

Welche Online-Ressourcen nutzen romanischsprachige DaF-Lernende bei der Verbesserung deutscher Texte?

*Martina Nied Curcio (Rom), Carolin Müller-Spitzer (Mannheim),
Idalete Maria Silva Dias (Braga), María José Domínguez Vázquez (Santiago
de Compostela) & Sascha Wolfer (Mannheim)*

1 Einleitung – Wörterbücher im Fremdsprachenunterricht

Wörterbücher haben eine lange historische Tradition und waren in der Fremdsprachendidaktik schon immer ein Thema. Allerdings hing ihr Einsatz – und hier v.a. der Gebrauch zweisprachiger Wörterbücher – im Fremdsprachenunterricht stets mit den Hypothesen in Bezug auf das Verhältnis von Erst- und Zweitsprachenerwerb und mit dem damit verbundenen Einsatz der Muttersprache und dem Sprachvergleich als Methode zusammen. War das zweisprachige Wörterbuch in der Grammatik-Übersetzungs-Methode ein selbstverständliches Hilfsmittel, so wurde es in der direkten, der audiolingualen und audiovisuellen Methode aus dem Unterricht verbannt. Auch in der kommunikativen Methode soll(t)en die Lernenden keinesfalls das zweisprachige Wörterbuch verwenden, sondern, wenn überhaupt, in einem einsprachigen Wörterbuch nachschlagen. Die Angst, in eine zu sehr auf Grammatik und Übersetzung konzentrierte Didaktik zurück zu fallen und einen kommunikativ ausgerichteten Unterricht zu behindern, ist bei DaF-Lehrenden auch heute noch sehr groß. Die Lernenden dagegen verwenden für sprachliche Fragen heutzutage ganz selbstverständlich Online-Wörterbücher, Übersetzungs-

tools und Wörterbuch-Apps (vgl. Domínguez Vázquez; Valcárcel Riveiro 2015, Müller-Spitzer 2014, Nesi 2012, Nied Curcio 2014).

Seit den 1990er Jahren und v.a. in Folge des Erscheinens des *Gemeinsamen Europäischen Referenzrahmens* (vgl. Europarat 2001) mit der Betonung auf Mehrsprachigkeit und lebenslanges Lernen wird die Verwendung von Wörterbüchern im Fremdsprachenunterricht wieder hoffähig. Ihre kompetente Benutzung wird für das erfolgreiche Erlernen einer Fremdsprache sogar als „unverzichtbar“ erklärt (Zöfgen 2010: 108).

Während die Wörterbucharbeit im muttersprachlichen Deutschunterricht curricular verankert ist (v.a. in der Grundschule und Sekundarstufe I, siehe Merten 2011: 348-352), scheint sie in der fremdsprachlichen Didaktik nicht unbedingt im Zentrum des Interesses zu stehen – und dies, obwohl man weiß, dass Fremdsprachenlernende immer schon Wörterbücher verwendet haben (vgl. Snell-Hornby 1987) und diese heutzutage insbesondere in elektronischer Form bzw. online benutzen.

Doch wie ist es um ihre Benutzungskompetenz bestellt? Welche Online-Ressourcen werden tatsächlich verwendet? Viele Studien aus dem Bereich der Wörterbuchbenutzungsforschung¹ beschäftigen sich mit der Frage, was die Benutzer an Wörterbüchern schätzen (vgl. Domínguez Vázquez; Mirazo Balsa; Vidal Perez 2013, Domínguez Vázquez; Valcárcel Riveiro 2015, Koplenig; Meyer; Müller-Spitzer 2014), welche Wörterbücher in welchen Situationen verwendet werden (vgl. Koplenig; Meyer; Müller-Spitzer 2014), welche Art der Datenpräsentation am benutzerfreundlichsten ist (vgl. Lew 2010, Lew; Grzelak; Leszkowicz 2013) oder welche Informationen am häufigsten in Online-Wörterbüchern nachgeschlagen werden (vgl. De Schryver; Joffe; Joffe; Hillewaert 2006, Hult 2012, Koplenig; Meyer; Müller-Spitzer 2014). Diese Studien setzen jedoch alle voraus, dass lexikografische Hilfsmittel tatsächlich benutzt werden.

Gleichzeitig haben viele Fremdsprachenlehrende und Dozierende der Sprachwissenschaften den Eindruck, dass in der alltäglichen Realität von Fremdsprachenlernenden zu wenig (gute) Wörterbücher verwendet werden. Noch stehen jedoch zu wenige empirische Daten zur Verfügung (Levy; Steel 2015: 194).²

Aus diesem Grunde haben wir uns empirisch der Frage genähert, wie oder ob bestimmte Gruppen heute überhaupt noch Wörterbücher nutzen und ob sie sie bewusst von anderen sprachbezogenen Daten im Web unterscheiden. Es sollten empirische Daten gesammelt werden, um zu erfahren, wie DaF-Lernende tatsäch-

¹ Die Wörterbuchbenutzungsforschung ist das jüngste Forschungsgebiet der Lexikographie. Mit der Ausarbeitung der theoretischen und methodischen Grundlagen v.a. durch Herbert Ernst Wiegand in den 1990er Jahren wurde der empirischen Forschung der Weg freigemacht (vgl. Wiegand 1998 und Wiegand; Beißwanger; Gouws; Kammerer; Storrer 2010). In den letzten 10-15 Jahren hat sie erhebliche Fortschritte erzielt (vgl. u.a. Töpel 2014, Welker 2013, Lew 2011). Auch die am Institut für Deutsche Sprache (IDS) durchgeführten Studien zu Onlinewörterbüchern (vgl. Müller-Spitzer 2014) haben den Kenntnisstand des Forschungsgebietes erweitert (Lew 2015: 251).

² Im Bereich Deutsch als Fremdsprache gibt es bisher vereinzelte Studien. (vgl. Roelcke 2015).

lich arbeiten (und nicht was sie dazu retrospektiv sagen), vor allem um eine bessere empirische Basis für den Unterricht zur Verfügung zu haben.

Zentrale Fragen dabei waren:

- Wie nutzen DaF-Lernende heutzutage lexikografische Ressourcen?
- Welche Suchstrategien wenden sie an?
- Differenzieren sie zwischen den unterschiedlichen Ressourcen?
- Welche Strategien erweisen sich als besonders erfolgreich?

Der Versuchsaufbau wurde gezielt sehr offen konzipiert, da die Benutzung bestimmter Nachschlagewerke nicht vorgeschrieben werden sollte. Im Gegenteil: Die Versuchssituation sollte so eng wie möglich an ein alltägliches Umfeld angenähert und die Recherchepraxis mit möglichst wenigen Vorannahmen beobachtet werden.³

Der vorliegende Beitrag ist folgendermaßen gegliedert: Im nächsten Abschnitt wird das Studiendesign und die Datenaufbereitung beschrieben, im 3. Abschnitt werden die Ergebnisse der Studie präsentiert, um danach in Abschnitt 4 die Konsequenzen für eine Wörterbuchdidaktik innerhalb des Fremdsprachenunterrichts zu diskutieren.⁴

2 Studiendesign und Datenaufbereitung

Vor diesem Hintergrund wurde ein internationales Kooperationsprojekt gestartet. Im Fokus des Projektes stand die Frage, welche Hilfsmittel romanischsprachige DaF-Lernende online zu Rate ziehen, wenn sie Interferenzfehler⁵ verbessern wollen⁶ und frei im Internet recherchieren können. Empirische Daten dieser Art können langfristig eine Grundlage dafür schaffen, Wörterbuchdidaktik besser in den Sprachunterricht zu integrieren.

Unsere Probandinnen und Probanden waren etwa auf dem A2/B1-Niveau und hatten alle eine vergleichbare Anzahl von Deutschstunden an der Universität absolviert. Die Erhebung wurde an den Orten Braga (Portugal), Santiago de Compostela (Spanien) und Rom (Italien) durchgeführt. Pro Ort wurden im März/April 2017 Daten von 14 Probandinnen und Probanden erhoben. Jede/r Teilnehmende

³ Es ist vorauszuschicken, dass unsere Studie noch keine umfassende Antwort auf diese Fragen gibt, sondern eher einen ersten Einblick. Es handelt sich um eine explorative Beobachtungsstudie.

⁴ Ein englischsprachiger Artikel zu dieser Studie, mit eher methodischer Ausrichtung, ist 2018 erschienen. (Wolfer; Nied Curcio; Silva Dias; Müller-Spitzer; Domínguez Vázquez 2018). Eine weitere ausführliche Publikation auf Englisch wurde in *Lexikos* (vgl. Müller-Spitzer; Domínguez Vázquez; Nied Curcio; Silva Dias; Wolfer 2018) veröffentlicht.

⁵ Bei der Studie haben wir uns auf Interferenzfehler konzentriert, die für DaF-Lernende aus Italien, Portugal und Spanien gleichermaßen zutreffen, sodass allen Studierenden die gleichen fehlerhaften Sätze zur Verbesserung vorgelegt werden konnten.

⁶ Die Aufgabe einer Korrektur orientierte sich an der Forschung von Wolfer; Bartz; Weber; Abel; Meyer; Müller-Spitzer; Storrer (2018).

erhielt 18 deutsche Sätze mit einem Interferenzfehler. Die Fehler wurden so konstruiert, dass sie drei Anforderungen erfüllten:

Der Fehler

- ist typisch für Deutschlernende mit geringen Kenntnissen (Niveaustufe A2) mit einer romanischen Erstsprache;
- kann nicht durch eine einfache Suche im Internet behoben werden;
- muss möglichst so schwer zu beheben sein, dass die Studierenden dafür Hilfsmittel benötigen und die Aufgabe nicht auf Grundlage ihrer eigenen Kompetenz lösen können.

Diesen Anforderungen gerecht zu werden, war eine größere Herausforderung als zunächst vermutet, denn nicht alle Fehler waren Interferenzfehler in allen drei Sprachen. Außerdem konnten viele der Standard-Interferenzfehler durch eine Google-Recherche gelöst werden. Durch mehrere Pretests abgesichert konnten dann von 26 Sätzen 18 Sätze ausgewählt werden, in denen die Fehler jeweils fett markiert waren (vgl. Tabelle 1).⁷

Der Ablauf der Studie sah folgendermaßen aus: Den Studierenden wurde im Vorfeld mitgeteilt, dass sie an einer Studie teilnehmen sollten, die allerdings keinerlei Einfluss auf ihre Kursbewertung haben würde. Bei der Durchführung der Studie war es außerdem wichtig, dass sich keine Lehrperson der Universität im Raum befand, da sonst die Gefahr bestanden hätte, dass die Probanden bestimmte Ressourcen, wie bspw. Übersetzungsprogramme nicht aufrufen, da sie wussten, dass diese evtl. von den Lehrpersonen nicht geschätzt würden. Die Probandinnen und Probanden sollten nun fehlerhafte deutsche Sätze möglichst so bearbeiten und verbessern, wie sie dies auch zuhause tun würden. Sie konnten also frei im Internet recherchieren. Gleichzeitig sollten sie laut denken, also kommentieren, ob sie den Fehler erkennen, warum sie welche Ressourcen aufrufen, ob sie zufrieden mit den Ergebnissen sind etc. (Think-Aloud-Protokoll (TAP), Ericsson; Simon 1993). Sie arbeiteten an einem von uns gestellten Laptop, damit die technischen Bedingungen für alle gleich waren.⁸ Zusätzlich zu den deutschsprachigen VersuchsleiterInnen war eine Person anwesend, die die Landessprache beherrschte, damit die Studierenden in ihrer Muttersprache Fragen stellen konnten. Diese zusätzliche Person hatte darüber hinaus die Aufgabe, die Studierenden regelmäßig daran zu erinnern, dass sie laut denken sollten. Keine der anwesenden Personen saß so im Raum, dass man auf den Laptop-Bildschirm schauen konnte. So sollte trotz der doch etwas künstlichen Atmosphäre eine möglichst geschützte Situation für die Probandinnen und Probanden geschaffen werden.

⁷ Die ursprüngliche Nummerierung blieb dabei erhalten.

⁸ Vorinstalliert waren als Browser Google Chrome und Mozilla Firefox, als Sprache wurde die jeweilige Landessprache eingestellt und der Browsercache und -verlauf wurden nach jeder Versuchsperson gelöscht, um Reihenfolgeeffekte zu vermeiden.

Tabelle 1: Stimulus-Sätze der Studie.

ID	Satz
1	Meine Nachbarin möchte immer alles wissen. Sie ist sehr kurios .
2	Leider kann ich heute nicht Tennis spielen. Ich bin zu besetzt .
3	Bist du bereit ? Wir müssen jetzt los, wir sind sowieso schon zu spät dran.
4	Kein Problem, wenn der Zucker beendet ist; ich nehme dann Honig.
7	Ich bin einverstanden mit dir.
9	Das erlaube ich dir nicht. Es ist außer Frage .
11	An unserem Forschungsinstitut ist Ihnen unsere Bibliothek 24 Stunden zur Verfügung .
12	Obwohl ich studiere, wohne ich noch mit meinen Eltern.
14	Wenn ich zur Schule ging, habe ich viel Sport gemacht.
15	Morgen habe ich einen Arzttermin und kann deshalb nicht in die Klasse gehen .
18	Ich vorbereite gerade meine letzte Prüfung.
19	Ich möchte ein Stipendium beim DAAD bewerben.
20	Ich habe die Hose viel zu klein gekauft. Jetzt muss ich nochmals ins Geschäft zurück und sie wechseln .
21	Obwohl sich der Junge beeilt hat, hat er die U-Bahn verloren .
22	Er wohnt seit Jahren in Berlin und trotzdem verliert er sich immer noch.
24	Um beim Kartenspielen zu gewinnen, musst du exakt die Regeln folgen .
25	Der Artikel handelt sich um die Migranten in Deutschland.
26	Ich möchte dir heute über einen interessanten Artikel sprechen .

Als Screenrecording-Software wurde die kostenlose Version der Software *Active Presenter* genutzt. Die Audiospur wurde über ein externes Mikrofon aufgenommen und nachträglich als Tonspur in den Bildschirmfilm integriert, um eine synchrone Untersuchung von Bildschirmaufnahmen und TAPs zu ermöglichen. Die Studie-

renden hatten jeweils 45 Minuten Zeit, die Stimulus-Sätze zu verbessern. Das Versuchsdesign wurde den ProbandInnen vorher erklärt; der Zeitrahmen war ihnen allerdings nicht bekannt, um keinen Zeitdruck zu erzeugen. Nach 30 Minuten wurde allerdings darauf hingewiesen, dass nun noch 15 Minuten verblieben. Diese zeitlichen Restriktionen mussten vorgegeben werden, um die Erhebung zeitlich planen zu können. Im Ergebnis entstand pro Proband/in ein ca. 45minütiger Bildschirmfilm. Dies sind in der Summe für die Studie etwa 1935 Minuten Filmmaterial plus TAPs mit mehr als 2.200 Suchaktionen. Um diese Daten sinnvoll auswerten zu können, war ein relativ aufwändiger Prozess der Datenaufbereitung notwendig, der im nächsten Abschnitt beschrieben wird.

Für eine statistische Auswertung mussten die Bildschirmfilme, die TAPs und die Lösungen der Studierenden aufbereitet werden. Die Bildschirmfilme wurden in Excel-Tabellen überführt, in denen eine Vielzahl von Spalten angelegt wurden, um verschiedenste Auswertungen zu ermöglichen wie den Namen der aufgerufenen Ressourcen, ob diese gezielt aufgerufen wurden, den genauen Suchstring oder die Sprache des Suchstrings. Die TAPs wurden in der jeweiligen Landessprache transkribiert und ins Deutsche übersetzt. Diese Daten wurden anschließend in die Tabellen zu den Bildschirmfilmen integriert. In einem dritten Schritt wurden die verbesserten Sätze der ProbandInnen dahingehend annotiert, ob sie richtig, falsch oder gar nicht verbessert wurden.⁹ Auch die Kategorisierung der Ergebnissätze wurde dann in die Gesamtauswertungstabelle überführt, sodass eine Gesamttabelle zur Auswertung der Studie vorlag. Alle weiteren statistischen Auswertungen wurden mit R¹⁰ durchgeführt.

3 Ergebnisse der Studie

3.1 Einige quantitative Ergebnisse

Im Durchschnitt korrigierten unsere Probanden 16 Sätze. Diese Zahl war an allen drei Orten nahezu gleich (siehe Abbildung 1).

⁹ Diese Annotation wurde zu zweit unabhängig voneinander durchgeführt. Im Anschluss wurden diese Annotationen automatisch verglichen. Das Inter-Annotator-Agreement bei den insgesamt 816 Sätzen war sehr gut: $\kappa_{\text{mean}} = 0.82$. Die unterschiedlich annotierten Sätze wurden dann mündlich durchgesprochen und man hat sich auf eine Annotation geeinigt.

¹⁰ R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (<https://www.R-project.org/> [27.09.2018]).

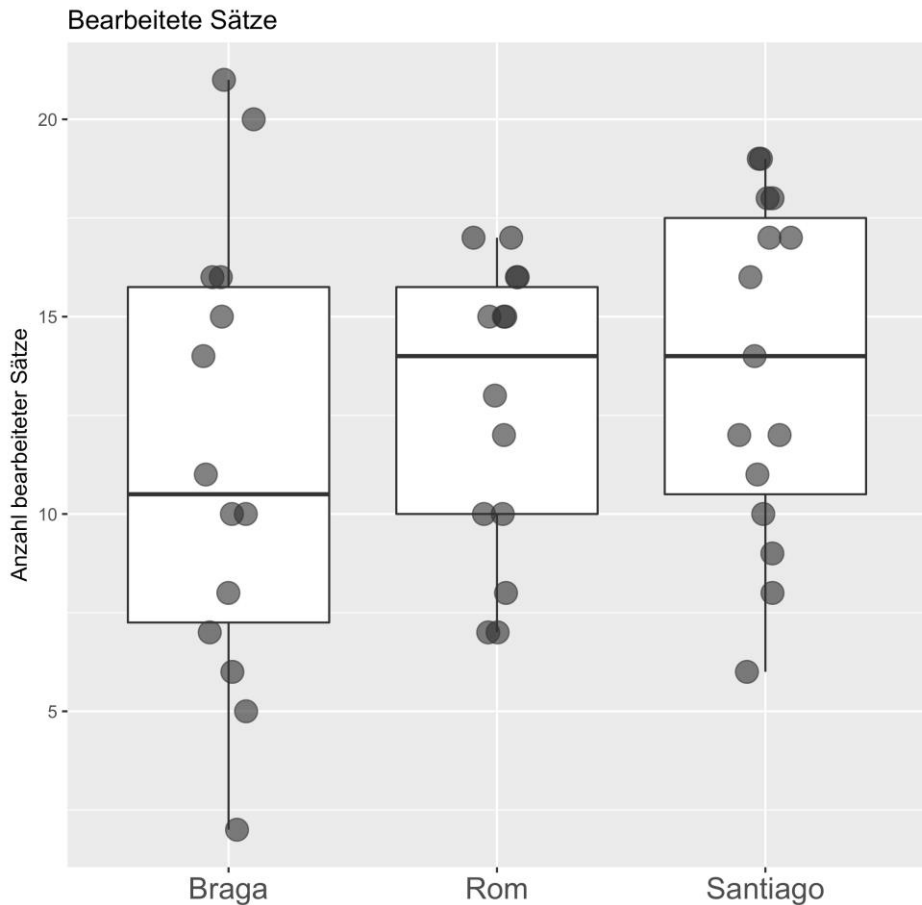


Abbildung 1: Boxplots für die Anzahl der bearbeiteten Sätze an den drei Standorten.¹¹

Allerdings unterschied sich die Anzahl der korrekt verbesserten Sätze zwischen den drei Standorten erheblich, wie in Abbildung 2 abzulesen ist.¹²

¹¹ Jeder Punkt zeigt die Anzahl der bearbeiteten Sätze für eine/-n Teilnehmer/-in. Die mittleren 50 % aller Punkte befinden sich in der Box (Interquartilsabstand). Die horizontale Linie innerhalb jeder Box gibt den jeweiligen Median an. Die vertikalen Linien ober- und unterhalb jeder Box erstrecken sich bis zum niedrigsten bzw. höchsten zulässigen Wert. Punkte außerhalb der vertikalen Linien gelten als Ausreißer (diese Beschreibungen gelten für jeden Boxplot in diesem Beitrag).

¹² Dieses Ergebnis zeigt bereits an dieser Stelle, was später durch andere Ergebnisse gestützt wird: Obwohl wir hofften, mit unseren Teilnehmern an allen drei Universitäten das gleiche Sprachniveau zu erreichen, war das tatsächliche Sprachniveau der Teilnehmer in Rom und Santiago de Compostela deutlich höher als in Braga.

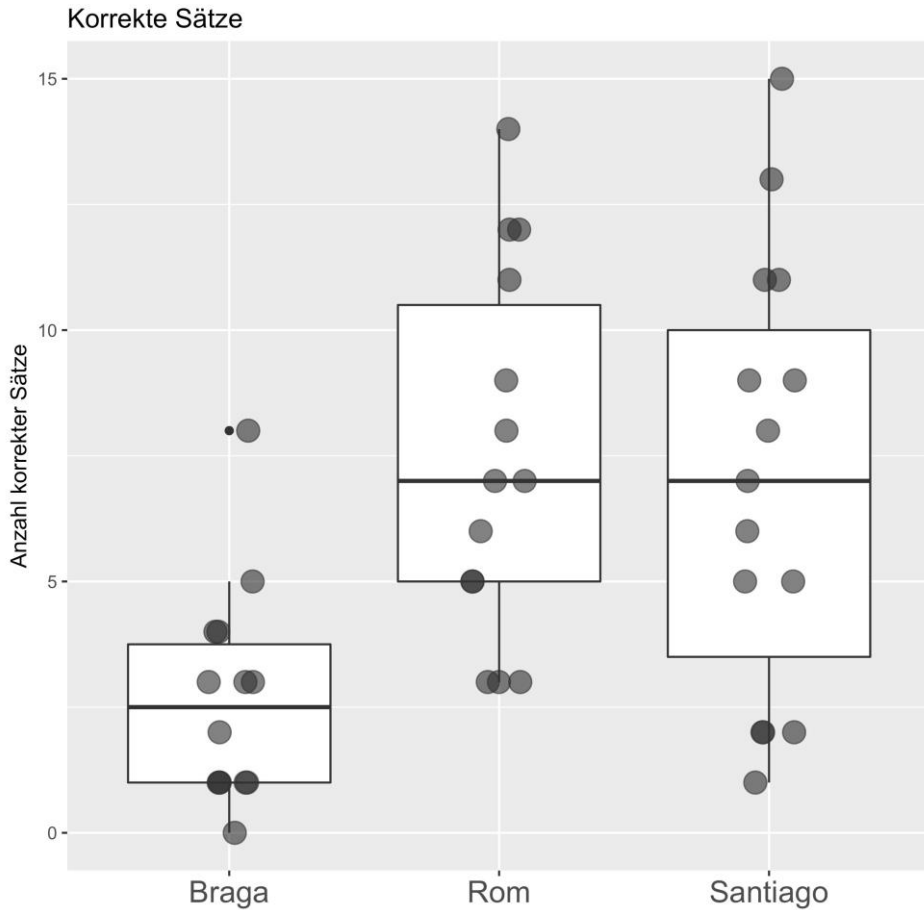


Abbildung 2: Boxplots für den Anteil korrekter Sätze an den drei Standorten.

Wie viele Teilnehmer einen Satz richtig verbessert haben, hängt auch sehr stark vom Satz selbst ab (vgl. Abbildung 3). Ein Satz wie „Leider kann ich heute nicht Tennis spielen. Ich bin zu besetzt.“ (richtig: „Leider kann ich heute nicht Tennis spielen. Ich bin zu beschäftigt“, Satz-ID 2) mit einem falschen Freund auf der Adjektivposition wurde in 70 % aller Fälle verbessert, während der Fehler im Satz „Kein Problem, wenn der Zucker beendet ist; ich nehme dann Honig“ (Satz-ID 4) offensichtlich sehr schwer zu identifizieren und in eine sinnvolle Online-Recherche zu verwandeln war. In diesem Fall wurden nur 11 % der Korrekturen als verbessert markiert.

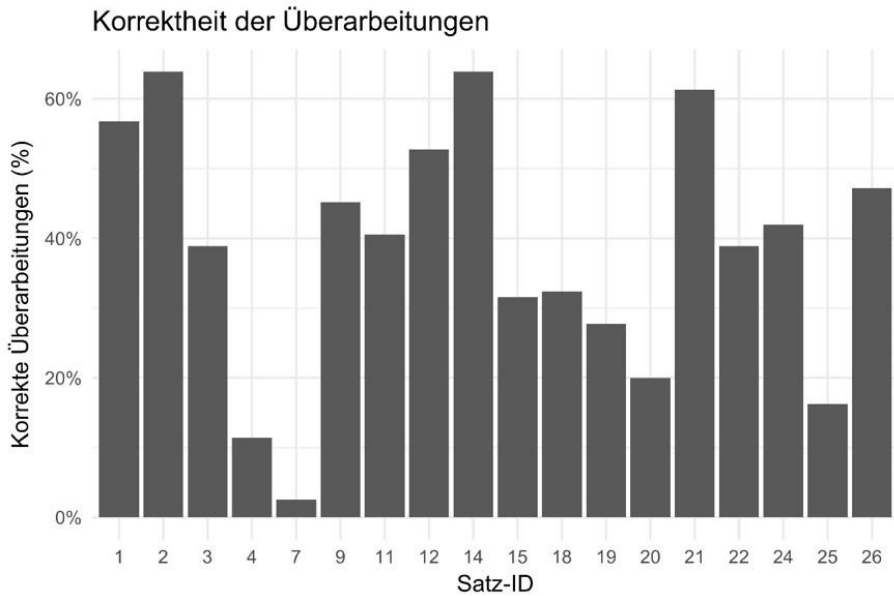


Abbildung 3: Korrektheit der einzelnen Sätze.

Für die Anzahl der Suchvorgänge pro Satz zeigt sich: In 108 Bearbeitungen wurde nur eine Suche durchgeführt, in 233 Bearbeitungen zwei bis drei Suchen und in 246 Fällen mehr als vier (vgl. Abbildung 4).

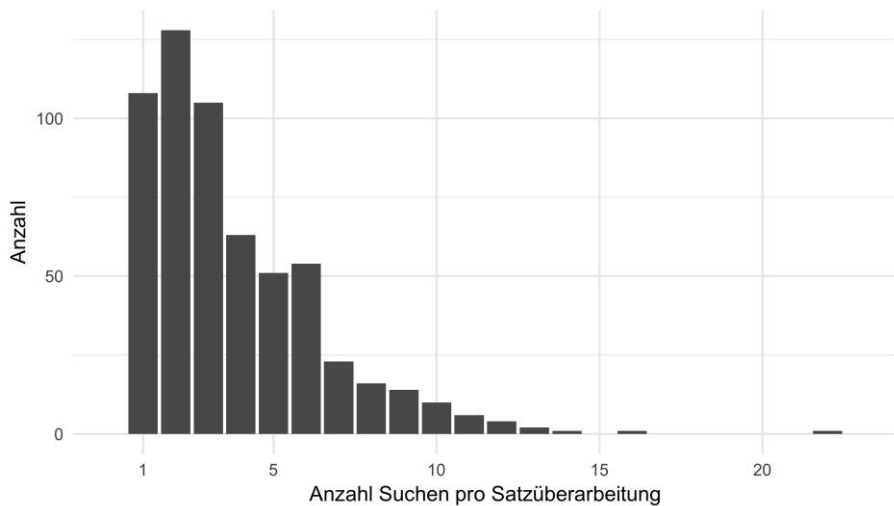


Abbildung 4: Anzahl der Suchvorgänge pro Satzüberarbeitung.

Interessant ist, dass nahezu nur die ersten vier Treffer der Suchergebnisliste aufgerufen wurden (vgl. Abbildung 5), was bedeutet, dass fast niemand gescrollt hat, denn 4 bis 5 Ergebnisse waren auf dem Laptopbildschirm direkt zu sehen, je nachdem, ob das Fenster für den Google Translator eingeblendet war oder nicht (vgl. Abbildung 6). Von diesen haben sich aber mehr als die Hälfte der ProbandInnen (63 %) auf den ersten Treffer konzentriert.

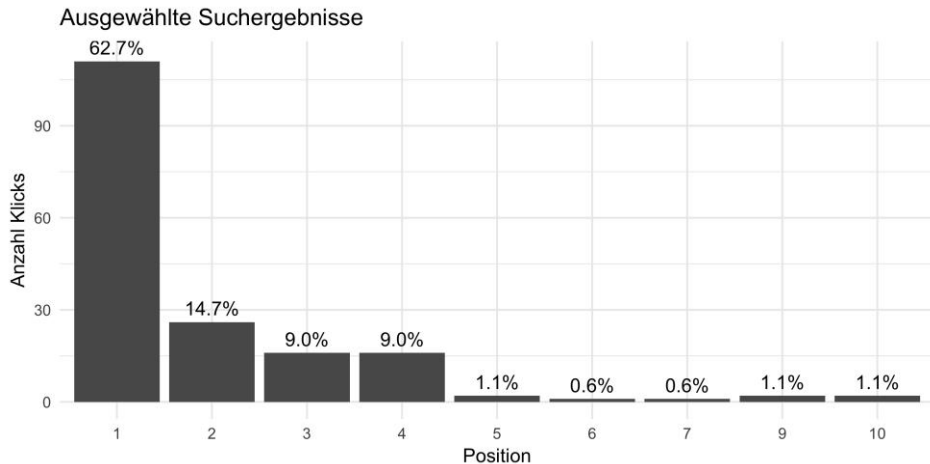


Abbildung 5: Verteilung der ausgewählten Google-Suchergebnisse nach Position des Suchergebnisses in der Ergebnisliste.

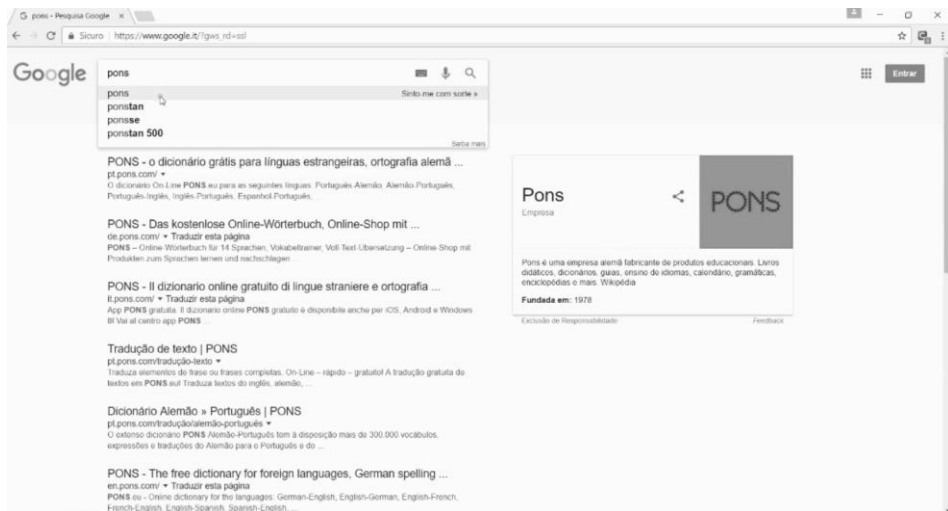


Abbildung 6: Screenshot des Laptopbildschirms (R-05, 00:01:03).

Zunächst einmal zeigt die Studie, dass die ProbandInnen eine große Anzahl von Ressourcen, und vor allem eine große Anzahl von verschiedenen Typen, genutzt haben (vgl. Abbildung 7).

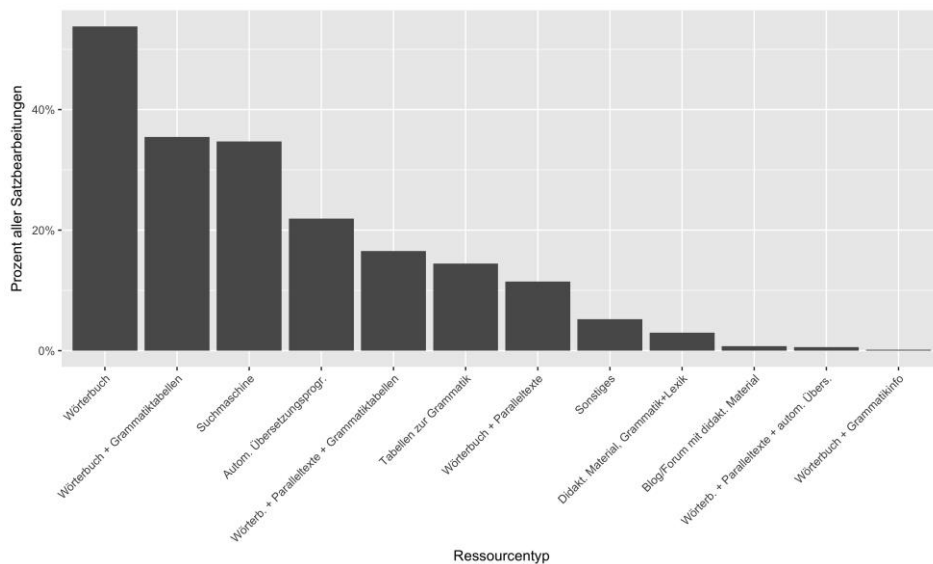


Abbildung 7: Anteil der Typen von Ressourcen an Satzbearbeitungen.¹³

Am meisten werden Wörterbücher oder Seiten genutzt, die Wörterbücher mit Grammatiktabellen erhalten, gefolgt von Suchmaschinen (die auch benutzt werden, um Ressourcen aufzurufen, in dem z.B. ‚Duden online‘ in Google eingegeben wurde). Zwar sind 42 Probandinnen und Probanden keine große Anzahl, dafür existieren – wie bereits erwähnt – etwa 1935 Minuten Filmmaterial mit mehr als 2.200 Suchaktionen. Die Daten sind aber auch in dem Sinne verlässlich, dass die Studierenden direkt bei der Arbeit beobachtet wurden, ohne dass man sich auf deren Selbstauskunft stützen muss, was im Kontext von Sprachunterricht durch den Faktor der sozialen Erwünschtheit stärker verzerrt sein könnte, da die Studierenden in der Regel wissen, was die Dozentinnen und Dozenten gerne hören würden. Es kann jedoch keine Aussage gemacht werden, inwieweit die Aufzeichnung das Arbeiten beeinflusst hat. Die Probanden blieben auf jeden Fall anonym.

¹³ Bei *Wörterbuch + Paralleltex.* handelt es sich um Ressourcen wie *Linguee*, bei denen die Suche in parallelen Übertexten möglich ist. Bei *Wörterb. + Paralleltex. + Grammatiktabellen* geht es um Online-Ressourcen wie *Reverso*. *Wörterb. + Paralleltex. + autom. Übers.* schließt Webseiten ein, die eine Suche im Wörterbuch und in parallelen Übertexten anbietet, aber auch eine zusätzliche Funktion der Textübersetzung bereitstellt, wie z.B. *Poms*.

3.2 Suchstrings

Unterschiedliche Typen von Ressourcen sind für verschiedene Arten von Suchanfragen konzipiert und ihre Benutzung muss deshalb unterschiedlich sein. So sollte man beispielsweise in den meisten Wörterbüchern keine ganzen Sätze eingeben, wohingegen automatische Übersetzungsprogramme umso besser funktionieren, je mehr Kontext gegeben ist. Vielen Fremdsprachenlernenden scheint dies jedoch nicht bewusst zu sein. Im Folgenden soll also der Frage nachgegangen werden, ob diese Bewusstheit bei den Probandinnen und Probanden ersichtlich werden konnte. Bei einer ersten Analyse fiel bereits auf, dass in Suchmaschinen oft der gesuchte sprachliche Ausdruck mit metasprachlichen Begriffen verbunden wurde, z.B. in folgenden Beispielen aus Rom (alles Suchstrings, die in ‚google.it‘ eingegeben wurden): „wenn significato“ (dt. wenn Bedeutung), „costruzione con wann“ (dt. Konstruktion mit wann), „inverstandeni frasi“ (dt. einverstanden Sätze), „handelni coniugazioni“ (dt. handeln Konjugationen), „wenn costruzione“ (dt. wenn Konstruktion), „beendet esempi“ (dt. beendet Beispiele). Um nun überprüfen zu können, ob dies wirklich systematisch so ist, mussten die Suchstrings entsprechend typologisiert und annotiert werden. Zunächst einmal wurden die Suchstrings in drei Hauptklassen unterteilt: Unter die *metasprachlichen Begriffe* fallen Anfragen wie „Konjunktiv 2 mit wenn“, unter die *nicht-sprachlichen Anfragen* fallen die Namen von Ressourcen und unter die *sprachlichen Anfragen* fallen Anfragen nach konkreten sprachlichen Strukturen. Die letzte Kategorie ist von Interesse, wenn analysiert werden soll, ob die Studierenden die Arten von sprachlichen Anfragen je nach Typen von Ressourcen unterschiedlich ausgestalten. Aus diesem Grund wurden die sprachlichen Anfragen in zwei Klassen eingeteilt: *einzelne Wortklassen* (mit entsprechenden Unterkategorien) vs. *Komplexe* (auch mit entsprechenden Unterkategorien).¹⁴ Mit Komplexen sind dabei sowohl ganze Sätze als auch Phrasen etc. gemeint. Sinnvoll wäre es dabei, wenn die Studierenden bei den Wörterbüchern eher nach einzelnen Wörtern suchen würden, d.h. nach einzelnen Wortklassen, wohingegen sie in automatischen Übersetzungsprogrammen eher nach umfangreicheren sprachlichen Strukturen, d.h. Komplexen nachschlagen sollten. Dass sie genau dies tun, zeigt die Auswertung aus Santiago in der folgenden Abbildung.

¹⁴ Als *Komplexe* wurden Mehrwortausdrücke bezeichnet.

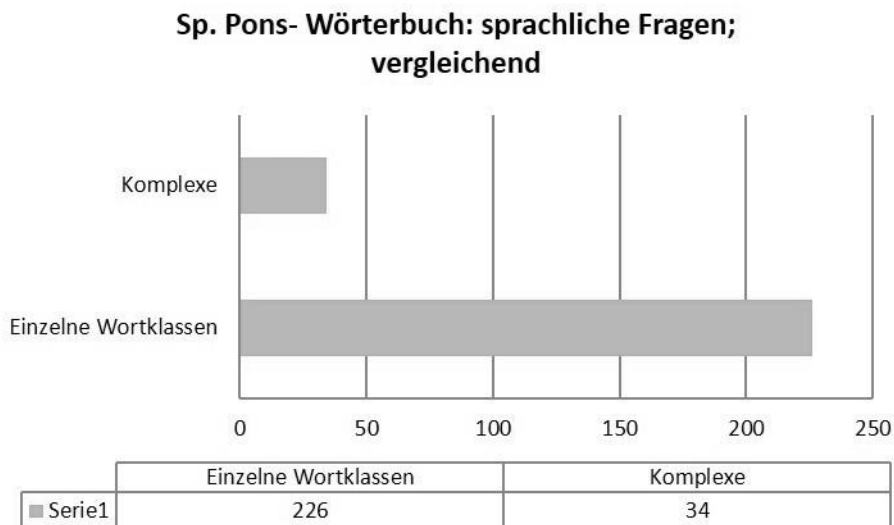


Abbildung 8: Suchanfragen nach einzelnen Wortklassen vs. Komplexen im Pons Wörterbuch.

Bei allen Probanden konnte festgestellt werden, dass sie sehr klar nach den Funktionalitäten der einzelnen Tools unterscheiden und wenig komplexe Phrasen im Wörterbuch und kaum einzelne Wörter in Übersetzungsprogrammen suchen.

3.3 Einflussfaktoren für eine korrekte Verbesserung der Sätze

3.3.1 Typ der Ressource

Einen deutlichen Einfluss auf die Korrektheitsrate hat der Typ der verwendeten Ressourcen. Dabei sind insbesondere zwei Aspekte auffällig. Erstens sind jene Teilnehmenden erfolgreicher, die mehr Wörterbuchressourcen benutzen. Der Zusammenhang ist in Abbildung 9 grafisch dargestellt und statistisch signifikant ($\beta = 0,47$; $p = 0,0001$). Es lässt sich außerdem ablesen, dass dieser Zusammenhang insbesondere von den Teilnehmenden aus Braga gespeist wird, die nur wenige Sätze korrekt überarbeiten konnten. Zweitens lässt sich zeigen, dass die Probandinnen und Probanden, die mehr auf den Einsatz automatischer Übersetzungsprogramme setzten, schlechtere Überarbeitungsergebnisse erzielt haben (ohne Abbildung). Auch dieser Zusammenhang ist statistisch signifikant ($\beta = -0,41$; $p = 0,002$). Zumindest im Kontext der vorliegenden Studie zeigt sich, dass die Verwendung von Wörterbuchressourcen messbar erfolgsversprechender ist – insbesondere im Gegensatz zu automatischen Übersetzungsprogrammen.

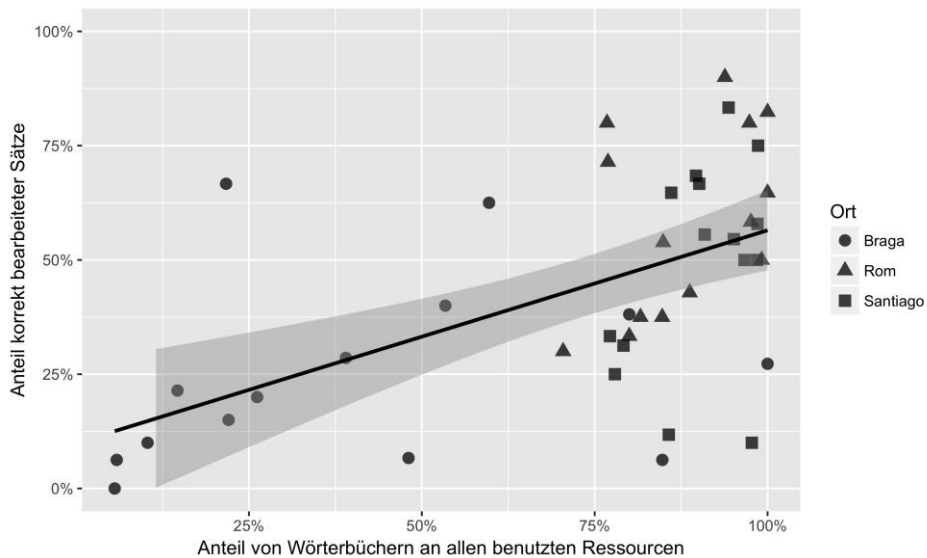


Abbildung 9: Zusammenhang zwischen dem Anteil von Wörterbüchern an allen konsultierten Ressourcen und dem Anteil korrekt verbesserter Sätze (ein Datenpunkt entspricht einer/-m Teilnehmenden). Die Form der Datenpunkte kodiert außerdem den Standort.

3.3.2 Zeit, Sprach(lern)bewusstheit und Motivation

Beim Betrachten der Daten entstand der Eindruck, dass Studierende mit weniger korrekten Sätzen weniger Zeit auf den einzelnen Ressourcen verbrachten. Dies war der Ausgangspunkt, die für eine Ressource aufgewendete Zeit mit der Genauigkeit der Satzkorrektur in Beziehung zu setzen. Es konnte herausgefunden werden, dass der durchschnittliche Zeitaufwand auf einer Ressource eine entscheidende Rolle für die Korrektheit des Satzes spielt (vgl. Abbildung 10). Bei korrekt gelösten Sätzen ist die zeitliche Verweildauer auf den Ressourcen statistisch signifikant höher als bei nicht korrekt verbesserten Sätzen. Mit 2,4 Sekunden ist der Durchschnitt zwar recht gering, doch muss beachtet werden, dass dieser Unterschied bedeutet, dass – im Durchschnitt – die Zeit, die für jede einzelne Ressource aufgewendet wird, 2,4 Sekunden länger ist, wenn sie einen Satz bearbeiten, der zu einem korrekten Satz führt.

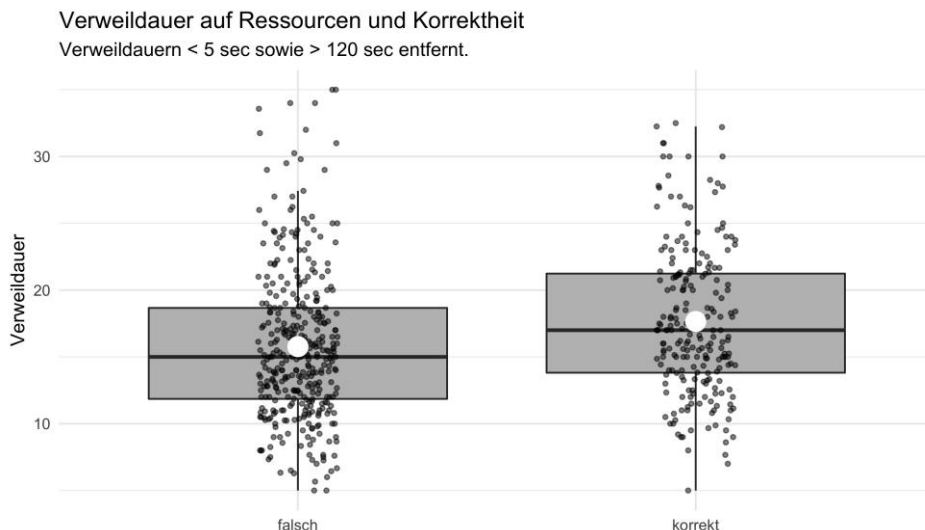


Abbildung 10: Boxplots für die Zeit, die mit Ressourcen für falsch und korrekt überarbeitete Sätze verbracht wurde.

Zu vermuten ist, dass die Verweildauer auf den Ressourcen ein Anzeichen für den Grad der Lösungskompetenz ist: Sowohl eine reflektierte Nutzung als auch ein starker ausgeprägter Wille zur Problemlösung könnten Ursache für erfolgreichere Verbesserungen sein.

Offensichtlich ist, dass die durchschnittliche Zeit, die auf einer Ressource verbracht wird, mit häufigem Zappen zwischen den Ressourcen zusammenhängt. Je weniger Zeit der/die Proband/in auf der Ressource verweilt und je öfters er/sie zwischen den Ressourcen zappt, umso geringer ist die Korrektheit der Sätze. Ein anschauliches Beispiel bietet Probandin R-04, die nur wenige Sätze korrigieren konnte und an Satz 1 nur 4'05" arbeitete und in dieser Zeit 20 Aktionen unternahm, was eine Durchschnittsdauer von 11'7" auf einer einzelnen Ressource ergibt (vgl. Tabelle 2).

Oft wird das kurze Verweilen auf einer Ressource und das schnelle Switchen zwischen den Ressourcen auch von den ProbandInnen versprachlicht. So lassen sich bei diesen Studierenden häufig Ausdrücke wie „un attimo“ oder „un attimino“ (italienisch, dt.: einen Augenblick, einen Moment, (mal) kurz) finden. Zudem wird die Verantwortung für eine fehlgeschlagene Recherche häufig von sich gewiesen, was sich in der 3. Person Singular oder Plural zeigt: „non so cosa intendano“ (Italienisch, dt: Ich weiß nicht, was sie meinen) oder „cerco su un'altra pagina se mi da degli esempi“ (Italienisch, dt.: ich suche auf einer anderen Seite ob sie mir andere Beispiele geben/zeigen).

Tabelle 2: Auszug aus der Transkription von Proband R-04 für den Satz 1.

	Timestamp	Sek./ Aktion	Name Res- source	Aktion
1	00:00:16		Excel	liest Satz 1
2	00:00:29	13	Google.it	öffnet Browser, öffnet Google, googelt „leo dizionario“
3	00:01:09	40	LEO Dictionary	öffnet Leo,
4	00:01:15	6	Excel	öffnet Excel
5	00:01:36	21	LEO Dictionary	öffnet Leo, sucht „forschung“
6	00:01:44	8	Excel	öffnet Excel
7	00:01:57	13	LEO Dictionary	öffnet Leo, sucht „verfügung“
8	00:02:09	12	LEO Dictionary	sucht „verfügung“
9	00:02:18	9	Excel	öffnet Excel
10	00:02:38	20	LEO Dictionary	öffnet Leo
11	00:02:41	3	Excel	öffnet Excel
12	00:03:02	21	LEO Dictionary	öffnet Leo
13	00:03:04	2	Excel	öffnet Excel
14	00:03:23	19	LEO Dictionary	öffnet Leo
15	00:03:26	3	Excel	öffnet Excel
16	00:03:59	33	LEO Dictionary	öffnet Leo
17	00:04:05	6	Excel	öffnet Excel
18	00:04:07	2	LEO Dictionary	öffnet Leo
19	00:04:09	3	Excel	öffnet Excel, keine Korrektur
20	00:04:21		Excel	liest Satz 2

In Tabelle 3 dagegen wird ein Auszug aus dem TAP einer Studierenden gezeigt, die bessere Ergebnisse erzielt hat. Sie ruft die Ressourcen oft gezielt auf, was bedeutet, dass sie weiß, wo sie was finden kann. Zudem reflektiert sie detailliert und verwen-

det dabei viele linguistische Begriffe wie bspw. Kollokation. Auch werden sprachliche Strategien verwendet, wie bspw. die Suche anhand von Synonymen. Führt der Suchvorgang noch zu keinem zufriedenstellenden Ergebnis, so gibt sie nicht auf, sucht weiter und benutzt andere Strategien, z.B. die Suche über das Englische, oder stellt die Sprachrichtung um und macht eine Gegenkontrolle.

Tabelle 3: Ausschnitt aus dem ins Deutsche übersetzten TAP von R-03 zum Satz „Kein Problem, wenn der Zucker beendet ist; ich nehme dann Honig“.

okay ahm # ich muss die geeignete Kollokation finden ahm ich suche sie auf Italienisch
okay # es gibt ein Beispiel mit (a quanto pare è finito lo zucchero) ahm aber die deutsche Version ist ein bisschen komplizierter die Tatsache das hier (a quanto pare) steht also ich such mal kurz was es ist aber (ist) ist nicht geändert also vielleicht muss ich es nicht korrigieren also ahm ich versuche es mit einer anderen Nahrung
es gibt überhaupt keine andere Formulierung wie (non c'è latte non c'è più zucchero) # vielleicht ausgegangen aber
ausgegangen scheint mir ein bisschen es scheint mir ein bisschen zu extrem ich weiss nicht warum wahrscheinlich weil ich eine Verbindung zum Italienischen und deshalb mit allem das was es bedeuten könnte mache aber wieso nicht ich suche es also auf Italienisch da finde ich es als (uscito) dann hat es bestimmt noch eine andere Bedeutungen aber # ich würde versuchen es im Englischen zu suchen weil es hat eine spezifischere Formulierung also
zu viele Wörter im Englischen deswegen findet es nichts vielleicht muss ich es umstellen ja okay
da steht ausgehen ausgehen ja wahrscheinlich ist es das da # ja weil hier ist es eben eindeutig in verschiedenen Kontexten ja vielleicht ist es das hier

3.3.3 Die richtige Ausgangshypothese für eine Korrektur

Neben diesen Faktoren scheint es aber auch sehr wichtig zu sein, dass die Studierenden überhaupt auf der richtigen Spur zur Suche des Fehlers sind, also die richtige Hypothese haben. Dies soll hier an einem Beispiel aus Braga gezeigt werden.

Einer der Stimulus-Sätze in unserer Studie war: „An unserem Forschungsinstitut *ist* Ihnen unsere Bibliothek 24 Stunden *zur Verfügung*“ (Satz Nr. 11). Der Proband B-08 geht von der Annahme aus, dass das Verb „gibt“ und nicht das Verb „ist“ in Verbindung mit „Verfügung“ korrekt ist. Er sagt dazu: „hier ist die Verwendung von (ist) mit (Verfügung) nicht korrekt # ich glaube es bleibt (gibt) anstatt (ist) #“. Alle weiteren Recherchen gehen von dieser Annahme aus. Der Proband gibt als Suchstring „gibt zur Verfügung“ in Google (google.pt) ein. Der erste

Treffer der Google-Suchergebnisliste lautet: „zur Verfügung gibt – Englisch Übersetzung – Linguee Wörterbuch“. Darauf sagt er: „exakt es ist (gibt) das benutzt werden soll denn es ist <...> korrekt #“ und korrigiert den Ausgangssatz entsprechend (falsch): „An unserem Forschungsinstitut gibt Ihnen unsere Bibliothek 24 Stunden zur Verfügung.“ Obwohl das erste Suchergebnis auf das Online-Wörterbuch Linguee verweist und die nachfolgenden Treffer andere Kombinationen aus „Verb + zur Verfügung“ enthalten, sucht der Proband nicht weiter. Seine komplette Recherche ist also nur dieser einen Frage gewidmet – in diesem Kontext wird das Verb „gibt“ in Verbindung mit „zur Verfügung“ verwendet – und führt in diesem Fall zu keinem zufrieden stellenden Ergebnis.

In anderen Beispielen zeigt sich sogar, dass eine falsche Ausgangshypothese dazu führen kann, dass – obwohl die korrekte Lösung im Wörterbuch steht – sie nicht gesehen wird.

4 Konsequenzen für die Fremdsprachendidaktik

Zusammenfassend lässt sich also sagen, dass mehrere Einflussfaktoren herausgefunden werden konnten, die die Korrektheit der Verbesserungen beeinflussen:

Erstens ist der *Typ der verwendeten Ressource* wichtig: Bei der Verwendung von Wörterbüchern ist eine korrekte Verbesserung in der vorliegenden Studie wahrscheinlicher als bei der Verwendung von automatischen Übersetzungsprogrammen. Zudem hat sich gezeigt, dass die Probandinnen und Probanden eine große Bandbreite an Ressourcen benutzen, obwohl sie sich noch nicht auf einem fortgeschrittenen Lernerniveau befinden. Wörterbücher werden deutlich häufiger benutzt als automatische Übersetzungsprogramme. Dabei unterscheiden die Studierenden zwischen den einzelnen Typen von Ressourcen, verfügen also anscheinend über ausreichendes Basiswissen für einen differenzierten Gebrauch der Ressourcen.

Zum zweiten ist der *Zeitfaktor* entscheidend. Bei korrekt verbesserten Sätzen wurde auf den aufgerufenen Ressourcen länger verweilt. Dies scheint ein Indikator für ein kompetenteres oder reflektiertes Vorgehen (*language (learning) awareness*) zu sein oder schlicht den Willen zu zeigen, eine Lösung finden zu wollen und zu können. Damit verbunden ist sicherlich das Niveau der Fremdsprachenkenntnisse. Da leider kein Einstufungstest vorgenommen wurde, kann darüber keine genauere Angabe gemacht werden.

Darüber hinaus ist es wichtig, zu wissen, in welcher Ressource man für ein bestimmtes Problem nachschlagen sollte und welche Informationen man dort findet. Je besser also die *Wörterbuchbenutzungskompetenz* ist, umso wahrscheinlicher ist es, den Satz korrigieren zu können – vorausgesetzt, man hat *die richtige Ausgangshypothese*.

Welche Konsequenzen sollte man aus diesen Ergebnissen ziehen? Die Wörterbuchdidaktik ist in den Fremdsprachenunterricht sowohl explizit als auch impli-

zit zu integrieren. Die Reflexion über das eigene Benutzungsverhalten spielt dabei eine sehr wichtige Rolle.¹⁵ Wie Zöfgen bereits 2010 fordert, sollte „die Wörterbuchdidaktik im Prinzip auf die Ergebnisse der Metalexikographie, insbesondere der Benutzungsforschung, zurückgreifen.“ (ebd.: 108) Er bedauert zudem, dass sich die Erkenntnisse bisher noch nicht auf die Unterrichtspraxis ausgewirkt haben. Die Vermittlung von Wissen über verschiedene Ressourcen, die Kenntnis darüber, wo man was nachschlagen kann oder sollte, und konstruktive Hinweise zum Lesen von lexikografischen Informationen sollten die Studierenden aber immer dort abholen, wo sie stehen.

Eine erste Möglichkeit könnte sein, diese Studie und ihre Ergebnisse als Anlass zu nehmen, mit den eigenen Studierenden und Sprachlernenden über *ihr* Wörterbuchbenutzungsverhalten zu diskutieren, welche Ressourcen sie nutzen, wann, warum, zu welchem Zweck, wie viel Zeit sie sich nehmen, ob sie häufig switchen, usw. und welche Strategien sie dabei verwenden. Zum Einstieg könnte die Studie – ohne Videoaufzeichnung, aber evtl. zu zweit mit gegenseitigem Protokollieren – durchgeführt werden.

Zweitens sollten über die von den Lernenden verwendeten Sprachressourcen diskutiert werden. Dabei können nicht nur pauschal Stärken und Schwächen festgestellt werden, sondern es könnte auch darüber diskutiert werden, welche Ressource für welche Art von sprachlichen Schwierigkeiten geeignet ist. Ein zweisprachiges Online-Wörterbuch kann bspw. hervorragende Dienste für die Rezeption leisten, auch wenn es bei der Produktion in die Fremdsprache versagt. Mit den von den Lernenden häufig verwendeten Übersetzungsprogrammen kann experimentiert werden, um ihre Funktionsweise zu ergründen. Ein zentraler Teil der Didaktik sollte dabei immer das Reflektieren über bestimmte Handlungen, v.a. über die eigenen Handlungen sein.

Um diese Forderungen tatsächlich und nachhaltig umsetzen zu können, ist es nicht nur zu empfehlen, die eigenen Lernenden in den Fokus der Diskussion und der Wörterbuchdidaktik zu stellen. Ein weiterer wichtiger Schritt wäre es, in der Lehreraus- und -fortbildung und gleichzeitig auf bildungspolitischer Ebene anzusetzen, denn solange Lehrkräfte selbst nur geringe Vorkenntnisse in Bezug auf Online-Ressourcen vorweisen können und oft nicht wissen, wo ihre Lernenden recherchieren und wo sie „abzuholen“ sind, und solange in den Prüfungen nur gedruckte (einsprachige) Wörterbücher zugelassen und Online-Ressourcen ausgeblendet werden, solange ist der Weg für eine konkrete und sinnvolle, an den Be-

¹⁵ Im November 2018 wurden von 18 SpezialistInnen aus der Wörterbuchforschung, der praktischen Lexikografie, dem Bereich Deutsch als Fremdsprache, der Italianistik, den Translationswissenschaften und der Empirischen Linguistik die *15 Villa-Vigoni-Thesen zur Lexikografie* verabschiedet. In These 14 wird spezifisch auf die Forderung nach einer Wörterbuchdidaktik eingegangen: „14. Das digitale Datenangebot in den Informationssystemen der Zukunft muss als wichtiges Hilfsmittel des *lifelong learning* angesehen werden, so dass die **kritische Benutzung der Ressourcen** als **strategische Schlüsselkompetenz** etabliert wird. Dies muss auch in der **Aus- und Fortbildung von LehrerInnen** verankert werden.“ (Hervorhebung im Original) (<https://www.emlex.phil.fau.de/15-villa-vigoni-thesen-zur-lexikografie-verabschiedet/> [30.03.2019]).

nutzerInnen orientierten Wörterbuchdidaktik im Fremdsprachenunterricht versperst. Das sollten wir ändern.

Literatur

- De Schryver, Gilles-Maurice; Joffe, David; Joffe, Pitta; Hillewaert, Sarah (2006): Do Dictionary Users Really Look up Frequent Words? – On the Overestimation of the Value of Corpus-Based Lexicography. In: *Lexikos* 16, 67-83.
- Domínguez Vázquez, María José; Mirazo Balsa, Mónica; Vidal Pérez, Vanessa (2013): Wörterbuchbenutzung: Erwartungen und Bedürfnisse. Ergebnisse einer Umfrage bei Deutsch lernenden Hispanophonen. In: Domínguez Vázquez, María José (Hrsg.): *Trends in der deutsch-spanischen Lexikographie*, Frankfurt am Main: Peter Lang, 135-172.
- Domínguez Vázquez, María José; Valcárcel Riveiro, Carlos (2015): Hábitos de uso de los diccionarios entre los estudiantes universitarios europeos: ¿nuevas tendencias? In: Domínguez Vázquez, María José; Gómez Guinovart, Xavier; Valcárcel Riveiro, Carlos (Hrsg.): *Lexicografía de las lenguas románicas II. Aproximaciones a la lexicografía contemporánea y contrastiva*. Berlin/New York: de Gruyter, 165-189.
- Ericsson, K. Anders; Simon, Herbert A. (1993): *Protocol Analysis: Verbal Reports as Data*. A Bradford Book. London: The MIT Press.
- Europarat (2001): *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. Deutsche Übersetzung von J. Quetz. Straßburg: Europarat.
- Hult, Ann-Kristin (2012): Old and New User Study Methods Combined – Linking Web Questionnaires with Log Files from the Swedish Lexin Dictionary. In: Torjusen, Julie Matilde; Vatvedt Fjeld, Ruth (Eds.): *Proceedings of the 15th EURALEX International Congress 2012*. Oslo, Norway, 922-928. (http://www.euralex.org/elx_proceedings/Euralex2012/pp922-928%20Hult.pdf [27.09.2018])
- Koplenig, Alexander; Meyer, Peter; Müller-Spitzer, Carolin (2014): Dictionary users do look up frequent words. A log file analysis. In: Müller-Spitzer, Carolin (Hrsg.): *Using Online Dictionaries*. Berlin/Boston: de Gruyter, 229-250.
- Levy, Mike; Steel, Caroline (2015): Language learner perspectives on the functionality and use of electronic language dictionaries. In: *ReCALL* 27, 177-196.
- Lew, Robert (2010): Users Take Shortcuts: Navigating Dictionary Entries. In: Dykstra, Anne; Schoonheim, Tanneke (Hrsg.): *Proceedings of the XIV Euralex International Congress*. Ljouwert: Afûk, 1121-1132.

- Lew, Robert (2011): Studies in Dictionary use. Recent developments. In: *International Journal of Lexicography* 24/1. Oxford: Oxford University Press, 1-4.
- Lew, Robert (2015): Research into the use of online dictionaries. In: *International Journal of Lexicography* 28, 232-253.
- Lew, Robert; Grzelak, Marcin; Leszkowicz, Mateusz (2013): How Dictionary Users Choose Senses in Bilingual Dictionary Entries: An Eye-Tracking Study. In: *Lexikos* 23, 228-254.
- Merten, Stephan (2011): Arbeit mit Wörterbüchern. In: Pohl, Inge; Ulrich, Winfried (Hrsg.): *Wortschatzarbeit*. Baltmansweiler: Schneider Verlag Hohengehren, 348-360.
- Müller-Spitzer, Carolin (2014) (Hrsg.): *Using Online Dictionaries*. Berlin/Boston: de Gruyter.
- Müller-Spitzer, Carolin; Domínguez Vazquez, Maria Jose; Nied Curcio, Martina; Silva Dias, Idalete Maria; Wolfer, Sascha (2018): Correct Hypotheses and Careful Reading Are Essential: Results of an Observational Study on Learners. Using Online Language Resources. In: *Lexikos* 28, 287-315.
- Nesi, Hillary (2012): Alternative e-dictionaries: Uncovering dark practices. In: Granger, Sylvaine; Paquot, Magali (Hrsg.): *Electronic Lexicography*. Oxford: Oxford University Press, 357-372.
- Nied Curcio, Martina (2014): Die Benutzung von Smartphones im Fremdsprachenerwerb und -unterricht. In: Abel, Andrea; Vettori, Chiara; Ralli, Natascia (Hrsg.): *EURALEX International congress: The user in focus. 15-19 luglio 2014*, Bolzano: Eurac, 263-280.
- Roelcke, Thorsten (Hrsg.) (2015): Themenreihe: Wörterbücher für Deutsch als Fremdsprache – Probleme und Perspektiven. In: *Info DaF* 5. München: Iudicium.
- Snell-Hornby, Mary (1987): Towards a learner's bilingual dictionary. In: Cowie, Anthony (Hrsg.): *The Dictionary and the Language Learner. Papers from the EURALEX Seminar at the University of Leeds, 1-3 April 1985*. Tübingen: Niemeyer, 159-170.
- Töpel, Antje (2014): Review of research into the use of electronic dictionaries. In: Müller-Spitzer, Carolin (Hrsg.): *Using online dictionaries*. Berlin/Boston: de Gruyter, 13-54.
- Welker, Herbert Andreas (2013): Empirical research into dictionary use since 1990. In: Gouws, Rufus H.; Heid, Ulrich; Schweickard, Wolfgang; Wiegand, Herbert Ernst (Hrsg.): *Dictionaries. An international encyclopedia of lexicography*. Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography. Berlin/New York: de Gruyter, 531-540.

- Wiegand, Herbert Ernst (1998): *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie*. Berlin/New York: de Gruyter.
- Wiegand, Herbert Ernst; Beißwanger, Michael; Gouws, Rufus H.; Kammerer, Matthias; Storrer, Angelika; Wolski, Werner (Hrsg.) (2010): *Wörterbuch zur Lexikographie und Wörterbuchforschung. Dictionary of Lexicography and Dictionary Research*. Band 1. Berlin New/York: de Gruyter.
- Wolfer, Sascha; Nied Curcio, Martina; Silva Dias, Idalete Maria; Müller-Spitzer, Carolin; Domínguez Vázquez, María José (2018): Combining Quantitative and Qualitative Methods in a Study on Dictionary Use. In: Čibej, Jaka; Gorjanc, Vojko; Kosem, Iztok; Krek, Simon (Hrsg.): *Lexicography in Global Contexts. Proceedings of the XVIII EURALEX International Congress 17-21 July 2018*. Ljubljana, Ljubljana University Press, 101-112.
- Wolfer, Sascha/Bartz, Thomas; Weber, Tassja; Abel, Andrea; Meyer, Christian M.; Müller-Spitzer, Carolin; Storrer, Angelika (2018): The effectiveness of lexicographic tools for optimising written L1-texts. In: *International journal of lexicography* 31(1). S. 1-28. (<https://doi.org/10.1093/ijl/ecw038> [25.01.2021])
- Zöfgen, Eberhard (2010): Wörterbuchdidaktik. In: Königs, Frank G.; Hallet, Wolfgang (Hrsg.): *Handbuch Fremdsprachenunterricht*. Seelze-Velber: Klett Kallmeyer, 107-110.