphrase-Eintrag hat dann die folgende Form:

|  |  |
| --- | --- |
| Nominalphrasenarten | Beispiel |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

Tabelle

In diesem Projekt ging es hauptsächlich darum, die Kongruenz der einzelnen Nominalphrasen zu bestimmen, die aus bestimmten Lernerkorpustexten extrahiert wurden.

Nominalphrasen bestimmen zu können setzt voraus, dass der Eingangstext schon entsprechend tokenisiert und nach Wortklasse getaggt wurde. Diese Aufgabe wurde von dem ParZu-Tagger übernommen.

Nach dem erfolgreichen Parsen wurden die die daraus stammenden Conll-Dateien nach Nominalphrasen untersucht. Diese wurden extrahiert und in einer gesonderten Datei gespeichert.

Ein Nominalphrase-Eintrag hat dann die folgende Form:

|  |  |
| --- | --- |
| **Nominalphrase** | meine Fahrkarte |
| **Morphologische Information** | meine ART Fem|Acc|Sg,  Fahrkarte N Fem|Acc|Sg |
| **Satz** | Ich gebe Dir meine Fahrkarte . |

Tabelle 1: morphologischer Eintrag

Nominalphrasen können verschiedene Formen annehmen und können somit Kategorien zugeordnet.

Die Nominalphrasen wurden erst in unterschiedlichen Kategorien eingeteilt. Innerhalb dieser Kategorien wurden einfache Heuristiken angewandt Um die Nominalphrasen auf Kongruenz untersuchen zu können,

Evaluation

Da es sich hier um ein regelbasiertes Model und nicht um ein statistisches Model handelt, musste das Programm an sich nicht trainiert werden. Jedoch wurden die Textdaten so aufgeteilt, dass 80 % als Trainigsdaten dienten, um die Algorithmen zu verbessern und die restlichen 20% als tatsächliche Testdaten.

Trainingsphase

Während der Trainingsphrase wurden die

|  |  |  |  |
| --- | --- | --- | --- |
|  | A1 | A2 |  |
|  |  |  |  |
|  |  |  |  |

Testphase

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Ergebnisse

Frankophonen und Anglophone wurden mit einander verglichen. Es hat sich herausgestellt, dass

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| French |  |  |  |
| English |  |  |  |

Diskussion

Im Allgemeinen lässt sich schwer sagen, ob so eine Methode geeignet ist, um Differenz

Fazit

Literaturverzeichnis

Boyd, A., Hana, J., Nicolas, L., Meurers, D., Wisniewski, K., Abel, A., Schöne, K., Štindlová, B., & Vettori, C. (2014). The MERLIN corpus : Learner language and the CEFR. *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC’14)*, 1281‑1288. http://www.lrec-conf.org/proceedings/lrec2014/pdf/606\_Paper.pdf

Corbett, G. (2003). Agreement : Terms and boundaries. *SMG conference papers*.

Długosz, K. (2022). Processing gender agreement in an additional language : The more languages the better? *Second Language Research*, 02676583221113333. https://doi.org/10.1177/02676583221113333

Gonen, H., Kementchedjhieva, Y., & Goldberg, Y. (2019). How Does Grammatical Gender Affect Noun Representations in Gender-Marking Languages? *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*, 463‑471. https://doi.org/10.18653/v1/K19-1043

Hofherr, P. C. (2020). *Agreement Morphology*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.500

Ravfogel, S., Goldberg, Y., & Tyers, F. (2018). Can LSTM Learn to Capture Agreement ? The Case of Basque. *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, 98‑107. https://doi.org/10.18653/v1/W18-5412

Sá-Leite, A. R., Luna, K., Fraga, I., & Comesaña, M. (2020). The gender congruency effect across languages in bilinguals : A meta-analysis. *Psychonomic Bulletin & Review*, *27*(4), 677‑693. https://doi.org/10.3758/s13423-019-01702-w

Spinner, P., & Juffs, A. (2008). L2 grammatical gender in a complex morphological system : The case of German. *Iral-international Review of Applied Linguistics in Language Teaching - IRAL-INT REV APPL LINGUIST*, *46*, 315‑348. https://doi.org/10.1515/IRAL.2008.014

Wisniewski, K., Abel, A., Vodičková, K., Plassmann, S., Meurers, D., Woldt, C., Schöne, K., Blaschitz, V., Lyding, V., Nicolas, L., Vettori, C., Pečený, P., Hana, J., Čurdová, V., Štindlová, B., Klein, G., Lauppe, L., Boyd, A., Bykh, S., & Krivanek, J. (2018). *MERLIN Corpus | Resources for research and practice related to foreign language learning* (1.1) [jeu de données]. https://www.merlin-platform.eu/

(Boyd et al., 2014; Corbett, 2003; Długosz, 2022; Gonen et al., 2019; Hofherr, 2020; Ravfogel et al., 2018; Sá-Leite et al., 2020; Spinner & Juffs, 2008; Wisniewski et al., 2018)

Eigenständigkeitserklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Hilfsmittel verwendet habe. Ich erkläre weiterhin, dass ich alles gedanklich, inhaltlich oder wörtlich von anderen (z.B. aus Büchern, Zeitschriften, Zeitungen, Lexika, Internet usw.) Übernommene als solches kenntlich gemacht, d.h. die jeweilige Herkunft im Text oder in den Anmerkungen belegt habe. Dies gilt ggf. auch für Tabellen, Skizzen, Zeichnungen, bildliche Darstellungen usw.

Ich nehme zur Kenntnis, dass die nachgewiesene Unterlassung der Herkunftsangabe als versuchte Täuschung bzw. als Plagiat („geistiger Diebstahl“) gewertet wird. Unkenntnis der in der Wissenschaft gebräuchlichen Regeln gilt nicht als Entlastung.

Ich anerkenne hiermit, dass bei Vorliegen eines Plagiats die Arbeit nicht als selbständige Leistung gewertet wird mit der Folge, dass mein Anspruch auf einen Leistungsnachweis (auch Teilnahmenachweis) in der Lehrveranstaltung und die Möglichkeit einer Nachbesserung oder Wiederholung der Arbeit entfällt.

Christopher Michael Chandler

Ort, Datum

Danksagung

Dieses Projekt wurde im Rahmen des computerlinguistischen Kurses

*Korpuslinguistische Analysen der Nominalflexion im Deutschen (050041-SoSe23)*an der Ruhr-Universität Bochum in Deutschland entwickelt. Ideen und Code von den Mitkommilitonen und der Kursleiterin Frau Dr. Ronja Laarman-Quante wurden hier verwendet und waren bei der Planung und Ausführung des Projekts ausschlaggebend.