

Lunds universitet
Språk- och litteraturcentrum

Einfach Übersetzen

Eine Studie der schwedischen Verständlichkeit des
Deutschen

TYSK02 - Tyska: Kandidatkurs
C-Aufsatz
Eingereicht von: Hannes Essfors
Eingereicht am: 2/5 - 2022

Inhaltsverzeichnis

Inhaltsverzeichnis	2
1. Einleitung	3
2. Frühere Forschung	4
3. Methodik	5
3.1 Übersetzungsübung als quantitative Studie	6
3.1.1 Begründung einer Übersetzungsübung	6
3.1.2 Die Auswahl von Wörtern für die Übersetzungsübung	6
3.1.3 Geschlossene und Offene Wortarten	7
3.1.4 Ergänzung der Wortliste	7
3.1.5 Ausarbeiten des Fragebogens	8
3.2 Linguistische und extralinguistische Variable	9
3.2.1 Lexikalischer Abstand	9
3.2.2 Orthografischer Abstand	10
3.2.3 Exposition	12
3.3 Bewertung und Analyse der Daten	12
4. Ergebnisse und Analyse	14
4.1 Ergebnisse der Studie	14
4.2 Regressionsanalyse der Variablen	20
4.3 Analyse von schweren Wörtern	22
5. Diskussion	24
5.1 Fragestellung 1	24
5.2 Fragestellung 2	25
6. Zusammenfassung	26
Literaturverzeichnis	27
Anlagen	30
Anlage A	30
Anlage B	34

1. Einleitung

“Die Speisekarte? Jaha, du har en karta över spisen.”¹

“Eine Tasse Kaffee? Ja, jo, du har taskigt Kaffee ja.”²

“Das ist möglich! ja, dasset är mögligt jo.”³

Diese drei Sätze stammen aus dem Sketch *Språkligt*, geschrieben von dem schwedischen Komiker Nils “Nisse” Ahlroth. Er spielt in seinem Sketch einen „typischen Schweden“, der versucht bei einem Besuch im Nachbarland Deutschland, Abendessen und Zimmer für die Nacht zu bekommen. Dies führt aber nur zu Missverständnissen. Wegen schlechten Deutschkenntnissen kommuniziert Ahlroth vor allem mit seinem Schwedisch, während die Kellnerin Deutsch spricht. Wenn man auf diese Art erfolgreich kommunizieren kann, wird das in der Sprachwissenschaft *gegenseitige Verständlichkeit* genannt. Das ist vor allem den Schweden nicht fremd, denn es ist für sie eine recht erfolgreiche und häufig verwendete Strategie, um mit Dänisch- und Norwegischsprachigen zu kommunizieren (Delsing & Lundin, 2005). Wenn die Verwendung der Methode im Sketch versucht wird, scheitert dies, als der Charakter Ahlroths beispielsweise das Wort *Speisekarte* als eine Karte über dem Ofen (Sv: *spis*) interpretiert, wodurch es unterhaltsam für das Publikum wird. Dies ist aber ein konstruiertes Beispiel und kein Urteil darüber, ob es für Schwedischsprachige tatsächlich möglich wäre, durch gegenseitige Verständlichkeit in Deutschland zu kommunizieren. Dem Sketch zugrunde liegt eine korrekte Beobachtung, dass recht viele Ähnlichkeiten in den zwei Sprachen vorkommen, oder wie die emeritierte Professorin für Germanistik an der Universität Umeå Astrid Stedje es in ihrem Buch *Deutsche Sprache gestern und heute* (2007, S.136) zum Ausdruck bringt: “Weder früher noch später hat das Deutsche andere Sprachen so stark beeinflusst wie das Mittelniederdeutsche die nordischen Sprachen.” Nach wie vor gibt es einen großen Austausch zwischen den beiden Sprachen, denn Deutschland ist der wichtigste Handelspartner Schwedens (SCB, 2021) und Deutsch wird den Schülern in schwedischen Schulen sogar als eine von drei Fremdsprachen angeboten (Skolförordning, 2011:185, kap. 9 5§). Die Frage, die durch den Sketch von Nils Ahlroth beleuchtet wird, wie viel Deutsch Schwedischsprachige verstehen können und ob Kommunikation allein mithilfe gegenseitiger Verständlichkeit möglich wäre, ist deshalb als interessant zu betrachten, und damit wird sich diese Arbeit beschäftigen. Sie nimmt ihren Ausgangspunkt in der Durchführung einer quantitativen Studie, die der Beantwortung der folgenden Fragestellungen zugrunde liegen soll:

¹ Deutsche Übersetzung: Ach so, du hast eine Karte über dem Ofen.

² Deutsche Übersetzung: Ja, doch, du hast ja schlechten Kaffee.

³ Deutsche Übersetzung: Ja, das Klo ist verschimmelt.

- Wie gut können Schwedischsprachige Deutsch anhand ihrer Schwedischkenntnisse verstehen und wäre es möglich zu behaupten, dass gegenseitiges Verständnis zwischen Deutsch und Schwedisch vorliegt?
- Inwieweit spielen linguistische und extralinguistische Faktoren eine Rolle bei der Verständlichkeit zwischen den Sprachen und welche Faktoren sind von Bedeutung? Welche Rolle spielt die große Menge Lehnwörter aus dem Deutschen in das Schwedische?

Der Aufbau der Arbeit sieht wie folgt aus: Zuerst werden einige frühere wissenschaftliche Arbeiten zum Thema kurz präsentiert. Danach folgt die Methode der Arbeit, in der die Durchführung der Studie präsentiert wird, und die für die Analyse der Ergebnisse wichtigen Begriffe und Methoden erklärt. Dann werden die Ergebnisse der Studie präsentiert und analysiert, und letztlich wird eine Diskussion mit den Fragestellungen als Ausgangspunkt geführt.

2. Frühere Forschung

Die Forschung zum Thema Verständlichkeit hat ihren Ausgangspunkt in der Dialektologie. Der Grund dafür ist, dass es schwer ist, zwischen Sprache und Dialekt zu unterscheiden. Schon früh wurde behauptet, dass der Unterschied nur etwas rein Politisches war, was durch das sehr bekannte Zitat von Max Weinreich (1944) deutlich exemplifiziert werden kann: „A sprach is a dialekt mit an armej un flot“. Ein Vorschlag dafür, wie man die Begriffe linguistisch trennen kann, entstand mit der Einführung des *Verständlichkeitskriteriums*. Durch das Verständlichkeitskriterium wird eine Sprache als „eine Sammlung von gegenseitig verständlichen Dialekten“ definiert (Chambers & Trudgill, 1980, s.3). Wenn zwei Sprecher durch ihre Sprache nicht kommunizieren können, sprechen sie folglich unterschiedliche Sprachen. Um diese Definition von Sprachen zu verwenden, wäre es notwendig, Verständlichkeit zu bemessen, weswegen man begonnen hat, über Verständlichkeit zu forschen. Das Ziel wurde dann, und das war es in den letzten Jahrzehnten (Swarte, 2016, s.4), Variablen zu finden, die Verständlichkeit zwischen Sprachen voraussagen können, weil das zu einer objektiven Klassifizierung von Sprachen führen könnte. Einen genaueren Einblick in solche Variablen gibt es in 3.3 *Linguistische und extralinguistische Variable*.

Viel Verständlichkeitsforschung wurde in den skandinavischen Sprachen Dänisch, Norwegisch und Schwedisch durchgeführt, wodurch gezeigt wurde, dass Kommunikation durch gegenseitige Verständlichkeit zwischen den Sprechern üblich ist (Haugen, 1966; Maurud, 1976; Bø, 1978; Börestam Uhlmann, 1991; Delsing & Lundin, 2005; Gooskens, 2007; Gooskens et al. 2008; Küschner et al, 2008; Dotjes & Gooskens, 2009, Swarte 2016). Weniger Forschung gibt es aber über die Verständlichkeit zwischen Deutsch und Schwedisch. Es gibt die Arbeit von Swarte (2016) und die 2014 durchgeführte Arbeit von Vanhove, in der Vanhove untersucht, wie gut Deutschsprachige schwedische Kognaten übersetzen können. Seine Ergebnisse zeigen, dass Levenshtein Distanz am besten voraussagt, wie einfach ein Kognat zu übersetzen ist, und dass Wörterlänge kein Prädiktor ist. Eine wichtige Arbeit, die auch der in dieser Arbeit verwendeten Methode zugrunde liegt, ist die Studie von Verständlichkeit zwischen Niederländisch und Hoch- und Niederdeutsch von Gooskens et al. (2011).

Die Arbeit von Swarte (2016) ist eine der umfassendsten Arbeiten über gegenseitiges Verständnis zwischen germanischen Sprachen, die es bisher gibt. In ihrer Studie wird die Verständlichkeit zwischen Sprechern von Deutsch, Englisch, Niederländisch, Dänisch und Schwedisch mithilfe einer Übersetzungsübung, eines Bild-tests und eines Lückentests bemessen. Die 100 frequentesten Substantive der englischen Sprache lagen diesen drei Tests zugrunde, und zwischen Schwedischsprachigen, die in der Schule Deutsch gelernt und nicht gelernt hatten, wurde kein Unterschied gemacht. Ihre Ergebnisse werden in ihrer Arbeit mit linguistischen und extralinguistischen Variablen korreliert und ihre wichtigsten Schlussfolgerungen sind, dass die Exposition von einer Sprache die wichtigste Variable überhaupt ist, während Lexikalischer Abstand unter den linguistischen Variablen am wichtigsten ist.

3. Methodik

In diesem Teil der Arbeit wird die Methode beschrieben. Erst wird erklärt, wie die Daten eingesammelt wurden, die dem Ergebnis zugrunde liegen. Danach folgen eine Präsentation und eine Erklärung von den linguistischen Variablen, die im Verhältnis zu den eingesammelten Daten gesetzt werden und schließlich die für die Behandlung der Daten verwendeten statistischen Methoden.

3.1 Übersetzungsübung als quantitative Studie

In den folgenden Abschnitten werden die einzelnen Schritte meiner Studie präsentiert und diskutiert.

3.1.1 Begründung einer Übersetzungsübung

Um die Verständlichkeit zu bemessen, wurde eine Übersetzungsübung verwendet. Übersetzungsübungen kommen in der Verständlichkeitsforschung häufig vor und bemessen wie gut die Sprecher von Sprache X die Wörter von Sprache Y verstehen können (Swarte 2016, S.50). Die Ergebnisse der Übung gelten dann in weiterem Sinn als Indikator, wie verständlich Sprache Y für Sprecher von Sprache X sind. Eine Übersetzungsübung ist auch in der einzigen früheren Untersuchung von gegenseitiger deutsch-schwedischer Verständlichkeit verwendet worden (Swarte, 2016). In ihrer Studie gehen Swarte von 100 hochfrequenten englischen Substantiven aus und die Ergebnisse meiner Studie werden in *4.1 Ergebnisse der Studie* den Ergebnissen Swartes gegenübergestellt. Meine Arbeit kann deshalb zum Teil als eine Replikation davon betrachtet werden. Die Arbeit als eine Replikation zu betrachten ist besonders sinnvoll, weil behauptet wird, dass eine laufende Krise der Replikation in den Sozialwissenschaften vorliegt (Peels, 2019). Und deshalb, wenn ich ähnliche Ergebnisse wie Swarte erhalte und ähnliche Schlussfolgerungen ziehen kann, trägt das zur Erforschung der Verständlichkeit von Schwedisch und Deutsch bei.

3.1.2 Die Auswahl von Wörtern für die Übersetzungsübung

Weil die Übersetzungsübung für die Verständlichkeit so repräsentativ wie möglich sein soll, wurden in früheren Untersuchungen hochfrequente Wörter verwendet. Deshalb gehe ich in meiner Studie von den 100 hochfrequentesten Wörtern der deutschen Sprache aus. Die Begründung ist die, dass hochfrequente Wörter wichtiger sind. Beispielsweise machen die 30 häufigsten Wortformen in deutschen Texten 31,8 % aller Wörter aus, und schon die 100 häufigsten Wortformen machen 47,1 % aller Wörter aus (König, 2004). Die logische Implikation davon ist, dass jemand, der nur die 100 häufigsten Wörter verstehen kann, etwa die Hälfte eines zufällig gegebenen Textes verstehen kann, und folglich sollte eine Übersetzungsübung viel über Verständlichkeit aussagen können. Die in meiner Studie verwendeten Wörter stammen aus dem Häufigkeitswörterbuch von Wilhelm Kaeding (1898). In diesem Wörterbuch hat Kaeding 10 910 777 Wörter aus Texten des 19. Jh. eingesammelt und nach Häufigkeit geordnet, aus seinem Korpus wurden die 100 frequentesten Substantive, Verben, und Adjektive für die Übersetzungsübung ausgewählt. Der Grund dafür, dass ein altes Wörterbuch wie das von Kaeding verwendet wird, ist, dass es das einzige Häufigkeitswörterbuch ist, das in der Universitätsbibliothek zu finden war. Es wäre besser mit einem aktuelleren Buch, das

die heutige Sprache besser widerspiegelt. Die Stärke meiner Studie ist aber, dass die häufigsten *deutschen* Wörter verwendet werden, im Gegensatz zu Swarte (2016), die die häufigsten englischen Wörter für ihre Übersetzungsübung verwendet hat. Der Grund dafür ist, dass Swarte mehrere Sprachen, unter anderem das Englische, anhand derselben Wörter untersuchen wollte.

3.1.3 Geschlossene und Offene Wortarten

Unter den 100 häufigsten deutschen Wörtern kommen sowohl die *geschlossenen* Wortarten Determinanten, Pronomina, Konjunktionen und Präpositionen, aber auch die *offenen* Wortarten Substantive, Adjektive, Verben und Adverbien vor. Die Geschlossenen heißen so, weil sehr wenige Wörter von diesen Wortarten neugebildet werden; die Anzahl dieser Wörter ist ziemlich konstant im Gegensatz zu den Offenen. Obwohl die Anzahl an geschlossenen Wortarten geringer ist – sie machen nur 2,54 % aller Einzelwörter aus – kommen sie viel häufiger vor, denn sie machen 55,1 % aller Wortformen aus (König, 2004). Die geschlossenen Wortarten sind deshalb nicht unwichtig, aber in meiner Arbeit wird der Fokus auf die Offenen gelegt. Ich entschied mich dafür, weil frühere Verständlichkeitsforschung sich mit geschlossenen Wörtern beschäftigt hat, vor allem Substantive (bspw: Küschner et al, 2008, Dotjes & Gooskens, 2009, Gooskens et al, 2011, Swarte, 2016). Der Grund dafür sei, dass offene Wortarten reicher an semantischem Inhalt sind und deshalb für die Verständlichkeit wichtiger sind (Swarte, 2016, s. 117). Ein Unterschied liegt auch beim neuronalen Verarbeiten von offenen und geschlossenen Wortarten vor (Münste et al, 2001), deshalb habe ich mich dazu entschlossen, nur die offenen Wortarten in meiner Studie zu verwenden. In dieser Hinsicht ähnelt meine Studie früheren Arbeiten. Der Unterschied besteht darin, dass ich im Gegensatz zu bspw. Swarte nicht nur von Substantiven ausgehe, sondern auch von Adjektiven und Verben.

3.1.4 Ergänzung der Wortliste

Nach der Auswahl an Wörtern durch die oben erläuterten Gründe, enthält die Wortliste für die Übersetzungsübung die 34 häufigsten Adjektive, 33 häufigsten Substantive und 33 häufigsten Verben. Die ganze Liste, mit allen Wörtern, die in der Studie verwendet wird, ist in Anlage A zu finden. Vor der Erstellung der endgültigen Wortliste, wurden zwei Wörter aus der Liste gestrichen, weil sie veraltet sind. Die zwei Wörter sind “Korps” und “Preußischen”. Die Wörter in der Liste sind in der deklinierten Form, in der sie im Häufigkeitswörterbuch am häufigsten auftauchten und wurden auch so in der Übersetzungsübung den Probanden präsentiert. Das steht in Kontrast zu früheren Studien, in der die Wörter in nicht-deklinierten Form vorkommen (Swarte, 2016) (Van Hoven, 2014). Ich entschloss mich dagegen mit der Begründung, dass die Ergebnisse sich dadurch in einer Nicht-Testsituation besser generalisieren lassen. Was jedoch dadurch entsteht, ist, dass

einige Wörter in mehreren Fällen als verschiedene Wortarten interpretiert werden können, z.B. kommt das Wort *macht* vor, was sich als das Substantiv “Macht” in Singular und das Verb “machen” in dritter Person Präteritum gleichzeitig interpretieren lässt.

3.1.5 Ausarbeiten des Fragebogens

Die Wörter wurden zufällig aufgereiht und in einen Fragebogen eingetragen. Der verwendete Fragebogen wurde im Fragebogenwerkzeug *Sunet Survey* konstruiert. Er besteht aus zwei Teilen. Der erste Teil enthält Fragen über die Probanden wie Geschlecht, Ausbildung und gesprochene Sprachen, dies zeigt, dass sämtliche Probanden auch englischsprachig sind. Andere Fragen versuchen herauszufinden, wie viel Exposition die Probanden von der deutschen Sprache bereits haben. Dies wurde, durch Fragen wie beispielsweise, wie viele Male sie im deutschen Sprachraum schon waren, oder ob sie bereits deutsche Medien wie Filme oder Zeitungen konsumiert haben, festzustellen versucht. Zwei Fragen zielen darauf ab, die Analyse der Daten zu vereinfachen. Eine Frage ist, ob sie in der Schule Deutsch gelernt haben, sodass die Einteilung von den Probanden in “haben Deutsch gelernt” und “haben kein Deutsch gelernt” möglich ist. Diejenigen, die Deutsch fließend können, werden aus den Analysen ausgeschlossen, was insgesamt 8 Probanden betrifft. Der zweite Teil des Fragebogens besteht aus den hochfrequenten deutschen Wörtern und unter jedem Wort gibt es ein freies Feld, in dem der Proband seine Übersetzung schreiben kann. Die Wörter kommen alle in der gleichen Reihe in jedem Fragebogen vor. Es wäre besser, wenn die Reihenfolge der Wörter in jedem Fragebogen zufällig vorkommt. Es würde das Risiko minimieren, dass Wörter, die am Ende vom Fragebogen vorkommen, schwerer zu übersetzen sind. Am Ende des Fragebogens könnten die Probanden unkonzentriert sein, was zu verzerrten Ergebnissen führen könnte. Dies ist leider mit dem Fragebogenwerkzeug nicht möglich, diese mögliche Verzerrung muss im Hinterkopf behalten werden. Der Fragebogen wurde durch Facebook verbreitet, um effizient so viele Probanden wie mögliche zu finden, aber das heißt, dass nur die Personen, die Teil der digitalisierten Gesellschaft sind, die Möglichkeit haben, an der Studie teilzunehmen. Laut einer Studie der Internetstiftelsen verwenden 93 % der Schweden im Jahr 2020 das Internet zumindest ein Mal pro Tag, und wegen dieser hohen Zahl ist es für die Repräsentativität der Studie als unproblematisch zu betrachten (Internetstiftelsen, 2020). Der Fragebogen, wie er den Probanden vorgestellt wird, ist in Anlage B zu besichtigen.

Die Studie lief etwa zwei Wochen und insgesamt wurden Antworten von 83 Teilnehmern registriert. Nach der Einsammlung der Daten wurden die Antworten als richtig und falsch markiert, was in 3.3 *Bewertung und Analyse der Daten* genauer beschrieben wird.

3.2 Linguistische und extralinguistische Variable

Mit genügend Daten mithilfe des Fragebogens, müssen diese im Verhältnis damit gesetzt werden, was hier als *linguistische und extralinguistische Variable* benannt wird, um die Fragestellung zu beantworten und die Ergebnisse erklären zu können. Diese Variablen, und wie sie in der Arbeit behandelt werden, werden in folgenden Absätzen genauer beschrieben. Linguistische Variablen werden als Unterschiede zwischen Sprachen auf linguistischer Ebene wie beispielsweise Lexikon oder Phonologie definiert, während extralinguistische Variablen sich auf die Unterschiede zwischen den Sprechern der Sprachen beziehen (Swarte, 2016, s. 5).

3.2.1 Lexikalischer Abstand

Lexikalischer Abstand ist eine Variable, die sehr oft in der Verständlichkeitsforschung vorkommt. Sie nimmt ihren Ausgangspunkt in dem Kognatbegriff. Wie der Kognatbegriff eigentlich definiert werden soll, wäre eine eigene Arbeit angemessen, denn es gibt noch keinen Konsens dazu. Deshalb wird in dieser Arbeit dieselbe Definition verwendet, wie in den Arbeiten von beispielsweise Swarte (2016) und Vanhove (2014, S.12). Sie gehen davon aus, dass es ein Kognatverhältnis gibt, wenn zwei Wörter einen gemeinsamen Ursprung und dieselbe Bedeutung haben. Dann sind sie Kognaten. Ein Beispiel aus den in meiner Studie verwendeten Wörtern wären *Name* und die entsprechende schwedische Übersetzung *Namn*. Die Wörter gehen auf die germanische Wurzel **naman* (DWDS) zurück und haben in den beiden Sprachen heutzutage die gleiche Bedeutung, nämlich das Kennzeichnungsmittel eines bestimmten Gegenstands. Was hier jedoch nicht als ein Kognat zu betrachten ist, sind Wörter, die denselben Ursprung haben, aber aufgrund Bedeutungswandel nun unterschiedliche Bedeutungen haben. Aus den Wörtern meiner Studie wäre *schnell* ein Beispiel dafür. Im Schwedischen gibt es das Wort *Snäll*, das aus derselben Wurzel wie *schnell* stammt, und dieselbe Bedeutung hatte, nämlich „hohe Geschwindigkeit“. Jetzt haben sie aber unterschiedliche Bedeutungen. *Snäll* kann im Deutschen mit dem Wort *nett* oder *freundlich* beschrieben werden, während das schwedische Wort, das *schnell* entspricht, *snabb* heißt. Deshalb können *schnell* und *snäll* nicht als Kognaten betrachtet werden, auf jeden Fall nicht nach der in dieser Arbeit verwendeten Definition des Kognatbegriffs.

Kognaten sind sehr entscheidend für die Verständlichkeit anderer Sprachen, denn durch Kognaten kann man erraten, welche Bedeutung ein Wort hat, ohne sie erst als eine Vokabel zu lernen. Es wird deshalb vermutet, dass dieses Kognatverhältnis der gegenseitigen Verständlichkeit zugrunde liegt (Gooskens, 2017). Das heißt, je mehr Wörter in zwei Sprachen Kognaten sind, desto gegenseitig verständlicher sollten die Sprachen sein. Dadurch ist der Begriff *lexikalischer Abstand* entstanden, und in dieser Arbeit wird der lexikalische Abstand, als die Prozentzahl, in der Übersetzungsübung verwendeten Wörter, die nicht Kognaten sind, definiert. Um den lexikalischen Abstand zwischen den deutschen und schwedischen Sprachen zu bestimmen, wurden nach schwedischen Übersetzungen, die Ähnlichkeiten im Aussehen zu den deutschen Wörtern aufzeigen, in der Studie gesucht. Solche Wörter, die auch aus denselben Wurzeln stammen werden als Kognaten bezeichnet. Die etymologischen Bücher, die dafür verwendet wurden, sind *Etymologisches Wörterbuch des Deutschen* (Pfeifer, 1993), *Svensk etymologisk ordbok* (Hellquist, 1992), *Svenska Akademiens Ordbok* (1893) und das vom schwedischen Sprachforscher Staffan Fridell herausgegebene etymologische Wörterbuch *Ordklok* (2021). Die Anzahl der Kognaten wurde summiert und dadurch ergibt sich, dass die Wörter in der Übersetzungsübung einen lexikalischen Abstand von 20 % haben. Während diesem Prozess wurde auch nachgeschaut, ob die schwedischen Wörter aus dem Deutschen geliehen sind. Dadurch kann untersucht werden, ob Lehnwörter einfacher zu übersetzen sind als andere Kognaten. Wenn das der Fall ist, dann könnte es einen Grund dafür geben, die Definition von lexikalischem Abstand zu ändern, sodass der Anteil von Lehnwörtern auch berücksichtigt wird. In meiner Arbeit gibt es 23 Lehnwörter, von 80 Kognaten sind dann 29 % der Kognaten Lehnwörter. Die entsprechenden schwedischen Kognaten und welche davon Lehnwörter sind, werden in Anlage A beschrieben.

3.2.2 Orthografischer Abstand

Um zu berücksichtigen, welchen Einfluss an Verständlichkeit die Ähnlichkeiten zwischen Wörtern haben können, wird in meiner Arbeit eine Bemessung von Ähnlichkeiten durch Levenshtein Distanz verwendet. Levenshtein Distanz ist ursprünglich ein Algorithmus aus der Informatik, der die geringste Anzahl an Veränderung findet, um ein String in einen anderen String zu verwandeln. Die möglichen Operationen, die Veränderungen darstellen, sind Insertionen, Deletionen und Substitutionen (Black, 2019). Insertion heißt die Einfügung eines Charakters, Deletion die Entfernung und Substitution der Austausch. Ein generelles Beispiel für die drei Operationen ergibt sich durch die Verwandlung des Strings *Lache* in *Klack*:

Lache → Klache ; „K“ wird zugeführt = Insertion

Klache → Klacke ; „h“ wird mit „k“ ausgetauscht = Substitution

Klacke → Klack ; „e“ wird entfernt = Deletion

Die Summe von Operationen entspricht dann einem Abstand von „3“ zwischen den Wörtern. Der Algorithmus wurde zu Beginn vor allem in der Dialektologie verwendet, um Unterschiede in der Aussprache zu bemessen (Heeringa, 2004). Später wurde diese Abstandsbemessung auch in der Verständlichkeitsforschung verwendet als Faktor, der linguistische Unterschiede modellieren kann und dadurch die Verständlichkeit voraussagt (Küschner et al, 2008).

In meiner Arbeit wird Verständlichkeit auf der geschriebenen Ebene untersucht, weshalb ein so genannter *orthographischer Abstand* zwischen kognaten Wörtern berechnet wurde. Deshalb werden phonetische Unterschiede nicht berücksichtigt, was dazu führt, dass beispielsweise die Wörter *Natur* und *natur* als identisch betrachtet werden, mit einem orthographischen Abstand von „0“. Rein lautlich wird *Natur* im Deutschen als [natu:ɐ] realisiert und im Schwedischen als [nat:ʉr], was bei *phonologischem Abstand* einem Unterschied von „2“ entsprechen würde. Das wäre aber sinnlos in meiner Studie zu betrachten, weil ein Schwedischsprachiger, der *Natur* liest, sollte es vernünftigerweise auch als das Schwedische *natur* interpretieren, denn rein orthographisch liegen keine Unterschiede vor. Zusätzlich werden auch Buchstaben mit diakritischen Zeichen (ö,ä,å,ü) als „0,3“ gewichtet, wenn sie mit einem Buchstaben ohne das Diakritikum ausgetauscht wurden, beispielsweise å → a oder ö → o, weil es das Verfahren von Swarte (2016, s.120) ist. Schließlich wurden die Abstände an die unterschiedlichen Wörterlängen der Wörter angepasst. Das heißt, dass die Summe von Insertionen, Deletionen und Substitutionen mit dem *Alignment* dividiert wird (Swarte, 2016, s. 115). *Alignment* ist der gesamte Abstand von zwei Wörtern, wenn sie untereinander platziert werden, sodass so viele Buchstaben wie möglich miteinander übereinstimmen. Diese Zahl wurde dann mit 100 multipliziert, damit der Abstand als Prozent ausgedrückt werden konnte. Ein theoretisches Beispiel für dieses Verfahren gibt es in Tabelle 1.

Tabelle 1: Berechnung von orthographischem Abstand

	Levenshtein Distanz	Alignment	Orthographischen Abstand
E I N F A C H E N F A L D I G 0 1 0 0 0 1 1 1 1	5	9	$5/9 * 100 = 55,5\%$

3.2.3 Exposition

Exposition bezeichnet wie viel man von einer Sprache erlebt hat und wird hier als eine extralinguistische Variable bezeichnet, weil es von den Eigenschaften der Sprecher abhängig ist. Um auszusagen, wie viel Exposition von Deutsch die Probanden haben, beziehen sich vier Fragen in dem Fragenbogen darauf. Diese Fragen befinden sich in Appendix B. Die Teilnehmer wurden gefragt, ob sie irgendwann im deutschsprachigen Raum für mindestens einen Tag waren, und wie oft es geschehen waren. Sie wurden auch gefragt, ob sie eine längere zusammenhängende Zeit, mindestens 3 Monate, in einem deutschsprachigen Land verbracht hatten. Auch gibt es eine Frage, ob sie irgendwann deutsche Medien konsumiert haben, beispielsweise deutsche Filme oder Bücher. Schließlich dürften sie selbst etwas über ihre Beziehung zu Deutsch schreiben. Für die Analyse bekommen die Teilnehmer Expositionspunkte nach ihren Antworten auf diese Fragen. Für jedes Mal, wenn man in einem deutschsprachigen Land gewesen ist, erhält man einen Punkt. Wenn man deutsche Medien konsumiert hat, erhält man noch einen Punkt. Wenn man eine längere Zeit in einem deutschsprachigen Land war, erhält man 20 Punkte, und was man selbst eingeschrieben hat, wird in jedem Fall individuell bewertet.

3.3 Bewertung und Analyse der Daten

Ein Muster wurde ausgearbeitet, nach dem ausgewertet wird, ob eine Übersetzung richtig ist, weil die Antworten so objektiv wie möglich bewertet werden sollten. Dieses Muster folgt zu großen Teilen einem von Van Hove (2014, s.57) verwendeten Muster. Der Ausgangspunkt ist, wenn jemand eine Antwort gegeben hat, und diese mit einer Übersetzung aus einem Schwedisch-Deutsche Wörterbuch übereinstimmt, dann gilt es als eine richtige Übersetzung. Das hierfür verwendete Wörterbuch war *Tysk-svensk Ordbok* (1980). Es wird nicht die Forderung gestellt, dass die Übersetzung mit dem deutschen Wort im Sinne von Deklination, Konjugation, Tempus oder Numerus übereinstimmen muss. Deshalb gilt beispielsweise *sidor* als eine richtige Übersetzung von *Seite*, obwohl *Seite* im Singular steht, aber *sidor* im Plural. Bei Wörtern, die als mehrere Wortarten interpretiert werden können, gilt eine Übersetzung als richtig, solange sie eine richtige Übersetzung

von einer der Interpretationen ist. Bei der Übersetzung von beispielsweise *Macht* werden deshalb sowohl Übersetzungen wie das Verb *göra*, als auch das Substantiv *makt* als richtig betrachtet. Wenn jemand aber ein Wort als ein Wort von einer Wortart übersetzt, der das Wort nicht grammatisch angehören kann, gilt das als eine falsche Übersetzung. Wenn jemand beispielsweise das Wort *sprach*, das eindeutig ein Verb sein muss, mit dem Substantiv *Språk* übersetzt, wird die Antwort als falsch markiert. Antworten, wo es deutlich ist, dass es sich um einen Schreibfehler handelt, gelten als richtige Übersetzungen, beispielsweise *Vesentlig* anstatt der richtig geschriebenen Übersetzung *väsentlig*. Die Ausnahme davon ist, wenn die Antwort durch den Schreibfehler als ein anderes Wort interpretiert werden kann, dies gilt dann als eine falsche Übersetzung.

Für die Bewertung der Antworten wurden sie in eine Matrix eingeordnet, wie in Tabelle 2, damit die Antworten nach sowohl Person als auch nach Wort summiert werden können. Die Summe pro Person besagt dann, wie viele richtige Übersetzungen eine Person hat, und die Summe pro Wort besagt, wie viele Personen ein Wort richtig übersetzt haben.

Tabelle 2: Beispiel für eine Antwortmatrix

	Wort A	Wort B	Wort C	Wort D	Wort E	Summe pro Person
Person 1	1	0	1	1	0	3
Person 2	0	0	0	1	1	2
Person 3	1	0	1	1	1	4
Person 4	0	1	0	0	0	1
Person 5	1	1	0	0	0	2
Summe pro Wort	3	2	2	3	2	12

Danach wurden die Ergebnisse statistisch analysiert, vor allem durch die Regressionsanalyse. Eine Regressionsanalyse wird generell dafür verwendet, um zu beurteilen, ob eine abhängige Variable *Y* von einer unabhängigen Variable *X* abhängig ist (Lantz, 2013 s. 387). In meiner Arbeit sind die Ergebnisse der Übersetzungsübung, also die Verständlichkeit von Deutsch, die abhängige Variable. Die unabhängigen Variablen sind in meiner Arbeit die linguistischen und die extralinguistischen Variablen. Durch Regressionsanalyse kann untersucht werden, inwieweit die Verständlichkeit von den Variablen abhängig ist. Außer der Regressionsanalyse wurden auch vergleichende Populationsanalysen durchgeführt, um die Ergebnisse von denen, die Deutsch gelernt hatten, mit

den Ergebnissen von denen, die kein Deutsch gelernt hatten, vergleichen zu können. Sie ermöglichen auch den Vergleich der Wörter von unterschiedlichen Wortarten. Die hierfür verwendeten statistischen Methoden wurden zu nicht parametrischen Tests begrenzt, weil es konstatiert wurde, dass die Ergebnisse keiner Normalverteilung folgten. Das Ziel solcher Tests ist zu zeigen, dass die Ergebnisse der unterschiedlichen Populationen voneinander statistisch signifikant unterschiedlich sind, weil es dann ausgeschlossen werden kann, dass die unterschiedlichen Ergebnisse zufällig entstanden sind.

4. Ergebnisse und Analyse

In dem folgenden Teil der Arbeit werden die wichtigsten Ergebnisse der Studie präsentiert und laufend in Bezug auf interessante Beobachtungen analysiert. Zuerst folgt eine Zusammenfassung der eingesammelten Daten in Tabellenform. Dann kommt ein zweiter Abschnitt, in dem die Korrelationsmatrix gezeigt wird, um einen Überblick zu geben. Zum Schluss werden einzelne Wörter aus der Studie genauer analysiert, die für die Teilnehmer besonders schwer zu übersetzen waren.

4.1 Ergebnisse der Übersetzungsübung

Wie in vorigen Kapiteln erläutert wurde, sind die Probanden in zwei Gruppen eingeteilt: Diejenigen, die Deutsch gelernt haben, und diejenigen, die kein Deutsch gelernt haben. Logischerweise gilt, dass jemand, der eine Sprache lernt, auch die frequentesten Wörter dieser Sprache lernt, weshalb zu behaupten ist, dass die Fähigkeit deutsche Wörter zu übersetzen bei dieser Gruppe besser sein soll, was dann einem höheren Anteil an richtigen Übersetzungen entsprechen muss.

Tabelle 3: Übersicht der Probanden und ihre Ergebnisse

	Haben kein Deutsch gelernt (n = 28)	Haben Deutsch gelernt (n = 47)	Alle Probanden (n = 75)
Geschlecht	M: 15 F: 13	M: 14: F: 33	M: 29 F: 46
Alter	32(15:84)	45(20:86)	40(15:86)
Jahre Deutsch gelernt	0	3,8	2,39
Expositionspunkte	7,1	12.4	9,8
Richtige Übersetzungen (%)	42,5	73,3	55,8

Das Ergebnis in Tabelle 3 zeigt, dass die Gruppe, die Deutsch gelernt hat, die Wörter viel besser übersetzen kann. Im Durchschnitt übersetzen sie 73,3 % der deutschen Wörter korrekt, während diejenigen, die nicht Deutsch gelernt haben, im Durchschnitt nur 42,5 % der Wörter richtig übersetzen können. Das ist ein großer Unterschied von etwa 30 Prozentpunkten, der nach einem Mann-Whitney Test auf der 0,001 Ebene statistisch signifikant ist ($U = 120$, $Z = -5,895$, $P < 0,001$)⁴, hier wird deutlich, dass Deutsch lernen natürlich einen Einfluss auf die Verständlichkeit deutscher Sprache hat. Der Mann-Whitney Test wird hier verwendet, weil ein Test der Normalität zeigte, dass die Ergebnisse nicht normalverteilt sind, weswegen parametrische Tests, wie ein t-Test, ausgeschlossen werden. Eine offenbare Frage ist folglich, ob man mehr Deutsch versteht, je länger man Deutsch gelernt hat. Eine Regressionsanalyse mit Anzahl an Jahren Deutsch gelernt als unabhängiger Variable und Prozent richtige Übersetzungen als abhängiger Variable ergibt eine moderate Korrelation ($r = 0,343$) zwischen Jahren, die Deutsch gelernt wurden und wie viel Deutsch man versteht. Die Korrelation ist signifikant auf der 0,05 Ebene ($p = 0,018$), also je längere Zeit man Deutsch lernt, desto mehr Deutsch versteht man. Diese Schlussfolgerung könnte banal scheinen, aber sie ist nicht unwichtig, vor allem weil die Korrelation nicht besonders stark ist. Das heißt, dass außer dem Deutschlernen müssen viele andere Faktoren eine Rolle für die Verständlichkeit spielen. Auch wenn man kein Deutsch gelernt hat, scheint man laut diesem Ergebnis durchschnittlich 42,5 % der frequentesten deutschen Wörter verstehen zu können. Das ist fast die Hälfte aller Wörter, und die versteht man als schwedischsprachige Person einfach so, aufgrund entweder der Ähnlichkeiten der Sprachen oder Exposition von der Sprache.

Hier wäre es des Weiteren auch sinnvoll, die Ergebnisse im Verhältnis mit der extralinguistischen Variable *Exposition* zu setzen. Grundsätzlich sollte gelten, dass je mehr Exposition von Deutsch, desto größer die Wahrscheinlichkeit ist, dass man Wörter aus der Sprache lernt. Anhand Tabelle 3 kann konstatiert werden, dass die Probanden, die Deutsch gelernt haben, auch mehr exponiert davon sind, denn sie haben durchschnittlich 12,4 Expositionspunkte, im Gegensatz zu denjenigen, die kein Deutsch gelernt haben, hier beträgt die Exposition 7,1. Das ist aber auch abgesehen davon, dass sie tatsächlich Deutsch gelernt haben, denn das könnte logischerweise als eine Art von Exposition betrachtet werden. Der Unterschied der Exposition ist nach einem Mann-Whitney Test signifikant auf der 0,05 Ebene ($U = 379,5$, $Z = -3,598$, $p = 0,002$), der verwendet wird, weil die Expositionspunkte auf keinen Fall einer Normalverteilung folgen. Dieses Ergebnis ergibt Sinn, weil

⁴ Die U, Z und P-Werte sind die Ergebnisse des Mann-Whitney Testes und aussagen, wie wahrscheinlich es ist, dass ein Unterschied durch Zufall entstanden ist. Das P-Wert bezeichnet die Wahrscheinlichkeit dieses Zufalls und der Z-wert bezeichnet die entsprechende Anzahl an Standardabweichungen. Der U-Wert ist das tatsächliche Ergebnis des Testes, wodurch die Z- und P-wert berechnet wird.

jemand, der Deutsch gelernt hat, einen höheren Ansporn haben dürfte, sich für die Sprache zu exponieren. Die Korrelation zwischen Exposition und richtigen Übersetzungen ergibt sich dann für alle Probanden als $r = 0,429$, $p < 0,01$. Für die Gruppe, die kein Deutsch gelernt hat, gilt $r = 0,353$, $p > 0,05$, und für die Gruppe, die Deutsch gelernt hat gilt $r = 0,372$ $p < 0,05$. Das heißt, dass Exposition ein wichtiger Teil ist, um Verständlichkeit zu erklären, weil sich diese ziemlich starke Korrelation ergibt. Es scheint tatsächlich auch unabhängig davon zu sein, ob man Deutsch gelernt hat oder nicht, aber das kann nicht mit Sicherheit gesagt werden, weil die Korrelation für die Gruppe, die kein Deutsch gelernt hat, nicht statistisch signifikant ist.

In Tabelle 4 sind die Ergebnisse der zwei Gruppen danach aufgeteilt, ob die Wörter Kognaten, Nicht-Kognaten oder Lehnwörter sind. Es wird durch das Ergebnis deutlich, dass es für Schwedischsprachige einfacher ist, Wörter, die Kognaten sind, zu übersetzen. Sie haben 60,0 % richtige Übersetzungen der Kognaten, aber nur 41,0 % der Wörter richtig, die nicht Kognaten sind. Auffällig ist, dass dieses Verhältnis sowohl für die Gruppe, die Deutsch gelernt hat (76,5 % bzw. 60,4 %) als für die Gruppe, die kein Deutsch gelernt hat (48,1 % bzw. 20,0 %) gilt. Um diese Gruppen in Bezug auf richtige Übersetzungen von Kognaten zu vergleichen, ist es sinnvoll, die Korrelation zwischen dem binär gecodeten lexikalischen Abstand und richtigen Übersetzungen zu studieren.

Tabelle 4: Durchschnittlich prozentualer Anteil richtiger Übersetzungen von Wörtern nach ihrem Status als Kognat und Lehnwort, und auch orthografischem Abstand und Wörterlänge

	Alle Wörter	Kognaten	Nicht-Kognaten	Lehnwörter	Nicht-Lehnwörter
Anzahl	100	80	20	23	57
Haben kein Deutsch gelernt (%)	42,5	48,1	20,0	70,7	39,0
Haben Deutsch gelernt (%)	73,3	76,5	60,4	87,1	72,2
Alle Probanden (%)	55,8	60,0	41,0	73,2	54,1
Orth. Abstand (%)	-	46,5	-	33,4	53,1
Lex. Abstand (%)	20	-	-	-	
Wörterlänge	5,74	5,58	6,40	6,96	5,52

Für alle Probanden gilt $r = 0,333$, $p < 0,01$, und für die Gruppe, die Deutsch gelernt hat, gilt $r = 0,269$, $p < 0,01$ aber für die Gruppe, die kein Deutsch gelernt hat, gilt $r = 0,372$, $p < 0,001$. Die Korrelationen sind also alle statistisch signifikant ($p < 0,05$). Das bedeutet, dass Kognaten wichtig sind, um deutsche Wörter zu verstehen, unabhängig davon, ob man Deutsch gelernt hat oder nicht. Weiter zeigt dieses Ergebnis, dass der lexikalische Abstand eine größere Rolle spielt, wenn man kein Deutsch gelernt hat, weil die Korrelation mit den Ergebnissen dieser Gruppe viel stärker ist. Diese Schlussfolgerung ist sehr logisch, weil es unmöglich sein dürfte, Nicht-Kognaten zu übersetzen, ohne, dass man sie durch Exposition oder durch Lernen gelernt hat.

Ein r -Wert von $0,333$ für lexikalischen Abstand reicht jedoch nicht, um die Ergebnisse ganz zu erklären. Wenn es so wäre, dass alle Kognaten übersetzbar sind, aber die Nicht-Kognaten nicht, dann würde $r = 1$ sein. Aber $r = 0,333$, bedeutet, dass manche Kognaten nicht erkannt werden. Der logische Erklärungsansatz ist, dass je ähnlicher der deutsche Kognat dem schwedischen Kognaten ist, desto einfacher zu erkennen und zu übersetzen ist er. Aus diesem Grund wurde der orthografische Abstand der Kognaten berechnet, was $46,5\%$ ergab. Um den oben erwähnten Ansatz zu überprüfen, können die Ergebnisse der Übersetzungsübung mit dem orthografischen Abstand korreliert werden. Das ergibt $r = -0,426$, $p < 0,001$ für alle Probanden und $r = -0,528$, $p < 0,001$ für diejenigen, die kein Deutsch gelernt haben und $r = -0,301$, $p < 0,01$ für diejenigen, die Deutsch gelernt haben. Die Korrelation ist stark und auch negativ. Die Negativität der Korrelation bedeutet, dass je größer der orthografische Abstand zwischen zwei Kognaten ist, desto schwerer das Wort zu übersetzen ist. Diese Schlussfolgerung gilt unabhängig davon, ob man Deutsch gelernt hat oder nicht, aber der Einfluss der Ähnlichkeiten der Wörter ist größer, wenn man kein Deutsch gelernt hat.

Was zusätzlich von dem Ergebnis in Tabelle 4 auffällig ist, ist die Tatsache, dass die Kognaten, die Lehnwörter sind, viel einfacher zu übersetzen sind, als die Kognaten, die nicht Lehnwörter sind. Dieser Unterschied ist nach einem Wilcoxon signed ranks Test ($Z = 2,481$, $p < 0,05$) statistisch signifikant. Der Wilcoxon signed ranks Test wurde hier verwendet, weil keine Normaldistribution vorliegt und die Stichproben voneinander abhängig sind. Die Schlussfolgerung ist deshalb, dass Lehnwörter einfacher zu verstehen sind als andere Kognaten. Es stellt sich die Frage, warum es so ist und eine Erklärung dafür, die hier zu finden ist, wäre, dass Lehnwörter einen niedrigeren orthografischen Abstand ($33,4\%$) haben, als die anderen Kognaten ($53,1\%$). Also, die Lehnwörter sind ähnlicher als andere Kognaten und sind deshalb einfacher zu erkennen. Das ist auch logisch, weil die Lehnwörter aus dem Deutschen viel später in die schwedische Sprache gekommen sind als die Erbwörter, das heißt die Kognaten, die aus derselben Ursprache stammen. Die Erbwörter sind

also eine längere Zeit in der Sprache gewesen, als die Lehnwörter, und können sich deshalb wahrscheinlich mehr verändert haben.

In Tabelle 5 sind die Ergebnisse nach Wortarten zu sehen. Man kann hier erkennen, dass die Anzahl an korrekten Übersetzungen bei den Verben und Adjektiven sehr ähnlich sind, aber bei den Substantiven sehr unterschiedlich. Für alle Probanden gilt es, dass im Durchschnitt 66,8 % aller Substantive richtig übersetzt werden können, für Verben und Adjektive nur 54,0 % bzw. 52,8 %. Es scheint also, als ob Substantive etwas einfacher zu übersetzen sind.

Tabelle 5: Einteilung der Wörter nach Substantiven, Adjektiven und Verben, Abstandberechnungen und Ergebnisse.

	Substantive	Adjektive	Verben
Anzahl	26	32	26
Kognaten	26	19	21
Lehnwörter	6	7	9
Lehnwörter (%)	24	36,8	42,9
Lex. Abstand (%)	0	37,1	19,2
Orth. Abstand (%)	39,9	53,8	53,0
Wörterlänge	4,9	7,1	5,3
Haben kein Deutsch gelernt (%)	57,8	39,4	35
Haben Deutsch gelernt (%)	83,5	71,9	72,3
Alle Probanden (%)	66,8	54,0	52,8

Ein Wilcoxon signed ranks Test zeigt, dass die unterschiedlichen Ergebnisse zwischen Substantiven auf der einen Seite, und Adjektiven auf der anderen Seite nicht signifikant sind

($Z = -1,829$, $p = 0,067$), aber der Unterschied zwischen Substantiven und Verben signifikant ist ($Z = -2,350$, $p < 0,05$). Der Unterschied zwischen Adjektiven und Verben ist klein und nicht signifikant ($Z = -0,357$, $p = 0,721$). Die Schlussfolgerung muss deshalb sein, dass ein Unterschied zwischen Substantiven und den anderen Wortarten vorliegt. Warum es so ist, kann anhand der Ergebnisse dieser Arbeit dadurch erklärt werden, dass die lexikalischen Abstände und auch die orthografischen Abstände viel niedriger für die Substantive sind. Ob es für das Deutsche und Schwedische generell gilt, dass Substantive öfter Kognaten sind und einen kleineren lexikalischen Abstand haben, geht aus dieser Studie nicht hervor. Es muss berücksichtigt werden, dass diese Ergebnisse nur für die 100 frequentesten Wörter gelten. Ob dieser Trend auch für die restlichen Wörter der Sprachen gilt, wäre ein mögliches Ziel weiterer Arbeiten.

Letztlich werden in Tabelle 6 die Ergebnisse meiner Studie mit der einzigen früheren Studie, die Verständlichkeit zwischen Deutsch und Schwedisch untersucht, nebeneinander präsentiert.

Tabelle 6: Vergleich der Ergebnisse mit Swarte (2016)

	Meine Studie (n = 75)	Swarte (2016) (n = 13-42)
Jahre Deutsch gelernt	2,39	1,88
Orth. Abstand (%)	46,5	36,7
Lex. Abstand (%)	20	13
Richtige Übersetzungen (%)	55,8	65,8

Meine Arbeit hat eine größere Anzahl an Probanden, was bedeutet, dass die Ergebnisse meiner Arbeit statistisch stärker und auch repräsentativer für die Population sein können. Die Ergebnisse sind ähnlich, Swarte zeigt eine Verständlichkeit von 65.8 % auf, im Gegensatz zu den 55,8 % in meiner Arbeit, ein Unterschied von 10 Prozentpunkten. Ob die Ergebnisse sich voneinander statistisch signifikant unterscheiden, lässt sich nicht anhand der zugänglichen Daten beurteilen. Die unterschiedlichen Ergebnisse können jedoch, anhand der Unterschiede der Studien erklärt werden. Swarte geht von anderen Wörtern aus, nämlich die 100 gewöhnlichsten Substantive der englischen Sprache, während die in meiner Studie verwendeten Wörter außer Substantiven auch Adjektive und Verben sind. Wie in Tabelle 5 zu sehen ist, haben Substantive einen niedrigeren lexikalischen und orthografischen Abstand, Faktoren, die mit höherer Verständlichkeit korrelieren. Deshalb ist es logisch, dass die Probanden in der Studie Swartes die Wörter besser übersetzen können. Diese

Studie ist deshalb als eine erfolgreiche Replikation der Ergebnisse der Studie Swartes zu betrachten, was auch die Schlussfolgerungen der beiden Studien verstärkt.

4.2 Regressionsanalyse der Variablen

In der Matrix sind sämtliche ausgerechneten Korrelationen zwischen Variablen überblickbar. Zellen mit $r = --$ und $p = --$ haben keine Ergebnisse, denn entweder war es irrelevant diese miteinander zu korrelieren, wie mit beispielsweise Wortlänge und Exposition, oder war eine Variable Teil einer anderen, wie mit den Ergebnissen von *Alle Probanden* und *Haben Deutsch gelernt*.

Tabelle 7: Matrix mit den Korrelationen und Signifikanzwerten der Ergebnisse und die Variablen

	Lex. Abstand	Orth. Abstand	Lehn- wörter	Wörter- länge	Expos- ition	Jahre Deutsch gelernt	Haben Deutsch Gelernt	Haben kein Deutsch gelernt
<i>Lexikalischen Abstand</i>								
<i>Orthographischen Abstand</i>	$r = --$ $p = --$							
<i>Lehnwörter</i>	$r = --$ $p = --$	$r = -,392$ $p = ,000$						
<i>Wörterlänge</i>	$r = -,170$ $p = ,091$	$r = ,006$ $p = ,955$	$r = ,460$ $p = ,000$					
<i>Exposition</i>	$r = --$ $p = --$	$r = --$ $p = --$	$r = --$ $p = --$	$r = --$ $p = --$				
<i>Jahre Deutsch gelernt</i>	$r = --$ $p = --$	$r = --$ $p = --$	$r = --$ $p = --$	$r = --$ $p = --$	$r = -,148$ $p = ,322$			
<i>Haben Deutsch gelernt</i>	$r = ,269$ $p = ,007$	$r = -,301$ $p = ,007$	$r = ,300$ $p = ,007$	$r = ,013$ $p = ,900$	$r = ,372$ $p = ,010$	$r = ,343$ $p = ,018$		
<i>Haben kein Deutsch gelernt</i>	$r = ,372$ $p = ,000$	$r = -,528$ $p = ,000$	$r = ,481$ $p = ,000$	$r = ,099$ $p = ,325$	$r = 0,353$ $p = ,065$	$r = --$ $p = --$	$r = ,770$ $p = ,000$	
<i>Alle Probanden</i>	$r = ,333$ $p = ,001$	$r = ,426$ $p = ,000$	$r = ,403$ $p = ,000$	$r = ,053$ $p = ,600$	$r = ,429$ $p = ,000$	$r = --$ $p = --$	$r = --$ $p = --$	$r = --$ $p = --$

In der Matrix sind Zellen mit Korrelationen, die statistisch signifikant $p < 0,05$ sind, grau gefärbt.

Statistisch signifikante Werten $p < 0,01$ sind mit einem helleren Grau gefärbt

Viele der Ergebnisse sind schon in *3.1 Ergebnisse der Studie* erwähnt und deshalb dient der Matrix dazu, die Ergebnisse überblickbarer zu machen. Außerdem können übrige noch nicht erwähnte aber interessante Korrelationen analysiert werden. Was noch nicht analysiert worden ist, ist die Wortlänge. Wortlänge scheint nach den Ergebnissen kein entscheidender Faktor zu sein. Keine signifikanten Korrelationen mit den Ergebnissen können gefunden werden, weshalb die Schlussfolgerung sein muss, dass Wortlänge kein entscheidender Faktor bei der Verständlichkeit von Deutsch ist. Jedoch korreliert die Wortlänge sehr stark mit Lehnwörtern, und das bedeutet, dass die Lehnwörter länger sind als die anderen Wörter. Warum es so ist, kann nicht anhand der Ergebnisse dieser Studie erklärt werden, aber eine mögliche Hypothese dafür wäre, dass diese Lehnwörter Tendenzen an Präfixierung und Suffixierung haben. In meiner Studie kommen beispielsweise diese Lehnwörter vor: *er-halten*, *wesent-lich*, *vollständ-ig*, *un-mög-lich*, *öffent-lich*, *plötz-lich*. Wenn es tatsächlich so ist, dass die Lehnwörter gern präfixieren und suffixieren, dann wäre es eine mögliche Erklärung der Korrelation.

Des Weiteren können die Korrelationen für die Gruppe, die Deutsch gelernt hat, und die Gruppe, die kein Deutsch gelernt hat, verglichen werden. Hier ist zu finden, dass die linguistischen Variablen, wie beispielsweise orthographischer Abstand, viel stärker mit den Ergebnissen der Gruppe korrelieren, die nicht Deutsch gelernt hat. Das bedeutet, dass die Ähnlichkeiten der Sprachen viel wichtiger sind, wenn man die Sprache nicht gelernt hat. Zusätzlich ist die extralinguistische Variable *Exposition* ungefähr gleich ähnlich für die beiden Gruppen, aber die Korrelationen sind nicht statistisch signifikant, weswegen wir nicht sicher sein können. Im Großen und Ganzen deuten diese Ergebnisse deshalb darauf hin, dass es wichtig ist, um Verständlichkeit voraussagen zu können, den Ausgangspunkt in den linguistischen Variablen zu nehmen. Hier ist wichtig zu beachten, ob ein substanzieller Anteil der Bevölkerung die Sprache gelernt hat, wie in Schweden, weil es eine Wirkung darauf hat, wie gut diese Variablen die Verständlichkeit voraussagen können. Die Beobachtung, dass dieselben Variablen mit den beiden Gruppen korrelieren, führt dazu, dass ihre Ergebnisse miteinander korrelieren, und das sehr stark ($r = 0,770$, $p < 0,01$). Es bedeutet, dass dieselben Wörter schwer, sowohl als einfach, zu übersetzen sind, unabhängig davon, ob man Deutsch gelernt hat oder nicht. Das sagt etwas sehr Interessantes über die Art aus, nach der genealogisch verwandten Sprachen gelernt werden. Nämlich, dass die Ähnlichkeiten der Sprache von großer Bedeutung sind, um die Sprache zu lernen und zu verstehen, und, dass die durch diese Ähnlichkeiten erworbene Kenntnisse nicht so leicht vergessen werden.

4.3 Analyse von schweren Wörtern

In 4.2 wurde darauf hingewiesen, dass dieselben Wörter für alle Schwedischsprachigen schwer zu übersetzen sind. Das ist unabhängig davon, ob man Deutsch gelernt hat oder nicht. Deshalb werden die 10 Wörter hier erörtert, die am schwersten zu übersetzen sind. Die Wörter und die Ergebnisse davon befinden sich in Tabelle 8.

Tabelle 8: Die schwersten Wörter der Studie nach prozentualen richtigen Übersetzungen und übliche Übersetzungsvorschläge der Probanden

<i>Wort</i>	Kognat-status	Orth. Abstand (%)	Alle Probanden (%)	Haben kein Deutsch gelernt (%)	Haben Deutsch gelernt (%)	Übersetzungsvorschläge
<i>Völlig</i>	Ja	67	13,3	3,5	21	füllig/villig/möjlig
<i>Tat</i>	Ja	77	13,3	0	23,4	tät/tagit
<i>Tief</i>	Ja	100	13,3	0	23,4	tjuv
<i>Rede</i>	Nein	---	12,0	0	21,3	rida/redo/reda/läsa/bo
<i>Darf</i>	Nein	---	15,6	0	27,6	dvärg/mörk/skarf/dart
<i>Gleichen</i>	Ja	75	12,0	0	21,3	glaciär/gissa/glad/glida
<i>Einzig</i>	Nein	---	12,0	10,7	14,9	cigarette/elva/enslig/enstaka
<i>Lage</i>	Ja	7,5	16,9	14,3	21,3	lag/laga/ligga/lägga/sjö
<i>Werden</i>	Nein	---	21,7	3,57	21,3	värd/väldigt/världen/vara
<i>Möchte</i>	Nein	---	25,3	7,41	36,1	möte/måste/mycket

Nach Tabelle 8 scheint es, dass der Kognatstatus nicht der Grund ist, warum die Wörter schwierig zu übersetzen sind. Sie verteilen sich sehr gut mit einer 5:5 Verteilung, was verwunderlich ist, denn vor allem bei denjenigen, die kein Deutsch gelernt haben, sollte es am schwersten sein, Nicht-Kognaten zu übersetzen. Bei dieser Gruppe ist es aber so, dass eine Mehrheit der Wörter, die niemand richtig übersetzen konnte, Kognaten sind. Diese Wörter sind *gleichen*, *tief* und *tat*. Bei diesen Wörtern ist auffallend, dass sie alle einen großen orthographischen Abstand aufweisen. Es weist deshalb wieder auf die starke Wirkung, die der orthographische Abstand bei der Verständlichkeit hat, und erklärt, weshalb der Kognatbegriff selbst nicht die Verständlichkeit erklären lässt. Das wird noch deutlicher, wenn man die Übersetzungsvorschläge betrachtet, die von den Probanden gegeben werden. Für *Tat* kommt *tät* (Deutsch: dicht) vor, was viel ähnlicher als der

richtige Kognat *dåd* ist. Man kann hier *tåt* nach dem etablierten Begriff *falschen Kognaten* definieren lassen. Das heißt, dass *tåt* ähnlicher zu *tat* ist, aber eine andere Bedeutung hat. Auch *tagit* (deutsch: genommen) ist orthographisch ähnlicher als *dåd* und deshalb ein falscher Kognat. Dass man nach einem ähnlichen Wort sucht und deshalb einen falschen Kognaten findet, könnte erklären, warum sämtliche Übersetzungsvorschläge, Wörter mit niedrigerem orthographischem Abstand sind. Deshalb wäre es logisch, die linguistische Variable *lexikalischer Abstand* mit der Bedingung zu ergänzen, dass ein Kognat nur zum lexikalischen Abstand zählt, wenn es auch das Wort mit dem geringsten linguistischen Abstand ist. So einer Ergänzung von lexikalischem Abstand könnte ergeben, dass die Variable die Verständlichkeit besser voraussagen kann, weil die Korrelation vermutlich besser wird, und diesen Ansatz zu überprüfen wäre ein relevantes Ziel weiterer Arbeiten.

Weiter ist auch bei *tief* die vorgeschlagene Übersetzung *tjuv* (Deutsch: Dieb) ähnlicher als der schwedische Kognat *djup*, aber nur marginal, weil nur der Buchstabe „t“ am Anfang der Gleiche ist. Wahrscheinlich ist jedoch, dass die Probanden beim Übersetzen an das englische Wort *thief* gedacht haben, das fast identisch zu *tief* ist, und dann einfach *thief* ins Schwedische übersetzt haben, was *tjuv* ergibt. Diese Tendenz, englische Wörter, die ähnlich sind zu verwenden, scheint ziemlich stark, denn bei drei von den zehn Wörtern in Tabelle 8 kommt das vor, nämlich bei *tief*, *rede* und *darf*. Und weil alle Probanden im Fragebogen angegeben haben, dass sie Englisch sprechen, und Englisch eine Sprache ist, die in der schwedischen Schule unterrichtet wird, wäre es logisch, bei den Berechnungen von linguistischem Abstand auch die Englischkenntnisse in Betracht zu ziehen. Möglicherweise sollte betrachtet werden, ob es auch englische Wörter gibt, die dem deutschen Wort so viel mehr ähneln, als die schwedischen, dass sie eine Wirkung auf das Ergebnis haben, was dazu führen könnte, dass die Verständlichkeit besser vorausgesagt werden kann.

Wenn die Übersetzungsvorschläge näher studiert werden, ist auch auffällig, dass die meisten Vorschläge dieselben oder ähnliche Anfangsbuchstaben haben, sowie beispielsweise *möte* als Übersetzungsvorschlag für *möchte*. Viel seltener kommt es vor, dass der Übersetzungsvorschlag einen anderen Anfangsbuchstaben hat, während der Rest des Wortes identisch ist, wie bei z.B. *skarf* als Übersetzungsvorschlag für *darf*. *Werden* wird nie von den Probanden mit dem ähnlicheren Wort *Herde* übersetzt, obwohl es einen geringeren orthographischen Abstand, als die tatsächlichen Übersetzungsvorschläge der Probanden aufweist. Hier wirft sich die Frage auf, ob deutsche Wörter eine Tendenz dazu haben, denselben Anfangsbuchstaben als ihr schwedischer Kognat zu haben. Wenn das der Fall ist, dann bedeutet es, dass diese üblich verwendete Übersetzungsmethode dazu führt, dass die deutschen Wörter leichter zu verstehen sind. Ein Vergleich der 80 Kognaten aus der Wörterliste ergibt, dass nur 16 Wörter unterschiedliche Anfangsbuchstaben als ihr schwedischer

Kognat haben und die Unterschiede dazwischen, bestehen am häufigsten, wenn phonetisch realisiert, nur darin, ob sie stimmhaft oder stimmlos sind, z.B. *Dag – Tag*. Das heißt, dass die Tendenz zu dieser Übersetzungsmethode vermutlich keine Wirkung auf das Ergebnis hat.

5. Diskussion

In diesem letzten Teil der Arbeit werden auf die aufgestellten Fragestellungen eingegangen. Dabei werden sie anhand der Analyse beantwortet.

5.1 Fragestellung 1

Wie gut können Schwedischsprachige Deutsch anhand ihrer Schwedischkenntnisse verstehen und wäre es möglich zu behaupten, dass gegenseitiges Verständnis zwischen Deutsch und Schwedisch vorliegt?

Die Ergebnisse zeigen, dass jemand, der kein Deutsch gelernt hat, nur anhand seiner Schwedischkenntnisse, schon etwa 40 Prozent der häufigsten deutschen Wörter beim Lesen verstehen kann. Das muss bedeuten, dass geschriebenes Deutsch unter Schweden ziemlich gut verstanden werden kann. Vielleicht reicht diese Verständlichkeit nicht, um eine Botschaft genau zu interpretieren, aber es wäre möglicherweise ausreichend, um den hauptsächlichen Inhalt zu verstehen. Wenn jemandem, der Schwedisch spricht, beispielsweise der Satz „Den Namen zu finden ist unmöglich“ präsentiert wird, versteht er laut der Ergebnisse der Arbeit mit hoher Wahrscheinlichkeit die Wörter *Name*, *finden* und *unmöglich*. Und durch dieses Verständnis könnte man annehmen, dass die Semantik des Satzes verstanden wird, auch wenn die Wörter *der*, *zu* und *ist* nicht verstanden werden. Die Schlussfolgerung muss deshalb sein, dass ein ziemlich großer Teil, von geschriebenem Deutsch, verständlich ist. Wie es sich mit gesprochenem Deutsch verhält, im Gegensatz zu geschriebenem, geht aus diesen Daten nicht hervor, weil nur geschriebenes Deutsch untersucht wurde. Eine logische Voraussetzung ist, dass gesprochenes Deutsch schwerer zu verstehen ist, weil man beim Lesen immer die Möglichkeit hat, sich Wörter mehrere Male anzuschauen, bis eine Übersetzung gefunden wird. Würde diese Studie mit dem Unterschied durchgeführt, dass den Probanden die Wörter einmal vorgespielt werden, wäre die Anzahl an richtige Übersetzungen vermutlich niedriger.

Um auszusagen, ob diese Ergebnisse darauf hindeuten, dass gegenseitige Verständlichkeit vorliegt, also, dass es möglich wäre, für Deutsch- und Schwedischsprachige in ihrer eigenen Sprache zu kommunizieren, müssen die Ergebnisse, mit anderen Ergebnissen von Studien über Sprachen,

verglichen werden, bei denen man weiß, dass gegenseitige Verständlichkeit vorkommt. Swarte (2016) untersuchte in ihrer Studie auch die Verständlichkeit zwischen Dänisch und Schwedisch, zwei Sprachen, die im Großen und Ganzen gegenseitig verständlich sind (Delsing & Lundin, 2005), anhand einer Übersetzungsübung. Ihre Ergebnisse zeigen, dass Schwedischsprachige, die kein Dänisch gelernt haben, 85,5 % der dänischen Wörter richtig übersetzen können. Weil meine Studie als eine Replikation von ihrer zu betrachten ist, können meine Ergebnisse mit ihren auch verglichen werden. Der Unterschied zwischen 85,5 % (Schwedisch-Dänisch) und 42,5 % (Schwedisch-Deutsch) ist sehr groß, weshalb die Schlussfolgerung deshalb sein muss, dass keine gegenseitige Verständlichkeit zwischen Schwedisch und Deutsch vorliegt. Wohlgedacht lässt sich, dass diese Schlussfolgerung nur anhand der Ergebnisse Swartes sich nicht beurteilen lässt, denn sie unterscheidet nicht, ob die Schwedischsprachigen Deutsch gelernt haben oder nicht. Weiter ist es auch interessant zu beobachten, dass Schwedischsprachige, die kein Dänisch gelernt haben, Dänisch besser oder ungefähr so gut⁵ verstehen (85,5 %), wie Personen, die durchschnittlich 3,8 Jahre Deutsch gelernt haben Deutsch verstehen (73,3 %). Das heißt, dass Dänisch so viel ähnlicher zu Schwedisch ist, dass auch Unterricht in Deutsch, was auch verwandt und ähnlich ist, diese Verständlichkeit zwischen den Sprachen nicht überwinden kann. Die Schlussfolgerung muss deshalb sein, dass die Ähnlichkeiten zwischen den Sprachen der mit Abstand wichtigste Faktor der Verständlichkeit sind.

5.2 Fragestellung 2

Inwieweit spielen linguistische und extralinguistische Faktoren eine Rolle bei der Verständlichkeit zwischen den Sprachen und welche Faktoren sind von größter Bedeutung? Welche Rolle spielt die große Menge Lehnwörter aus dem Deutschen in das Schwedische?

Was sich in der Analyse ergab, war, dass die untersuchten Faktoren fast alle eine Rolle bei der Verständlichkeit spielten. Das einzige, was keine Rolle spielte, war die Wörterlänge, die überhaupt nicht mit der Verständlichkeit korrelierte. Der Schluss ist deshalb, dass die Wörterlänge keine Rolle dafür spielt, ob ein Wort schwer oder einfach zu übersetzen ist. Weiter zeigt die Analyse, dass die Faktoren in unterschiedlichem Ausmaß korrelieren, abhängig davon, ob man Deutsch gelernt hat oder nicht. Für Personen, die in der Schule nie Deutsch gelernt haben, spielen die linguistischen Faktoren eine größere Rolle bei der Verständlichkeit als für Personen, die Deutsch gelernt haben. Für alle Probanden korrelieren die linguistischen Faktoren und der extralinguistische Faktor

⁵ Die Studie von Swarte geht nur von Substantiven aus, die sich als einfacher zu übersetzen gezeigt haben. Mit meiner Methodik sollte deshalb eine Studie zwischen Dänisch und Schwedisch ein etwas geringeres Ergebnis als 85,5% ergeben.

Exposition ähnlich gut, warum die Schlussfolgerung sein muss, dass linguistische und extralinguistische Faktoren von gleicher Bedeutung sind. Das heißt, dass sowohl die Ähnlichkeiten zwischen den Sprachen als auch die Exposition von der Sprache signifikante Wirkungen auf die Verständlichkeit haben. Die Ähnlichkeiten der Sprachen können bemessen und quantifiziert werden und dann generell für die Sprachen gelten. Was sich aus dieser Arbeit ergeben hat, ist, dass bei dem Voraussagen der Verständlichkeit berücksichtigt werden muss, ob es Kognaten gibt, aber auch wie ähnlich diese Kognaten sind, in diesem Fall durch orthographischen Abstand, ob sie Lehnwörter sind und ob es falsche Kognaten gibt, aber auch ob es Kognaten und falsche Kognaten im Englischen gibt. Wenn man, wie bei der Verständlichkeitskriterium (siehe 2. *Frühere Forschung*), davon ausgeht, dass Verständlichkeit das ist, was grundsätzlich Sprachen als unterschiedliche Sprachen definieren lässt, könnte eine objektive Bemessung von Verständlichkeit dazu führen, dass eine einheitliche und standardisierte Klassifizierung von Sprachen ermöglicht wird. Das Problem ist, dass *Exposition* wahrscheinlich ein fast gleich wichtiger Faktor bei der Verständlichkeit ist, wie die sprachlichen Ähnlichkeiten. Dann wirft sich die Frage auf, wie könnte man die Exposition einer ganzen Population von anderen Sprachen bemessen? Diese Frage lässt sich nicht einfach beantworten, aber sie ist eine der wichtigsten Fragen, womit die Verständlichkeitsforschung sich in der Zukunft beschäftigen muss.

6. Zusammenfassung

In dieser Arbeit wurde untersucht, wie viel Deutsch Schwedischsprachige verstehen können. Eine empirische Studie wurde durchgeführt, in der Schwedischsprachige 100 hochfrequente geschriebene deutsche Substantive, Adjektive und Verben ins Schwedische übersetzen sollten. Die Ergebnisse zeigten, dass Schwedischsprachige ohne Deutschunterricht im Durchschnitt etwa 40 % der deutschen Wörter übersetzen, was zu der Schlussfolgerung führte, dass keine gegenseitige Verständlichkeit zwischen den Sprachen zu vermuten ist. Es wurde festgestellt, dass Substantive statistisch signifikant einfacher zu übersetzen sind als Adjektive und Verben, was man bei zukünftigen Übersetzungsübungen beachten sollte. Auch Lehnwörter waren statistisch signifikant einfacher zu übersetzen als andere Kognaten, was dadurch erklärt wurde, dass die Lehnwörter ihren Kognaten ähnlicher sind und deshalb einfacher zu verstehen sind. Diesen beiden Faktoren, Wörterklassen und Lehnwörtern, ist wenig frühere Forschung gewidmet und meine Arbeit fängt an, diese Lücken der Kenntnis zu füllen. Ich bin zu dem Schluss gekommen, dass linguistische und extralinguistische Variablen, bei schwedischer Verständlichkeit von Deutsch, von ungefähr gleicher Bedeutung sind, aber, dass die Wortlänge dafür unwichtig ist. Schließlich wurde auch konstatiert, dass Englischkenntnisse vermutlich verwendet werden, um Deutsch zu übersetzen, weshalb die

weitverbreiteten Englischkenntnisse in Schweden, und die Ähnlichkeiten zwischen Deutsch und Englisch auch als Faktoren für die Verständlichkeit berücksichtigt werden müssen.

Literaturverzeichnis

1. Black, E. P, (2019) Algorithms and Theory of Computation Handbook, CRC Press LLC, 1999, "Levenshteindistance". <https://www.nist.gov/dads/HTML/Levenshtein.html>
2. Bø, I. (1978) Ungdom og naboland. En undersøkelse av skolens og fjernsynets betydning for nabospråksforståelsen. In: Putnam, M.T, Page B.R.(2020), The Cambridge Handbook of Germanic linguistics. Cambridge, UK, 773
3. Börestam Uhlmann, U. (1991) Språkmöten och mötesspråk i Norden. In: Putnam, M.T, Page B.R.(2020), The Cambridge Handbook of Germanic linguistics. Cambridge, UK, 773
4. Chambers, J.K, Trudgill, P. (1980), Dialectology. Cambridge, 3.
5. Delsing, L-O., & Lundin, K. (2005). Håller språket ihop Norden? : en forskningsrapport om ungdomars förståelse av danska, svenska och norska. (TemaNord). Nordiska ministerrådet. <https://lucris.lub.lu.se/ws/portalfiles/portal/5905032/625654.pdf>
6. Doetjes, G. & Gooskens, C.(2009), Skriftsprogets rolle i den dansk-svenske talesprogsforståelse. Språk & stil NF 19
http://www.let.rug.nl/~gooskens/pdf/publ_spraakstil_2009.pdf
7. Eberhard, M.D, Gary S.F, Charles F.D. (2021). Ethnologue: Languages of the World. Auflage 24. Dallas.
8. Esselte Studium. (1980) Tysk-svensk Ordbok, Esselte Herzogs, Nacka.
9. Fridell, S. (2021) Ordlok, Lettland
10. Gooskens, C. (2007) The contribution of linguistic factors to the intelligibility of closely related languages. In: Putnam, M.T, Page B.R.(2020), The Cambridge Handbook of Germanic linguistics. Cambridge, UK, 776

11. Gooskens, C. Heeringa, W. Beijering, K. (2008) Phonetic and lexical predictors of intelligibility. In: Putnam, M.T, Page B.R.(2020), The Cambridge Handbook of Germanic linguistics. Cambridge, UK, 776
12. Gooskens, C. Küschner, s. Van Bezooijen, R. (2011) INTELLIGIBILITY OF STANDARD GERMAN AND LOW GERMAN TO SPEAKERS OF DUTCH, *Dialectologia: revista electrònica*, [en línia], p. 35-63
https://www.let.rug.nl/gooskens/pdf/publ_dialectologia_accepted.pdf
13. Gooskens, C. van Heuven, V. Golubović, J. Schüppert, A. Swarte, F, Voigt, S. (2017) Mutual intelligibility between closely related languages in Europe. *International Journal of Multilingualism*, p. 169.193.
14. Haugen, E. (1966). Semicommunication – The Language Gap in Scandinavia, In: Putnam, M.T, Page B.R.(2020), The Cambridge Handbook of Germanic linguistics. Cambridge, UK, 771-772
15. Heeringa, W. J. (2004). Measuring Dialect Pronunciation Differences using Levenshtein Distance. <https://pure.rug.nl/ws/portalfiles/portal/9800656/thesis.pdf>
16. Hellquist, E. (1992) *Svensk etymologisk ordbok*. Digitaliserade Version durch Karp zugänglich,
<https://spraakbanken.gu.se/karp/#?mode=hellqvist&lang=swe&advanced=false&searchTab=special&hpp=25&extended=and%7Canything%7Cequals%7C&page=1>
17. Internetstiftelsen. (2020) 6 procent av befolkningen riskerar ett digitalt utanförskap som sällan- eller ickeanvändare av internet. <https://svenskarnaochinternet.se/rapporter/digitalt-utanforskap-2020/internetutveckling-och-uppkoppling/ar-2020-q1-anvander-96-procent-av-svenskarna-internet/>
18. Kaeding, W. (1898). *Häufigkeitswörterbuch der deutschen Sprache*, Berlin.
19. König, W. (2004). *dtv-Atlas Deutsche Sprache*, 14. Auflage, München, 114-115
20. Küschner, s. Gooskens, C. Van Bezooijen, R. (2008), LINGUISTIC DETERMINANTS OF THE INTELLIGIBILITY OF SWEDISH WORDS AMONG DANES, *International Journal of Humanities and Arts Computing* 2 (1–2), 83–100
http://www.let.rug.nl/gooskens/pdf/publ_IJHAC_2009a.pdf

21. Maurud, ø. (1976). Reciprocal comprehension of Neighbour languages in Scandinavia: An investigation of how well people in Denmark, Norway and Sweden understand each other's languages. In: Putnam, M.T, Page B.R.(2020), The Cambridge Handbook of Germanic linguistics. Cambridge, UK, 772
22. Münte, T. Wieringa, B. Weyerts H. Szentkuti, A. Matzke, M. Sönke, J. (2001). Differences in brain potentials to open and closed class words: class and frequency effects. *Neuropsychologia*, p. 91-102
<https://www.sciencedirect.com/science/article/abs/pii/S0028393200000956>
23. Peels, R. (2019). Replicability and replication in the humanities. *Res Integr Peer Rev* 4, 2.
<https://doi.org/10.1186/s41073-018-0060-4>
24. Pfeifer, W.(1993) Etymologisches Wörterbuch des Deutschen , digitalisierte und von Wolfgang Pfeifer überarbeitete Version im Digitalen Wörterbuch der deutschen Sprache,
<https://www.dwds.de/d/wb-etymwb>
25. SAOB. (1893) Ordbok över svenska språket, Svenska Akademien. Lund
26. SCB. (2021) Import från våra 30 största handelspartner. <https://www.scb.se/hitta-statistik/statistik-efter-amne/handel-med-varor-och-tjanster/utrikeshandel/utrikeshandel-med-varor/pong/tabell-och-diagram/import-fran-vara-30-storsta-handelspartner/>
27. SFS 2011:185. Skolförordning. kap 9, 5§. https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/skolforordning-2011185_sfs-2011-185
28. Stedje, A. (2007), Deutsche Sprache gestern und heute, 6. Auflage, Paderborn, 136.
29. Swarte, F. H. E. (2016). Predicting the mutual intelligibility of Germanic languages from linguistic and extralinguistic factors. Rijksuniversiteit Groningen.
30. Vanhove, J. (2014) Receptive multilingualism across the lifespan. Universität Freiburg in der Schweiz.
31. Wheinreich, M (1944) YIVO Bleter (Ausgabe 23).

Anlagen

Anlage A

ID	Wort	Kognat	Lehnwort	Orthographischer Abstand	Ergebnisse: Alle probanden	Ergebnisse: haben kein Deutsch gelernt	Ergebnisse: haben Deutsch gelernt
1	großen	-	-	-	73,49398	57,142857	95,74468085
2	ganze	-	-	-	28,91566	0	51,06382979
3	glaube	-	-	-	49,39759	14,285714	78,72340426
4	gleichen	Lika	Nein	75	12,04819	0	21,27659574
5	halten	hålla	Nein	55	75,90361	71,428571	91,4893617
6	lassen	Låta	Nein	71,66666667	36,14458	21,428571	51,06382979
7	frei	Fri	Ja	25	81,92771	78,571429	97,87234043
8	kleinen	-	-	-	71,08434	53,571429	93,61702128
9	leben	leva, liv	Nein	60	63,85542	50	82,9787234
10	bleiben	Bli	Ja	57,14285714	51,80723	35,714286	70,21276596
11	machen	-	-	-	54,21687	28,571429	78,72340426
12	sprach	Språkade	Ja	53,75	33,73494	28,571429	42,55319149
13	bringen	Bringa	Ja	28,57142857	72,28916	71,428571	85,10638298
14	deutschen	tyska, tyskarna	Nein	100	87,95181	92,857143	100
15	verschiedenen	-	-	.	27,71084	7,1428571	44,68085106
16	Haus	Hus	Nein	25	89,15663	100	97,87234043
17	Name	Namn	Nein	25	89,15663	96,428571	100
18	gegeben	Gett	Nein	57,14285714	38,55422	28,571429	51,06382979
19	Ende	Ände, ändar	Nein	25	83,13253	89,285714	93,61702128

20	einzig	-	-	-	12,04819	10,714286	14,89361702
21	denken	Tänka	Nein	66,66666667	55,42169	28,571429	80,85106383
22	bestimmten	Bestämda	Ja	40	79,51807	78,571429	93,61702128
23	geht	Går	Nein	75	39,75904	17,857143	59,57446809
24	öffentlichen	Offentliga	Ja	35,83333333	73,49398	71,428571	87,23404255
25	Tag	Dag	Nein	33,33333333	67,46988	42,857143	93,61702128
26	tief	Djup	Nein	100	13,25301	0	23,40425532
27	will	Vill	Nein	25	56,62651	57,142857	65,95744681
28	Natur	Natur	Ja	0	87,95181	96,428571	97,87234043
29	Mensch	Människa	Ja	62,5	73,49398	67,857143	89,36170213
30	einfach	Enfaldig	Nein	55,55555556	34,93976	25	46,80851064
31	Seite	Sida	Nein	60	44,57831	21,428571	65,95744681
32	Stadt	Stad	Ja	25	85,54217	92,857143	95,74468085
33	heißt	Heter	Nein	57,14285714	63,85542	35,714286	91,4893617
34	sein	vara, sin	Nein	25	48,19277	25	70,21276596
35	voll	Full	Nein	25	26,50602	7,1428571	42,55319149
36	möchte	-	-	-	25,3012	7,1428571	40,42553191
37	erhalten	Erhålla	Ja	53,75	59,03614	50	74,46808511
38	Herr	Herr	Ja	0	87,95181	96,428571	97,87234043
39	schönen	Sköna	Ja	57,14285714	77,10843	75	91,4893617
40	Hand	Hand	Nein	0	86,74699	96,428571	95,74468085
41	Stelle	Ställa, ställe	Ja	5	61,44578	57,142857	74,46808511
42	hat	Ha	Nein	33,33333333	54,21687	28,571429	78,72340426
43	rede	-	-	-	12,04819	0	21,27659574
44	Jahre	År	Nein	66	74,6988	53,571429	100
45	sehen	Se	Nein	60	62,6506	21,428571	97,87234043
46	steht	Stå	Nein	60	45,78313	10,714286	74,46808511

47	Wort	Ord	Nein	50	34,93976	14,285714	53,19148936
48	stellen	ställa, ställen	Ja	14,28571429	62,6506	53,571429	78,72340426
49	Sache	Sak	Nein	60	39,75904	17,857143	59,57446809
50	Herz	Hjärta	Nein	66,66666667	44,57831	28,571429	61,70212766
51	Frage	Fråga	Ja	26	83,13253	85,714286	95,74468085
52	besondere	-	-	-	31,3253	7,1428571	51,06382979
53	Welt	Värld	Nein	80	56,62651	35,714286	78,72340426
54	meisten	Mesta	Nein	42,85714286	32,53012	7,1428571	53,19148936
55	liegt	Ligger	Nein	57,14285714	50,60241	21,428571	76,59574468
56	schnell	-	-	-	69,87952	46,428571	95,74468085
57	werden	-	-	-	21,68675	3,5714286	36,17021277
58	König	Kung	Nein	40	46,98795	35,714286	61,70212766
59	wesentlich	Väsentlig	Ja	30	84,33735	85,714286	97,87234043
60	Lage	Läge	Nein	7,5	16,86747	14,285714	21,27659574
61	weniger	-	-	-	31,3253	0	55,31914894
62	Augen	Ögon	Nein	60	54,21687	25	80,85106383
63	Frau	Fru	Ja	25	87,95181	92,857143	100
64	ruhig	-	-	-	43,37349	10,714286	70,21276596
65	schwer	Svår	Nein	66,66666667	67,46988	39,285714	95,74468085
66	Paragraph	Paragraf	Ja	22,22222222	83,13253	82,142857	97,87234043
67	weiß	Vit, vet	Nein	75	68,6747	39,285714	97,87234043
68	darf	-	-	-	15,66265	0	27,65957447
69	unmöglich	Omöjlig	Ja	55,55555556	84,33735	85,714286	97,87234043
70	Macht	Makt	Ja	40	55,42169	42,857143	72,34042553
71	nächsten	Nästa	Nein	50	60,24096	35,714286	85,10638298
72	sagen	Säga	Nein	46	63,85542	39,285714	89,36170213
73	scheint	Skina	Nein	57,14285714	42,16867	25	59,57446809

74	einzelne	-	-	-	38,55422	46,428571	40,42553191
75	finden	Finna	Nein	50	85,54217	85,714286	100
76	weiteren	Vidare	Nein	75	39,75904	17,857143	59,57446809
77	Mutter	Moder	Nein	50	74,6988	64,285714	93,61702128
78	hoch	Hög	Nein	57,5	59,03614	32,142857	85,10638298
79	Recht	Rätt	Nein	60	66,26506	42,857143	91,4893617
80	Weise	vis, visa	Nein	60	28,91566	32,142857	31,91489362
81	später	-	-	-	50,60241	14,285714	80,85106383
82	gut	God	Nein	66,66666667	83,13253	78,571429	100
83	völlig	Fullt	Nein	66,66666667	13,25301	3,5714286	21,27659574
84	Grund	Grund	Nein	0	65,06024	67,857143	74,46808511
85	vollständig	Fullständig	Ja	18,18181818	78,31325	78,571429	91,4893617
86	Liebe	-	-	-	69,87952	57,142857	89,36170213
87	folge	följd, följa	Nein	46	42,16867	42,857143	48,93617021
88	Nacht	Natt	Nein	40	84,33735	85,714286	97,87234043
89	muss	Måste	Ja	60	51,80723	28,571429	74,46808511
90	Vater	Fader	Nein	40	57,83133	28,571429	85,10638298
91	plötzlich	Plötsligt	Ja	33,33333333	86,74699	89,285714	100
92	alten	-	-	-	50,60241	21,428571	76,59574468
93	Mann	Man	Nein	25	83,13253	82,142857	97,87234043
94	führen	Föra	Nein	66,66666667	26,50602	10,714286	40,42553191
95	neue	Nya	Nein	75	74,6988	67,857143	91,4893617
96	letzte	-	-	-	42,16867	14,285714	65,95744681
97	Zeit	Tid	Nein	75	53,01205	21,428571	80,85106383
98	soll	Ska	Nein	75	33,73494	3,5714286	57,44680851
99	kann	Kan	Nein	25	67,46988	57,142857	85,10638298
100	Tat	Dåd	Nein	76,66666667	13,25301	0	23,40425532

Anlage B

Tack för att du deltar denna i studie vilken kommer ligga till grund för mitt examensarbete i tyska vid Lunds universitet. Jag ämnar undersöka svensktalandes förmåga att förstå skriven tyska. Du kommer först att få fylla i ett antal uppgifter om dig själv och därefter kommer du att presenteras för 100 tyska ord som jag ber dig att efter bästa förmåga översätta till svenska. Tanken är inte att du ska söka upp det rätta ordet utan du ska helt enkelt bara gissa. Stort tack för ditt deltagande!

Vid frågor, hör vänligen av dig via mail eller telefon:

Mail: Hannes.essfors@gmail.com

Telefon: 0767675464

Så här fyller du i pappersenkäten

Nedan ser du hur du markerar ett svarsalternativ, och hur du avmarkerar ett redan

gjort val. Korrekt markerat svarsalternativ

Inkorrekt markerat svarsalternativ, krysset ska vara mitt i rutan

Inkorrekt markerat svarsalternativ, krysset är alltför

kraftigt Ångrat val, svarsalternativet räknas inte som markerat

Hur gammal är du?

Vänligen specificera din könstillhörighet.

Man

Kvinn

a

Annat/vill inte uppge

Vilken är din högsta avslutade utbildningsnivå?

Grundskola

Gymnasium

Yrkesutbildnin

g

Högskola eller universitet

Har du svenska som modersmål?

Ja

Nej

Talar du något/några andra språk än svenska? Skriv i så fall vilket/vilka.

Har du studerat tyska t.ex. i grundskolan, gymnasiet, universitetet eller på egen hand?

Ja

Nej

Om du svarade ja på ovanstående fråga, vänligen uppskatta hur många år du sammantaget studerat tyska

Har du tillbringat en längre sammanhängande tid (mer än 3 månader) i något tysktalande land?

Ja

Nej

Har du vid något tillfälle spenderat minst en dag i ett tysktalande land?

Ja

Nej

Om du svarade ja på ovanstående fråga, vänligen uppskatta hur många gånger du har besökt något tysktalande land. Lämna annars fältet blankt

Har du någonsin sett en tysk film eller konsumerat annan tyskspråkig media.

Ja

Nej

Vet

ej

Vänligen uppskatta hur mycket av en tysk tidningsartikel du tror du skulle kunna förstå.

I princip ingenting

Det huvudsakliga

budskapet Det mesta av

innehållet

Om du har någon annan koppling till Tyskland eller det tyska språket så specificera gärna det med några ord eller meningar nedan. Lämna annars fältet blankt.

Du är nu vid den andra delen av enkäten. Här består varje fråga av ett tyskt ord och jag ber dig att skriva en svensk översättning under varje ord. Om du inte vet den korrekta översättningen så vill jag att du gissar vad du tror att det kan betyda. Tanken är inte att du ska leta upp en korrekt översättning, utan bara själv översätta efter bästa förmåga. Här kommer första ordet:

Ganze

Glaube

Gleichen

Halten

Lassen

Frei

Kleinen

Leben

Bleiben

Machen



\$Förmamn\$

Sprach

Bringen

Deutschen

Verschiedenen

Haus

Name

Gegeben

Ende

Einzig

Denken

Bestimmten



\$Förmamn\$

Geht

Öffentlichen

Tag

Tief

Will

Natur

Mensch

Einfach

Seite

Stadt

Heißt



\$Förmamn\$

Sein

Voll

Möchte

Erhalten

Herr

Schönen

Hand

Stelle

Hat

Rede

Jahre



\$Förmamn\$

Sehen

Steht

Wort

Stellen

Sache

Herz

Frage

Besondere

Welt

Meisten

Liegt



Schnell

Werden

König

Wesentlich

Lage

Weniger

Augen

Frau

Ruhig

Schwer

Paragraph



Weiß

Darf

Unmöglich

Macht

Nächsten

Sagen

Scheint

Einzelne

Finden

Weiteren

Mutter



Hoch

Recht

Weise

Später

Gut

Völlig

Grund

Vollständig

Liebe

Folge

Nacht



Muss

Vater

Plötzlich

Alten

Mann

Führen

Neue

Letzte

Zeit

Soll

Kann



\$Förnamn\$

Tat

Stort tack för ditt medverkande i denna studie! Är du intresserad av att veta hur många ord du korrekt lyckades översätta så fyll i din mail-adress nedan så återkommer jag till dig. Bästa Hälsningar Hannes Essfors