universität wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Wahrnehmung der Übersetzungsqualität von Koch- bzw. Backrezepten verschiedener Systeme der neuronalen maschinellen Übersetzung durch Kochbegeisterte

verfasst von | submitted by

Melanie Leko BA

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Arts (MA)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 070 331 342
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Translation Deutsch Englisch
Betreut von | Supervisor:	Assoz. Prof. Mag. Dr. Dagmar Gromann BSc

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

1. Einleitung

Mit der Globalisierung und der damit verbundenen hohen Übersetzungsnachfrage wird Maschinelle Übersetzung (MÜ) sowohl von professionellen Übersetzer*innen als auch von vielen Menschen im Alltag, zum Beispiel als Kommunikationsmittel, auf Reisen, im Unterricht, zunehmend verwendet (vgl. Koehn 2020). Der neueste Ansatz von MÜ, die Neuronale Maschinelle Übersetzung (NMÜ), basiert auf Künstlicher Intelligenz (KI) beziehungsweise neuronalen Netzwerken und führte zu einer höheren Qualität der Ergebnisse maschineller Übersetzung (Forcada 2017: 292; Koehn 2020: 29; Schmalz 2019: 195). Dementsprechend wurde die NMÜ zum aktuellen Stand der Technik, was sich sowohl in den für die breite Öffentlichkeit online verfügbaren Tools als auch in der aktuellen Forschung auf dem Gebiet der MÜ widerspiegelt (Koehn 2020: 40).

Die Qualität maschineller Übersetzungen ist Gegenstand zahlreicher Studien, die in diesem Kontext zu unterschiedlichen Ergebnissen gelangen. So kamen etwa Hassan et al. (2018) zum Schluss, dass die Übersetzungsqualität von Ergebnissen der MÜ gleichwertig mit jener von Humanübersetzungen ist. Diese Schlussfolgerung konnte jedoch in anderen Studien (z. B. Läubli et al. 2018; Wiesmann 2019; Heinisch & Lušicky 2019) nicht bestätigt werden. Die Qualität von MÜ kommt in bestimmten Fachgebieten, zum Beispiel Recht, noch nicht an die Humanübersetzung heran (Wiesmann 2019: 118). Heinisch und Lušicky (2019: 42) kamen in ihrer Studie zum Ergebnis, dass die Übersetzungsqualität von NMÜ nicht so hoch ist, wie von angehenden Übersetzer*innen erwartet wurde. Neuere Untersuchungen deuten jedoch auf erhebliche Fortschritte der NMÜ hin. So gelangten Popel et al. (2020) zum Ergebnis, dass moderne NMÜ-Systeme bei der Übersetzung von Nachrichtentexten durchaus an die Qualität von Humanübersetzungen herankommen können. Hiebl und Gromann (2024) zeigen in diesem Kontext auf, dass die Qualität von DeepL in bestimmten Fachgebieten mit jener von Humanübersetzungen konkurrieren kann.

Mit der Entwicklung des Internets nahm die Menge an verfügbaren Texten, insbesondere in Form von Websites, rasant zu. Durch die Globalisierung ist der Übersetzungsbedarf entsprechend hoch. In diesem Zusammenhang scheint MÜ eine geeignete Lösung für die Bewältigung dieser steigenden Textmengen zu sein. MÜ muss dahingehend nicht als Gefahr gesehen werden (vgl. Koehn 2020: 21). Vielmehr ist sie ein Tool, das für Übersetzer*innen vor allem dann hilfreich sein kann, wenn sie es richtig einsetzen (Burchardt & Porsiel 2017: 17). Darüber hinaus bedarf es einer Klärung, ob der Einsatz von MÜ für ein bestimmtes Sachgebiet beziehungsweise eine spezifische Textsorte sinnvoll ist (Burchardt & Porsiel 2017: 17). In

einem nächsten Schritt ist es empfehlenswert, mehrere MÜ-Tools zu testen und deren Ergebnisse miteinander zu vergleichen, um unter Berücksichtigung der eigenen Bedürfnisse das beste Tool zu finden (Rossi & Carré 2022: 56). Demnach ist es sinnvoll, die Ergebnisse der bekanntesten Systeme maschineller Übersetzung wie DeepL (DeepL 2024a), Google Translate (Google Übersetzer 2024a), Microsoft Bing Translator (Microsoft 2024a) oder SYSTRAN (SYSTRAN 2024a) zu vergleichen (vgl. Rivera-Trigueros 2022: 612).

Bislang wurden zahlreiche Studien zur Bewertung der Qualität maschinell übersetzter Texte durchgeführt (z. B. Heinisch & Lušicky 2019; Wiesmann 2019). In diesen Studien handelte es sich bei den Teilnehmer*innen jedoch hauptsächlich um professionelle Übersetzer*innen oder Studierende der Translationswissenschaft. Da das Zielpublikum dieser Texte nicht unbedingt ausschließlich aus Translator*innen besteht, sondern aus einem sachgebietsspezifischen Publikum, das sowohl Fachexpert*innen als auch Personen mit privatem Interesse an einem bestimmten Sachgebiet umfassen kann, ist es sicherlich aufschlussreich zu untersuchen, wie dieses die Qualität maschineller Übersetzungen im Hinblick auf Richtigkeit und Angemessenheit der Inhalte bewertet. Zur Bewertung des allgemeinen Textverständnisses kann die Miteinbeziehung von Personen ohne sachgebietsspezifische Vorkenntnisse sicherlich von Nutzen sein.

1.1. Forschungsfrage und Annahmen

Aus den Überlegungen heraus, dass man klären sollte, ob der Einsatz von MÜ für ein bestimmtes Sachgebiet beziehungsweise eine spezifische Textsorte sinnvoll ist, und das entsprechende Zielpublikum in diesen Bewertungsprozess miteinbezogen werden sollte, leitet sich die folgende Forschungsfrage ab: Wie bewerten Kochbegeisterte die Übersetzung von Rezepten aus dem Englischen ins Deutsche durch zwei Systeme neuronaler maschineller Übersetzung? Zur weiteren Eingrenzung des Forschungsvorhabens sollen in diesem Zusammenhang folgende Unterfragen beantwortet werden:

- Welches NMÜ-System erzielt aus Sicht der Befragten die besten und welches die schlechtesten Ergebnisse bei der Übersetzung von Rezepten?
- Inwiefern können die maschinellen Übersetzungen von Rezepten eine erfolgreiche Zubereitung der entsprechenden Gerichte gewährleisten und welche Einschränkungen werden hierbei festgestellt?
- Inwiefern hat das kulinarische Hintergrundwissen der Befragten Einfluss auf die Bewertung?

Durch die kulturellen Besonderheiten von Rezepten als Textsorte und die angenommene terminologische Inkonsistenz von MÜ-Systemen kann hierbei angenommen werden, dass die Qualität von MÜ im Allgemeinen nicht ausreicht, um die Anleitungen aus den Rezepten ohne Missverständnisse zu befolgen. Darüber hinaus ist davon auszugehen, dass die Befragten, je ausgeprägter ihre Kenntnisse zu den gängigen Prozessen auf diesem Gebiet sind, die maschinellen Übersetzungen eher akzeptieren und davon Gebrauch machen können. Im Gegensatz dazu dürfte die Gesamtbewertung der MÜ bei einem höheren Grad des prozeduralen Wissens schlechter ausfallen.

1.2. Motivation und Zielsetzung

Bislang wurden zahlreiche Studien zur Bewertung der Qualität maschinell übersetzter Texte durchgeführt (z. B. Heinisch & Lušicky 2019; Wiesmann 2019). In diesen Studien handelte es sich bei den Teilnehmer*innen jedoch hauptsächlich um professionelle Übersetzer*innen oder Studierende der Translationswissenschaft. Da das Zielpublikum dieser Texte nicht unbedingt ausschließlich aus Translator*innen besteht, sondern aus einem sachgebietsspezifischen Publikum, das sowohl Fachexpert*innen als auch Personen mit privatem Interesse an einem bestimmten Sachgebiet umfassen kann, ist es sicherlich aufschlussreich zu untersuchen, wie dieses die Qualität maschineller Übersetzungen bewertet. Das entsprechende sachgebietsspezifische Publikum hat schließlich das nötige prozedurale Wissen, um die Richtigkeit und Angemessenheit der Inhalte zu beurteilen. Durch die Einbeziehung der tatsächlichen Zielgruppe in den Bewertungsprozess ist das Ziel der Masterarbeit, die Arbeit von Translator*innen im Bereich der Übersetzung von Rezepten aus dem Englischen ins Deutsche zu optimieren und produktiver zu gestalten sowie eine bestmögliche Nutzer*innen-Erfahrung zu gewährleisten. Da Koch- beziehungsweise Backrezepte strukturelle Ähnlichkeiten mit anderen Textsorten, zum Beispiel Bedienungsanleitungen, aufweisen, liefert eine derartige Untersuchung darüber hinaus interessante Aufschlüsse über die Fähigkeit von NMÜ-Systemen, prozedurale Textsorten so zu übersetzen, dass die beschriebenen Anleitungen nachvollziehbar sind. Im Hinblick auf Kontext und terminologische Konsistenz stellen die Ergebnisse der Bewertung ein bedeutendes Feedback für Entwickler*innen maschineller Übersetzungssysteme dar und können zugleich die Nutzer*innen solcher Systeme entsprechend sensibilisieren. Schließlich können derartige Untersuchungen auch für die Ausbildung von Translator*innen, wo verschiedene Textsorten sowohl human als auch mithilfe von MÜ übersetzt werden, hilfreich sein.

2. Grundlagen

Im folgenden Kapitel wird zum besseren Verständnis das Konzept der Übersetzungsqualität aus translationswissenschaftlicher Sicht näher beleuchtet. Anschließend werden Grundlagen zur MÜ, darunter ihre Geschichte und Arten sowie verfügbare Systeme neuronaler maschineller Übersetzung und Methoden zur Qualitätsanalyse beschrieben.

2.1. Übersetzungsqualität aus translationswissenschaftlicher Sicht

„Translation is a complex, cognitive, linguistic, social, cultural, and technological process." (Castilho et al. 2018: 10). Dementsprechend ist die Operationalisierung und Messung der Übersetzungsqualität durchaus schwierig (Castilho et al. 2018: 10). Die Bewertung der Übersetzungsqualität hat sowohl in der Translationswissenschaft als auch in der Sprachindustrie einen wesentlichen Stellenwert. Bevor man die Qualität von Übersetzungen bewerten kann, muss diese zunächst näher definiert werden. Dafür sollen in diesem Kapitel zentrale Theorien der Translationswissenschaft zur und Definitionen von Übersetzungsqualität vorgestellt werden.

Zunächst ist die sogenannte Skopostheorie von Reiß und Vermeer (1984) zu behandeln, die auch heute noch eine wesentliche Bedeutung in der Translationswissenschaft einnimmt und für modernere Qualitätsdefinitionen herangezogen wird (vgl. Koby et al. 2014). In dieser Theorie ist die Erfüllung des Zwecks einer Übersetzung (Skopos) das wichtigste Kriterium zur Messung der Übersetzungsqualität. Reiß und Vermeer (1984: 18) nehmen an, dass jede Übersetzung von einem Text ausgeht, den ein*e Produzent*in zu einem bestimmten Zweck für Rezipient*innen produziert. Der Text ist demzufolge eine Handlung zur Erfüllung eines bestimmten Zwecks, mit der Produzent*innen und Rezipient*innen in Interaktion treten (Reiß & Vermeer 1984: 18). Voraussetzung für jede Translation ist ein zuvor produzierter Ausgangstext, der von den Translator*innen verstanden und interpretiert wird und in einen Zieltext für eine andere Kultur übersetzt wird (Reiß & Vermeer 1984: 19). Somit stellt der Zieltext, wie auch der Ausgangstext, ein Informationsangebot an Rezipient*innen dar (Reiß & Vermeer 1984: 19). Ein Text beziehungsweise eine Handlung soll zur Änderung eines vorhandenen Zustandes ein bestimmtes Ziel erreichen (Reiß & Vermeer 1984: 95). Wenn man in der Translationstheorie also von einem Ausgangstext als sogenannte Primärhandlung ausgeht, stellt sich nicht die Frage, ob und wie gehandelt wird, sondern was und wie übersetzt werden soll (Reiß & Vermeer 1984: 95). „Die Dominante aller Translation ist deren Zweck." (Reiß & Vermeer 1984: 96). Dies bedeutet, dass die Erreichung des Skopos einer Übersetzung

der Translationsstrategie, also auf welche Weise eine Translation durchgeführt wird, übergeordnet ist (Reiß & Vermeer 1984: 95, 100). Der Skopos der Übersetzung kann dabei vom Skopos des Ausgangstexts abweichen (Reiß & Vermeer 1984: 103).

Die Skopostheorie stellt für den Ansatz des funktionalen Übersetzens die Grundlage dar. Nord (2006: 15f), eine wesentliche Vertreterin dieses Ansatzes, hebt neben der Funktionsgerechtigkeit die Loyalität der Übersetzer*innen gegenüber ihren Handlungspartner*innen als weiteres wichtiges Kriterium zur Messung der Übersetzungsqualität hervor. Die Loyalität soll Übersetzer*innen daran hindern, „deren Erwartungen an die Ähnlichkeitsbeziehung zwischen Übersetzung und Original zugunsten einer radikalen Funktionsgerechtigkeit einfach zu übergehen“ (Nord 2006: 16). Dies bedeutet jedoch nicht, dass ein Zieltext oder Teile davon nicht eine andere Funktion erhalten können oder müssen (Nord 2006: 16). Ein Übersetzungsfehler aus funktionaler Sicht stellt demzufolge keinen Verstoß gegen die Regeln der Zielsprache dar, sondern vielmehr eine mangelhafte Realisierung des Übersetzungsauftrags hinsichtlich gewisser funktionaler Aspekte (Nord 2006: 17f). Jede Übersetzungsleistung lässt sich also nur im Hinblick auf ein festgelegtes Übersetzungsziel bewerten, das dem*der Übersetzer*in bekannt sein muss (Nord 2006: 17). Nach Nord (2006: 23) sollten daher pragmatische Übersetzungsfehler, sprich Fehler im Hinblick auf Aspekte wie etwa die Textfunktion, am schwerwiegendsten gewichtet werden. An zweiter Stelle stehen kultur- und sprachpaarspezifische Übersetzungsfehler, zu denen im professionellen Kontext eher Verstöße gegen kulturspezifische Verhaltensnormen, unter anderem Textsorten- oder Übersetzungskonventionen, gehören (Nord 2006: 23). Während die Einhaltung kulturspezifischer Konventionen in manchen Fällen nicht wesentlich erscheint, sind die pragmatischen Aspekte unter allen Umständen zu beachten (Nord 2006: 23). Zielsprachliche Fehler, zum Beispiel morphologische Fehler, Kohäsionsfehler, Komma- und Rechtschreibfehler, die für gewöhnlich in der Übersetzungsausbildung bewertet werden und an mangelnden zielsprachlichen Kenntnissen liegen, sieht Nord (2006: 23f) nicht als Übersetzungsfehler im eigentlichen Sinne, sondern als sogenannte Performanz-Fehler.

Darüber hinaus sind die Kriterien für die Bewertung von Fachübersetzungen nach Ahrend (2006), auch im Hinblick auf das in der vorliegenden Masterarbeit behandelte Thema, durchaus relevant. Ein Rezept ist schließlich auch eine Textsorte aus dem Fachbereich der Kulinarik. Ahrend (2006) betont, dass es keine einheitliche Antwort auf die Frage, wann eine Übersetzung gut ist, gibt. Vielmehr gibt es dafür zahlreiche verschiedene Antworten, die vor

allem von der Perspektive der jeweiligen Rezipient*innen abhängen und dementsprechend unterschiedliche Maßstäbe beziehungsweise Anforderungen voraussetzen.

In diesem Zusammenhang unterscheidet Ahrend (2006) zwischen allgemeinen, praxisorientierten und fachtextspezifischen Kriterien. Zu den allgemeinen zählen Vollständigkeit, Genauigkeit, Übersetzungstreue und Verständlichkeit der Übersetzung, die alle auf eine gewisse akademische Betrachtungsweise schließen lassen (Ahrend 2006: 31f). Auf die Vollständigkeit und Genauigkeit einer Übersetzung wird etwa vor allem bei ersten Übersetzungsübungen in der Ausbildung bestanden (Ahrend 2006: 31). Diese Übersetzungstreue setzt möglichst tiefgehende Kenntnisse sowohl der Ausgangs- als auch der Zielsprache voraus (Ahrend 2006: 31). Die Verständlichkeit einer Übersetzung hängt von Faktoren wie Stilsicherheit, Kenntnis des behandelten Themas und Bereitschaft zu einer freieren Übersetzung ab (Ahrend 2006: 31f). Die zuletzt genannten Fähigkeiten entwickeln sich mit steigender Erfahrung und Auseinandersetzung mit verschiedensten Fachbereichen und Textsorten. Diese allgemeinen Kriterien werden im Arbeitsleben von Übersetzer*innen mit der Zeit als Voraussetzung angesehen und oft durch praxisorientierte Kriterien ergänzt oder womöglich ersetzt (Ahrend 2006: 32). Die praxisorientierten Kriterien umfassen Sicherstellung des Leseverständnisses der Zielgruppe, Unsichtbarkeit der Übersetzung sowie Berücksichtigung des Übersetzungszwecks (Ahrend 2006: 32ff). Das Leseverständnis der Zielgruppe ist neben der Richtigkeit ein relevantes Kriterium für eine erfolgreiche Übersetzung beziehungsweise Kommunikationssituation. Der Ausgangstext kann unter Umständen unverständlich oder mehrdeutig sein. Sofern eine solche Mehrdeutigkeit beziehungsweise Unklarheit nicht vom*von der Autor*in beabsichtigt ist, gilt es, eine möglichst verständliche Übersetzung zu gewährleisten und sich keinesfalls genau an das Original zu halten (Ahrend 2006: 32f).

Eine Übersetzung wird außerdem als gut erachtet, wenn sie unsichtbar ist und sich wie ein Original liest (Ahrend 2006: 33). In der Praxis ist die Berücksichtigung des Übersetzungszwecks ein wichtiges Qualitätskriterium. Je nach dem, welche Funktion eine Übersetzung erfüllen soll, werden andere Ansprüche an diese gestellt (Ahrend 2006: 35). Wenn ein Text etwa übersetzt werden soll, um den Inhalt nach Relevanz beurteilen zu können, reicht eine schnelle Rohübersetzung oder sogar nur eine kurze Zusammenfassung aus. Wenn die Übersetzung jedoch veröffentlicht werden soll, sind besondere Detailtreue und Genauigkeit erforderlich (Ahrend 2006: 35). Zu den fachtextspezifischen Kriterien zählen Verwendung der korrekten Fachterminologie, Kenntnis der zu behandelnden Materie, Formalitäten und Verwendung der korrekten Maßeinheiten (Ahrend 2006: 37f). In diesem Zusammenhang

sollten sich Fachübersetzer*innen mit der jeweiligen Materie auseinandersetzen und beispielsweise Fachliteratur über das zu behandelnde Thema in der Zielsprache heranziehen (Ahrend 2006: 38).

Auch die Berücksichtigung von gewissen sprachspezifischen Formalitäten ist für eine idiomatisch korrekte Übersetzung unabdingbar. Während im Deutschen in prozeduralen Texten wie etwa Bedienungsanleitungen der Infinitiv so gesehen als Befehlsform verwendet wird, wirken diese in anderen Sprachen, wie zum Beispiel Französisch, wie eine Bitte oder Einladung (Ahrend 2006: 38). Auch die Verwendung von korrekten Maßeinheiten ist ein wichtiger Aspekt. Abhängig von der jeweiligen Zielgruppe können Maßeinheiten von anderen Sprachräumen beibehalten oder sollten umgerechnet beziehungsweise übersetzt werden (Ahrend 2006: 38). Schließlich geht Ahrend (2006) noch auf außersprachliche Kriterien ein, die nicht direkt mit der Übersetzungsqualität verbunden sind, jedoch von zentraler Bedeutung sein können. Diese umfassen das Verhältnis von Quantität und Qualität, Effizienz, Übersetzungskosten sowie -verfügbarkeit (Ahrend 2006: 40ff). So müssen Übersetzer*innen für sich herausfinden, wie schnell sie qualitativ ausreichende Übersetzungen anfertigen können, um genug Geld zu verdienen. Hierbei gilt es, ein eigenes Verständnis von Qualität zu entwickeln und eigene Anforderungen zu bestimmen (Ahrend 2006: 40). Außerdem ist größtmögliche Effizienz sicherzustellen, indem der Kosten- und Zeitaufwand für Recherche, Revision und Gestaltung mit dem zu erzielenden Nutzen in ein bestmögliches Verhältnis gebracht wird (Ahrend 2006: 41). In bestimmten Situationen sind für Auftraggeber*innen auch die Übersetzungskosten ein relevanter Aspekt. In vielen Fällen kann statt einer Humanübersetzung eine maschinelle Übersetzung verlangt werden, wobei man hier mit Qualitätsabstrichen zu rechnen hat (Ahrend 2006: 41). Damit verbunden ist auch die Übersetzungsverfügbarkeit, das heißt in gewissen Situationen wird die Verfügbarkeit einer Übersetzung vor die Qualität gestellt (Ahrend 2006: 41f).

2.2. Maschinelle Übersetzung

Das zweite Grundlagenkapitel behandelt in jeweils eigenen Unterkapiteln Geschichte, Arten und verfügbare Tools der MÜ. Darüber hinaus werden die Methoden zur Qualitätsanalyse näher beschrieben.

2.2.1. Geschichte der maschinellen Übersetzung

Die Anfänge der maschinellen Übersetzung gehen bis in die 1930er zurück (Burchardt & Porsiel 2017: 12; Kenny 2018: 429). Die ersten Ideen zu Universal- beziehungsweise

Geheimsprachen, die später zur Entwicklung der MÜ führten, stammen jedoch bereits aus dem 17. Jahrhundert (Stein 2018: 5). Vor allem die Universalsprache als Idee zur grenzenlosen, unmissverständlichen Kommunikation mittels logischen Prinzipien und Symbolen kann als Vorreiter für die in manchen Ansätzen der MÜ vorgesehenen Interlinguas betrachtet werden (Hutchins & Somers 1992: 5; Kenny 2018: 429).

Nichtsdestotrotz werden die Anfänge der MÜ auf die Patentanmeldungen für elektromechanische Geräte von Georges Artsrouni in Frankreich und Petr Trojanskij in Russland im Jahr 1933, die als mehrsprachige Wörterbücher dienen sollten, zurückgeführt (Kenny 2018: 429). Die Ideen von Trojanskij trugen aus heutiger Sicht mehr Früchte. Er stellte eine sogenannte dreistufige mechanische Übersetzung vor. Zuerst sollte eine einsprachige Person einen Ausgangstext mit einem universellen Schema analysieren, das etwaige grammatikalische Funktionen von Wörtern erfassen konnte. In einem weiteren Arbeitsschritt sollte diese Person jeweils die einzelnen Wörter des Ausgangstextes nach und nach im Übersetzungswörterbuch der Maschine abrufen und den entsprechenden grammatikalischen Code für die jeweilige Verwendung des Wortes hinzufügen. Die Maschine sollte dann das entsprechende zielsprachliche Wort mit dem grammatikalischen Code ausgeben. Eine einsprachige Person der Zielsprache sollte diese Informationen in einem dritten Schritt schließlich zur Zieltexterstellung verwenden (Kenny 2018: 429f). Darüber hinaus sah Trojanskij ursprünglich noch eine dritte Person in diesem Prozess vor, eine*n zweisprachige*n Revisor*in, die sich um die letzten Feinschliffe kümmern sollte (Kenny 2018: 430). Damit war er auch ein Vorreiter in der Debatte um die Rolle von zweisprachigen Personen in MÜ-Arbeitsprozessen (Kenny 2018: 430).

Unabhängig davon und ohne Kenntnis der Überlegungen von Trojanskij trafen sich 1946 und 1947 Warren Weaver von der Rockefeller-Stiftung und der britische Kristallograph Andrew Booth und diskutierten erste Ideen zum Einsatz der damals neuen Erfindung der Computer für die Übersetzung von natürlichen Sprachen (Hutchins 2007: 1; Hutchins & Somers 1992: 5). 1948 erforschte Booth nach seiner Rückkehr nach London die Mechanisierung eines zweisprachigen Wörterbuchs und arbeitete schließlich mit Richard H. Richens aus Cambridge, der Lochkarten für die Erstellung von Wort-für-Wort-Übersetzungen für Abstracts nutzte, an der morphologischen Analyse für dieses Wörterbuch (Hutchins 2007: 1; Hutchins & Somers 1992: 5). Ausgangspunkt für die Forschung im Bereich der MÜ in den USA war jedoch ein Memorandum von Weaver aus dem Jahr 1949, in dem er auf Forschungsmethoden zur Bewältigung des Problems der Mehrdeutigkeit einging. Er schlug die

Verwendung von kryptographischen Methoden aus dem Zweiten Weltkrieg, statistischer Analyse und der Informationstheorie von Claude Shannon sowie die Erforschung der Logik und Universalsprache vor (Hutchins & Somers 1992: 6; Hutchins 2007: 1f; Kenny 2018: 430).

Die MÜ erweckte also gegen Ende des Zweiten Weltkrieges und in der Zeit des Kalten Krieges größere wissenschaftliche Aufmerksamkeit, als die betroffenen Nationen – vorwiegend Vereinigtes Königreich und USA – versuchten, Funksprüche und andere Kommunikationsinhalte der Gegenseite – Deutschland, Russland – möglichst zeitgleich zu übersetzen (Burchardt & Porsiel 2017: 12). In diesem Zusammenhang ist das Georgetown-University-IBM-Experiment zu nennen, aus dem 1954 die erste öffentliche Präsentation eines MÜ-Systems für das Sprachpaar Russisch-Englisch hervorging (Burchardt & Porsiel 2017: 12; Hutchins 2007: 2; Kenny 2018: 430f). Dafür wurden 49 russische Sätze ausgewählt und ins Englische übersetzt, wobei ein stark begrenzter Wortschatz von 250 Wörtern und lediglich sechs Grammatikregeln Anwendung fanden (Hutchins 2007: 2). Die Präsentation des Systems resultierte vor allem in den USA in einer Welle der medialen Aufmerksamkeit und Begeisterung (vgl. Hutchins 2007: 2; Kenny 2018: 430f). Das System erlangte laut Hutchins (2007: 2) zwar nur einen geringen wissenschaftlichen Wert, seine Ergebnisse waren jedoch für eine umfangreiche Finanzierung der MÜ-Forschung in den USA und die Einleitung von MÜ-Projekten in anderen Ländern, vor allem in der Sowjetunion, ausreichend vielversprechend.

Im darauffolgenden Jahrzehnt fokussierte sich die MÜ-Forschung auf die RegelBasierte MÜ (RBMÜ) (vgl. Kapitel 2.2.2.1.) und ihre drei Ansätze der direkten, Interlingua- und Transfer-Übersetzung (vgl. Hutchins 2007: 2f). Da die Forschung in den USA im Hinblick auf den Kalten Krieg vorwiegend von der CIA und dem Militär finanziert wurde, konzentrierte sie sich auf die Übersetzung aus dem Russischen ins Englische (Kenny 2018: 431). Die Forschung in der Sowjetunion hingegen beschäftigte sich neben der Übersetzung aus dem Englischen ins Russische durch die mehrsprachige Politik des Landes mit einem breiteren Spektrum an Sprachen (Hutchins 2007: 3). Laut Hutchins und Somers (1992: 6) war die Forschung dieser Zeit nicht nur für die MÜ von erheblicher Bedeutung, sondern auch für die Computerlinguistik und künstliche Intelligenz. In diesem Zusammenhang sind vor allem die Entwicklung von automatischen Wörterbüchern und Techniken für die syntaktische Analyse, aber auch die theoretische Arbeit im Bereich der linguistischen Analyse hervorzuheben (Hutchins & Somers 1992: 6).

Die von Begeisterung und Optimismus in Bezug auf vollautomatische Systeme geprägten Jahre 1954 bis 1959 wurden von einer Zeit der Ernüchterung, in der die Qualität der

Übersetzungen etwa durch komplexe linguistische Probleme und den fehlenden Umgang der Maschine mit Makrokontext an ihre Grenzen stieß, abgelöst (Hutchins 2007: 5; Hutchins & Somers 1992: 6; Kenny 2018: 431). 1960 hob Yehoshua Bar-Hillel, der 1951 als erster Vollzeitforscher im Bereich der MÜ am Massachusetts Institute of Technology (MIT) eingestellt wurde, in einer Studie zum Stand der MÜ-Forschung hervor, dass die von der MÜ-Forschung anvisierte Erstellung von Systemen vollautomatischer Übersetzung mit hoher Qualität (engl. fully automatic high-quality translation, FAHQT), die nicht von Humanübersetzungen unterscheidbar ist, nicht realistisch sei (vgl. Bar-Hillel 1960; Hutchins 2007: 5; Hutchins & Somers 1992: 6f). Er argumentierte, dass kein System jemals ein so umfangreiches enzyklopädisches Wissen über die reale Welt erwerben könnte, um die semantischen Herausforderungen hinsichtlich der Mehrdeutigkeit zu bewältigen (vgl. Bar-Hillel 1960; Hutchins & Somers 1992: 7; Kenny 2018: 431). Vielmehr war Bar-Hillel (1960) der Ansicht, dass entweder eine vollautomatische Übersetzung mit geringer Qualität oder eine teilautomatische Übersetzung mit hoher Qualität anzustreben ist, wobei die hohe Qualität in letzterer nur unter Hinzuziehung menschlicher Post-Editor*innen erreicht werden kann. 1964 wurde in den USA zur Untersuchung und Bewertung des MÜ-Standes schließlich das Automatic Language Processing Advisory Committee (ALPAC) gegründet (Hutchins 2007: 5). 1966 kam ALPAC in seinem Bericht zum Schluss, dass MÜ langsamer, ungenauer und wesentlich teurer als die Humanübersetzung ist (vgl. ALPAC 1966). Darüber hinaus führte der Bericht zu folgendem Ergebnis: „[…] there is no immediate or predictable prospect of useful machine translation“ (ALPAC 1966: 32). Statt weiterer Investitionen in die MÜ-Forschung sah das Komitee die Entwicklung von maschinellen Hilfsmitteln für Übersetzer*innen als sinnvollere Lösung an (vgl. ALPAC 1966). Der ALPAC-Bericht führte in den USA und auch anderen Ländern wie der Sowjetunion zu einem Rückgang der Forschung und Finanzierung im Bereich der MÜ (Burchardt & Porsiel 2017: 12f; Hutchins 2007: 6).

Ab diesem Zeitpunkt erfolgte die MÜ-Forschung vorwiegend in Kanada und Europa (Hutchins & Somers 1992: 7). Während in den USA der Fokus auf der Übersetzung wissenschaftlicher und technischer Texte aus dem Russischen ins Englische lag, bestand durch die Zweisprachigkeit in Kanada und die Mehrsprachigkeit in der damaligen Europäischen Wirtschaftsgemeinschaft (EWG), heute Europäische Union (EU), eine große Übersetzungsnachfrage für amtliche Dokumente in der Sprachrichtung Englisch-Französisch sowie aus und in alle Amtssprachen, die die Kapazitäten des Marktes weitgehend sprengte (Hutchins & Somers 1992: 7). Die Übersetzung diente somit nicht mehr ausschließlich der Überwachung

von rivalisierenden Mächten, sondern war vielmehr eine „politische Notwendigkeit“ und „Frage der Bürgerrechte“ (Kenny 2018: 432). So bestand etwa im Fall der EWG Übersetzungsbedarf für wissenschaftliche, technische, aber auch Verwaltungs- und Rechtstexte aus und in alle Amtssprachen (Hutchins & Somers 1992: 7). Im Bereich der Übersetzung von Wettervorhersagen war die MÜ in Kanada besonders erfolgreich. Das von der Gruppe Traduction Automatique de l’Université de Montréal (TAUM) entwickelte und 1977 veröffentlichte System Météo erzielte durch die Beschränkung auf die Subsprache der Wettervorhersagen, die von einer einfachen Sprache, einem begrenzten Vokabular und grammatikalischen Strukturen charakterisiert ist, hochqualitative Ergebnisse mit einer Genauigkeit von 90% (Hutchins & Somers 1992: 207f, 219). Andere Subsprachen, wie etwa jene von der TAUM-Gruppe erforschte der Luftfahrthandbücher, haben sich für die MÜ als weniger geeignet erwiesen (Hutchins & Somers 1992: 219). Die MÜ fand darüber hinaus 1976 in den Europäischen Gemeinschaften Verwendung. Die damalige Kommission der Europäischen Gemeinschaften setzte erstmals das MÜ-System SYSTRAN (vgl. Kapitel 2.2.3.) für die Übersetzung aus dem Englischen ins Französische und später auch für weitere Sprachpaare ein (Hutchins & Somers 1992: 7). In den darauffolgenden Jahren wurden weitere MÜ-Systeme für die Kommission, etwa für die Sprachpaare Französisch-Englisch, Englisch-Italienisch und Englisch-Deutsch, entwickelt (Hutchins & Somers 1992: 7). 1978 strebte man schließlich ein System an, das mit allen Amtssprachen der Europäischen Gemeinschaften – zu diesem Zeitpunkt Dänisch, Niederländisch, Englisch, Französisch, Deutsch und Italienisch, später zusätzlich Griechisch, Portugiesisch und Spanisch – umgehen konnte (Hutchins & Somers 1992: 239). Dafür wurde das Forschungsprojekt EUROTRA mit Forschungsgruppen in allen Mitgliedsstaaten ins Leben gerufen, das zwischen allen neun Amtssprachen in allen 72 Sprachpaaren übersetzen können sollte (Hutchins & Somers 1992: 239). Es blieb zwar noch Vieles offen, das Projekt leistete vor allem in den damaligen Mitgliedsstaaten jedoch einen wichtigen Beitrag in der MÜ-Forschung (vgl. Hutchins & Somers 1992: 255). 2010 stellte die Europäische Kommission schließlich die Verwendung von SYSTRAN ein und erstellte ihr eigenes MÜ-System, das 2017 durch ein anderes internes System, eTranslation, ersetzt wurde, das mittlerweile gute maschinelle Rohübersetzungen für alle EU-Amtssprachen sowie für Arabisch, Chinesisch, Isländisch, Japanisch, Norwegisch, Russisch, Türkisch und Ukrainisch liefert (Europäische Kommission 2025; Kenny 2018: 432).

Die 1980er Jahre waren von der Entwicklung kommerzieller Systeme, vorwiegend in Nordamerika und Japan, geprägt (Kenny 2018: 432). In diesem Zusammenhang kooperierten

gewisse Forschungsgruppen mit Herstellern (Poibeau 2017: 44). So entwickelte die University of Texas mit Siemens das Übersetzungssystem METAL, welches zuerst auf das Sprachpaar Deutsch-Englisch angelegt und später um weitere Sprachen erweitert wurde (Poibeau 2017: 44). In Japan konzentrierten sich Unternehmen aus der Software- beziehungsweise Hardwareindustrie, zum Beispiel Fujitsu und Toshiba, auf die Entwicklung von Systemen im Sprachpaar Japanisch-Englisch, wobei man sich auch auf andere asiatische Sprachen wie Chinesisch oder Koreanisch fokussierte (Hutchins & Somers 1992: 8; Poibeau 2017: 44). Dabei wurde eine halbautomatische Übersetzung von kurzen technischen Texten und Produktbroschüren angestrebt (Poibeau 2017: 44). Die Texte sollten mit dem automatischen Übersetzungssystem übersetzt und die MÜ-Ergebnisse schließlich manuell korrigiert werden (Poibeau 2017: 44). Die Übersetzungsqualität dieser Systeme galt als grob und nachbearbeitungsbedürftig, wobei sie unter Umständen kosteneffizient eingesetzt werden konnten (Hutchins & Somers 1992: 8; Kenny 2018: 432). Neben den kommerziellen Systemen wurden einige Systeme entwickelt, die speziell auf mehrsprachige Organisationen oder bestimmte Kund*innen zugeschnitten waren (Hutchins & Somers 1992: 8f). In diesem Kontext sind unter anderem die englischen und spanischen Systeme der Panamerikanischen Gesundheitsorganisation SPANAM und ENGSPAN sowie die SYSTRAN-Installationen bei Organisationen wie Aérospatiale und NATO zu nennen (Hutchins 2007: 7; Hutchins & Somers 1992: 8f). Die meisten Systeme setzten für die Generierung von akzeptablen Ergebnissen eine Nachbearbeitung, sogenanntes Post-Editing, voraus (Kenny 2018: 433). Die Vorbearbeitung, das Pre-Editing, war jedoch auch weit verbreitet. Dabei wurde der Ausgangstext unter Verwendung von kontrollierter Sprache und Regeln hinsichtlich Satzlänge oder Vokabular bearbeitet (Kenny 2018: 433).

Ende der 1980er Jahre entwickelte ein Forschungsteam von IBM rund um Peter Brown einen alternativen MÜ-Ansatz, nämlich jenen der Statistischen MÜ (SMÜ) (vgl. Kapitel 2.2.2.2.) (Kenny 2018: 433). Damit distanzierte man sich von den linguistisch-orientierten, regelbasierten Systemen (Kenny 2018: 433). Das Forschungsteam belegte, dass probabilistische Übersetzungsmodelle unter Hinzuziehung von Statistik und Informationstheorie direkt mit Paralleltexten ohne linguistisches Vorwissen trainiert werden können, etwa anhand der zweisprachigen Protokolle des kanadischen Parlaments (Kenny 2018: 433). Trotz anfänglicher skeptischer Stimmen gewann die SMÜ ab den 2000er Jahren an Relevanz. Dies wurde vor allem durch die Zunahme elektronischer Paralleltexte und Open-Source-Toolkits sowie die steigende Computerleistung und Datenspeicherung ermöglicht (Kenny 2018: 433). Geopolitische und wirtschaftliche Faktoren, wie der erhöhte Übersetzungsbedarf im Sprachpaar

Englisch-Arabisch nach den Terroranschlägen vom 11. September 2001 in New York oder die schrittweise EU-Erweiterung und die damit verbundene Zunahme der Amtssprachen, trieben die MÜ-Forschung weiter voran (Kenny 2018: 433).

Das Internet spielte im Kontext der MÜ-Forschung ebenfalls eine Schlüsselrolle. Einerseits bewirkte die steigende sprachliche Vielfalt im Internet eine Vielzahl an mehrsprachigen Websites, mit denen statistische Systeme trainiert werden konnten. Andererseits bot das World Wide Web eine Plattform für die Bereitstellung von MÜ an Millionen von Nutzer*innen, darunter auch Privatpersonen (Kenny 2018: 433). So wurde schließlich 2006 das Übersetzungssystem Google Translate (vgl. Kapitel 2.2.3.) online gestellt, das mit der Zeit multimodale Eingabemöglichkeiten, zum Beispiel Text, Bild, Sprache, anbot (Kenny 2018: 433). In den 2000er Jahren stellten auch andere Anbieter wie Microsoft (vgl. Kapitel 2.2.3.) MÜ-Systeme mit ähnlichen Möglichkeiten zur Verfügung (Kenny 2018: 434).

Trotz der zunehmenden Verfügbarkeit von kostenlosen Online-Übersetzungssystemen Mitte bis Ende der 1990er Jahre stieß die MÜ bei professionellen Übersetzer*innen vor allem wegen der unzureichenden Qualität auf wenig Interesse (Kenny 2018: 434). Gleichzeitig dominierte die computergestützte Übersetzung mit Translation-Memory-Tools, die eine effiziente Wiederverwendung von Humanübersetzungen und eine gewisse Automatisierung ermöglichten und somit die Produktivität der Übersetzer*innen steigerten, den Markt (Kenny 2018: 434). Die Translation-Memory-Tools ermöglichten die umfangreiche Erstellung von Paralleltexten als Trainingsmaterial für SMÜ-Systeme. Darüber hinaus stellten sie eine Umgebung dar, in der MÜ-Ergebnisse mit den aus der Translation Memory vorgeschlagenen Übersetzungstreffern professionellen Übersetzer*innen zur Bearbeitung bereitgestellt wurden (Kenny 2018: 434). Dadurch verschmolzen die Humanübersetzung und MÜ miteinander (Kenny 2018: 434).

Die SMÜ galt ab Mitte der 2000er Jahre bis 2015 als Stand der Technik (Kenny 2018: 434). In den Jahren 2015 und 2016 kam es im Bereich der MÜ zu einer Neuerung. Die Neuronale MÜ (NMÜ) (vgl. Kapitel 2.2.2.3.) übertraf die SMÜ vor allem hinsichtlich der Qualität (Kenny 2018: 434). In weiterer Folge integrierten 2016 große MÜ-Unternehmen wie Google, Microsoft und SYSTRAN neuronale Modelle (vgl. Kenny 2018: 434). Die Entwicklung und zeitgemäße Möglichkeiten dieser NMÜ-Systeme werden in Kapitel 2.2.3. näher erläutert.

2.2.2. Arten der maschinellen Übersetzung

Wie in Kapitel 2.2.1. beschrieben, wurden im Laufe der Jahrzehnte verschiedenste Ansätze der MÜ entwickelt. Im Allgemeinen werden hierbei drei Arten unterschieden: RegelBasierte MÜ (RBMÜ), Statistische MÜ (SMÜ) und Neuronale MÜ (NMÜ) (vgl. Burchardt & Porsiel 2017: 12f; Kenny 2022: 35f, 39). In den folgenden Unterkapiteln soll auf deren Funktionsweisen näher eingegangen werden.

2.2.2.1. Regelbasierte maschinelle Übersetzung

Die RBMÜ kann als klassischer Ansatz gesehen werden und dominierte trotz kritischer Prognosen wie etwa dem ALPAC-Bericht aus dem Jahr 1966 und dem daraus resultierenden Rückgang in der Forschung und Finanzierung im Bereich der MÜ bis zum Anfang des 21. Jahrhunderts (Stein 2018: 9; Kenny 2022: 35; Burchardt & Porsiel 2017: 12f). RBMÜ-Systeme werden mit linguistischen Daten wie etwa grammatikalischen und lexikalischen Regeln trainiert (Kenny 2018: 434). Abhängig von Art und Schwierigkeitsgrad des Ausgangstextes stehen hier drei Arten der Übersetzung zur Verfügung: direkte, Transfer- und Interlingua-Übersetzung (Stein 2018: 9). RBMÜ-Systeme erstellen Übersetzungen je nach Komplexität in drei aufeinanderfolgenden Schritten: Analyse, Transfer und Generierung/Synthese (Stein 2018: 9).

Systeme der direkten Übersetzung produzieren im Allgemeinen Wort-für-Wort-Übersetzungen und stützen sich nach einer morphologischen Analyse der Wörter auf bilinguale Wörterbücher (Hutchins & Somers 1992: 72; Stein 2018: 9). Somit werden bei dieser Übersetzungsart die Schritte Analyse, Transfer und Generierung/Synthese ausgelassen (Stein 2018: 9). In einem weiteren Schritt werden Änderungen in der Satzstellung entsprechend den Syntaxregeln der Zielsprache vorgenommen (Stein 2018: 9f). Der direkte Ansatz der RBMÜ hat hinsichtlich der Qualität seine Grenzen. Die Ergebnisse enthalten oft lexikalische Fehler und eine fehlerhafte Syntax, die jener der Ausgangssprache zu treu ist (Hutchins & Somers 1992: 72). Die Umsetzung der direkten Übersetzung ist trotz ihrer niedrigen Komplexität zeit- und arbeitsintensiv, da jeweils immer nur ein Sprachpaar herangezogen werden kann und dafür Wörterlisten mit sämtlichen Wortformen zu jedem Eintrag erstellt werden müssen (Werthmann & Witt 2014: 88). Die Ergebnisse finden wegen der Mehrdeutigkeit gewisser Wörter nur beschränkt Anwendung (Stein 2009: 8). Außerdem stellen nicht wörtlich zu übersetzende Mehrwortlexeme eine Herausforderung für die direkte Übersetzung dar (Stein 2009: 8).

Die Transfer-Übersetzung berücksichtigt morphologische und semantische Informationen, wobei die syntaktische Komponente im Vergleich zur direkten Übersetzung ausgefeilter ist (Stein 2009: 8). Die Übersetzung erfolgt dabei nicht mehr auf Wort-, sondern auf Satzebene (Werthmann & Witt 2014: 89). Wie zu Beginn dieses Unterkapitels ausgeführt, werden die Übersetzungen in den drei aufeinanderfolgenden Schritten Analyse, Transfer und Generierung/Synthese angefertigt. Bei der Analyse wird der Ausgangstext segmentiert und morphologisch, syntaktisch sowie semantisch analysiert (Werthmann & Witt 2014: 89). Die analysierten Segmente beziehungsweise Sätze werden anschließend in eine abstrakte Struktur transferiert (Werthmann & Witt 2014: 89). Diese werden im nächsten Schritt in abstrakte Strukturen der Zielsprache umgewandelt (Werthmann & Witt 2014: 89). Schließlich werden aus den zielsprachlichen abstrakten Strukturen in der Zielsprache anwendbare Segmente generiert (Werthmann & Witt 2014: 89). Die Komplexität der dabei verwendeten Regeln erscheint grenzenlos (Stein 2018: 10). So können Regeln und Kombinationen in unendlicher Zahl festgelegt werden (Stein 2009: 8). Ab einem gewissen Stadium führt die höhere Komplexität jedoch nicht zu einer höheren Übersetzungsqualität (Stein 2009: 8). In der Praxis entstehen vielmehr aufgrund interner Konflikte und unstimmiger Regeln neue Fehler (Stein 2009: 8). RBMÜ-Systeme können zum überwiegenden Teil der Transfer-Übersetzung zugeordnet werden (Stein 2018: 10).

Die Interlingua-Übersetzung ist der komplexeste Ansatz der RBMÜ, der von einer universellen und sprachunabhängigen Kodierungsart sprachlicher Informationen ausgeht (Stein 2009: 8). Die Übersetzung erfolgt in diesem Ansatz in zwei Schritten: Analyse und Generierung (Werthmann & Witt 2014: 90). Der Ausgangstext wird segmentiert und analysiert (Werthmann & Witt 2014: 90). Danach wird eine universelle Sprache, die Interlingua, erstellt, die die Wiedergabe von Informationen in jeder Sprache ermöglicht (Stein 2018: 10; Werthmann & Witt 2014: 90). Durch die Verwendung einer gemeinsamen Zwischensprache wird der Sinn beider Sprachen auf eindeutige Art und Weise wiedergegeben, ohne ständig zwischen zwei Sprachen wechseln zu müssen (Stein 2018: 10). Aus dieser Zwischensprache wird schließlich die Übersetzung in die Zielsprache generiert (Werthmann & Witt 2014: 90). Die Interlingua soll somit die Zielsprache für jede Übersetzung darstellen und gleichzeitig die Grundlage für die Übersetzung in die Zielsprache sein (Stein 2018: 10). Eine solche Metasprache wurde bislang jedoch trotz zahlreicher Bemühungen nicht realisiert (Stein 2018: 10).

2.2.2.2. Statistische maschinelle Übersetzung

Die SMÜ wurde in den 1980er Jahren vom IBM-Wissenschaftler Peter Brown entwickelt und geht von der Annahme aus, „dass Übersetzungsentscheidungen anhand von bedingten Wahrscheinlichkeiten getroffen werden können“ (Stein 2009: 9). Um diese bedingten Wahrscheinlichkeiten bestimmen zu können, sind große Parallelkorpora erforderlich (Stein 2018: 11). Komplexe Regelsätze, wie sie in anderen MÜ-Systemen (RBMÜ) verwendet werden, finden hier keine Anwendung (Stein 2018: 11).

Ziel der SMÜ ist, die wahrscheinlichste Übersetzung aus allen möglichen und unmöglichen Sätzen einer Sprache zu finden (Stein 2009: 9f). Dabei muss jedoch berücksichtigt werden, dass der Zugang zu allen Sätzen einer Sprache unrealistisch ist, weswegen die SMÜ mit Modellen – einem Übersetzungsmodell und einem Sprachmodell – arbeitet (Stein 2018: 11). Das Übersetzungsmodell umfasst ein zweisprachiges aligniertes Korpus, „das alle möglichen Übersetzungen zwischen beiden Sprachen repräsentiert“ (Stein 2009: 10). Daraus ergibt sich, dass die Größe des Übersetzungsmodells in direktem Zusammenhang mit der Qualität der Ergebnisse steht (Stein 2018: 11). Jeder Satz repräsentiert dabei eine mögliche Übersetzung für alle anderen, wobei die alignierten Sätze die höchste Wahrscheinlichkeit aufweisen (Stein 2009: 10, 2018: 11). Beim Übersetzungsmodell der SMÜ wird darüber hinaus zwischen einem Lexikonmodell und einem Alignierungsmodell unterschieden (Stein 2009: 10). Das Lexikonmodell bezieht sich auf die Richtigkeit der Übersetzung von Wörtern oder Wortsequenzen, das Alignierungsmodell auf die Richtigkeit der Satzstellung (Stein 2009: 10). Ein einsprachiges Korpus in der Zielsprache wird als Sprachmodell angesehen, das alle gültigen Sätze einer Sprache umfasst (Stein 2009: 10). Zur besseren Anwendung wird hierbei auf Wort- oder Wortsequenzebene gearbeitet (Stein 2018: 11). Die wahrscheinlichste Übersetzung wird in weiterer Folge aus dem höchsten Ergebnis, das die Multiplikation der Werte von Satzgültigkeit, Wortübersetzung und Satzstellung ergibt, bestimmt (Stein 2009: 10). Diese Modelle werden zunächst direkt aus den Daten trainiert (Kenny 2022: 37). In der Tuning-Phase werden in weiterer Folge die einem Modell zuzuweisenden Gewichtungen zur Erzielung bestmöglicher Ergebnisse herausgearbeitet (Kenny 2022: 37). Nach Abschluss dieser zwei Phasen kann das System zuvor unbekannte Sätze aus der Ausgangssprache übersetzen (Kenny 2022: 37). Der Übersetzungsprozess läuft dabei in zwei Dekodierungsschritten ab (Werthmann & Witt 2014: 95). Im ersten Dekodierungsschritt werden über das Übersetzungsmodell die zielsprachlichen Übersetzungsoptionen aller ausgangssprachlichen Wörter sowie die entsprechenden Wahrscheinlichkeiten determiniert (Werthmann & Witt 2014: 95). Der zweite

Dekodierungsschritt umfasst die Bestimmung der dem Kontext entsprechenden wahrscheinlichsten Übersetzungen, die in weiterer Folge an die Satzstellung der Zielsprache angepasst werden (Werthmann & Witt 2014: 95).

Da ein Korpus nie alle möglichen Sätze einer Sprache umfassen kann, ist es in der SMÜ weniger praktikabel, komplette Sätze zu analysieren (Stein 2009: 11). Einheiten werden daher vor der Analyse eingegrenzt (Stein 2009: 11). Abhängig von der Ebene der Textanalyse differenziert man hierbei zwischen wort- und phrasenbasierter SMÜ (Stein 2009: 11). Bei der wortbasierten SMÜ erfolgt eine Betrachtung der Daten auf Wortebene, wobei jedes Wort in der Ausgangssprache mit einem eigenen Wort in der Zielsprache zu übersetzen ist (Stein 2009: 11, 2018: 12). In einigen Fällen kann ein Wort in der Ausgangssprache durch eine Mehrwortbenennung in der Zielsprache ersetzt werden (Stein 2009: 11). Die Ersetzung einer Mehrwortbenennung in der Ausgangssprache durch ein einziges Wort in der Zielsprache ist hingegen nicht durchführbar (Stein 2009: 11). Ähnlich verhält es sich bei der Übersetzung von zusammengehörenden Wörtern, die an unterschiedlichen Stellen in einem Satz positioniert sind, zum Beispiel Klammerverben (Stein 2018: 12). Dies äußert sich vorwiegend in Sprachen mit großen Unterschieden in der Syntax (Stein 2009: 12).

Mit der phrasenbasierten SMÜ wollte man diesen Herausforderungen entgegenwirken (Stein 2009: 12). Dabei werden die Trainings- und Testdatensätze in Wortsequenzen, im Englischen n-grams, die jedoch keine Phrasen im herkömmlichen Sinne darstellen, aufgegliedert (Stein 2009: 12, 2018: 13). Die Betrachtung von Wortsequenzen erlaubt somit auch die Übersetzung von Mehrwortbenennungen mit einem einzigen Wort und umgekehrt (Stein 2009: 12). Darüber hinaus kann mit der phrasenbasierten SMÜ der erweiterte Kontext betrachtet werden, wodurch gewisse mehrdeutige Wörter einfacher übersetzt werden können (Stein 2009: 12). Schließlich können je nach Größe der Wortsequenzen Schwierigkeiten bezüglich der Satzstellung beziehungsweise Syntax überwunden werden (Stein 2018: 13).

Ein großer Vorteil der SMÜ im Gegensatz zur RBMÜ liegt in der Kosten- und Zeitersparnis, da linguistische Regeln nicht in umfangreichen Modellen eingespeist werden müssen (Stein 2009: 12). SMÜ-Systeme eignen sich auch für wenig repräsentierte Sprachen, deren Ressourcen für ein RBMÜ-System nicht ausreichen würden (Stein 2009: 12). Die einzige Voraussetzung hierfür sind ausreichend alignierte mehrsprachige Korpora (Stein 2009: 12). Darüber hinaus produzieren SMÜ-Systeme bezüglich lexikalischer Ambiguitäten und Wortwahl bei genügender Repräsentation im Trainingsmaterial bessere Ergebnisse (Stein 2009: 12f, 2018: 14). Die Nachteile der SMÜ stehen in engem Zusammenhang mit den Vorteilen (Stein

2018: 14). So wird die Feststellung gewisser Fehlerquellen durch die Übersetzungserstellung anhand von nicht nachvollziehbaren Berechnungen und unübersichtlichen Textmengen erheblich erschwert (Stein 2009: 13). Eine Korrektur solcher systematischer Fehler gestaltet sich im Vergleich zur RBMÜ durchaus schwierig (Stein 2009: 13). Darüber hinaus gerät die SMÜ bei Sprachpaaren mit unterschiedlicher Struktur an ihre Grenzen (Stein 2009: 13). Die Notwendigkeit großer Korpora zur Erzielung brauchbarer Ergebnisse stellt kleine Sprachen beziehungsweise Sprachen mit niedrigen Ressourcen vor erhebliche Herausforderungen (Stein 2009: 13). Da die SMÜ-Systeme hauptsächlich mit zweisprachigen Korpora aus fachsprachlichen Texten, etwa aus dem rechtlichen Bereich, trainiert sind, haben diese in der Fachsprache im Gegensatz zur Standardsprache eine weitaus höhere Erfolgsquote (Stein 2009: 13). Schließlich ist in der SMÜ der Erstellungsaufwand von großen Textkorpora für valide Wahrscheinlichkeitsberechnungen sehr hoch (Werthmann & Witt 2014: 96).

Wie in Kapitel 2.2.1. ausgeführt, stellte die SMÜ ab Mitte der 2000er Jahre bis 2015 den Stand der Technik dar (Kenny 2018: 434, 2022: 37). Dementsprechend arbeiteten die Systeme großer MÜ-Unternehmen in dieser Zeit nach dem statistischen Ansatz (vgl. Kapitel 2.2.3.). 2015 löste schließlich die NMÜ die SMÜ vor allem aufgrund ihrer besseren Qualität ab (Kenny 2018: 434). Diese soll im folgenden Unterkapitel näher beleuchtet werden.

2.2.2.3. Neuronale maschinelle Übersetzung

Die NMÜ als modernster Ansatz von MÜ basiert auf neuronalen Netzwerken und weist eine höhere Qualität der Ergebnisse maschineller Übersetzung auf (Forcada 2017: 292; Koehn 2020: 29; Schmalz 2019: 195). Ähnlich wie bei der SMÜ werden NMÜ-Systeme mit riesigen Korpora aus Übersetzungseinheiten, also ausgangssprachlichen Segmenten und ihren Übersetzungen, trainiert (Forcada 2017: 292). Diese Korpora können als riesige Übersetzungsspeicher (engl. translation memories) aus Hunderttausenden oder Millionen Übersetzungseinheiten betrachtet werden (Forcada 2017: 292). Im Gegensatz zur SMÜ werden bei diesem MÜ-Ansatz neuronale Netzwerke verwendet (Forcada 2017: 292). Anders als die Bezeichnung neuronale maschinelle Übersetzung vermuten lässt, hat diese mit der Funktionsweise des menschlichen Gehirns keine Gemeinsamkeiten (Forcada 2017: 292). Vielmehr umfassen die neuronalen Netzwerke Tausende künstliche Einheiten, genauer gesagt verbundene Neuronen (Forcada 2017: 292). Die Erregung oder Hemmung dieser Einheiten wird ähnlich wie bei herkömmlichen Neuronen von den Reizen anderer Einheiten und von der Stärke des Reiztransfers aktiviert (Forcada 2017: 292). Diese Einheiten werden zur Erstellung von sogenannten verteilten Repräsentationen (engl. distributed representations) in Schichten geordnet (Forcada 2017: 293). Dabei werden

die Aktivierungsstatus jedes Neurons einer Neuronengruppe zur Erstellung von verteilten Repräsentationen von Wörtern oder Zeichen beziehungsweise Zeichensequenzen und deren Kontexten generiert (Forcada 2017: 293). Dies betrifft sowohl die Verarbeitung des ausgangssprachlichen Satzes als auch die Erstellung eines zielsprachlichen Satzes (Forcada 2017: 293f). Bei einer Repräsentation handelt es sich in diesem Zusammenhang um eine Momentaufnahme eines jeden Neurons in einer Gruppe beziehungsweise Schicht (Forcada 2017: 294). Diese Momentaufnahme enthält einen Vektor, eine Aufstellung von Zahlenwerten mit festgelegter Größe, aus der schließlich eine Übersetzung generiert wird (Forcada 2017: 294). Die Repräsentationen kann man sich in einem Raum vorstellen: ähnliche Wörter beziehungsweise Sätze sollen möglichst nah aneinander sein und ähnliche Koordinaten aufweisen, während unterschiedliche Begriffe weit auseinander liegen und unterschiedliche Koordinaten aufweisen (Forcada 2017: 295). Die Repräsentationen werden schrittweise generiert und sind über eine Vielzahl von Schichten in der Architektur des neuronalen Netzwerks verteilt (Forcada 2017: 295).

Das Ziel der NMÜ ist die Erstellung einer der Humanübersetzung des Trainingsmaterials möglichst ähnlichen Übersetzung durch Analyse ausgangssprachlicher Sätze und Bildung verteilter Repräsentationen (Forcada 2017: 295). Dafür werden neuronale Netzwerke trainiert, indem Gewichtungen und Stärken aller Neuronen-Verbindungen iterativ und langsam gelernt werden (Forcada 2017: 295). Wie zu Beginn dieses Unterkapitels erläutert, benötigt die NMÜ zur Erzielung bestmöglicher Ergebnisse sehr große Trainingskorpora, die wiederum hohe technische Standards bei der Hardware voraussetzen (Forcada 2017: 295). Beim Training wird für eine möglichst geringe Abweichung des MÜ-Ergebnisses von der Humanübersetzung bei den Gewichtungen ein Optimierungsprozess angewandt (Forcada 2017: 295f). Hierbei werden zahlreiche Trainingsdurchläufe durchgeführt, in denen Gewichtungen iterativ und schrittweise solange optimiert werden, bis die Abweichung des MÜ-Ergebnisses von der Humanübersetzung minimiert ist (Forcada 2017: 296).

Bei der NMÜ kann die Übersetzung als Vorhersage angesehen werden. Aufgrund der Repräsentation des Ausgangssatzes werden mithilfe der gelernten Muster und Repräsentationen im neuronalen Netz mögliche Übersetzungen berechnet (Forcada 2017: 296). In diesem Zusammenhang erstellt ein Encoder kontextbezogene Repräsentationen, während ein Decoder auf Basis der Repräsentation des Ausgangssatzes einen Zielsatz generiert (Forcada 2017: 296). Sprachmodelle werden in diesem Kontext zunächst im Pre-Training mit umfangreichen, meist multilingualen Korpora ohne konkrete Aufgabenstellung trainiert. Hierbei lernen die Modelle

kontextbezogene Repräsentationen (Devlin et al. 2019: 4174f). Die vortrainierten Modelle können in einem weiteren Schritt für spezifische Aufgabenstellungen, zum Beispiel MÜ, verfeinert werden. Die gelernten Repräsentationen werden dabei adaptiert, sodass das Modell die Aufgabenstellung bestmöglich erfüllen kann (Devlin et al. 2019: 4175).

Damit ging man von der Aneinanderreihung übersetzter Wörter oder Wortsequenzen bei der SMÜ zu einer Übersetzung unter Heranziehung des ganzen ausgangssprachlichen Segments bei der NMÜ über (Forcada 2017: 301). Darüber hinaus werden bei der NMÜ umfangreiche Repräsentationen generiert (Kenny 2022: 43). Durch die Berücksichtigung ganzer Sätze gestaltet sich bei der NMÜ der Umgang mit schwierigen linguistischen Merkmalen, zum Beispiel diskontinuierlichen Abhängigkeiten, und gewissen Übereinstimmungsaspekten einfacher (Kenny 2022: 43). Es ist daher nicht weiter verwunderlich, dass die NMÜ bisher die qualitativ besten Ergebnisse im Bereich der MÜ erzielt (Kenny 2022: 43). Die Ergebnisse der NMÜ stellen jedoch gleichzeitig eine große Herausforderung dar. Die NMÜ kann nämlich flüssige Übersetzungen erstellen, die jedoch gleichzeitig fehlerhaft sind (Kenny 2022: 43). Durch solche natürlich klingenden Ergebnisse können gewisse Fehler unentdeckt bleiben (Kenny 2022: 43). Außerdem können durch das Training mit riesigen Textkorpora bei der NMÜ gewisse Vorurteile, wie etwa durch die übermäßige Verwendung männlicher Formen, verstärkt werden (Kenny 2022: 43).

Der Fortschritt der NMÜ ist sowohl in den für die breite Öffentlichkeit online verfügbaren Tools als auch in der aktuellen Forschung auf dem Gebiet der MÜ zu beobachten (Koehn 2020: 40). Demzufolge ist es durchaus nachvollziehbar, dass große MÜ-Unternehmen, wie zum Beispiel Google, DeepL, SYSTRAN oder Microsoft, bereits neuronale Modelle in ihre MÜ-Technologien integriert haben (vgl. Rivera-Trigueros 2022: 596).

2.2.3. Verfügbare Systeme neuronaler maschineller Übersetzung

Heutzutage sind online zahlreiche MÜ-Systeme verfügbar. Zu den bekanntesten Systemen maschineller Übersetzung gehören DeepL, Google Translate, Microsoft Bing Translator und SYSTRAN (vgl. Rivera-Trigueros 2022: 612). Diese werden im vorliegenden Unterkapitel näher beschrieben.

Google veröffentlichte im Jahr 2007 erstmals sein selbstlernendes Übersetzungssystem Google Translate, im Deutschen auch Google Übersetzer genannt (Schmalz 2019: 198). Dabei handelte es sich anfangs um ein System statistischer maschineller Übersetzung (Schmalz 2019: 198). 2016 brachte Google schließlich das erste Übersetzungssystem auf den Markt, das auf

neuronalen Netzwerken basierte (vgl. Wu et al. 2016). Damit erhoffte man sich in erster Linie, an eine möglichst natürliche, menschenähnliche Übersetzung heranzukommen (Wu et al. 2016: 1). Google Translate ist sowohl online im Browser als auch als App für Android und iOS verfügbar (Google Übersetzer 2024b). Heute unterstützt Google Translate nach einer Erweiterung um mehr als 110 neue Sprachen im Juni 2024 insgesamt 249 Sprachen (Google 2024; Google Übersetzer 2024a). Diese bis jetzt umfangreichste Erweiterung wurde unter Verwendung von neuronalen Sprachmodellen, wie etwa dem großen Sprachmodell PaLM 2, ermöglicht, wodurch man dem Ziel, die 1.000 meistgesprochenen Sprachen weltweit zu unterstützen, einen Schritt näher kam (Google 2024). Google Translate ermöglicht die Übersetzung verschiedener Texte, wie zum Beispiel von Dokumenten oder Websites (Google Übersetzer 2024b). Darüber hinaus kann man mit Google Translate mit der Kamera übersetzen, mit anderssprachigen Menschen kommunizieren, gesprochene Inhalte simultan übersetzen lassen, direkt in allen Apps übersetzen oder „Text durch Tippen, Sprechen oder per Handschrift eingeben“ (Google Übersetzer 2024a).

DeepL stellt einen weiteren Onlinedienst für maschinelle Übersetzung dar. Das von Jaroslaw Kutylowski gegründete Unternehmen brachte den DeepL Übersetzer 2017 auf den Markt (DeepL 2021: 2). Die Firma vertrieb zuvor unter dem Namen Linguee GmbH die Übersetzungssuchmaschine Linguee, mit deren Datensatz das MÜ-System schließlich trainiert wurde (DeepL 2021: 2). Mithilfe von neuronalen Netzen wurde eine bis dahin unbekannte Übersetzungsqualität erreicht (DeepL 2021: 2). Seit 2018 wurde der Übersetzungsdienst durch die kostenpflichtige Version DeepL Pro ergänzt, die vor allem interne Abläufe in Unternehmen durch höhere Datensicherheit sowie unbegrenztes Textvolumen effizienter gestaltet (DeepL 2021: 2, 2024b). Darüber hinaus umfasst die kostenpflichtige Version eine API, die es Entwickler*innen ermöglicht, DeepL in ihre Anwendungen, Websites oder Software zu integrieren (DeepL 2021: 2). Heute unterstützt DeepL 30 Sprachen und ermöglicht etwa die Übersetzung von Dokumenten, Websites, E-Mails und Bildern (DeepL 2024a, 2024c). 2020 wurde DeepL mit der Glossar-Funktion der erste kostenlose Übersetzungsdienst, bei dem man genau festlegen kann, wie bestimmte Wörter beziehungsweise Ausdrücke übersetzt werden sollen (DeepL 2021: 2). DeepL ist außerdem als App für Windows und macOS verfügbar, mit der man direkt in allen Apps, die sich auf dem Gerät befinden, übersetzen kann, sowie als mobile App für Android und iOS, auf die Nutzer*innen im Alltag jederzeit zugreifen können (DeepL 2021: 2).

SYStem und TRANslation (SYSTRAN) gilt als Pionier in der maschinellen Übersetzung (vgl. SYSTRAN 2024b). Das Unternehmen wurde 1968 von Peter Toma auf Grundlage seiner Forschung an der Georgetown University zur Gewährleistung der Kommunikation zwischen Personen in mehreren Sprachen durch Software gegründet (Hutchins & Somers 1992: 175f; SYSTRAN 2024b, 2024c). Während des Kalten Krieges arbeitete SYSTRAN mit der US Air Force zusammen, um die erste Übersetzungssoftware für das Sprachpaar Russisch-Englisch zu entwickeln (SYSTRAN 2024c). Im Laufe der Zeit folgten weitere Sprachpaare, zum Beispiel Englisch-Französisch und Englisch-Italienisch, die bei verschiedenen militärischen beziehungsweise staatlichen Institutionen, wie etwa der Europäischen Kommission, Anwendung fanden (Hutchins & Somers 1992: 176). 1997 entwickelte SYSTRAN schließlich Babel Fish, den ersten kostenlosen Online-Übersetzungsdienst für AltaVista, nach dem regelbasierten Ansatz (Schmalz 2019: 194; SYSTRAN 2024c). SYSTRAN richtet sich an ein breites Spektrum von Nutzer*innen – von Privatpersonen, die schnelle Übersetzungen benötigen, bis hin zu Unternehmen, die auf maßgeschneiderte Lösungen angewiesen sind (vgl. SYSTRAN 2024d). Heute ermöglicht SYSTRAN mit dem kostenlosen Online-Übersetzungsdienst SYSTRAN Translate, der auf neuronalen Netzwerken basiert, die Übersetzung von Texten in 55 Sprachen und mit der kostenpflichtigen Cloud-basierten Version SYSTRAN Translate Pro die Übersetzung von ganzen Dokumenten mit fachgebietsspezifischen Übersetzungsfunktionen (SYSTRAN 2024a, 2024d). Mit SYSTRAN model Studio wird Unternehmen etwa die Möglichkeit geboten, eigene Übersetzungsmodelle zu erstellen, die mit eigenen Translation Memories oder zweisprachigen Dateien zur Verbesserung der Übersetzungsqualität sowie zur Steigerung der Produktivität trainiert werden können (SYSTRAN 2024c).

Schließlich stellt das Unternehmen Microsoft einen weiteren bedeutenden Dienstleister für MÜ dar. Die ersten maschinellen Übersetzungssysteme von Microsoft sind inzwischen über 20 Jahre alt (Doss Mohan & Skotdal 2021). 2003 erreichte Microsoft einen Meilenstein, indem die gesamte Microsoft Knowledge Base mit einem maschinellen Übersetzungssystem aus dem Englischen ins Spanische, Französische, Deutsche und Japanische übersetzt und die Ergebnisse auf der Website veröffentlicht wurden. Dies stellte damals die umfangreichste öffentlich zugängliche maschinelle Rohübersetzung im Internet dar (Doss Mohan & Skotdal 2021). Zunächst wurden die MÜ-Systeme nach dem statistischen Ansatz entwickelt, mit der Zeit stieg Microsoft wie die anderen Dienstleister auf NMÜ um, genauer gesagt auf die Transformer-Architektur, was zu einer beachtlichen Steigerung der Übersetzungsqualität führte (Doss

Mohan & Skotdal 2021). Die Transformer-Architektur ermöglichte es darüber hinaus, Modelle für ressourcenarme Sprachen leichter zu erstellen, da sie mit kleineren Datenmengen arbeitet und Trainingsdaten aus verwandten Sprachen einbeziehen kann (Doss Mohan & Skotdal 2021). Dementsprechend unterstützen sowohl der im Browser zugängliche Microsoft Bing Translator als auch die im App Store, Google Play Store und auf Amazon erhältliche App Microsoft Translator über 130 Sprachen (vgl. Microsoft 2024a, 2024b, 2024c). Im Microsoft Bing Translator kann lediglich Text übersetzt werden (Microsoft 2024a). Den Microsoft Translator hingegen kann man für persönliche Zwecke im Alltag, etwa für die Übersetzung von Gesprächen, Straßenschildern, Speisekarten, Websites oder ganzen Dokumenten, für geschäftliche Zwecke im Sinne einer mehrsprachigen Kundeninteraktion sowie für die Ausbildung im Klassenzimmer verwenden (Microsoft 2024c).

2.2.4. Methoden zur Qualitätsanalyse

Bei der MÜ stellt sich immer die Frage nach ihrer Qualität (vgl. Koehn 2020: 41). In diesem Zusammenhang wurden zahlreiche automatische sowie menschliche Methoden zur Qualitätsanalyse entwickelt. Diese sollen in den folgenden Unterkapiteln näher erläutert werden.

2.2.4.1. Automatische Evaluierungsmethoden

Automatische Evaluierungsmethoden wurden aus Effizienzgründen eingeführt und sind für das Training von NMÜ-Systemen unerlässlich. Außerdem können diese häufiger und kostengünstig die Qualität der MÜ beurteilen (Koehn 2020: 53). Bei vielen gängigen Evaluierungsmetriken werden Humanübersetzungen als Referenz herangezogen und mit der MÜ verglichen (Koehn 2020: 53). Zu den bekanntesten automatischen Evaluierungsmetriken der Qualität von MÜ gehören etwa Bilingual Evaluation Understudy (BLEU) (Papineni et al. 2002), Metric for Evaluation of Translation with Explicit Ordering (METEOR) (Banerjee & Lavie 2005) und Translation Error Rate (TER) (Snover et al. 2006).

Bei BLEU werden nicht nur die mit der Referenz übereinstimmenden Wörter gezählt, sondern auch Übereinstimmungen von N-Grammen beziehungsweise Wortsequenzen (Papineni et al. 2002: 312). Die Punktzahl liegt dabei zwischen 0 und 1, wobei eine höhere Punktzahl eine größere Übereinstimmung und eine Punktzahl von 1 eine hundertprozentige Übereinstimmung mit der Referenzübersetzung bedeutet (Papineni et al. 2002: 315). In diesem Zusammenhang gibt es mehrere Möglichkeiten zur Verfeinerung der Ergebnisse. Man kann etwa mehrere Referenzübersetzungen heranziehen; wenn die MÜ mit einer von diesen hohe N-Gramm-Übereinstimmungen aufweist, kann von hoher Qualität gesprochen werden (Papineni

et al. 2002: 312). Bei der BLEU-Metrik wird die Genauigkeit, das heißt das Verhältnis der Wörter der MÜ, die mit der Referenzübersetzung übereinstimmen, gemessen (Papineni et al. 2002: 312). Die Satzlänge stellt dabei eine gewisse Herausforderung dar. In einigen Fällen können zu kurze Sätze trotz fehlender Informationen eine hohe Punktzahl erreichen, weil ein System öfter vorkommende Wörter übersetzt (Papineni et al. 2002: 314). In diesem Kontext wird diese Genauigkeit in Metriken mit einer sogenannten Trefferquote verbunden, bei dem das Verhältnis der mit der MÜ übereinstimmenden Wörter in der Referenzübersetzung gemessen wird (Papineni et al. 2002: 314). Da BLEU jedoch mehrere Referenzübersetzungen heranzieht, erweist sich eine Trefferquote als durchaus schwierig; schließlich kann ein Satz auf verschiedene Weisen korrekt übersetzt werden (Papineni et al. 2002: 314). Referenzübersetzungen variieren oft in ihrer Länge, wodurch bei kürzeren maschinellen Übersetzungen und längeren Referenzübersetzungen adäquatere Optionen schlechter abschneiden können (Papineni et al. 2002: 315). Um dieser Problematik entgegenzuwirken, wurde ein Strafabzug für die Kürze (engl. brevity penalty) eingeführt (Papineni et al. 2002: 315). Für eine möglichst hohe Bewertung sollen Länge, Wortwahl und Wortfolge einer MÜ mit jenen der Referenzübersetzungen übereinstimmen (Papineni et al. 2002: 315). Dabei wird nur die Länge der zielsprachlichen Referenzübersetzungen und nicht direkt die Länge des Ausgangstextes miteinbezogen (Papineni et al. 2002: 315). Der Strafabzug wird bei gleicher MÜ- und Referenzübersetzungslänge auf 1 gesetzt (Papineni et al. 2002: 315). Die beste Übereinstimmungslänge bezieht sich daher auf jene Referenzübersetzung, die der MÜ im Hinblick auf die Länge am nächsten kommt (Papineni et al. 2002: 315). Der Strafabzug wird nicht satzweise, sondern über den gesamten Korpus als exponentielle Funktion berechnet (Papineni et al. 2002: 315).

METEOR griff die Idee von BLEU auf und entwickelte sie unter Beachtung ihrer Schwächen, beinahe Übereinstimmungen nicht in die Bewertung miteinzubeziehen, weiter (Banerjee & Lavie 2005: 66f). METEOR ermöglichte nun die Berücksichtigung von Synonymen und morphologischen Abweichungen, indem die Oberflächenformen der Wörter abgeglichen werden und anschließend auf Wortstämme und semantische Klassen zurückgegriffen wird (Banerjee & Lavie 2005: 67ff). Die Berechnung der METEOR-Punktzahl erfolgt dabei auf Grundlage der Anzahl der exakten Wort-für-Wort-Übereinstimmungen zwischen der MÜ und der Referenzübersetzung (Banerjee & Lavie 2005: 67). Anders als bei BLEU kommen bei METEOR nicht direkt mehrere Referenzübersetzungen zur Anwendung. METEOR wählt jene Referenzübersetzung aus, die bei der jeweiligen MÜ-Bewertung das beste Ergebnis erzielt

(Banerjee & Lavie 2005: 67). Wörter ohne Übereinstimmung werden in einem nächsten Schritt im Hinblick auf Synonyme und morphologische Abweichungen geprüft (Banerjee & Lavie 2005: 67). Darüber hinaus wird bei METEOR auch eine Strafe (engl. penalty) angesetzt. Damit sollen fragmentierte Übersetzungen mit unzusammenhängenden Teilen und abweichender Wortstellung bestraft und zusammenhängende Übersetzungen bevorzugt werden (Banerjee & Lavie 2005: 68f). Die finale Punktzahl setzt sich schließlich aus dem Zusammenspiel der Genauigkeit einzelner Wörter, deren Trefferquote und der Strafe zusammen (Banerjee & Lavie 2005: 69). Daraus ergibt sich aber auch ein großer Nachteil dieser Methode: Die Formel zur Berechnung der Punktzahl ist um Einiges komplizierter als jene von BLEU (vgl. Banerjee & Lavie 2005).

TER wurde als Verbesserung der Word Error Rate (WER) eingeführt, bei der die Anzahl von hinzugefügten, gelöschten und ersetzten Wörtern gezählt wird (Snover et al. 2006: 225f). Da die WER Verschiebungen von Phrasen innerhalb eines Satzes nicht berücksichtigt, wurde ihr mit TER eine entsprechende Komponente hinzugefügt (Snover et al. 2006: 225f). Somit bezeichnet TER „the minimum number of edits needed to change a hypothesis so that it exactly matches one of the references, normalized by the average length of the references“ (Snover et al. 2006: 225). Folglich gilt, je niedriger die TER, desto besser ist die MÜ (Snover et al. 2006: 228). TER erlaubt zwar die Verwendung mehrerer Referenzen, die Anzahl der nötigen Änderungen wird jedoch nur unter Berücksichtigung der besten Referenzübersetzung und der MÜ berechnet (Snover et al. 2006: 226). Die Akzeptabilität einer MÜ kann allerdings nicht vollständig von der TER-Bewertung abgedeckt werden, da diese semantische Äquivalenzen wie Synonyme außer Acht lässt (Snover et al. 2006: 226). Um dieser Problematik entgegenzuwirken und zuverlässigere Ergebnisse zu erzielen, sind durch menschliche Annotator*innen angepasste Referenzen erforderlich (Snover et al. 2006: 226). In diesem Kontext haben Snover et al. (2006: 226) TER durch eine menschliche Komponente HTER erweitert. Snover et al. (2006: 230) kamen zu dem Schluss, dass TER für Forschungszwecke eingesetzt werden kann. TER erzielt darüber hinaus mit nur einer Referenzübersetzung eine ebenso hohe Korrelation mit menschlichen Bewertungen wie BLEU mit mehreren Referenzübersetzungen (Snover et al. 2006: 223). Daher ist bei der Bewertung von einzelnen Sätzen die TER anderen Methoden wie etwa BLEU vorzuziehen (vgl. Snover et al. 2006: 230f).

Neben diesen automatischen Qualitätsmetriken, die Referenzübersetzungen heranziehen, gibt es alternative Methoden der automatischen Evaluierung. Da viele maschinelle Übersetzungen bestimmte Qualitätsansprüche noch nicht erfüllen können und automatische

Qualitätsmetriken wie BLEU, METEOR oder TER durch den Bedarf an Referenzübersetzungen hinsichtlich der zu generierenden Datenmenge eingeschränkt sind, hat sich etwa die Quality Estimation (QE) als automatische Evaluierungsmethode zur Vorhersage der Qualität von MÜ etabliert (Specia et al. 2010: 40; Specia & Shah 2018: 201). Ziel der QE ist die Qualitätseinschätzung von Ergebnissen eines MÜ-Systems ohne Heranziehung von Referenzübersetzungen (Specia et al. 2010: 40). In diesem Zusammenhang fokussiert sich die QE nicht auf die Bewertung der Gesamtleistung eines MÜ-Systems, sondern auf die Leistung in Bezug auf einzelne Übersetzungen (Specia & Shah 2018: 202). Dafür werden referenzunabhängige Merkmale aus den eingegeben Sätzen und ihren Übersetzungen ausgelesen (Specia et al. 2010: 39). Anhand von eigens für die direkte Bewertung trainierten Modellen wird schließlich die Qualität eines anderen MÜ-Modells bewertet (Specia et al. 2010: 39). Der hauptsächliche Grund für die Entwicklung einer solchen Methode war die Gestaltung einer schnelleren und effizienteren Qualitätskontrolle der MÜ-Ausgabe, ohne dass professionelle Übersetzer*innen sowohl den Ausgangstext als auch die MÜ lesen müssen oder Referenzübersetzungen erforderlich sind (Specia et al. 2010: 40). Darüber hinaus wollte man gewisse Anwendungen für den Einsatz in bestimmten Situationen, wo die Qualität einzelner MÜ-Ausgaben ohne zugängliche Referenzübersetzungen bewertet werden muss, effizienter gestalten (Specia & Shah 2018: 202). QE-Methoden können auf mehrere Arten angewendet werden. Zu den geläufigsten Anwendungsbeispielen zählen das Herausfiltern von schlechteren Übersetzungsergebnissen zur MÜ-Nachbearbeitung und die Auswahl des besten MÜ-Systems bei Übersetzung eines Segments durch mehrere Systeme (Specia et al. 2010: 48). Weiters kann die QE etwa zur Aufwandseinschätzung der Nachbearbeitung eines Segments, zur Bewertung, ob eine MÜ für das eigenständige Training von MÜ-Systemen geeignet ist, sowie zu einer stichprobenartigen Übersetzungsauswahl für die manuelle Bewertung eingesetzt werden (Specia & Shah 2018: 203). Neben diesen Anwendungsmöglichkeiten, die vorwiegend auf Satz- beziehungsweise Segmentebene verwendet werden, kann die QE auch auf Wort- und Dokumentenebene eingesetzt werden (Specia & Shah 2018: 231ff). Auf Wortebene wird die QE zur Fehleridentifikation bei Wörtern eines übersetzten Segments verwendet (Specia & Shah 2018: 231f). In diesem Zusammenhang können folgende Anwendungsmöglichkeiten genannt werden: Kennzeichnung von bei der Nachbearbeitung zu korrigierenden Wörtern; Sensibilisierung von Benutzer*innen hinsichtlich nicht authentischer Satzteile einer Übersetzung; Auslese der besten Übersetzungsoptionen für Wörter beziehungsweise Phrasen aus mehreren MÜ-Systemen; Unterstützung der automatischen Nachbearbeitung (Specia & Shah 2018: 232). Bei der QE auf Dokumentenebene geht es hauptsächlich um eine grobe Bewertung

der Gesamtqualität ganzer Dokumente (Specia & Shah 2018: 232). Durch die Betrachtung des gesamten Textes kann gewährleistet werden, dass eine Übersetzung in sich kohärent ist (Specia & Shah 2018: 232). Die QE auf Dokumentenebene wird in diesem Kontext vor allem für das grobe Textverständnis ohne weitere Nachbearbeitung genutzt (Specia & Shah 2018: 232). Schließlich erlaubt die QE eine flexiblere Modellierung des Qualitätskonzepts in Abhängigkeit der Bedürfnisse der jeweiligen Benutzer*innen (Specia & Shah 2018: 202). Im Vergleich zu anderen Evaluierungsmetriken wie BLEU, METEOR oder TER korreliert diese automatische Evaluierungsmethode besser mit der humanen Evaluierung (Specia et al. 2010: 39). Dies trifft auch bei Modellen, die auf mehreren MÜ-Systemen beziehungsweise mit unterschiedlichen Sprachpaaren und Texten trainiert wurden, zu (Specia et al. 2010: 39).

2.2.4.2. Menschliche Evaluierungsmethoden

Zu den geläufigsten menschlichen Evaluierungsmethoden gehört die Bewertung maschineller Übersetzung anhand der beiden Kriterien Adäquatheit (engl. adequacy) und Sprachfluss (engl. fluency) (Koehn 2020: 45f). Adäquatheit beschreibt, inwiefern die Übersetzung den Sinn des Ausgangstextes wiedergibt (Castilho et al. 2018: 18). Sprachfluss hingegen ist auf die Zielsprache ausgelegt und beschreibt, inwiefern die Übersetzung den Regeln der Zielsprache entspricht (Castilho et al. 2018: 18). Die beiden Kategorien werden normalerweise anhand einer Ordinalskala in Form von Likert-Skalen auf Segmentebene bewertet (Castilho et al. 2018: 18). Darüber hinaus werden maschinelle Übersetzungen anhand verschiedener Kriterien sowohl auf Makro- als auch auf Mikroebene evaluiert (vgl. Castilho et al. 2018; Van Slype 1979). Auf Mikroebene ist etwa die Methode des Rankings durchaus üblich (vgl. Castilho et al. 2018: 21). Dabei reihen Evaluator*innen mehrere zielsprachliche Sätze anhand gewisser Kriterien, wie zum Beispiel Sprachfluss (Castilho et al. 2018: 21). Dieser relativ schnelle und effiziente Ansatz wird üblicherweise in der Forschung herangezogen und wird bei umfangreichen Vorgängen von maschinellen Übersetzungssystemen eingesetzt (Castilho et al. 2018: 21).

In der vorliegenden Masterarbeit soll eingehender auf die Evaluierung auf Makroebene eingegangen werden, da diese in Bezug auf die vorgesehene Methodik am relevantesten ist. Laut Van Slype (1979: 12) hat die Evaluierung von Übersetzungen auf Makroebene verschiedene Ziele. Einerseits wird das Ausmaß, inwieweit ein Übersetzungssystem akzeptable Ergebnisse erbringt, untersucht. Andererseits können auch ein Qualitätsvergleich von zwei Übersetzungssystemen oder zwei Varianten desselben Systems sowie eine Bewertung der Brauchbarkeit der Übersetzungen eines Systems durchgeführt werden. In diesem Zusammenhang unterscheidet Van Slype (1979: 57) auf den vier Ebenen kognitiv, ökonomisch,

linguistisch und operational zwischen verschiedenen Kriterien. Für diese Masterarbeit sind vor allem die Kriterien auf kognitiver Ebene interessant. Diese sind: Verständlichkeit (engl. intellegibility), Wiedergabetreue (engl. fidelity), Kohärenz (engl. coherence), Brauchbarkeit (engl. usability) und Akzeptabilität (engl. acceptability) (Van Slype 1979: 57). Unter Verständlichkeit versteht man in diesem Zusammenhang das Ausmaß der Klarheit der Übersetzung für die Rezipient*innen (Van Slype 1979: 62). Bei der Wiedergabetreue wird bewertet, inwieweit ein Segment in der Übersetzung die im Ausgangstext enthaltenen Informationen unverfälscht wiedergibt (Van Slype 1979: 72). Die Kohärenz beschreibt laut Van Slype (1979: 78) die Qualität einer Übersetzung in Bezug auf ihren inneren Zusammenhang, ohne dass ihre Richtigkeit im Vergleich zum Originaltext untersucht werden muss. Brauchbarkeit bezieht sich darauf, wie gut eine Übersetzung für die praktische Nutzung geeignet ist, das heißt die tatsächliche Verwendbarkeit der Übersetzung in realen Kontexten (Van Slype 1979: 79). Akzeptabilität definiert Van Slype (1979: 92) als subjektive Bewertung, inwieweit eine Übersetzung für die Endbenutzer*innen akzeptabel ist. Hinsichtlich der Akzeptabilität betont Van Slype (1979: 13), dass diese nur durch eine Befragung der Endbenutzer*innen effektiv gemessen werden kann. Diese Kriterien werden, wie Van Slype (1979) in diversen Studien beobachtet hat, auf verschiedenste Weise bewertet, vor allem aber anhand von Skalen. Mit Ausnahme der Wiedergabetreue können all diese Kriterien der kognitiven Ebene laut Van Slype (1979) lediglich anhand der Übersetzung ohne Hinzunahme des Ausgangstextes bewertet werden. Auf ökonomischer Ebene unterscheidet Van Slype (1979: 13) zwischen Lesezeit, Korrekturzeit und Übersetzungszeit. Diese Ebene berücksichtigt dabei nicht die Kosten der MÜ (Van Slype 1979: 94). Die linguistische Ebene umfasst die Rekonstruktion semantischer Beziehungen, syntaktische und semantische Kohärenz, „absolute“ Qualität, lexikalische Bewertung, syntaktische Bewertung und Fehleranalyse (Van Slype 1979: 13). Zuletzt gehören zur operationalen Ebene die automatische Spracherkennung und die Überprüfung der Herstelleransprüche (Van Slype 1979: 13).

Schließlich ist noch die Multidimensional Quality Metrics (MQM), die als Weiterentwicklung eines Versuchs zur Standardisierung der Evaluierung in Form des LISA QA Models gilt, zu nennen (vgl. Lommel et al. 2014: 458; Lommel 2018: 113f). Der MQM-Framework ist hierarchisch aufgebaut und beinhaltet seit einer Aktualisierung im Jahr 2024 sieben Fehlertypen, die sich auf bis zu drei weiteren Ebenen mit zunehmender Spezifität erstrecken (Lommel 2018: 116f; The MQM Council 2024a). Mit Stand 2024 umfasst die Metrik insgesamt 135 Problemtypen (vgl. The MQM Council 2024b). Zu den sieben Hauptfehlertypen

auf erster Ebene gehören Terminologie (engl. terminology), Genauigkeit (engl. accuracy), linguistische Konventionen (engl. linguistic conventions), Stil (engl. style), lokale Konventionen (engl. locale conventions), Angemessenheit für die Zielgruppe (engl. audience appropriateness) sowie Gestaltung und Markup (engl. design and markup). Ein Terminologiefehler liegt vor, wenn ein verwendeter Terminus nicht den festgelegten Terminologiestandards des Fachbereichs oder der Organisation entspricht oder wenn der Terminus im Zieltext nicht die angemessene und standardisierte Entsprechung des Terminus im Ausgangstext darstellt (The MQM Council 2024b). Fehler des Typs Genauigkeit treten auf, wenn der Inhalt des Zieltextes durch Verzerrungen, Auslassungen oder Hinzufügungen nicht den Inhalt des Ausgangstextes wiedergibt (The MQM Council 2024b). Unter linguistische Konventionen, ehemals Sprachfluss (engl. fluency), fallen Fehler der sprachlichen Textform. Dazu zählen etwa Grammatik- beziehungsweise Idiomatikfehler oder maschinell bedingte Probleme (The MQM Council 2024a, 2024b). Die Kategorie Stil bezieht sich auf Fehler, die nicht grammatikalischer Natur sind, sich jedoch durch Abweichung von organisatorischen Stilrichtlinien oder Verwendung eines unangemessenen Sprachstils auszeichnen (The MQM Council 2024b). Gegen die lokalen Konventionen wird verstoßen, wenn die Übersetzung lokale Inhalts- oder Formatierungsanforderungen für Datenelemente nicht erfüllt, wie zum Beispiel Verwendung falscher Maßeinheiten (The MQM Council 2024b). Angemessenheit für die Zielgruppe, ehemals Realitätsbezug (engl. verity), bezieht sich auf Inhalte, die für das Zielgebiet und die Zielgruppe inadäquat sind. Konkret geht es dabei um die Verwendung von im Zieltext schwer verständlichen kulturspezifischen Referenzen oder von Inhalten, die von der Zielgruppe als offensiv empfunden werden könnten (The MQM Council 2024a, 2024b). Unter den Fehlertyp Gestaltung und Markup fallen Fehler hinsichtlich der äußerlichen Gestaltung eines Übersetzungsprodukts. Dazu gehören die Formatierung und Auszeichnung von Zeichen, Absätzen und UI-Elementen, die Einheit von Text und grafischen Elementen sowie das Gesamtlayout (The MQM Council 2024b). Schließlich ist noch ein allgemein individuell anpassbarer Fehlertyp (engl. custom) zu nennen, auf den man zurückgreifen kann, wenn ein Fehler von keinem der sieben Hauptfehlertypen abgedeckt wird (The MQM Council 2024b).

Der MQM-Framework kann dabei je nach Anforderungen und Spezifikationen der individuellen Übersetzungsprojekte durch relevantere Fehlertypen ergänzt beziehungsweise im Allgemeinen angepasst werden, um eine zielgerichtetere Bewertung zu erhalten (Lommel 2018: 120). Bei der Evaluierung einer Übersetzung reicht die Feststellung der Fehleranzahl jedoch nicht immer aus. Ebenso ist laut Lommel (2018: 120) die Kenntnis über den Schweregrad und

die Relevanz eines Fehlertyps im Kontext einer spezifischen Übersetzungsaufgabe von Bedeutung. In diesem Zusammenhang wurden vier verschiedene Schweregrade, die mit dem MQM-Framework kompatibel sind, definiert: kritische Fehler (engl. critical errors), schwere Fehler (engl. major errors), leichte Fehler (engl. minor errors) und keine Fehler (engl. null) (Lommel 2018: 120f). Der Schweregrad wirkt sich dabei direkt auf die Brauchbarkeit der Übersetzung für die Nutzer*innen aus und ist nur auf individuelle Fehler anwendbar (Lommel 2018: 120f). Kritische Fehler verhindern auch nur in geringstem Ausmaß die Zweckerfüllung einer Übersetzung. Sie können etwa in Bedienungsanleitungen die Befolgung von Anweisungen durch Nutzer*innen in erheblichem Ausmaß erschweren und sicherheitsrelevante beziehungsweise rechtliche Auswirkungen mit sich bringen (Lommel 2018: 120). Als Beispiel führt Lommel (2018: 120) die fälschliche Übertragung der Gewichtsgrenze für industrielle Zentrifugen *2 pounds* durch *2 Kilogramm* statt *0,9 Kilogramm* an. Derartige Fehler können fatal werden, wenn sie aus dem Text nicht deutlich hervorgehen. Schwere Fehler haben hingegen keine schädlichen Auswirkungen, führen aber dazu, dass der beabsichtigte Sinn des Textes unklar und für die Nutzer*innen nicht verständlich ist (Lommel 2018: 120). Diese Fehler sollten vor der Veröffentlichung korrigiert werden. Ansonsten ziehen sie außer einer möglichen Verärgerung der Nutzer*innen keine schwerwiegenden Folgen nach sich (Lommel 2018: 120). Leichte Fehler wirken sich auf die Brauchbarkeit einer Übersetzung nicht negativ aus. Dabei handelt es sich um leichte Rechtschreib- oder Grammatikfehler, die durch die Nutzer*innen beim Durchlesen automatisch beziehungsweise unbewusst korrigiert werden. Solche Fehler müssen vor der Veröffentlichung nicht zwingend behoben werden (Lommel 2018: 120f). Der Schweregrad null bezieht sich auf Änderungen, die keine Fehler darstellen. Darunter fallen etwa nachträgliche Änderungen in der Terminologie, die durch Revisor*innen als Fehler mit diesem Schweregrad markiert werden können (Lommel 2018: 121). Jedem Schweregrad ist eine gewisse Anzahl an Fehlerpunkten zuzuordnen: 100 Punkte für kritische Fehler, 10 Punkte für schwere Fehler, 1 Punkt für leichte Fehler und 0 Punkte für keine Fehler (Lommel 2018: 121).

Darüber hinaus kann die Relevanz von gewissen Fehlerkategorien durch Zuordnung von relativen Werten unterschiedlich gewichtet werden. Die Gewichtung bezieht sich dementsprechend nicht auf individuelle Fehler, sondern auf ganze Kategorien (Lommel 2018: 121). Je nach Kontext können bestimmte Fehlertypen, wie zum Beispiel Stil, als weniger relevant gewichtet werden, solange der beabsichtigte Sinn nicht beeinträchtigt wird (Lommel 2018: 121). Bei MQM beträgt die Standard-Gewichtung 1,0, wobei höhere Werte eine höhere Relevanz und niedrigere Werte eine niedrigere Relevanz der Fehler darstellen. Wenn in einem

Projekt zum Beispiel Terminologie von höherer Relevanz ist, könnte man dem Fehlertyp Terminologie einen Wert von 1,5 zuordnen. Dies würde bedeuten, dass ein Terminologiefehler um 50% mehr als der Standardwert gewichtet wird (Lommel 2018: 121). Weniger relevanten Fehlern, in diesem Fall etwa Stilfehlern, kann in weiterer Folge ein Wert von 0,5 zugewiesen werden (Lommel 2018: 121). Gewichtungen finden in Software weniger Anwendung. Es besteht jedoch die Möglichkeit, derartige Gewichtungen im Rahmen der Bewertung einzubeziehen (Lommel 2018: 121).

Der MQM-Framework ermöglicht sowohl eine holistische als auch eine analytische Bewertung (Lommel 2018: 122). Während sich die holistische Bewertung auf die Übersetzung als Ganzes konzentriert und die Qualitätsbewertung anhand allgemeiner Kriterien erfolgt, fokussiert sich die analytische Bewertung auf einzelne Fehler (Lommel 2018: 122). Demzufolge eignet sich die holistische Bewertung für eine überblicksmäßige Überprüfung der Erfüllung der Übersetzungsvorgaben, liefert jedoch keine detaillierten Rückmeldungen zu spezifischen Fehlern (Lommel 2018: 122). Mit der analytischen Bewertung hingegen lassen sich konkrete Probleme feststellen, der Gesamteindruck der Übersetzung wird allerdings nicht berücksichtigt. Außerdem ist diese Bewertungsart zeitaufwendig und erfordert für eine gewisse Bewertungskonsistenz entsprechende Schulungen (Lommel 2018: 122). Für die Verwendung des MQM-Frameworks nach dem holistischen Ansatz muss eine Metrik angelegt werden, die die Dimensionen der analytischen Metrik abdeckt. Zu diesem Zweck sollten Ersteller*innen der Metrik die Hauptfehlertypen aus der analytischen Metrik heranziehen und Evaluator*innen sensibilisieren, diese bei der gesamten Übersetzung zu beachten (Lommel 2018: 122f). Wenn also eine analytische Metrik die Kategorien Grammatik, Interpunktion und Rechtschreibung unter dem Hauptfehlertyp linguistische Konventionen berücksichtigt, würde die entsprechende holistische Metrik nur linguistische Konventionen beinhalten (vgl. Lommel 2018: 123). Eine Bewertung würde in diesem Zusammenhang anhand einer Likert-Skala, zum Beispiel auf einer Skala von 0 (vollkommen inakzeptabel) bis 5 (vollkommen akzeptabel), oder einem ähnlichen Bewertungssystem erfolgen (Lommel 2018: 123). Eine derartige Befragung zu jedem Hauptfehlertyp ermöglicht den Evaluator*innen die Überprüfung der Erfüllung der Übersetzungsvorgaben, ohne einzelne Fehler hervorzuheben (Lommel 2018: 123). Führt die holistische Bewertung zu keinem negativen Ergebnis, ist der Bedarf nach einer detaillierten analytischen Bewertung nicht gegeben, außer die Folgen unentdeckter Fehler wären gravierend (Lommel 2018: 123). Fällt das Ergebnis der holistischen Bewertung negativ aus, ist eine detaillierte analytische Bewertung hinfällig. Bevor die Übersetzung zur Korrektur übermittelt wird,

müssten Revisor*innen lediglich Fehlerbeispiele zu Dokumentationszwecken aufzeichnen (Lommel 2018: 123). Eine vollständige analytische Bewertung ist nur dann erforderlich, wenn die Akzeptabilität der Übersetzung gemäß den Vorgaben in Frage gestellt wird (Lommel 2018: 123).

2.3. Textsorte der Rezepte

„Kochrezepte gibt es, seit Menschen kochen." (Donalies 2012: 25). Sie sind im Wesentlichen Gebrauchstexte und werden zur Beschreibung der Zubereitung von Gerichten und ihren Komponenten verwendet, sowohl für das eigene Gedächtnis als auch zur Erklärung für andere, oder zur schriftlichen Dokumentation einer bestimmten Küchentradition (Wolańska-Köller 2010: 1). Die ersten schriftlich überlieferten deutschen Rezepte gehen auf das späte Mittelalter zurück, wobei 1542 das erste deutschsprachige Kochbuch als Übersetzung aus dem Lateinischen erschien (Cölfen 2007: 87; Donalies 2012: 25). Kochrezepte können durch ihre starke Normierung, ihre vergleichsweise niedrige Komplexität und ihre hohe Bekanntheit als prototypische Textsorte betrachtet werden (vgl. Cölfen 2007: 85). Es handelt sich dabei um instruktive Texte mit deskriptiver Themenentfaltung und dominierender appellativer Funktion (Cölfen 2007: 85f). Das klassische Rezept besteht aus einer Zutatenliste und einem Anleitungstext zur Zubereitung. Heutzutage ist das klassische Rezept in den erhältlichen Kochbüchern beziehungsweise Rezeptsammlungen Teil eines großen Ganzen und wird etwa durch Bilder, narrative Elemente, Hintergrundinformationen wie zum Beispiel Nährwertangaben und Tipps aller Art ergänzt (Cölfen 2007: 86). Trotz verständlicher Gestaltung der Anleitungen scheitern viele Personen, die ein Rezept nachkochen wollen, oft an fehlendem Wissen, Erfahrung und Kenntnissen (Cölfen 2007: 86). Um solchen Problemen entgegenzuwirken und gewisse Kochkenntnisse verständlicher zu vermitteln, kann zusätzliches Bild- beziehungsweise Videomaterial weiterhelfen (vgl. Cölfen 2007). Durch die Entwicklung von sozialen Medien existieren in diesem Kontext auf Plattformen wie YouTube, TikTok oder Instagram Rezeptvideos jeglicher Art. Auf diesen veröffentlichen sowohl professionelle Köch*innen beziehungsweise Bäcker*innen als auch sich darauf spezialisierende Youtuber*innen beziehungsweise Content Creator*innen Videos unterschiedlicher Längen – von wenigen Sekunden bis hin zu einer Stunde oder mehr – und Schwierigkeitsgrade. Durch visuelles Vorzeigen der Zubereitungsschritte kann auch Lai*innen das Kochen beziehungsweise Backen auf einfache Weise näher gebracht werden (vgl. Cölfen 2007).

Rezepte sind auch in sprachlicher Hinsicht durchaus besonders. So werden im Zubereitungsteil unterschiedliche Verbformen verwendet (vgl. Donalies 2012). In diesem

Zusammenhang unterscheidet Donalies (2012) zwischen anweisenden und beschreibenden Kochrezepten. Anweisende Kochrezepte zeichnen sich durch ihren auffordernden Ton aus. Sie fordern Adressat*innen direkt zu einer Handlung auf (Donalies 2012: 26). Dabei werden Verbformen wie der adhortative Konjunktiv (*Man nehme…*), Man-Konstruktionen mit Modalverben wie *sollen* oder *müssen*, Imperativ, Sie-Distanzform und die zweite Person Singular Indikativ Aktiv verwendet (Donalies 2012: 26f). Beschreibende Kochrezepte hingegen zeichnen sich durch einen moderaten Ton aus, der die Adressat*innen „an der Kochhandlung teilhaben lässt" (Donalies 2012: 28). Diese umfassen Verbformen wie Infinitiv, Vorgangspassiv, dritte Person Singular oder Plural mit einem gegenständlichen Subjekt (*Dazu kommt etwas Salz.*), Man-Konstruktionen ohne Modalverb, erste Person Singular oder Plural Indikativ Aktiv (Donalies 2012: 28f). Trotz des prototypischen Charakters der Textsorte Rezept besteht eine große Vielfalt an sprachlichen Ausdrucksformen, wobei sich heutzutage der beschreibende Infinitiv als gängigste Verbform etabliert hat (Donalies 2012: 30). Koch- beziehungsweise Backrezepte zeichnen sich darüber hinaus sprachlich durch ihre kulinarische Fachsprache aus. Die kulinarische Terminologie setzt sich aus Begriffen zu beim Kochen vorkommenden Handlungen, Produkten, Zutaten oder Bestandteilen eines Menüs zusammen (Cölfen 2007: 89). Die deutsche Terminologie dieses Bereichs beinhaltet darüber hinaus zahlreiche Entlehnungen aus anderen Sprachen, vor allem aus dem Französischen, Italienischen und Englischen (vgl. Cölfen 2007; Zenjob 2022). In diesem Zusammenhang listet Tabelle 1 häufig vorkommende Fachbegriffe im Bereich der Kulinarik, hauptsächlich Handlungstermini, auf, die ebenso Beispiele für Entlehnungen aus den drei genannten Sprachen beinhaltet.

Tabelle 1: Fachbegriffe aus dem Bereich der Kulinarik (vgl. Cölfen 2007; Zenjob 2022)

Ablöschen	Coulis	Frittieren	Passieren / passer
Abschäumen	Croutons	Glasieren / Glacieren	Pochieren
Abschrecken	Dämpfen	Gratinieren / Überbacken	Reduzieren
Al dente	Deglacieren	Grillen	Sautieren / Sauté
Anschwitzen	Dünsten	Jus	Schmoren
Backen	Einkochen	Klären	Sieden
Binden	Entfetten	Legieren	Tournieren / tourner
Blanchieren	Entree	Marinieren	Überbacken
Blindbacken	Flambieren	Nappieren / napper	Well done
Braten	Fond	Panieren / Panade	

Als dritter Aspekt im Kontext der Sprache von Koch- beziehungsweise Backrezepten ist die Sprachvarietät zu nennen, die vor allem als Identitätsmerkmal eine wesentliche Rolle spielt. Die Bedeutung der Sprachvarietät wird etwa durch den Vertrag zum Beitritt Österreichs in die Europäische Union verdeutlicht (vgl. EU 1995). Dieser enthält nämlich mit dem Protokoll Nr. 10 eine Akte, in der sich die Republik Österreich von der EU zusichern ließ, dass spezifische österreichische Ausdrücke der deutschen Sprache, die in einer angehängten Liste angeführt wurden, mit den entsprechenden in Deutschland verwendeten Ausdrücken gleichgestellt werden (EU 1995). In der deutschen Version neuer Rechtsakte der EU sollten also die im Anhang des Protokolls aufgelisteten Ausdrücke den deutschen Entsprechungen „in geeigneter Form hinzugefügt“ werden (EU 1995). Die 23 ausgewählten Begriffe stammen dabei ausschließlich aus dem Bereich der Kulinarik und werden in Abb. 1 mit den jeweiligen in Deutschland üblichen Begriffen aufgelistet (vgl. EU 1995). Die Liste wurde vom Gesundheits- und Landwirtschaftsministerium ausgearbeitet und sollte mit Begriffen aus dem Bereich der Kulinarik jene Ausdrücke enthalten, die den Menschen besonders naheliegen und Teil der nationalen Identität sind (vgl. Der Standard 2005; Langeder o. J.). Österreichische Ausdrücke aus anderen relevanten Bereichen wurden nicht in das Protokoll aufgenommen. Kritiker*innen sahen darin dementsprechend keine sprachpolitische Maßnahme, vielmehr wollte man damit der Bevölkerung Zuversicht vermitteln und der „Angst um den Verlust der eigenen Sprache“ entgegenwirken (Langeder o. J.).

ANHANG

Österreich	*Amtsblatt der Europäischen Gemeinschaften*
Beiried	Roastbeef
Eierschwammerl	Pfifferlinge
Erdäpfel	Kartoffeln
Faschiertes	Hackfleisch
Fisolen	Grüne Bohnen
Grammeln	Grieben
Hüferl	Hüfte
Karfiol	Blumenkohl
Kohlsprossen	Rosenkohl
Kren	Meerrettich
Lungenbraten	Filet
Marillen	Aprikosen
Melanzani	Aubergine
Nuß	Kugel
Obers	Sahne
Paradeiser	Tomaten
Powidl	Pflaumenmus
Ribisel	Johannisbeeren
Rostbraten	Hochrippe
Schlögel	Keule
Topfen	Quark
Vogerlsalat	Feldsalat
Weichseln	Sauerkirschen

Abbildung 1: Anhang von Protokoll Nr. 10 (EU 1995)

3. Aktueller Stand der Forschung

Wie in der Einleitung beschrieben, wurden bislang zahlreiche Studien zur Qualitätsbewertung maschinell übersetzter Texte durchgeführt. Im folgenden Kapitel sollen daher mehrere Ansätze beziehungsweise wissenschaftliche Publikationen zu dieser Thematik vorgestellt und deren Thema, Methode beziehungsweise Forschungsdesign näher beschrieben werden.

Wiesmann (2019) beschäftigte sich in ihrer Studie mit der Übersetzung von Rechtstexten aus dem Italienischen ins Deutsche mithilfe neuronaler maschineller Übersetzung. Ziel des Versuchs war es herauszufinden, inwieweit NMÜ in der Lage ist, Rechtstexte oder zumindest bestimmte Textsorten aus diesem Bereich in eine andere juristische Sprache zu übersetzen, so dass der Post-Editing-Aufwand begrenzt ist. In diesem Kontext wurden häufig übersetzte und repräsentative Textsorten aus den drei Hauptbereichen juristischer Tätigkeit, welche sich auf Gesetz, Rechtspraxis und Rechtstheorie bezogen, ausgewählt. Dafür wurden sechs Texte unterschiedlicher Länge mit insgesamt 9.918 Wörtern und 261 Sätzen herangezogen. Aufgrund der raschen Entwicklung der MÜ-Industrie wurden nicht nur die Ergebnisse verschiedener Systeme verglichen, sondern der Versuch wurde auch nach vier Monaten wiederholt, um vermeintliche Verbesserungen in der Qualität festzustellen. Für den Vergleich der Systeme wurden DeepL und MateCat, ein CAT-Tool, das MÜ integriert und die Verwendung einer Kombination aus verschiedenen MÜ-Systemen ermöglicht, ausgewählt. Alle Übersetzungen wurden von der Autorin manuell analysiert und im Anschluss bewertet. Für die Bewertung wurden dabei die Kategorien Verständlichkeit/Sinnhaftigkeit (engl. comprehensibility/meaningfulness) und Übereinstimmung zwischen Ausgangs- und Zieltext (engl. correspondence between source and target text) herangezogen, wobei letztere die spezifische Übersetzungssituation berücksichtigt. Diese entsprechen den Kategorien Sprachfluss und Genauigkeit der MQM-Fehlertypologie (vgl. Lommel et al. 2014). Jeder Satz wurde zur Bestimmung der Übersetzungsqualität in beiden Kategorien mit einer Note von 0 bis 3 beurteilt. Die Ergebnisse der Analyse zeigten, dass DeepL im Allgemeinen besser abschnitt als MateCat. Nach Wiederholung des Versuches konnten bei beiden Systemen Verbesserungen und Verschlechterungen festgestellt werden, sodass kein eindeutiger Entwicklungstrend zu erkennen war. Die Bewertung ergab, dass die maschinellen Übersetzungen an sich verständlich, aber nicht immer korrekt waren. In diesem Zusammenhang kam es bei der Verwendung von juristischen Fachtermini und -begriffen häufig zu Inkonsistenzen. Darüber hinaus wurde festgestellt, dass die Qualität der MÜ von verschiedenen Faktoren, wie etwa der Textlänge und dem Schwierigkeitsgrad, abhängt. Insgesamt wurden die Resultate von MÜ als unzureichend

angesehen, um der Nachbearbeitung von maschinell übersetzten Rechtstexten einen größeren Platz in der Übersetzungspädagogik einzuräumen. Da die Bewertung der Übereinstimmung zwischen Ausgangs- und Zieltext grundsätzlich schlechter ausfiel als die Bewertung der Verständlichkeit beziehungsweise Sinnhaftigkeit des Zieltextes, kam die Autorin zu dem Schluss, dass die Übersetzungspädagogik Bewusstsein für die Unterschiede zwischen maschineller und menschlicher Übersetzung in diesem Bereich schaffen sowie den Übersetzungsansatz verbessern und die juristische Expertise stärken sollte.

Heinisch und Lušicky (2019) befassten sich in ihrer Arbeit mit den Erwartungen von Nutzer*innen an die MÜ. Zwar handelt es sich dabei um keinen Vergleich mehrerer maschineller Übersetzungssysteme, sie enthält allerdings eine Bewertung der MÜ durch Studierende des Masterstudiums Translation. Ihre Fallstudie verfolgte mehrere Ziele. Zuerst wurden die Erwartungen von 47 Studierenden des Masterstudiums Translation an die MÜ sowie die Frage, inwiefern Vorerfahrungen mit MÜ ihre Erwartungen an die Gesamtqualität und Fehlertypen im MÜ-Output beeinflussen, untersucht. Darüber hinaus sollte die Übereinstimmung beziehungsweise Nicht-Übereinstimmung dieser Erwartungen mit der Bewertung von zwei Outputs maschineller Übersetzung untersucht werden. Zu diesem Zweck wandten die Autorinnen einen Mixed-Method-Ansatz an, bei dem ein Fragebogen sowie zwei sich wiederholende Bewertungen von rohen MÜ kombiniert wurden. Der Fragebogen bestand aus drei Teilen. Beim ersten Teil ging es um die Übersetzungserfahrung, die Arbeitssprachen und die Berufserfahrung der Befragten. Beim zweiten Teil ging es um die Vorerfahrungen der Befragten in der Nutzung von MÜ, einschließlich der Häufigkeit und der Gründe für die Nutzung. Im dritten Teil sollten die Befragten den englischen Ausgangstext lesen und ihre Erwartungen an die Qualität des entsprechenden MÜ-Outputs anhand einer fünfstufigen Notenskala angeben: ausgezeichnet, gut, befriedigend, genügend, unbrauchbar. Die Bewertungen beinhalteten die Identifizierung und Klassifizierung von Fehlern im NMÜ-Output gemäß dem MQM-Framework nach Lommel et al. (2014). Den Studierenden wurden dafür zwei englische Ausgangstexte, und zwar Auszüge aus britischen Zeitungsartikeln zu einem in Österreich relevanten Thema, und ihre deutschen maschinellen Übersetzungen des EU-Ratspräsidentschaftsübersetzers (EU Council Presidency Translator) vorgelegt. Die deutschen Sätze wurden auf Segmentebene nach dem MQM-Framework bewertet. Alles in allem erwarteten die Befragten eine eher geringere Qualität der MÜ-Ergebnisse, jedoch lag die Qualität des NMÜ-Outputs unter den Erwartungen. Außerdem wurde eine höhere Häufigkeit gewisser Fehlertypen im Vergleich zur tatsächlich gemeldeten Häufigkeit nach der Bewertung

festgestellt. Die Fallstudie unterstrich schließlich die Bedeutung der Berücksichtigung der Erwartungen der Nutzer*innen von MÜ und nicht nur der MÜ-Evaluierung. Weiters argumentieren Heinisch und Lušicky (2019), dass die Erfahrungen und Erwartungen der Nutzer*innen die Nutzung und Bewertung der MÜ beeinflussen.

Larassati et al. (2019) untersuchten und verglichen in ihrer Forschungsarbeit die Ergebnisse maschineller Übersetzung von prozeduralen Texten aus dem Englischen ins Indonesische von den MÜ-Systemen Google Translate und Instagram. Der Fokus lag dabei auf den Fehlern, die in den Übersetzungen auftraten. Dafür wurden drei prozedurale Texte, genauer gesagt Kochrezepte vom Instagram-Account @howtofoodprep aus Gründen der umfassenden Verfügbarkeit auf Instagram herangezogen und gezielt im Hinblick auf Struktur und sprachliche Merkmale ausgewählt. Die Verwendung prozeduraler Texte wurde damit begründet, dass die Leser*innen im Unterschied zu anderen Textsorten zur Erreichung eines bestimmten Ziels gewisse Anweisungen befolgen müssen. Jeder Text wurde kopiert, in Google Translate eingefügt und anschließend ins Indonesische übersetzt. Auf Instagram hingegen wurden die Übersetzungen durch Betätigung der Schaltfläche *Übersetzung anzeigen* generiert. Die Übersetzungen wurden von den Autor*innen manuell auf Fehler untersucht. Dabei wurde die Fehlertypologie der American Translators Association (ATA), einem Berufsverband von Übersetzer*innen und Dolmetscher*innen in den Vereinigten Staaten, herangezogen und Fehler in den Zieltexten auf Segmentebene identifiziert und klassifiziert. Die Fehleranalyse ergab insgesamt 95 Fehler beim MÜ-System von Instagram und 37 Fehler bei Google Translate. Bei beiden Systemen kamen am häufigsten die Fehlerkategorien Terminologie (engl. terminology), Syntax (engl. syntax) und wörtliche Übersetzung (engl. literalness) vor, die im engen Zusammenhang zueinander stehen. Demnach wurden etwa Terminologiefehler durch die wörtliche Übersetzung von Phrasen beziehungsweise Sätzen verursacht, wodurch in den Übersetzungen Begriffe verwendet wurden, die nicht dem Kontext entsprachen. Andere Fehlerkategorien kamen im Vergleich dazu nur in sehr geringem Maße vor. Diese umfassten Verwendung (engl. usage), Wortformen/Wortarten (engl. word forms/parts of speech), Ergänzung (engl. addition), Textsorte (engl. text type), Auslassung (engl. omission) und sonstige Fehler – Bedeutungstransfer (engl. other errors – meaning transfer).

Trotz der begrenzten Anzahl an Sprachpaaren, MÜ-Systemen, Datenmengen und Textsorten gelangten Larassati et al. (2019) zu dem Schluss, dass sich insbesondere die Übersetzungsqualität von Google Translate erheblich erhöht hat. Dies ließ sich vor allem daraus erkennen, dass dieses MÜ-System Fachbegriffe und Abkürzungen identifiziert und auch

Grammatik, Wortstellung und Wortwahl bei den Übersetzungen berücksichtigt. Darüber hinaus traten keine textsortenbezogenen Fehler bei der MÜ von Google Translate auf, woraus zu schließen ist, dass dieses MÜ-System textsortenspezifische Konventionen befolgt. Google Translate wurden nach dieser Studie vielversprechende Prognosen für die Zukunft gestellt, obwohl es an das Qualitätslevel einer professionellen Übersetzung nicht herankam. Beim MÜ-System von Instagram wurde hingegen Einiges an Verbesserungsbedarf festgestellt. Die erhöhte Fehlerhäufigkeit war vor allem auf wörtliche Übersetzung der Ausgangstexte zurückzuführen. Dadurch traten vermehrt Wortwahl- und Grammatikfehler in der Zielsprache auf.

4. Methodik

Zur Untersuchung der Forschungsfrage in Bezug auf die Qualitätsbewertung der maschinell erstellten Rezeptübersetzungen durch Kochbegeisterte soll eine Online-Umfrage erstellt werden, da es sich beim Forschungsgegenstand der Masterarbeit um die Qualität maschineller Übersetzung handelt und die tatsächliche Zielgruppe der gewählten Textsorte direkt befragt werden soll. Diese Methode ist im Gegensatz zu anderen, wie etwa Interviews, sinnvoller, weil es sich hierbei um eine Textbewertung ohne Beobachtung mit ausreichend zur Verfügung gestellter Zeit handelt. Weiters spricht für die gewählte Methode, dass damit eine größere Anzahl an Personen erreicht werden kann und somit repräsentativere Daten ausgewertet werden können. Im Rahmen der Masterarbeit sollen Kochbegeisterte zur Übersetzungsqualität von Koch- beziehungsweise Backrezepten durch zwei Systeme neuronaler maschineller Übersetzung befragt werden. In den folgenden Unterkapiteln soll daher die Methodik durch Beschreibung des Zielgruppenprofils, der System- und Textauswahl, der Fragebogenerstellung einschließlich des genaueren Aufbaus der Umfrage und der Auswertungsmethode näher erläutert werden.

4.1. Profil der Zielgruppe

Die Zielgruppe der Befragung waren Kochbegeisterte, also Personen, die zu einem ausreichenden Ausmaß mit den Prozessen im Kochen und Backen vertraut sind. Damit sollte das tatsächliche Zielpublikum der Texte erreicht werden, um die Qualität der maschinellen Übersetzungen im Hinblick auf Richtigkeit und Angemessenheit der Inhalte entsprechend bewerten zu können. Das sachgebietsspezifische Publikum konnte dabei sowohl aus Translator*innen als auch aus Fachexpert*innen beziehungsweise Personen mit privatem Interesse am Bereich der Kulinarik bestehen. Die Umfrageteilnehmer*innen hatten im allgemeinen Teil der Umfrage (vgl. Kapitel 4.4.) die Möglichkeit zu spezifizieren, ob sie professionelle Translator*innen beziehungsweise Studierende im Bereich der transkulturellen Kommunikation beziehungsweise Translation oder generell mit den Prozessen des Kochens beziehungsweise Backens vertraut sind. Damit sollte eine differenzierte Betrachtung der Ergebnisse ermöglicht werden. Eine weitere Voraussetzung für die Teilnahme an der Umfrage und in weiterer Folge für die Beurteilung der Richtigkeit der Übersetzungen waren ausreichende Deutschkenntnisse. Die Zielgruppe der Umfrage wurde über verschiedene Kanäle kontaktiert und mit einem Link zur Umfrageteilnahme eingeladen. Die Teilnehmer*innen setzten sich aus Personen zusammen, die der Einladung in unterschiedlichen Gruppen auf WhatsApp, sowohl in der Gruppe von Studierenden des Masters Translation als auch in privaten Gruppen, sowie in den Facebook-

Gruppen Transkulturelle Kommunikation und Master Dolmetschen und Übersetzen nachgekommen sind. Darüber hinaus wurden Mitarbeiter*innen aus einem Übersetzungsbüro und Personen im Bekanntenkreis um Teilnahme an der Umfrage und Weiterleitung der Einladung gebeten.

4.2. Systemauswahl

Für die Übersetzung der Rezepte wurden die frei verfügbaren maschinellen Übersetzungssysteme DeepL und Google Translate (vgl. Kapitel 2.2.3.) ausgewählt. Wie in diesem Kontext bereits in Kapitel 1.2. beschrieben, bedarf es zunächst einer Klärung, ob der Einsatz von MÜ für den Bereich der Kulinarik beziehungsweise für die Textsorte der Koch- und Backrezepte sinnvoll ist, und in einem nächsten Schritt unter Berücksichtigung der eigenen Bedürfnisse eines Ergebnisvergleichs mehrerer MÜ-Tools, um das beste Tool zu finden. Demzufolge war das Ziel der Umfrage einerseits herauszufinden, ob die Nutzung von MÜ-Systemen in diesem Bereich sinnvoll ist, und andererseits, welches der beiden Systeme bessere Ergebnisse erzielt. So wurden für den Bewertungsteil insgesamt acht Rezepte ausgesucht, aus denen vier Sets erstellt wurden. Pro Set wurde jeweils ein Rezept mit DeepL und eines mit Google Translate aus dem Englischen ins Deutsche übersetzt. Die MÜ-Systeme DeepL und Google Translate wurden in diesem Zusammenhang wegen ihres hohen Bekanntheitsgrads unter Translator*innen und Lai*innen ausgewählt. Darüber hinaus erzielen die beiden Systeme im Vergleich zu anderen bessere Ergebnisse, wie etwa auch in diversen Studien (z. B. Larassati et al. 2019; Wiesmann 2019) festgestellt werden konnte.

4.3. Textauswahl

Für den Bewertungsteil wurden insgesamt acht Rezepte von den berühmten britischen Köchen Gordon Ramsay und Jamie Oliver, die in Tabelle 2 als Text 1 bis Text 8 aufgelistet sind, herangezogen. Aus diesen acht Rezepten wurden vier Sets mit je zwei in Bezug auf Schwierigkeit, Zubereitung und Gericht vergleichbaren Rezepten erstellt. Hierbei wurde darauf geachtet, dass Rezepte eines Sets ungefähr die gleiche Länge aufweisen. Um die Bewertung so repräsentativ wie möglich zu gestalten und nicht zu beeinflussen, wurde sichergestellt, dass gewisse Sets Rezepte von beiden Köchen und andere jeweils nur von einem enthalten (siehe Tabelle 2). Im Folgenden werden die einzelnen Sets genauer vorgestellt. Dabei wird auf die individuellen Rezepte und deren mögliche Herausforderungen näher eingegangen.

Tabelle 2: Übersicht der ausgewählten Texte unter Angabe des jeweiligen Kochs und des verwendeten MÜ-Systems

Text	Koch	MÜ-System
Text 1: Sweet potato, coconut & cardamom soup	Jamie Oliver	DeepL
Text 2: Cauliflower soup and ham cheese toastie	Gordon Ramsay	Google Translate
Text 3: Harissa tomato jam toast with poached egg, feta, preserved lemon	Gordon Ramsay	Google Translate
Text 4: Spanish-style avocado toast with chorizo and tomatoes	Gordon Ramsay	DeepL
Text 5: Beef Wellington	Jamie Oliver	Google Translate
Text 6: Beef Wellington	Gordon Ramsay	DeepL
Text 7: Pineapple upside-down cake	Jamie Oliver	Google Translate
Text 8: Vegan toffee apple upside-down cake	Jamie Oliver	DeepL

4.3.1. Set 1

Das erste Set besteht aus den in Tabelle 2 aufgelisteten Texten 1 und 2. Dabei handelt es sich um Suppenrezepte, und zwar eine Suppe aus Süßkartoffeln, Kokosnuss und Kardamom von Jamie Oliver (siehe Anhang A, Text 1) und eine Karfiolsuppe von Gordon Ramsay (siehe Anhang A, Text 2). Beide Rezepte sind von mittlerer Schwierigkeit.

In Text 1, der mit DeepL übersetzt wird (siehe Tabelle 2), sollten die Zutaten im Hinblick auf eine MÜ und die anschließende Bewertung keine großen Herausforderungen darstellen. Die einzige Zutat, die den Umfrageteilnehmer*innen unter Umständen unbekannt sein könnte, ist *poppadoms*, auf Deutsch *Papadam*, ein aus Indien stammender dünner frittierter Teigfladen aus Linsenmehl (vgl. PONS 2025). Die Anweisungen sind ziemlich klar, erfordern aber eine präzise Übersetzung. Die Übersetzung des mehrdeutigen Verbs *cook* im Sinne von *braten* könnte bei der MÜ ebenfalls zu Missverständnissen führen. Auch in Text 2, der mit Google Translate übersetzt wird (siehe Tabelle 2), sollten die Zutaten und Anweisungen keine größeren Übersetzungsschwierigkeiten und damit verbundene Verständnisprobleme in der

Zielsprache verursachen. Lediglich die Zutat *Montgomery Cheddar* könnte einem deutschsprachigen Publikum nicht geläufig sein.

4.3.2. Set 2

Set 2 enthält mit den Texten 3 und 4 zwei Toast-Rezepte von Gordon Ramsay (siehe Tabelle 2). Bei Text 3 (siehe Anhang A, Text 3) handelt es sich um ein Rezept für einen Toast mit pikanter Tomatenmarmelade, pochiertem Ei, Feta und eingelegter Zitrone. Bei Text 4 (siehe Anhang A, Text 4) geht es um die Zubereitung eines Avocado-Toasts mit Chorizo und Tomaten. Die Rezepte sind relativ unkompliziert, auch wenn sie gewisse anspruchsvolle Kochtechniken, wie zum Beispiel *Eier pochieren*, als dem Zielpublikum bekannt voraussetzen und nicht in eigenen Schritten beschreiben.

Die Mengenangaben in der Zutatenliste stellen sowohl in Text 3 als auch in Text 4 eine potentielle Herausforderung für die MÜ dar. Sie beinhalten nämlich nichtmetrische Mengenangaben in den Einheiten *cups*, *pounds* und *ounces*, die einem deutschsprachigen Publikum eher weniger geläufig sind. Darüber hinaus enthält Text 3, der mit Google Translate übersetzt wird (siehe Tabelle 2), mit der *Harissa-Paste* eine Zutat, die nicht unbedingt alle Umfrageteilnehmer*innen kennen könnten. Auch der Begriff *Artisan-Brot* als Benennung für handwerkliches Brot könnte zu Missverständnissen führen. In Text 4, der mit DeepL übersetzt wird (siehe Tabelle 2), kommt der spanische *Manchego-Käse* vor, der im deutschsprachigen Raum nicht überall erhältlich ist. In diesem Fall wäre bei einer Humanübersetzung eine Erklärung notwendig, die Alternativen aufzeigt.

4.3.3. Set 3

Im dritten Set finden sich mit den Texten 5 und 6 (siehe Tabelle 2) Rezepte zu einem eher festlichen und durchaus komplexeren Gericht, nämlich Beef Wellington. Bei Text 5 handelt es sich um ein Rezept zu Beef Wellington mit Madeira-Sauce von Jamie Oliver (siehe Anhang A, Text 5). Text 6 ist ein Beef-Wellington-Rezept von Gordon Ramsay (siehe Anhang A, Text 6). Das Gericht setzt gewisse anspruchsvolle Prozesse voraus und besteht aus mehreren Komponenten, die schließlich zu einem Ganzen vereint werden. Dementsprechend haben beide Rezepte dieses Sets einen hohen Schwierigkeitsgrad.

Aufgrund der Komplexität der Rezepte und der darin enthaltenen Anweisungen ist davon auszugehen, dass die MÜ-Resultate in diesem Set gewisse Schwierigkeiten aufweisen könnten. Die Zutatenliste in Text 6 ist sehr allgemein gehalten, wodurch hinsichtlich der MÜ durch DeepL (siehe Tabelle 2) keine gravierenden Fehler zu erwarten sind. Auch in Text 5, der

mit Google Translate übersetzt wird (siehe Tabelle 2), sollte die Übersetzung der Zutaten problemlos ablaufen. Die einzige enthaltene Zutat, die einem deutschsprachigen Publikum unter Umständen nicht bekannt sein könnte, ist die Rotweinsorte Madeirawein.

4.3.4. Set 4

Das letzte Set enthält mit den in Tabelle 2 angeführten Texten 7 und 8 zwei Dessertrezepte, und zwar zu sogenannten Upside-Down-Kuchen beziehungsweise gestürzten Kuchen von Jamie Oliver. Während Text 7 (siehe Anhang A, Text 7) die Zubereitung einer klassischen Variante des in Amerika und dem Vereinigten Königreich beliebten Kuchens mit Ananas und kandierten Kirschen beschreibt, handelt es sich bei Text 8 (siehe Anhang A, Text 8) um eine vegane Variante mit Äpfeln und Walnüssen. Sowohl Text 7 als auch Text 8 wurden von Jamie Oliver mit dem mittleren Schwierigkeitsgrad gekennzeichnet.

Die Zutatenliste von Text 7 ist allgemein verständlich gehalten. Die Zubereitungsanweisungen hingegen stellen für die MÜ von Text 7, die mit Google Translate erfolgt (siehe Tabelle 2), eine Herausforderung dar. So können die Syntax, welche zum Teil lange Sätze beinhaltet, sowie Benennungen wie *roasting tray* oder *pan juices* für ein Backblech beziehungsweise eine Kuchenform und den in der Pfanne zurückgebliebenen Saft eine Fehlerquelle darstellen. In Text 8, der mit DeepL übersetzt wird (siehe Tabelle 2), hingegen sind die Anweisungen ziemlich klar und sollten keine gravierenden Probleme verursachen. Die Zutatenliste enthält jedoch *muscovado sugar* und *mixed spice*, die in dieser Form einem deutschsprachigen Zielpublikum vermutlich weniger geläufig sind und eine nähere Erklärung beziehungsweise Anpassung erfordern.

4.4. Erstellung des Fragebogens

Nach Auswahl der Rezepte für den Bewertungsteil und ihrer Übersetzung durch Google Translate und DeepL wurde der Fragebogen erstellt. Dafür kamen zunächst verschiedene Online-Umfragetools, wie etwa LimeSurvey (LimeSurvey 2025), Survio (Survio 2025), UmfrageOnline (UmfrageOnline 2025) oder LamaPoll (LamaPoll 2025a), infrage. Die Auswahl fiel schließlich nicht zuletzt wegen ihrer Benutzerfreundlichkeit auf die Plattform LamaPoll. Auch das Angebot einer dreimonatigen kostenlosen Studierendenversion mit Freischaltung der für die zu erstellende Umfrage nützlichsten Funktionen, wie zum Beispiel jene der bedingten Fragen, war für die Auswahlentscheidung maßgeblich (LamaPoll 2025b).

In einem nächsten Schritt wurde ein erster Entwurf des Fragebogens erstellt und mehrmals auf Einheitlichkeit, Design, Rechtschreibung und Fragenfunktionen geprüft. Dieser wurde

in einer ersten Pilotstudie durch die Betreuer*in dieser Masterarbeit absolviert und die entsprechenden Änderungsvorschläge wurden eingearbeitet. Die vorläufig finale Version wurde schließlich einer dritten Person, die die Umfrage noch nicht kannte, für einen weiteren Testdurchlauf übersandt. Die Testperson füllte die Umfrage aus und meldete in Bezug auf die Fragestellungen keine Verständnisprobleme. Sie merkte allerdings an, für die Umfrage insgesamt 50 Minuten benötigt zu haben, was potenziellen Umfrageteilnehmer*innen als zu hoher Zeitaufwand erscheinen könnte. Um dieser Problematik entgegenzuwirken, wurde in Absprache entschieden, eines der ursprünglich fünf Sets aus dem Bewertungsteil herauszunehmen. Somit wurde der Bewertungsteil auf vier Sets mit insgesamt acht Rezepten reduziert. Nach kleineren gestaltungstechnischen Korrekturen wurde die finale Version der Umfrage schließlich auf verschiedenen Kanälen (vgl. Kapitel 4.1.) verbreitet, um einen möglichst großen Personenkreis zu erreichen.

Die Online-Umfrage (siehe Anhang B) besteht aus insgesamt fünf Teilen. Im ersten Teil wurden die Umfrageteilnehmer*innen allgemein zu ihrer Person befragt. Dabei wurden zunächst Fragen zu ihrem Alter, Geschlecht und höchsten Bildungsabschluss gestellt. Im Hinblick auf den Bildungsabschluss standen folgende Antwortmöglichkeiten zur Auswahl: kein Schulabschluss, Pflichtschule, Lehrabschluss, Matura, Universität bzw. Hochschule (Bachelor), Universität bzw. Hochschule (Master), Universität bzw. Hochschule (Magister), Universität bzw. Hochschule (Doktorat), anderer Abschluss nach Matura. Bei der Auswahl *Anderer Abschluss nach Matura* wurde zudem ein Textfeld für eine genauere Angabe des Bildungsabschlusses eingefügt. Um eine gewisse Differenzierung der Zielgruppe zu erzielen, wurden die Umfrageteilnehmer*innen gefragt, ob sie professionelle Translator*innen sind oder im Bereich der transkulturellen Kommunikation beziehungsweise Translation studieren. Abhängig von der Antwort wurden die Fragen im weiteren Verlauf des ersten Umfrageteils adaptiert. Professionelle Translator*innen wurden dementsprechend nach ihren Arbeitssprachen und weiteren Sprachkenntnissen befragt. Personen, die nicht im Bereich der Translation tätig sind, sollten ihre Branche angeben. Zur Auswahl standen dabei folgende Optionen:

- Tourismus, Gastgewerbe
- Bildung
- Forschung
- Umwelt
- Soziales, Gesundheit

- Medien, Grafik, Design, Kunst
- IT
- Handel, Logistik
- Marketing
- Psychologie
- Finanz, Wirtschaft
- Recht
- Lebensmittel
- Baugewerbe
- Sprachen, Kultur
- Sonstiges.

Da ausreichend gute Deutschkenntnisse Voraussetzung für die Teilnahme an der Umfrage waren, wurden die Umfrageteilnehmer*innen im ersten Teil schließlich noch gefragt, ob Deutsch ihre Erstsprache ist. Im Fall, dass Deutsch nicht die Erstsprache ist, wurde noch um Angabe der tatsächlichen Erstsprache und Einschätzung des Sprachniveaus auf Deutsch in A1, A2, B1, B2, C1 oder C2 gemäß dem Gemeinsamen Europäischen Referenzrahmen für Sprachen (GER) (vgl. GER 2025) gebeten.

Der zweite Teil der Umfrage diente der Einschätzung der Koch- und Backerfahrung der Teilnehmenden. Hierbei wurde zunächst gefragt, wie oft sie nach Rezept kochen beziehungsweise backen. Dabei standen die Häufigkeitsangaben *täglich*, *mehrmals in der Woche*, *mehrmals im Monat* und *mehrmals im Jahr* zur Auswahl. Im weiteren Verlauf wurden die Umfrageteilnehmer*innen zur Einschätzung ihrer Expertise auf diesem Gebiet nach ihren Kenntnissen zu Koch- und Backtechniken befragt und konnten diese auf einer Skala von 1 (= nicht vertraut) bis 5 (= sehr vertraut) spezifizieren. Auf Anraten der Betreuerin der vorliegenden Masterarbeit wurde noch die Frage *Aus welcher Küche bereiten Sie besonders gerne Gerichte zu?* hinzugefügt, da dies in Bezug auf die Rezepte im Bewertungsteil und den darin vorkommenden Zutaten durchaus relevant sein könnte. Dabei hatten die Teilnehmenden die Möglichkeit, die jeweils zutreffenden Optionen aus den folgenden Antwortmöglichkeiten auszuwählen: Österreichisch, Italienisch, Asiatisch, Amerikanisch, Balkan, Orientalisch, Sonstiges. Bei Auswahl der Option *Sonstiges* wurde um genauere Angaben gebeten. Professionelle Translator*innen wurden darüber hinaus gefragt, ob sie Erfahrungen mit der Übersetzung von Koch- und Backrezepten aufweisen.

Im dritten Teil ging es um die Bewertung der MÜ-Erfahrung der Teilnehmer*innen. In diesem Zusammenhang wurden sie gefragt, ob sie generell MÜ-Systeme nutzen. Bei Beantwortung dieser Frage mit Ja folgten weitere Fragen zum Nutzungsverhalten. Diese umfassten die Angabe der verwendeten MÜ-Systeme, die Nutzungshäufigkeit von *täglich*, *mehrmals in der Woche*, *mehrmals im Monat*, bis zu *mehrmals im Jahr* und die Nutzungszwecke in Form von *private Zwecke*, *berufliche Zwecke*, *Ausbildung*, *Kommunikation* oder *Sonstiges*. Darüber hinaus sollte die allgemeine Zufriedenheit mit den Ergebnissen von MÜ-Systemen auf einer Skala von 1 (= nicht zufrieden) bis 5 (= sehr zufrieden) bewertet werden. Dieser Teil der Umfrage wurde mit der Frage, ob die Umfrageteilnehmer*innen bereits MÜ-Systeme für Koch- und Backrezepte verwendet haben, und wenn ja, wie zufrieden sie mit der diesbezüglichen MÜ-Qualität auf einer Skala von 1 (= nicht zufrieden) bis 5 (= sehr zufrieden) waren, abgeschlossen.

Der vierte Teil der Umfrage umfasste die Bewertungsaufgabe. Dafür wurden insgesamt acht Rezepte (vgl. Kapitel 4.3. sowie Anhang A) von Gordon Ramsay und Jamie Oliver ausgesucht und in vier Sets geteilt. Pro Set wurde je ein Rezept mit DeepL und eines mit Google Translate aus dem Englischen ins Deutsche übersetzt. Den Umfrageteilnehmer*innen wurden nur die deutschen Versionen, das heißt ohne die Originaltexte, mit dem Hinweis, dass es sich um maschinelle Übersetzungen handelt, vorgelegt. Die Teilnehmer*innen wussten jedoch nicht, dass die Rezepte pro Set mit unterschiedlichen MÜ-Systemen übersetzt wurden. Dadurch sollte die Qualitätsbewertung nicht weiter beeinflusst werden. Die Befragten wurden gebeten, die maschinellen Übersetzungen anhand von allgemeineren Aussagen auf einer Likert-Skala von 1 (= stimme überhaupt nicht zu) bis 5 (= stimme völlig zu) zu bewerten. In diesem Zusammenhang wurden folgende neun Aussagen zur Bewertung formuliert:

- Die Übersetzung ist akzeptabel.
- Mit diesem Rezept kann man dieses Gericht zubereiten.
- Die Kochanleitung ist verständlich und könnte leicht umgesetzt werden. // Die Backanleitung ist verständlich und könnte leicht umgesetzt werden.
- Alle Zutaten sind Ihnen bekannt.
- Die verwendeten Mengenangaben beziehungsweise Maßeinheiten sind verständlich und nachvollziehbar.
- Die richtigen Fachbegriffe werden verwendet.
- Die Übersetzung enthält sprachliche beziehungsweise idiomatische Fehler.
- Die Übersetzung enthält grammatikalische Fehler.

- Bestimmte Begriffe sind für Sie im Deutschen nicht üblich.

Die Aussagen wurden mit dem Ziel formuliert, relevante Aspekte der Textsorte Rezepte (vgl. Kapitel 2.3.), Qualitätsansprüche aus translationswissenschaftlicher Sicht (vgl. etwa allgemeine und fachtextspezifische Kriterien nach Ahrend (2006)) sowie Qualitätskriterien menschlicher Evaluierungsmethoden (vgl. Kapitel 2.2.4.2.) abzudecken. Da es bei der Bewertung um die Frage ging, ob die übersetzten Rezepte aus Sicht der Teilnehmer*innen verständlich und in der Praxis umsetzbar sind, und somit im Allgemeinen die Brauchbarkeit der Übersetzungen von MÜ-Systemen bewertet werden sollte, wurden die Aussagen unter Berücksichtigung der Kriterien zur Bewertung auf Makroebene nach Van Slype (1979) nach Verständlichkeit, Brauchbarkeit und Akzeptabilität sowie der Fehlerkategorien Terminologie, linguistische Konventionen, lokale Konventionen und Angemessenheit für die Zielgruppe des MQM-Frameworks (The MQM Council 2024a) formuliert. Auch die fachtextspezifischen Qualitätskriterien nach Ahrend (2006), und zwar Verwendung der korrekten Fachterminologie, Formalitäten und Verwendung korrekter Maßeinheiten (vgl. Kapitel 2.1.) spielten hierbei eine wichtige Rolle. Um detailliertere Einblicke in die Bewertung zu erhalten, wurden noch zwei Freitextfragen eingefügt, die in Abhängigkeit zur Bewertung der Aussagen *Alle Zutaten sind Ihnen bekannt* und *Bestimmte Begriffe sind für Sie im Deutschen nicht üblich* standen. So wurden die Befragten bei einer Bewertung der ersten Aussage mit einem Wert, der kleiner als 5 ist, gefragt, welche Zutaten sie bisher nicht kannten. Bei Bewertung der zweiten Aussage mit einem Wert, der größer als 1 ist, wurde um Angabe der im Deutschen unüblichen Begriffe gebeten. Schließlich hatten die Umfrageteilnehmer*innen bei jeder Übersetzung die Möglichkeit, weitere Anmerkungen hinzuzufügen, und sollten beantworten, ob der Titel des jeweiligen Rezeptes ansprechend ist und sie ein solches Rezept zum Nachkochen beziehungsweise Nachbacken einladen würde.

Im Schlussteil wurden die Umfrageteilnehmer*innen um Einschätzung gebeten, wie leicht ihnen die Bewertungsaufgabe gefallen ist und ob ihre Koch- beziehungsweise Backkenntnisse hierfür ausreichend waren. Abschließend hatten sie noch die Möglichkeit, ihre Gedanken zur Umfrage und zum Thema im Allgemeinen mitzuteilen.

4.5. Auswertungsmethode

Die Umfrage war drei Wochen für die Teilnahme online, bevor die Ergebnisse ausgewertet werden konnten. Die Auswertung erfolgte daraufhin mithilfe der entsprechenden Funktion im Umfragetool LamaPoll. Hierfür wurden nur vollständig ausgefüllte Umfragen herangezogen.

Die statistischen Daten wurden den Ergebnissen in LamaPoll entnommen. Die berechneten Umfrageergebnisse wurden entsprechend auf ihre Richtigkeit geprüft. Darüber hinaus wurden die Ergebnisse bestimmter Fragen zur besseren Veranschaulichung in Excel-Arbeitsmappen übertragen und daraus in weiterer Folge Diagramme erstellt.

Die Ergebnisse der Bewertungsaufgabe wurden zunächst in ihrer Gesamtheit betrachtet und in Übersichtstabellen dargestellt. Für eine detailliertere Betrachtung der Ergebnisse wurden Tabellen mit den berechneten arithmetischen Mitteln, Standardabweichungen und Modalwerten erstellt. Die Bewertungsergebnisse wurden nach den Kriterien der MÜ-Systeme, Koch- und Backbegeisterung sowie der Köche beleuchtet und analysiert. Damit sollte herausgefunden werden, welches MÜ-System besser bewertet wurde und ob die Rezepte eines bestimmten Kochs bei den Umfrageteilnehmer*innen besser angekommen sind. Weiters sollte untersucht werden, wie sich das Bewertungsverhalten bei Teilnehmer*innen mit einer höheren Koch- und Backbegeisterung von jenem bei Befragten mit einer niedrigeren Koch- und Backbegeisterung beziehungsweise der gesamten Teilnehmer*innengruppe unterscheidet. Die Antworten zu Freitextfragen wurden dabei inhaltlich geordnet und entsprechend zusammengefasst.

5. Ergebnisse

An der Umfrage nahmen insgesamt 33 Personen teil, wovon 14 Personen die Umfrage vollständig ausgefüllt haben. In diesem Zusammenhang wurden die vollständig beendeten Umfragen herausgefiltert und die Ergebnisse entsprechend ausgewertet (vgl. Kapitel 4.5.). Die Ergebnisse der Umfrage werden in den folgenden Unterkapiteln präsentiert. Zudem erfolgt im Unterkapitel 5.3. eine manuelle Analyse der MÜ-Ergebnisse im Hinblick auf potentielle Verständnisprobleme bei den Umfrageteilnehmer*innen.

5.1. Demographische Daten

Unter den Umfrageteilnehmer*innen fanden sich Personen unterschiedlicher Altersgruppen. Drei Personen (21%) gehörten der Gruppe der 18- bis 24-Jährigen an. Die Gruppe der 25- bis 34-Jährigen war mit sieben Personen (50%), die Gruppen der 45- bis 54-Jährigen und 55- bis 64-Jährigen mit jeweils zwei Personen (je 14%) vertreten. Somit war die Altersgruppe der 25- bis 34-Jährigen die größte. Die Gruppen *unter 18*, *35-44* sowie *65 und älter* waren in dieser Umfrage nicht vertreten.

Der Großteil der Teilnehmer*innen fühlte sich dem weiblichen Geschlecht zugehörig. Demzufolge nahmen zwölf weibliche Personen (86%) und zwei männliche Personen (14%) an der Umfrage teil. Die Kategorie *Divers* wurde von keiner Person (0%) ausgewählt.

Bei der Frage zum höchsten Bildungsabschluss gaben 64% der Umfrageteilnehmer*innen einen Bachelor-Abschluss an. 14% haben einen Matura-Abschluss und je 7% einen Lehrabschluss, Master-Abschluss beziehungsweise einen anderen Abschluss nach der Matura. Die Person mit einem anderen Abschluss nach der Matura absolvierte eine Akademie zur Europasekretärin. Es waren keine Personen ohne Schulabschluss, mit einem Pflichtschul-, Magister- oder Doktorat-Abschluss vertreten.

Der Teilnehmer*innenkreis umfasste sowohl professionelle Translator*innen beziehungsweise Studierende der transkulturellen Kommunikation oder Translation als auch Personen aus anderen Branchen (im Nachstehenden als Lai*innen bezeichnet). Wie aus Abb. 2 zu entnehmen ist, gaben neun Personen (64%) an, professionelle Translator*innen zu sein; fünf Personen (36%) sind in einer anderen Branche tätig. Die Personen aus dem Bereich Translation waren ausschließlich weiblich.

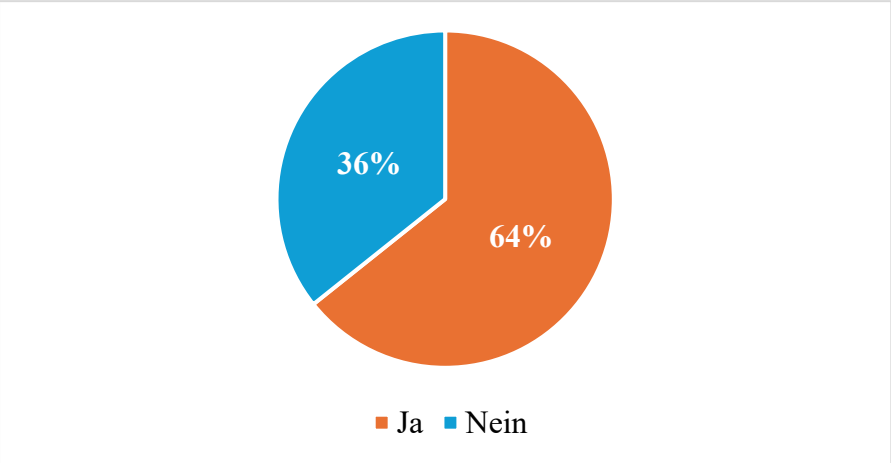

Abbildung 2: Anzahl der Personen aus dem Tätigkeitsbereich der Translation (Ja)

Bei der Fragestellung zu den Arbeitssprachen gaben acht Translatorinnen (89%) Deutsch als A-Sprache und eine Translatorin (11%) Japanisch als A-Sprache an. Darüber hinaus hat der Großteil der Translatorinnen (sechs von neun) Englisch als B-Sprache und je eine Person führte Deutsch, Italienisch beziehungsweise Russisch als B-Sprache an. Zwei Translatorinnen hatten keine C-Sprache; drei listeten Spanisch und je eine Englisch, Französisch, Russisch und Ungarisch auf. Wie in Tabelle 3 abgebildet, beherrschen die befragten Translatorinnen weitere Sprachen, die sie nicht unbedingt im Beruf verwenden, auf verschiedenen Niveaus. Lediglich zwei beherrschen keine weiteren Sprachen.

Tabelle 3: Weitere Sprachkenntnisse der professionellen Translatorinnen

Sprachen (Niveau)	**Anzahl der professionellen Translatorinnen**
Keine	2
Chinesisch (B1), Serbisch (B1)	1
Englisch (A2)	1
Englisch (C1)	1
Französisch (A2)	1
Französisch (n.a.), Spanisch (n.a.)	1
Französisch (B2), Italienisch (B2)	1
Serbisch (C2)	1
Gesamt	9

Die fünf Lai*innen kommen aus verschiedenen Branchen: Zwei Personen sind im Bereich Soziales, Gesundheit tätig, eine im Bereich Bildung und eine im Bereich Handel, Logistik. Eine Person gab unter Sonstiges den Bereich Internationale kriminalpolizeiliche Zusammenarbeit an.

Die Frage, ob Deutsch ihre Erstsprache ist, beantworteten elf Personen (79%) mit Ja und drei Personen (21%) mit Nein. Die Personen, deren Erstsprache nicht Deutsch ist, gaben Japanisch, Serbisch und Serbisch/Kroatisch als ihre Erstsprachen an. Sie schätzen Ihr Sprachniveau auf Deutsch jedoch alle mit C2 ein.

Im zweiten Teil der Umfrage spezifizierten die Befragten ihre Koch- und Backerfahrung. Wie aus Abb. 3 ersichtlich, kocht der Großteil der Teilnehmer*innen mehrmals im Monat nach Rezept. Vier Teilnehmer*innen gaben an, mehrmals im Jahr nach Rezept zu kochen. Weitere drei Personen kochen mehrmals in der Woche und eine Person täglich nach Rezept. In Abb. 4 wird dargestellt, wie oft die Befragten nach Rezept backen. In diesem Kontext gab die Hälfte der Befragten an, mehrmals im Monat nach Rezept zu backen; sechs Personen kreuzten *mehrmals im Jahr* und eine Person *mehrmals in der Woche* an.

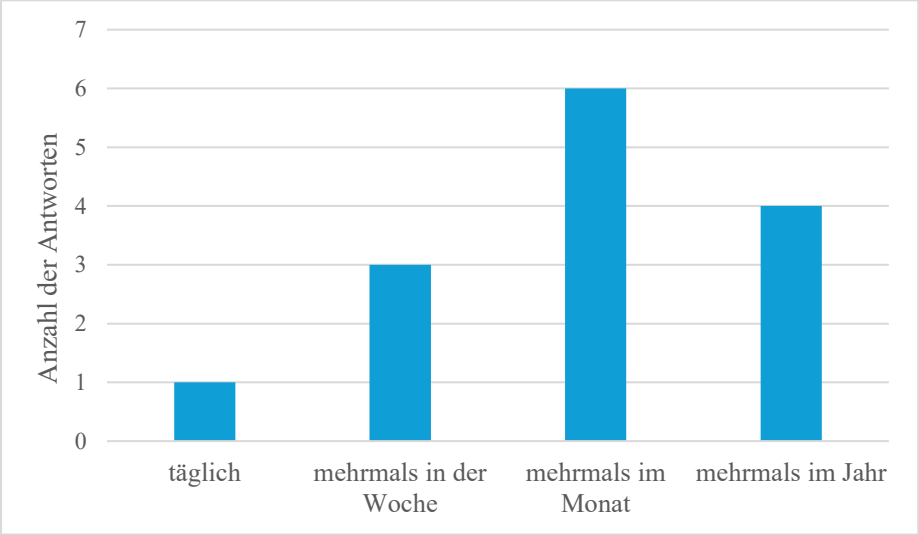

Abbildung 3: Kochverhalten der Umfrageteilnehmer*innen

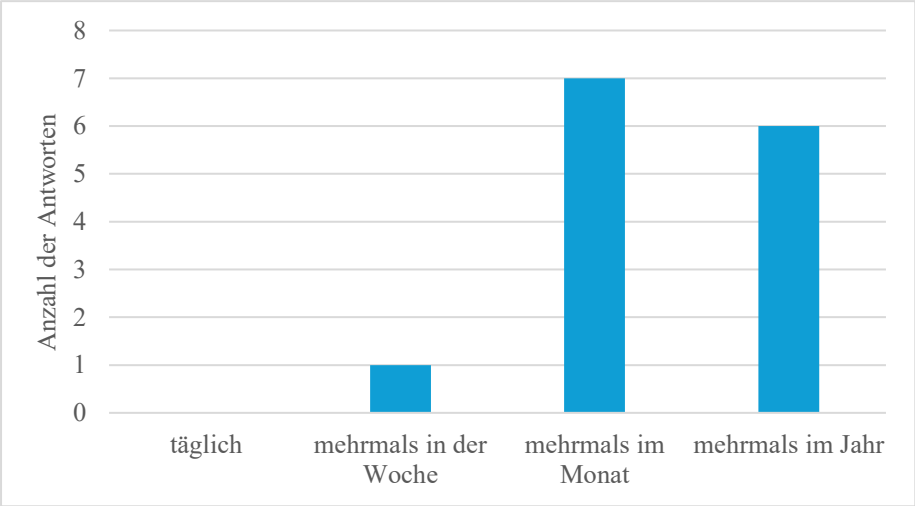

Abbildung 4: Backverhalten der Umfrageteilnehmer*innen

Die Umfrageteilnehmer*innen wurden schließlich gebeten, ihre Koch- beziehungsweise Backexpertise auf einer Skala von 1 (= nicht vertraut) bis 5 (= sehr vertraut) einzuschätzen. Bei der Vertrautheit mit Kochtechniken wurde ein arithmetisches Mittel von 3,64 berechnet, wobei der Wert 4 am häufigsten (neunmal) ausgewählt wurde. In diesem Zusammenhang kreuzte eine Person den Wert 5, je zwei Personen die Werte 2 und 3 und keine Person den Wert 1 an. Die Befragten schätzten im Gesamtbild ihre Backkenntnisse etwas höher ein. So wurde ein arithmetisches Mittel von 3,71 ermittelt, wobei auch hier der Wert 4 am häufigsten (siebenmal) ausgewählt wurde. In diesem Kontext wurde der Wert 5 dreimal, der Wert 3 zweimal und die Werte 2 beziehungsweise 1 je einmal angekreuzt. Dementsprechend besteht in beiden Fällen eine Tendenz in Richtung einer höheren Vertrautheit mit Koch- und Backprozessen.

Wie aus Abb. 5 ersichtlich, wählten bei der Frage, aus welcher Küche die Teilnehmer*innen besonders gerne Gerichte zubereiten, zehn Personen (71%) Italienisch, sieben (50%) Asiatisch und sechs (43%) Österreichisch aus. Darüber hinaus kreuzten je drei Teilnehmer*innen (21%) Amerikanisch, Balkan, Orientalisch und Sonstiges an. Bei der Antwortmöglichkeit *Sonstiges* wurden Japanisch, Deutsch und Divers angegeben. Am Ende dieses Umfrageteils berichteten alle neun professionellen Translatorinnen (100%), mindestens einmal ein Koch- beziehungsweise Backrezept übersetzt zu haben.

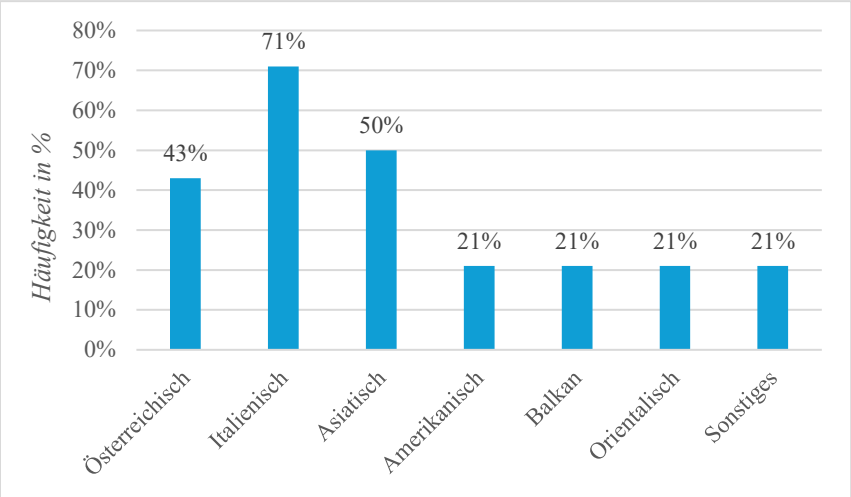

Abbildung 5: Bevorzugte Küchenrichtungen der Umfrage-teilnehmer*innen

5.2. Erfahrung mit maschineller Übersetzung

Von den 14 Umfrageteilnehmer*innen nutzen zwölf (86%) MÜ-Systeme, darunter alle neun professionellen Translatorinnen sowie drei der fünf Lai*innen. Dementsprechend gaben lediglich zwei Teilnehmer*innen (14%) an, keine MÜ-Systeme zu verwenden. Wie in Tabelle 4 aufgelistet, ist unter den MÜ-Nutzer*innen dieser Umfrage Google Translate mit 92% am beliebtesten, wobei 100% der befragten Lai*innen und acht von neun Translatorinnen dieses System angegeben haben. Darauf folgt DeepL mit 67% der MÜ-Nutzer*innen. Während DeepL von fast allen Translatorinnen genutzt wird, ist dieses MÜ-System unter den Lai*innen weniger beliebt. Zwei Personen listeten ChatGPT und je eine Person den PONS-Textübersetzer und den Online-Übersetzungsdienst der Europäischen Kommission eTranslation auf. Von einer Person wurde das Online-Wörterbuch LEO angegeben, obwohl es kein MÜ-System darstellt.

Tabelle 4: Nutzung maschineller Übersetzungssysteme

MÜ-System	Anzahl der Antworten	Anzahl der prof. Translatorinnen	Anzahl der Lai*innen
DeepL	8 (67%)	7 (78%)	1 (33%)
Google Translate	11 (92%)	8 (89%)	3 (100%)
ChatGPT	2 (17%)	1 (11%)	1 (33%)
LEO	1 (8%)	0	1 (33%)
PONS	1 (8%)	0	1 (33%)
eTranslation	1 (8%)	1 (11%)	0
Gesamt	12 Teilnehmer*innen	9 Translatorinnen	3 Lai*innen

Wie in Abb. 6 dargestellt, verwenden die Umfrageteilnehmer*innen relativ häufig MÜ-Systeme. So gaben zwei Personen an, täglich MÜ-Systeme zu nutzen. In beiden Fällen handelt es sich dabei um professionelle Translatorinnen. Jeweils vier Personen verwenden solche Systeme mehrmals in der Woche oder mehrmals im Jahr, darunter jeweils drei Translatorinnen und ein*e Lai*in. Zwei Personen wählten die Antwortmöglichkeit *mehrmals im Jahr* aus. Die Verwendungszwecke für die MÜ-Systeme sind, wie in Abb. 7 erkennbar, vielfältig. An oberster Stelle stehen mit neun Personen (75%) die privaten Zwecke. Darauf folgen mit je der Hälfte der MÜ-Nutzer*innen berufliche Zwecke und Ausbildung, wobei diese Gründe in beiden Fällen von fünf professionellen Translatorinnen und einem*einer Lai*in angekreuzt wurden. Ein Drittel der befragten MÜ-Nutzer*innen verwendet solche Systeme zur Kommunikation. Darüber hinaus führte eine Person unter *Sonstiges* Lernzwecke an. Die Teilnehmer*innen sind im Allgemeinen mit den Ergebnissen maschineller Übersetzungssysteme zufrieden. Die Hälfte der MÜ-Nutzer*innen wählte auf einer Skala von 1 (= nicht zufrieden) bis 5 (= sehr zufrieden) den Wert 4 aus. Vier Personen kreuzten den Wert 3 an und jeweils eine Person die Werte 2 und 5. Daraus ergibt sich ein arithmetisches Mittel von 3,58, wobei professionelle Translatorinnen mit einem Mittelwert von 3,67 etwas zufriedener sind als Lai*innen mit einem Mittelwert von 3,3.

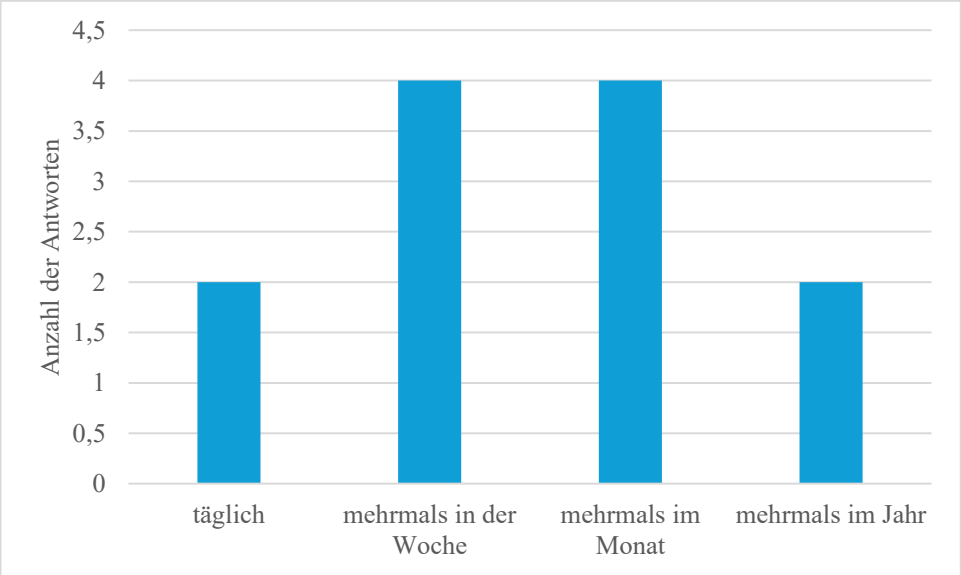

Abbildung 6: Häufigkeit der Nutzung maschineller Übersetzungssysteme

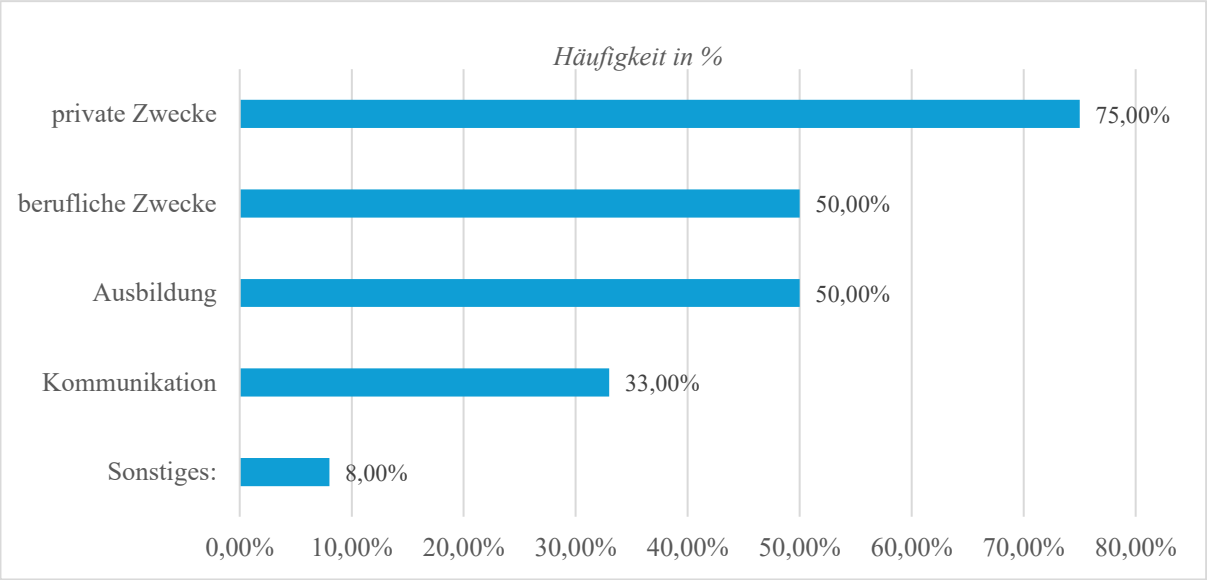

Abbildung 7: Nutzungszwecke maschineller Übersetzungssysteme

Vier von 14 Personen gaben an, bereits einmal MÜ-Systeme für Koch- beziehungsweise Backrezepte verwendet zu haben. Dabei handelte es sich bei drei Befragten um professionelle Translatorinnen. Bei der Zufriedenheit mit der Qualität der MÜ von Koch- und Backrezepten wurde ein arithmetisches Mittel von 2,75 berechnet, wobei drei Personen auf der Skala den Wert 3 (= zufrieden) und eine Person den Wert 2 (= eher nicht zufrieden) gewählt haben.

5.3. Manuelle Analyse der Übersetzungen

Im Vorfeld der Präsentation der Ergebnisse der Bewertungsaufgabe werden die Übersetzungen im folgenden Unterkapitel manuell analysiert. Hierbei erfolgt eine Einschätzung der im Zuge der MÜ entstandenen Fehler, die bei den Umfrageteilnehmer*innen zu Verständnisschwierigkeiten führen könnten.

Übersetzung 1 enthält wenige Unstimmigkeiten, diese können jedoch zu Verständnisbeeinträchtigungen und in weiterer Folge einer schlechteren Bewertung führen. So wurden etwa für *red pepper topping*, den Belag aus roter Paprika, verschiedene Übersetzungsvarianten verwendet: *roter Pfeffer als Belag*, *Paprikaaufstrich*, *Paprika-Topping*. Die Anweisung *Für den Paprikaaufstrich das Öl in einer Pfanne erhitzen, die Koriandersamen grob zerstoßen und die Chiliflocken hinzufügen.* kann ebenfalls missverstanden werden, da aus der Übersetzung nicht eindeutig hervorgeht, dass man die Koriandersamen hinzufügen soll. Darüber hinaus enthält die MÜ zwei terminologische Fehler. Der Spinat für den Belag soll demnach gebraten werden, bis er *verwelkt* ist und nicht bis er *zusammenfällt*. Weiters soll die Suppe mit einem Klecks des Paprika-Belags *abgeschlossen* statt *garniert* werden.

Die Zutaten und Anweisungen in Übersetzung 2 sollten zum Großteil verständlich sein. Lediglich *Montgomery Cheddar* könnte einem deutschsprachigen Publikum unbekannt sein, wobei der Zusatz *oder anderer kräftiger traditioneller Cheddar-Käse* zu einem besseren Verständnis beiträgt. Die Anweisungen enthalten eine Inkonsistenz im Stil, die für das Publikum durchaus irritierend sein kann. So werden der Infinitiv und die Sie-Form abwechselnd verwendet. Der abschließende Zubereitungsschritt *top each [bowl] with a spoonful of the ceps* wurde mehrdeutig mit *auf jede Schüssel einen Löffel Steinpilze geben* übersetzt. In diesem Zusammenhang wäre eine Formulierung wie *jede Schüssel mit einem Löffel Steinpilze garnieren* angebrachter.

Die Zubereitungsanweisungen in Übersetzung 3 enthalten keine Fehler, die das Verständnis gravierend beeinträchtigen könnten. Die Infinitiv- und die Sie-Form werden abwechselnd verwendet. In Übersetzung 4 wurde das Verb *cook* mit *kochen* statt richtigerweise mit *braten* übersetzt. In der Anweisung *simmer until the liquid has reduced a bit, check for seasoning* wurde der Teil *check for seasoning* wörtlich mit *Gewürze überprüfen* übersetzt, was für das Zielpublikum irritierend sein könnte. Eine korrekte idiomatische Übersetzung sollte in diesem Zusammenhang eher das Verb *abschmecken* beinhalten.

In Übersetzung 5 könnte das MÜ-Ergebnis zur Mengenangabe *knob* (2 knobs of butter) zu Irritationen führen. Hierfür wurde *Stück* verwendet, worunter man sich möglicherweise mehr vorstellt, als tatsächlich benötigt wird. Genauere Angaben, wie etwa *Esslöffel*, wären in diesem Kontext zielführender. Darüber hinaus enthalten die Anweisungen gewisse Formulierungen beziehungsweise Ungereimtheiten, die zu Verständnisproblemen führen können. So wurde der Satz *Rub the beef all over with sea salt and black pepper.* wörtlich mit *Das Rindfleisch rundherum mit Meersalz und schwarzem Pfeffer einreiben.* übersetzt. In diesem Fall wäre eine Formulierung wie *Das Rindfleisch großzügig…einreiben.* verständlicher und idiomatischer. Die Phrase *taste and season to perfection* wurde von Google Translate mit *abschmecken und perfekt würzen* übersetzt. An dieser Stelle ist zu erwarten, dass die Umfrageteilnehmer*innen den Sinn der Anweisung zwar verstehen, den Ausdruck *perfekt würzen* allerdings befremdlich finden. In diesem Kontext wären Formulierungen wie *abschmecken und nachwürzen* geläufiger. Darüber hinaus enthält Übersetzung 5 auch einzelne terminologische Fehler. So wurde, wie auch in anderen Sets, fälschlicherweise für das Verb *cook* von Google Translate teilweise *kochen* statt *braten* verwendet. Auch in Schritt 9 *snugly wrap the pastry around the beef, pinching the ends to seal.* ist der Teil mit dem Zusammendrücken der Teigenden fehlerhaft. Anstatt die Enden zusammenzudrücken, um sie zu *verschließen*, sollen diese laut

Übersetzung 5 zusammengedrückt werden, um sie *abzudichten*. Zwar versteht man auch in diesem Fall den Sinn der Aussage, dennoch ist die Verwendung des Verbs *abdichten* in diesem Kontext nicht korrekt.

Übersetzung 6 enthält keine für das Zielpublikum unbekannten Zutaten. Die Übersetzung der Zutatenliste weist jedoch zwei terminologische Fehler auf, die das Verständnis und in weiterer Folge die Ausführung durchaus beeinträchtigen können. So wurde in der MÜ bei der Zutat für die Rotweinsauce *shallots, peeled and sliced* für das mehrdeutige Partizip *sliced* die spezifische Übersetzung *in Scheiben geschnitten* verwendet. Hierbei wäre eine allgemeinere Formulierung wie *geschnitten* angemessener. In der Anmerkung zu den Rindfleischabschnitten, engl. *beef trimmings*, soll der Metzger gebeten werden, diese beim Abschneiden des Filets *zu reservieren,* anstatt sie richtigerweise *aufzuheben*. Ähnlich wie bei Übersetzung 5 könnten die Anweisungen auch in diesem Fall das Zielpublikum vor größere Herausforderungen stellen. Diese enthalten gewisse irreführende Formulierungen, terminologische Fehler beziehungsweise Inkonsistenzen. Die Verwendung von teilweise längeren Sätzen führte bei der MÜ von Schritt 10 zum Beispiel dazu, dass ein irreführendes Pronomen hinzugefügt wurde. Im Satz *Wenn Sie bereit sind, die Wellingtons zuzubereiten, ritzen Sie den Teig leicht ein und bepinseln Sie ihn erneut mit der Eimasse, dann backen Sie ihn […]* wurde im letzten Teil in der Übersetzung angegeben, dass der Teig gebacken werden soll. Eigentlich sollen die Wellingtons als Einheit mit dem Teig gebacken werden, wodurch die Verwendung des Pronomens *ihn* irreführend ist. Außerdem wurde der letzte Satz *Serve the wellingtons sliced, with the sauce as an accompaniment.* in der MÜ etwas verdreht. Während das Original aussagt, dass die Wellingtons in Scheiben geschnitten und mit der Sauce als Beilage serviert werden sollen, setzt die MÜ bereits voraus, dass die Wellingtons unter Umständen in einem vorherigen Schritt in Scheiben geschnitten wurden. Dies ist zwar kein Fehler, der die Ausführung beeinträchtigt, eine solche Umstellung des Satzes macht jedoch einen wesentlichen Unterschied aus. Weiters soll gemäß des MÜ-Resultats in der Phrase *cook […] until all the excess moisture has evaporated* die Masse in der Pfanne gebraten werden, bis die *überschüssige Feuchtigkeit verdampft ist*. Dies ist in diesem Kontext jedoch nicht korrekt. Vielmehr sollte hier statt von *Feuchtigkeit* von *Flüssigkeit* die Rede sein. Schritt 9 enthält im Originaltext die Anweisung *Pour in the vinegar and let it bubble for a few minutes until almost dry.*, die im Übersetzungstext mit *Den Essig dazugießen und einige Minuten sprudelnd kochen lassen, bis er fast trocken ist.* übertragen wurde. Die idiomatisch richtige Übersetzung sollte an dieser Stelle *bis er eingekocht ist* lauten. Zudem sollen in Schritt 10 die Wellingtons gebacken werden, bis der Teig *gar* ist. Die

Verwendung des Adjektivs *gar* ist unter Berücksichtigung der kulinarischen Fachsprache in diesem Kontext unangemessen. Stattdessen wäre der Terminus *durchgebacken* angebracht. Schließlich ist als letzte terminologische Ungenauigkeit die Übersetzung von *carving* im Satz *Rest for 10 minutes before carving.* zu erwähnen. Das Verb wurde von DeepL mit *tranchieren* übersetzt. *Tranchieren* stammt aus dem Französischen und bedeutet die Zerteilung beziehungsweise das Aufschneiden eines Bratens, Wildes beziehungsweise Geflügels in Scheiben (Duden 2025). Voraussetzung hierfür ist, dass es sich dabei um im Ganzen gebratenes Fleisch handelt. Beef Wellington enthält zwar im Inneren ein ganzes Rinderfilet, das Fleisch ist jedoch nur Teil des Gerichts. Aus diesem Grund wäre hier eine allgemeinere Formulierung, wie zum Beispiel *anschneiden*, angebrachter. Schließlich wurde in Übersetzung 6 die Zutat *beef trimmings* für die Rotweinsauce in der Zutatenliste und den Anweisungen unterschiedlich übersetzt. Während *Rindfleischabschnitte* in der Zutatenliste korrekt ist, ist die Verwendung von *Rinderfilets* in den Anweisungen nicht angemessen und kann zu Missverständnissen führen.

Darüber hinaus weisen sowohl Übersetzung 5 als auch Übersetzung 6 gewisse terminologische Inkonsistenzen auf. So wird für *mushrooms* in beiden Übersetzungen *Champignons* und *Pilze* verwendet. Da in beiden Rezepten eine Mischung von Pilzen in den Zutaten gefordert wird, wäre die alleinige Verwendung der allgemeineren Variante *Pilze* angemessen. *Eggwash* aus Text 5 wurde einmal mit *Ei* und einmal mit *Eigelb* übersetzt. Da in den Zutaten ein Ei aufgelistet ist und die Anweisungen keine Trennung des Eigelbs beinhalten, ist davon auszugehen, dass das ganze Ei verquirlt und der Teig damit bestrichen werden soll. Somit kann die Übersetzung mit *Eigelb* bei der Leserschaft zu Missverständnissen führen. Die inkonsistente Verwendung von *Ei* und *Eimasse* in Übersetzung 6 hingegen ist nicht irreführend, da in den Zutaten bereits erklärt wird, dass das Ei mit Wasser und Salz verquirlt werden soll. Dennoch wäre die ausschließliche Nutzung einer Übersetzungsvariante hilfreicher. Schließlich weisen beide MÜ-Resultate eine Inkonsistenz im Stil auf, die für das Zielpublikum durchaus irritierend sein kann. Wie in den Sets 1 und 2 werden auch hier Infinitiv und Sie-Form abwechselnd verwendet.

In Übersetzung 7 wurde *roasting tray* mit *Bratpfanne* statt richtigerweise mit *Backblech* oder *Kuchenform* übersetzt. Auch die Verwendung von *Bratensaft* für *pan juices* kann irreführend sein, da es sich dabei eigentlich um den übriggebliebenen Ananassaft aus der Pfanne handelt. Die Übersetzung weist darüber hinaus gewisse terminologische Ungereimtheiten auf, die zu Missverständnissen führen können. *Desiccated coconut* wurde in der Zutatenliste irrtümlicherweise mit *getrocknete Kokosnüsse* und in den Anweisungen schließlich korrekt mit

Kokosraspeln übersetzt. Während im dritten Schritt die Form mit *Backpapier* ausgelegt wird, wird im letzten Schritt *Backfolie* vom Kuchen abgezogen. Außerdem übersetzte Google Translate die Überschrift *pineapple upside-down cake* mit *Ananaskuchen mit Deckel*. Diese Übersetzung ist nicht korrekt, sollte aber aufgrund der Bereitstellung eines Fotos bei der Bewertung zu keinen schwerwiegenden Verständnisproblemen führen.

Übersetzung 8 sollte in den Anweisungen keine Verständnisprobleme bereiten. Lediglich *bicarbonate of soda* wurde in der Zutatenliste und den Anweisungen inkonsistent übersetzt, nämlich einmal mit der chemischen Bezeichnung *Natriumbikarbonat* und einmal mit der in Rezepten üblicheren Variante *Natron*. Für die Zutat *mixed spice*, eine in Großbritannien erhältliche Gewürzmischung aus Zimt, Piment, Muskatnuss, Ingwer, Nelke und Koriander (vgl. Kochwiki 2016), wurde hier die Bezeichnung *Gewürzmischung* verwendet, mit der ein deutschsprachiges Publikum nicht viel anfangen kann. Hierbei wäre eine nähere Erklärung beziehungsweise zusätzliche Auflistung der entsprechenden Zutaten und einzelnen Gewürze gängige Praxis. Auch die Übersetzung von *muscovado sugar* mit *Muscovado-Zucker* ist im deutschsprachigen Raum eher unbekannt, was die Auflistung vergleichbarer Produkte, wie etwa *Vollrohrzucker*, für eine erfolgreiche Umsetzung des Rezepts unverzichtbar macht.

5.4. Ergebnisse der Bewertungsaufgabe

In diesem Unterkapitel werden zunächst die Gesamtergebnisse der Bewertungsaufgabe näher beleuchtet. In einem nächsten Schritt wird detaillierter analysiert, wie die Übersetzungen in den einzelnen Sets bewertet wurden. Anschließend wird die Bewertung der MÜ-Ergebnisse nach den Kriterien der MÜ-Systeme und Koch- beziehungsweise Backbegeisterung näher betrachtet. Schließlich soll festgestellt werden, ob ein Koch besser bewertet wurde als der andere.

Insgesamt fiel die Bewertung der maschinellen Übersetzungen tendenziell positiv aus. Wie in Tabelle 5 dargestellt, erzielten alle Rezepte bei der Bewertung positiver Aussagen, also der Aussagen zur Akzeptabilität, Rezeptzubereitung, Verständlichkeit der Koch- beziehungsweise Backanleitung, Bekanntheit der Zutaten, Verständlichkeit der Mengenangaben beziehungsweise Maßeinheiten und richtigen Verwendung von Fachbegriffen, Mittelwerte, die größer als 3 sind. Dementsprechend stimmten die Umfrageteilnehmer*innen diesen Aussagen im Gesamtbild zu, wobei die Übersetzungen 3 und 4 mit Mittelwerten von 3,3 und 3,28 am schlechtesten und Übersetzung 7 mit einem Mittelwert von 4,24 am besten abschnitten. Auch die Bewertung negativer Aussagen, also der Aussagen zu sprachlichen beziehungsweise idiomatischen Fehlern, Grammatikfehlern und Verwendung im Deutschen unüblicher Begriffe,

erzielte in der Mehrheit positive Ergebnisse (siehe Tabelle 5). Mit Ausnahme der Übersetzungen 4 und 5 wurde diesen Aussagen insgesamt nicht zugestimmt, wobei erstere am schlechtesten abschnitt. Übersetzung 6 erreichte den günstigsten Mittelwert (siehe Tabelle 5).

Tabelle 5: Übersicht der Gesamtergebnisse nach Rezept, MÜ-System DeepL oder Google Translate (GT), durchschnittlichem Mittelwert der Bewertung positiver Aussagen (Pos), durchschnittlichem Mittelwert der Bewertung negativer Aussagen (Neg), Prozentsatz der zum Nachkochen beziehungsweise Nachbacken motivierten Umfrageteilnehmer*innen und Koch Jamie Oliver (JO) oder Gordon Ramsay (GR)

Übersetzung	**MÜ**	**Pos**	**Neg**	**Nachkochen/-backen**	**Koch**
Übersetzung 1	DeepL	3,8	2,93	64%	JO
Übersetzung 4	DeepL	3,28	3,24	86%	GR
Übersetzung 6	DeepL	3,82	2,19	50%	GR
Übersetzung 8	DeepL	3,88	2,57	64%	JO
Übersetzung 2	GT	3,97	2,93	43%	GR
Übersetzung 3	GT	3,3	2,88	71%	GR
Übersetzung 5	GT	3,68	3,12	43%	JO
Übersetzung 7	GT	4,24	2,67	43%	JO

In diesem Zusammenhang ist auch der Anteil der Teilnehmer*innen, die diese Rezepte nachkochen beziehungsweise -backen würden, zu berücksichtigen. Wie aus Tabelle 5 ersichtlich, würden die Umfrageteilnehmer*innen die Mehrheit der Rezepte ausprobieren. Lediglich drei Gerichte würden von weniger als der Hälfte der Teilnehmer*innen zubereitet werden (siehe Tabelle 5). Hierbei ist die Diskrepanz zwischen der Aussagenbewertung gewisser Übersetzungen und dem Anteil der zum Nachkochen beziehungsweise Nachbacken motivierten Personen zu nennen. So wurde Übersetzung 4 insgesamt am schlechtesten bewertet, erzielte aber mit 86% die höchste Nachkochrate (siehe Tabelle 5). Im Gegensatz dazu würde Übersetzung 7 mit der besten Bewertung der positiven Aussagen und zweitbesten Bewertung der negativen Aussagen nur von 43% der Teilnehmer*innen ausprobiert werden (siehe Tabelle 5).

Die Betrachtung der Gesamtergebnisse nach dem Aspekt der MÜ-Systeme lässt den Rückschluss zu, dass beide Systeme bei der Bewertung relativ ähnliche Ergebnisse erreichten und sich keines mit seiner Leistung klar hervorheben konnte (siehe Tabelle 5). Insgesamt konnte im Falle von DeepL bei den positiven Aussagen ein Mittelwert von 3,7 und bei den

negativen Aussagen ein Mittelwert von 2,73 berechnet werden. Google Translate erzielte in diesem Kontext bei den positiven Aussagen einen Mittelwert von 3,8 und bei den negativen Aussagen einen Mittelwert von 2,9. Dementsprechend liegt Google Translate bei den positiven Aussagen und DeepL bei den negativen Aussagen knapp vorne.

Unter Berücksichtigung der Frage, von welchem Koch die jeweiligen Rezepte stammen, gelangt man zum Ergebnis, dass die Rezepte von Gordon Ramsay im direkten Vergleich der Mittelwerte zur Bewertung der positiven und negativen Aussagen bessere Ergebnisse erzielten als jene von Jamie Oliver (siehe Tabelle 5). Auch der Prozentsatz der zum Nachkochen motivierten Umfrageteilnehmer*innen zeigt, wie in Tabelle 5 dargestellt, eine deutliche Tendenz zu Gordon Ramsay.

In Set 1 fiel die Bewertung der Übersetzungen tendenziell positiv aus. Wie Tabelle 6 zu entnehmen ist, fanden die Teilnehmer*innen im Schnitt, dass die Übersetzungen akzeptabel und die Rezepte damit in der Praxis umsetzbar sind. Bei der Aussage Alle Zutaten sind Ihnen bekannt. waren die Antworten unterschiedlich auf der Skala von 1 bis 5 verteilt. So wählten bei Übersetzung 1 insgesamt sieben Personen die Werte 1 beziehungsweise 2 (Wert 1: zweimal; Wert 2: fünfmal) und sieben Personen die Werte 4 und 5 (Wert 4: zweimal; Wert 5: fünfmal). Zu Übersetzung 2 kreuzten je sechs Personen die Werte 1 und 2 (Wert 1: einmal; Wert 2: fünfmal) beziehungsweise 4 und 5 (Wert 4: einmal; Wert 5: fünfmal) an; zwei Personen gaben den Wert 3 an. Die in diesem Kontext gestellte Frage nach den unbekannten Zutaten wurde dementsprechend von je neun Teilnehmer*innen beantwortet. Bei Übersetzung 1 gaben, wie aus der Gesamtübersicht zu den nicht gängigen oder unbekannten sprachlichen Aspekten in Tabelle 7 ersichtlich, 100% an, *Poppadoms* nicht zu kennen. Vier Personen (44%) führten *Kardamomkapseln* und je eine Person (je 11%) *Koriandersamen*, *gebratene Paprikaschoten aus dem Glas* und *Kokosnussflocken* an (siehe Tabelle 7). Dabei spielte die Tatsache, ob die Teilnehmer*innen regelmäßig asiatisch kochen, keine gravierende Rolle. Zutaten wie *Poppadoms* und *Kardamomkapseln* wurden auch in diesen Fällen als unbekannt eingeordnet. Bei Übersetzung 2 wurden *Bayonner-Schinken* (89%), *Kastanienpilze* (78%), *Montgomery Cheddar* (56%) und *Thymian* (11%) angegeben (siehe Tabelle 7). Die Verständlichkeit und Nachvollziehbarkeit der Mengenangaben und Maßeinheiten wurde mit Mittelwerten von 4,21 und 4,5 ebenfalls gut beurteilt. Bei der Freitextfrage zu unüblichen Begriffen gab gemäß Tabelle 7 je eine Person an, dass die Zentimeter-Angabe beim Ingwer in Übersetzung 1 sowie die Mengenangabe *Klecks* nicht ganz nachvollziehbar und ungewöhnlich sind. Bei Übersetzung 2 merkte eine Person an, dass in der Zutatenliste zu den Pilzen unterschiedliche

Angaben gemacht wurden, und zwar Steinpilze in Gramm und Kastanienpilze in Stück (siehe Tabelle 7). Die richtige Verwendung von Fachbegriffen wurde, wie aus Tabelle 6 ersichtlich, in der Gesamtheit mittelmäßig mit einer eher positiven Tendenz beurteilt. Den Aussagen zu sprachlichen beziehungsweise idiomatischen und grammatikalischen Fehlern wurde eher nicht zugestimmt. Die Bewertung der Üblichkeit gewisser Begriffe im Deutschen fiel mittelmäßig aus. Die Bewertungen der letzten vier Aussagen variieren bei beiden Übersetzungen, was nicht zuletzt durch die relativ hohen Standardabweichungen bestätigt wird (siehe Tabelle 6). Bei den Freitextfragen zu unüblichen Begriffen und weiteren Kommentaren wurde für beide Übersetzungen eine Vielzahl an Anmerkungen vorgenommen, wobei manche positiver ausfielen als andere.

Tabelle 6: Bewertung von Set 1 mit DeepL und Google Translate (GT) nach dem Durchschnittswert und Modalwert pro zu bewertender Aussage durch 14 Personen

Aussage	**Übersetzung 1 (DeepL)**		**Übersetzung 2 (GT)**	
	Ø (σ)	**Modal-wert**	**Ø (σ)**	**Modal-wert**
Die Übersetzung ist akzeptabel.	3,93 (0,88)	4	3,86 (0,99)	3 & 5
Mit diesem Rezept kann man dieses Gericht zubereiten.	4,29 (0,7)	4 & 5	4,43 (0,62)	5
Die Kochanleitung ist verständlich und könnte leicht umgesetzt werden.	3,64 (1,04)	3	4,14 (1,06)	5
Alle Zutaten sind Ihnen bekannt.	3,21 (1,57)	2 & 5	3,29 (1,44)	2 & 5
Die verwendeten Mengenangaben bzw. Maßeinheiten sind verständlich und nachvollziehbar.	4,21 (1,01)	5	4,5 (0,73)	5
Die richtigen Fachbegriffe werden verwendet.	3,5 (1,18)	4	3,57 (1,35)	5
Die Übersetzung enthält sprachliche bzw. idiomatische Fehler.	3 (1,46)	4	2,93 (1,49)	4
Die Übersetzung enthält grammatikalische Fehler.	2,29 (1,39)	1 & 2	2,86 (1,55)	1
Bestimmte Begriffe sind für Sie im Deutschen nicht üblich.	3,5 (1,5)	4	3 (1,65)	1 & 4

Tabelle 7: Übersicht der nicht gängigen oder unbekannten sprachlichen Aspekte nach Zutaten, Vorgängen und Sonstigem, wie etwa Mengen- oder Zeitangaben sowie sonstigen sprachlichen Auffälligkeiten, mit der Personenanzahl in Klammern und Koch Jamie Oliver (JO) oder Gordon Ramsay (GR)

Übersetzung	Zutaten	Vorgänge	Sonstiges	Koch
Übersetzung 1	Poppadoms (9) Kardamomkapseln (4) Koriandersamen (1) gebratene Paprika-schoten aus dem Glas (1) Kokosnussflocken (1) roter Pfeffer als Belag (3) die Paprika (1)	mit roter Paprika und Spinat überbacken (3) Zwiebel kochen (2) Paprikaaufstrich kochen (1) Samen herausschütteln (3) Suppe mit einem Klecks Paprika-Topping abschließen (3) sanft köcheln lassen (1) kochen, bis der Spinat verwelkt ist (5)	Zentimeter-Angabe beim Ingwer (1) Klecks (1) trockene Pfanne (2) Paprikaaufstrich (3) Methode (2)	JO
Übersetzung 4	Pedro Ximenez (12) Frisée-Salat (6) Sherry-Essig (5) Manchego-Käse (5) Chorizo (2) Tomate (1) gelbe Zwiebel (1)	Salat in Essig geschwenkt (1) Gewürze überprüfen und bei Bedarf anpassen (5) um alle gebräunten Stücke aus der Pfanne zu kratzen (4) Tomaten kochen, bis sie Blasen werfen (3) mit Salz und einem Spritzer Zitronensaft bestreuen (3) köcheln, bis die Flüssigkeit etwas eingekocht ist (1) Käse großzügig darüber raspeln (1) Zwiebeln hinzufügen (1)	Unzen (6) Tasse (1) 1 Avocado plus Salz und Zitronensaft (1) Anweisungen zum Garen (1) Spritzer (1) die Ränder von der Chorizo (1) Schnittfläche der Tomate (1)	GR
Übersetzung 6	Schalotten (2) Rindfleisch-abschnitte (2/2)*	unter häufigem Rühren (2) Essig dazugießen (1)	Gas 6 (1) 1 x 750ml Flasche Rotwein (2)	GR

	Lorbeerblatt (1) Thymianzweig (1) Waldpilze (1) Rinderfilet vs. Rindfleisch (1)	sprudelnd kochen lassen (1) Essig kochen, bis er fast trocken ist (1)	Rindfleisch-Wellington (2) Methode (1) dreifache Lage Frischhaltefolie (2) ein Blatt Frischhaltefolie (1) Duxelle (1) Schreibweise Duxelle (1) Verwendung von Infinitiv und Sie-Form (1) Schreibweise Soße vs. Sauce (1) Verwendung der Pluralform Wellingtons (1)	
Übersetzung 8	Muscovado-Zucker (9) Dessert-Äpfel (2/1)* Gewürz-mischung (1/4)* Natriumbikarbonat (1) Natron (1) normales Mehl (2) feuchte Zutaten (1) Natriumbikarbonat vs. Natron (3)	Beschreibung des Kuchens als klebrig-süß (1) backen, bis ein Spieß sauber herauskommt (2) Zucker und Margarine in einer Pfanne schmelzen (1) Walnüsse gießen (1) die trockenen Zutaten schnell, aber gründlich mit den feuchten Zutaten vermischen (1)	Gas 4 (1) Größenangabe 23 cm große, quadratische Kuchenform (1) plus extra zum Einfetten (1) Methode (1) Nachmittagstee (1)	JO
Übersetzung 2	Bayonner-Schinken (8) Kastanienpilze (7) Montgomery Cheddar (5) Thymian (1) Blumenkohl (2)	Suppe durch ein Sieb passieren (1) Steinpilze braten, bis sie gebräunt sind (1) Blumenkohl in einen Mixer geben und pürieren (1) bei Bedarf mehr Salz und Pfeffer hinzufügen (1) auf jede Schüssel einen Löffel Steinpilze geben (1)	unterschiedliche Mengenangaben bei Pilzen (1) Toastie (3) Sandwich (1) schwere Bratpfanne (3) schwach geheizter Backofen (2) etwas Gewürz (1) bis die Blumenkohl ist weich (7)	GR

		entfernen Sie fast alle Blätter vom Blumenkohl (1)	Schreibweise Bayonner-Schinken (1) Garnitur (1) Kochflüssigkeit (1) zur Zubereitung der Suppe (1) zum Servieren (1)	
Übersetzung 3	Harissa-Paste (8/1)* Artisan-Brot (6/4)* eingelegte Zitronen (3/1)* Flockensalz (3/1)* gepresster Zucker (2/2)* koscheres Salz (2/1)* Tomaten-marmelade (1) Kreuzkümmel (1) gehobelte Zwiebeln (2) gelbe Zwiebel (1) natives Olivenöl (1) geröstetes Brot (2) Salz- und Pfefferflocken (1)	sprudelnd köcheln (3) mit 3EL Zitronensaft kochen (1) Beschreibung der Marmelade als dick (1)	Pfund (4) Tasse (2) Zusammen-stellung (3) Kochanleitung (1) Verwendung von Infinitiv und Sie-Form (1) Zeitangabe 55 Minuten bis 1 Stunde 15 Minuten (2)	GR
Übersetzung 5	Madeirawein (4/1)* Worcestershire-Soße (4) pariertes Rinderfilet (2/1)* Blätterteigblock (1/3)* englischer Senf (1/1)* Thymian (1) ungesalzene Butter (1) gemischte Champignons (1) geputzte Leber (1) Semmelbrösel wie diese (1)	auf hoher Flamme erhitzen (1) Rosmarinblätter abstreifen (1) Rosmarinblätter kochen (1) den Inhalt auf ein großes Brett kippen (2) rustikale Konsistenz (4) Teig mit Pfannkuchen auslegen (3) Blech erhitzen (1) bis Sie ein rötliches, saftiges Rindfleisch bekommen (1) die beiden Endstücke sind dann stärker	Stück Butter (1) die Lebern (2) Mehl zum Bestäuben (1) Teignaht (1) kühlschrankkalt (1) Schreibweise Soße vs. Sauce (2) plus (1) nach und nach die Brühe (1) Kante des Teigs (2)	JO

		gegart, aber manche Leute bevorzugen das normalerweise (2) Sauce köcheln lassen, bis sie glänzt und ziemlich dunkel ist (1) Sauce mit einem Streichholz anzünden (2) Sauce durch ein Sieb passieren (1) perfekt würzen (1)		
Übersetzung 7	getrocknete Kokosnüsse (1/2)* kandierte Kirschen (1) getrocknete Kokosnüsse vs. Kokosraspeln (5)	Kokosraspeln darüber streuen (1) braten, bis der Honig fast verschwunden ist (2) in einer Lage verteilen (1) Eier auflockern (1) Teig darübergießen (1)	Gas 4 (1) Dosen Ananasringe in Saft (à 425 g) (1) die größte Bratpfanne (3) Bratensaft (3) Backfolie (2) antihaftbeschichtete Pfanne (2) Bratpfanne (1) Ananaskuchen mit Deckel (1) Formulierung Ananaskuchen mit Deckel (4)	JO

* Bestimmte Zutaten wurden sowohl als unbekannt (erste Personenanzahl) bewertet als auch als unübliche Begriffe (zweite Personenanzahl) thematisiert.

In diesem Zusammenhang wurde hinsichtlich Übersetzung 1 von zwei Personen angemerkt, dass das Rezept verständlich und nicht so schwer nachzumachen ist. Dabei gab eine Person an, lediglich die *Poppadoms* extra nachschlagen zu müssen. Nichtsdestotrotz stießen viele Teilnehmer*innen bei gewissen Begrifflichkeiten beziehungsweise Ausdrücken auf Unverständnis. So war etwa der Untertitel vom Rezept *Mit roter Paprika und Spinat überbacken* für drei Personen verwirrend (siehe Tabelle 7). Hier wurde mit *garniert* eine verständlichere Alternative vorgeschlagen. Darüber hinaus wurde die Verwendung des Verbs *kochen* in den Schritten 3 (*Die Zwiebel […] kochen, bis sie weich und süß sind.*) und 7 (*Für den Paprikaaufstrich das Öl in einer Pfanne erhitzen […] 1 Minute lang kochen […].*) mehrfach thematisiert (siehe Tabelle 7). Da diese Schritte in einer Pfanne erfolgen und nach Meinung der

Teilnehmer*innen gebraten werden muss, wurden Korrekturen wie *Zwiebel glasig andünsten*, *Zwiebel dünsten* und *Paprikaaufstrich anbraten* vorgeschlagen. Wie Tabelle 7 zu entnehmen ist, wurden die Phrasen *die Samen herausschütteln* und *Suppe […] mit einem Klecks Paprika-Topping abschließen* von je drei Personen als nicht idiomatisch wahrgenommen. Im letzteren Fall wurden Verben wie *krönen*, *anrichten* oder *garnieren* als korrektere Varianten genannt. Eine Person empfand den Ausdruck *sanft köcheln lassen* als seltsam und korrigierte diesen mit *auf schwacher Hitze kochen*. Die am häufigsten genannte (fünf Personen) Antwort bei der Frage zu unüblichen Begriffen stellte die Phrase *bis der Spinat verwelkt ist* dar (siehe Tabelle 7). Diese wurde von einigen Teilnehmer*innen mit *bis der Spinat zusammenfällt* ausgebessert. Außerdem war für zwei Personen nicht ganz klar, was mit einer *trockenen Pfanne* gemeint ist (siehe Tabelle 7). Schließlich gab eine Person die Verwendung des femininen Artikels beim Nomen *Paprika* als unüblich an (siehe Tabelle 7). Dies lässt auf regionale Unterschiede im deutschsprachigen Raum schließen, da sowohl der maskuline als auch der feminine Artikel zulässig sind. Darüber hinaus wurde in den Kommentaren die inkonsistente Verwendung von Begriffen zur Beschreibung des Paprika-Toppings thematisiert. Mehrere Teilnehmer*innen waren von der Angabe *roter Pfeffer als Belag* in der Zutatenliste verwirrt (siehe Tabelle 7). Hierbei wurde von einer professionellen Translatorin die Vermutung geäußert, dass in Anbetracht der angegeben Zutaten und der Beschreibung der Zubereitung die *Paprika*, engl. *pepper*, scheinbar nicht korrekt übersetzt worden war. Gemäß Tabelle 7 war für drei Teilnehmer*innen die Verwendung von *Paprikaaufstrich* irritierend und wurde im Kontext des zuzubereitenden Gerichts als unlogisch und ungeeignet empfunden. Weiters war für eine Person nicht klar, ob bei der Zutat *2 gebratene Paprikaschoten aus dem Glas* eingelegte Paprika gemeint ist. Darüber hinaus wurde zweimal angemerkt, dass die Überschrift *Methode* für den Zubereitungsteil eher unüblich ist (siehe Tabelle 7).

In sprachlicher Hinsicht sind den Teilnehmer*innen ebenfalls gewisse Aspekte aufgefallen. So wurde in Schritt 3 für die Zwiebel im Singular an späterer Stelle fälschlicherweise der Plural verwendet. Außerdem geht aus Schritt 7 nicht klar hervor, dass alle Zutaten in einer Pfanne landen sollen. Eine Teilnehmerin merkte in diesem Kontext Folgendes an: „Würde ich mich exakt an den Text halten, würde ich das Öl in einer Pfanne erhitzen, dann die Koriandersamen getrennt davon in einem Mörser zerstoßen und zu den zerstoßenen Samen auch die Chiliflocken geben“. Bei der Zutatenliste wurde vorgeschlagen, die Angabe *4cm Stück Ingwer* zu *4 Zentimeter langes Stück Ingwer* zu ändern und bei *800ml Bio-Gemüse- oder Hühnerbrühe* den Bindestrich nach Bio wegzulassen. Für den Titel des Rezepts wurde ebenfalls

eine Alternative vorgeschlagen: *Süßkartoffelsuppe mit Kokosnuss und Kardamom*. Nichtsdestotrotz fanden neun Teilnehmer*innen (64%) den Titel ansprechend und würden das Rezept nachkochen.

Zu Übersetzung 2 wurde ebenso von einer Person angemerkt, dass das Rezept leicht verständlich und nachkochbar ist. Andere Teilnehmer*innen hingegen fanden gewisse Begriffe beziehungsweise Anweisungen unüblich und verwirrend. So gab gemäß Tabelle 7 von den neun zu den unüblichen Begriffen Befragten ein Drittel an, dass der Begriff *Toastie* im Deutschen eher nicht gebräuchlich ist, wobei in den Kommentaren *Toast* als geeignetere Variante vorgeschlagen wurde. Die Verwendung von *Sandwich* bei der Zubereitung der *Toasties* wurde von einer Person als unüblich und in diesem Kontext unpassend wahrgenommen. Darüber hinaus wurde von zwei Personen ein regionaler Unterschied zwischen Deutschland und Österreich thematisiert: *Blumenkohl* wurde zwar verstanden, aber als in Österreich nicht geläufig erachtet (siehe Tabelle 7). Ein Drittel der neun Teilnehmer*innen beurteilte den Ausdruck *schwere Bratpfanne* als ungewöhnlich. Zwei Personen führten aus, dass unter *schwach geheizter Backofen* eher ein defekter Backofen verstanden werden könnte. Außerdem wurden von je einer Person die Phrasen *Suppe durch ein Sieb passieren* und *Steinpilze […] braten, bis sie gebräunt sind* als nicht gebräuchlich empfunden (siehe Tabelle 7). Bei letzterer Phrase wurde in den Kommentaren mit *goldbraun braten* eine gängigere Alternative angeboten. In Anbetracht der Tatsache, dass in den Zutaten nur Salz und Pfeffer als Gewürze angegeben werden, verwirrte die Erwähnung von *etwas Gewürz* im Zubereitungsteil der Suppe eine Person. Die Anweisung, den Blumenkohl zum Pürieren in einen Mixer zu geben, wurde ebenfalls als ungewöhnlich erachtet (siehe Tabelle 7). In diesem Zusammenhang wurde in den Kommentaren angemerkt, dass Pürieren nur mit einem Stabmixer gut funktionieren kann.

Weitere Kommentare, die noch nicht behandelt wurden, waren sprachlicher, grammatikalischer, stilistischer und inhaltlicher Natur. So wurde geschrieben, dass das Rezept keine ganzen Sätze enthält, da keine Punkte gesetzt wurden. Außerdem wurden laut einer Person nicht immer die richtigen Pronomen in aufeinanderfolgenden Sätzen verwendet, wodurch das Leseverständnis des Rezepts erschwert wurde. Die Phrasen *bei Bedarf mehr Salz und Pfeffer hinzufügen* und *auf jede Schüssel einen Löffel Steinpilze geben* wurden, wie Tabelle 7 zu entnehmen ist, als nicht idiomatisch erachtet. In diesem Kontext wurden Korrekturen wie *bei Bedarf nachwürzen* oder *in jede Schüssel einen Löffel Steinpilze geben* empfohlen. Sieben von elf Teilnehmer*innen, die einen Kommentar hinterlassen haben, fiel der grammatikalische Fehler *köcheln, bis die Blumenkohl ist weich* auf (siehe Tabelle 7). An dieser

Stelle wurde sowohl der falsche Artikel als auch eine fehlerhafte Satzstellung verwendet. In inhaltlicher Hinsicht wurde kommentiert, dass die Schreibweise *Bayonner-Schinken* nicht korrekt ist und die Schreibweise *Bayonne-Schinken* verbreiteter ist. Ebenso seien die Begriffe *Garnitur* in den Zutaten und *Kochflüssigkeit* bei der Zubereitung der Suppe nicht korrekt beziehungsweise ungewöhnlich (siehe Tabelle 7). In diesen Fällen wären *Garnierung* beziehungsweise *Garnier* und *Sud* dem Kontext entsprechend passender. Die Anweisung *Entfernen Sie fast alle Blätter vom Blumenkohl* wurde gemäß Tabelle 7 von einer Person als irritierend angesehen. Hier sei unklar, welche Blätter nicht entfernt werden sollten. Die Erwähnung der genauen Anzahl der Brotscheiben bei der Zubereitung des Toasties sei darüber hinaus nicht nötig. Schließlich wurde stilistisch gesehen angemerkt, dass bei den Zubereitungsüberschriften, zum Beispiel *Zur Zubereitung der Suppe*, die Präposition *zur* nicht unbedingt nötig ist (siehe Tabelle 7). Bei der Überschrift *Zum Servieren* wurde mit dem Vorschlag *Anrichten* ebenso eine stilistische Korrektur empfohlen. Im Gegensatz zu Übersetzung 1 würden im Falle von Übersetzung 2 trotz einer deutlich besseren Bewertung hinsichtlich der Aussage zur Verständlichkeit und Umsetzung der Kochanleitung nur sechs Teilnehmer*innen (43%) das Rezept nachkochen. Daraus erschließt sich, dass mehr als die Hälfte (53%) der Teilnehmer*innen das Rezept nicht ansprechend fand. Diese Diskrepanz ergibt sich daraus, dass die Frage zur Bereitschaft zum Nachkochen des Rezepts lediglich mit Ja oder Nein zu beantworten war und die diesbezüglichen Antworten nicht in direktem Zusammenhang mit der Verständlichkeit und Umsetzung der Kochanleitung stehen. Darüber hinaus hinterließ eine Teilnehmerin folgenden Kommentar: „Obwohl die Übersetzung flüssiger klingt als die erste, wirkt sie etwas holprig und nicht so vertrauenswürdig. Gerade bei Rezepten sollte der Text möglichst fehlerfrei sein; je besser der Text, desto vertrauenswürdiger wirkt er und desto eher kocht man es nach. Man möchte schließlich kein Rezept ausprobieren, von dem man von Anfang an nicht sehr überzeugt ist."

Unter Berücksichtigung der Bewertung beider Übersetzungen (siehe Tabelle 6) kann der Rückschluss gezogen werden, dass in Set 1 Google Translate mit Übersetzung 2 nach dem reinen Mittelwert aller Antworten besser abgeschnitten hat. Nichtsdestotrotz spricht die geringere Anzahl der Teilnehmer*innen, die dieses Rezept nachkochen würden, doch insgesamt für Übersetzung 1. Die Übersetzung von DeepL wurde bei den Aussagen Die Übersetzung ist akzeptabel., Die richtigen Fachbegriffe werden verwendet. und Die Übersetzung enthält grammatikalische Fehler. besser bewertet. Hierbei ist anzumerken, dass Übersetzung 1 zwar bei der Aussage zur Verwendung richtiger Fachbegriffe ein niedrigeres

arithmetisches Mittel aufweist, die Werte 1 und 2 auf der Skala jedoch weniger oft angekreuzt wurden als bei Übersetzung 2 und die Bewertung dementsprechend weniger um den Mittelwert streut (siehe Tabelle 6).

Die Übersetzungen in Set 2 zu den Toastrezepten wurden im Allgemeinen schlechter bewertet als jene in Set 1. Wie in Tabelle 8 dargestellt, wurden beide Übersetzungen im Schnitt als eher akzeptabel erachtet. Die Rezepte und die darin vorkommenden Anweisungen sind laut den Umfrageteilnehmer*innen in der Praxis umsetzbar, wobei Übersetzung 4 dazu bessere Ergebnisse erzielte. Die Zutaten hingegen waren den Umfrageteilnehmer*innen eher nicht bekannt. Beide Übersetzungen erzielten bei der Bewertung diesbezüglicher Aussagen einen Mittelwert, der geringer als 3 ist (siehe Tabelle 8). Sowohl bei Übersetzung 3 als auch bei Übersetzung 4 wurde auf der Skala der Wert 2 von je sechs Personen und somit am häufigsten gewählt (siehe Tabelle 8). Fünf Personen gaben im Fall von Übersetzung 3 und vier Personen im Fall von Übersetzung 4 die Werte 4 beziehungsweise 5 an. Von den zehn Teilnehmer*innen, die die Freitextfrage zu unbekannten Zutaten bei Übersetzung 3 zu beantworten hatten, gaben gemäß Tabelle 7 acht die *Harissa-Paste* an. Davon bereiten drei Personen regelmäßig asiatische Gerichte zu. Sechs Personen hatten zuvor noch nichts von *Artisan-Brot* gehört. Je ein Drittel gab *eingelegte Zitronen* und *Flockensalz* als unbekannte Zutaten an; je zwei Personen erwähnten den *gepressten Zucker* und das *koschere Salz*. *Tomatenmarmelade* und *Kreuzkümmel* wurden je einmal aufgelistet (siehe Tabelle 7). Bei Übersetzung 4 wurden insgesamt zwölf Teilnehmer*innen nach unbekannten Zutaten befragt. Davon führten 100% *Pedro Ximenez* an (siehe Tabelle 7). Die Hälfte sprach den *Frisée-Salat* und je fünf Personen den *Sherry-Essig* und *Manchego-Käse* an. Lediglich zwei Personen nannten die *Chorizo*. Wie Tabelle 8 zu entnehmen ist, wurde die Verständlichkeit und Nachvollziehbarkeit der Mengenangaben beziehungsweise Maßeinheiten eher negativ mit arithmetischen Mitteln von 2,79 und 2,36 beurteilt. In diesem Kontext wählten bei Übersetzung 3 insgesamt acht und bei Übersetzung 4 neun Personen die Werte 1 und 2 auf der Skala aus, wobei der Wert 2 in beiden Fällen am häufigsten angekreuzt wurde. So fanden bei Übersetzung 3, wie aus Tabelle 7 ersichtlich, vier Personen die Maßeinheit *Pfund* und zwei Personen die Mengenangabe *Tasse* im deutschsprachigen Raum ungeeignet. Die Verwendung von *Unzen* in Übersetzung 4 wurde von sechs Teilnehmer*innen thematisiert; die Mengenangabe *Tasse* wurde einmal angesprochen (siehe Tabelle 7). In diesem Zusammenhang wurde in den Kommentaren als Korrekturmaßnahme die Umrechnung von Pfund in Kilogramm und von Unzen in Gramm vorgeschlagen; auch die Mengenangabe *Tasse* sollte an den deutschsprachigen Raum angepasst werden. Gemäß

Tabelle 8 fiel die Bewertung der Aussagen zur Verwendung der richtigen Fachbegriffe und zu grammatikalischen Fehlern positiver aus, während die Teilnehmer*innen mehr sprachliche beziehungsweise idiomatische Fehler und unübliche Begriffe feststellten.

Tabelle 8: Bewertung von Set 2 mit DeepL und Google Translate (GT) nach dem Durchschnittswert und Modalwert pro zu bewertender Aussage durch 14 Personen

Aussage	**Übersetzung 3 (GT)**		**Übersetzung 4 (DeepL)**	
	Ø (σ)	**Modal-wert**	**Ø (σ)**	**Modal-wert**
Die Übersetzung ist akzeptabel.	3,57 (1,18)	4	3,43 (0,98)	4
Mit diesem Rezept kann man dieses Gericht zubereiten.	3,64 (1,11)	4 & 5	4 (0,76)	4
Die Kochanleitung ist verständlich und könnte leicht umgesetzt werden.	3,36 (0,81)	3	3,79 (0,67)	4
Alle Zutaten sind Ihnen bekannt.	2,93 (1,49)	2	2,86 (1,19)	2
Die verwendeten Mengenangaben bzw. Maßeinheiten sind verständlich und nachvollziehbar.	2,79 (1,26)	2	2,36 (1,11)	2
Die richtigen Fachbegriffe werden verwendet.	3,5 (1,12)	4	3,21 (1,08)	4
Die Übersetzung enthält sprachliche bzw. idiomatische Fehler.	3,21 (1,21)	4	3,5 (1,12)	4
Die Übersetzung enthält grammatikalische Fehler.	1,86 (0,99)	1	2,21 (1,08)	1
Bestimmte Begriffe sind für Sie im Deutschen nicht üblich.	3,57 (0,98)	4	4 (0,93)	4

So wurde gemäß Tabelle 7 bei Übersetzung 3 der Ausdruck *sprudelnd köcheln* von drei Personen als unpassend wahrgenommen, wobei eine Person als Alternative *sieden* vorgeschlagen hat. Von den 13 zu unüblichen Begriffen Befragten nannten drei die Überschrift *Zusammenstellung* in den Zutaten (siehe Tabelle 7). Hierzu wurde in den Kommentaren erläutert, dass *Zusammensetzung* in diesem Fall adäquater wäre. Die meisten Angaben zu unüblichen Begriffen bezogen sich auf Zutaten. Dementsprechend nannten, wie in Tabelle 7 dargestellt, vier Personen *Artisan-Brot* und zwei Personen merkten an, nichts mit *gepresstem braunen Zucker* anfangen zu können. Ebenso listeten zwei Personen *gehobelte Zwiebeln* als nicht gängigen Begriff auf. Je einmal wurden die Ausdrücke *Harissa-Paste*, *koscheres Salz*, *gehobelte Zwiebeln*, *Flockensalz*, *gelbe Zwiebel*, *eingelegte Zitronen* und *natives Olivenöl* als

im deutschsprachigen Raum nicht gebräuchlich bewertet (siehe Tabelle 7). Beim Zubereitungsteil wurde von einer Person thematisiert, dass die Überschrift *Kochanleitung* nicht oft in Rezepten zu sehen ist; Varianten wie *Zubereitungsanleitung* seien gängiger. Da die Zutaten keine anderen Flüssigkeiten als drei Esslöffel Zitronensaft enthalten, kommentierte eine Person, dass es unüblich sei, etwas mit so wenig Flüssigkeit zu *kochen* und vermutete, dass damit eigentlich *andünsten* gemeint ist (siehe Tabelle 7). Zudem wurde die Beschreibung der Marmelade mit der Eigenschaft *dick* als befremdlich befunden. Auch die Bezeichnung *geröstetes Brot* wurde von zwei Personen als ungewöhnlich erachtet; eine Person gab schließlich noch an, den Ausdruck *Salz- und Pfefferflocken* nicht angemessen zu finden (siehe Tabelle 7).

In weiterer Folge wurden in den zusätzlichen Kommentaren verschiedene Meinungen zur Übersetzung geteilt. So schrieb eine Person, dass das Rezept leicht nachkochbar und ansprechend ist. Im Gegensatz dazu wurde angemerkt, dass die Anleitung viel Text am Stück enthält beziehungsweise in dieser Übersetzung Grammatikfehler und unübliche Begrifflichkeiten besonders aufgefallen sind. In diesem Kontext schrieb eine weitere Person, dass die Begriffe nicht zu 100% richtig verwendet wurden, dies aber das Verständnis des Rezepts nicht beeinträchtigt. Laut einer Person ist die Zubereitung der Marmelade relativ einfach, jedoch wird für den Rest der Zubereitung viel vorausgesetzt, zum Beispiel ein pochiertes Ei. Hier wurde auch von einer anderen Person kritisiert, dass das Einlegen von Zitronen gar nicht beschrieben wurde. Darüber hinaus wurden, wie in Tabelle 7 aufgelistet, im Zubereitungsteil sowohl der Infinitiv als auch die Sie-Form verwendet, was die Leserlichkeit deutlich erschwerte. In diesem Kontext fügte eine Teilnehmer*in an, dass gerade ein Rezept verständlich sein und möglichst wenig Zeit für den Lese- und Verstehensprozess verbracht werden sollte. In den weiteren Antworten fanden sich schließlich ein paar detailliertere Anmerkungen. So wurde die Beschreibung des Artisan-Brots mit *herzhaft* als überflüssig erachtet. Die Zeitangabe *55 Minuten bis 1 Stunde 15 Minuten* wurde mehrfach als ungewöhnlich und holprig angesehen (siehe Tabelle 7); eine Formulierung wie *55 bis 75 Minuten* wurde als adäquatere Lösung vorgeschlagen. Beim ersten Zubereitungsschritt sollen alle Zutaten in einen Topf gegeben werden. Eine Teilnehmerin merkte hierzu an, dass man klarstellen sollte, dass damit nur die Zutaten für die Marmelade gemeint sind, dies aber wahrscheinlich im Originaltext ebenso geschrieben wurde. Auf die Frage, ob die Teilnehmer*innen das Rezept nachkochen würden und den Titel ansprechend finden, antworteten zehn Personen (71%) mit Ja und vier Personen (29%) mit Nein. In diesem Kontext schrieb eine Person, dass der zweite Teil zur Zusammenstellung

unverständlich war und sie vom Nachkochen abhalten würde. Außerdem wurde der Titel des Rezepts als sehr lang empfunden, wobei zwei Personen anmerkten, dass man zwischen *Feta* und *eingelegter Zitrone* statt dem Komma ein *und* einfügen sollte.

Bei Übersetzung 4 beantworteten insgesamt 13 Teilnehmer*innen die Frage zu unüblichen Begriffen und acht Teilnehmer*innen hinterließen weitere Kommentare. Im Allgemeinen wurde angemerkt, dass viele Zutaten nicht so leicht erhältlich beziehungsweise ziemlich spezifisch sind und außergewöhnlich wirken. Wie bereits oben beschrieben, wurde in diesem Zusammenhang etwa die Zutat *Pedro Ximenez* als Beispiel angeführt. Darüber hinaus wurde mehrfach angesprochen, dass die Vorgänge im übersetzten Rezept nicht richtig beschrieben beziehungsweise einige für die Art von Zutaten unübliche Begriffe verwendet wurden. Hierbei merkte eine Person zusätzlich an, dass die Übersetzung zwar ein paar unglückliche Formulierungen enthält, diese aber trotzdem verständlich sind. Auch im Falle von Übersetzung 4 wurde von einer Person bei den unüblichen Begriffen die Sprachvarietät im Deutschen angesprochen, indem *Tomate* als Bezeichnung aus dem deutschen Deutsch genannt wurde (siehe Tabelle 7). Die Zutat *gelbe Zwiebeln* wurde wie auch bei Übersetzung 3 von einer Person als im deutschsprachigen Raum nicht gebräuchlich angesehen. Die Formulierung *1 Avocado […] plus Salz und Zitronensaft* wurde ebenso von einer Person als ungewöhnlich erachtet (siehe Tabelle 7). Statt *plus* wurden in diesem Fall *dazu* und *mit* als geeignetere Alternativen vorgeschlagen. Schließlich hob gemäß Tabelle 7 eine Person die Anweisung in den Zutaten, dass der Salat in Essig *geschwenkt* werden soll, als unüblich hervor.

Wie bereits oben beschrieben, stellten die Beschreibungen der Anweisungen in dieser Übersetzung laut den Teilnehmer*innen die größte Problematik dar. So merkte eine Person an, dass die Überschrift des Zubereitungsteils *Anweisungen zum Garen* nicht korrekt ist, da im Rezept gar nicht gegart wird (siehe Tabelle 7). Die Anweisung *Gewürze überprüfen und bei Bedarf anpassen* wurde, wie aus Tabelle 7 ersichtlich, von fünf der zu unüblichen Begriffen befragten 13 Teilnehmer*innen als nicht gebräuchlich empfunden. Vielmehr wäre eine einfachere Formulierung wie *abschmecken* üblicher. Vier Personen sahen die Phrase *um alle gebräunten Stücke aus der Pfanne zu kratzen* als befremdlich und dementsprechend verwirrend an (siehe Tabelle 7). Ebenso wurde die Anweisung *Tomaten […] kochen, bis sie Blasen werfen* von drei Personen als ungewöhnlich erachtet, wobei eine Person zusätzlich anmerkte, dass das Verb *kochen* in diesem Fall auch nicht korrekt ist und damit wahrscheinlich *andünsten* gemeint ist. Ebenso drei Personen schrieben, dass die Formulierung *mit Salz und einem Spritzer Zitronensaft bestreuen* inhaltlich nicht korrekt ist (siehe Tabelle 7). Man kann zwar mit Salz

bestreuen, aber mit Zitronensaft nicht. Hierbei wären Ausdrücke wie *mit Zitronensaft bespritzen* adäquater. Generell wurde auch die Verwendung der Mengenangabe *Spritzer* in Verbindung mit diversen Flüssigkeiten im Rezept als unüblich erachtet (siehe Tabelle 7); *Schuss* wird in diesem Kontext als gängigere Variante vorgeschlagen. Außerdem wurden die Vorgangsbeschreibungen *köcheln […], bis die Flüssigkeit etwas eingekocht ist*, *Käse großzügig darüber raspeln* und *Zwiebeln hinzufügen* gemäß Tabelle 7 von je einer Person als nicht gebräuchlich hervorgehoben. Zu letzterer Formulierung wurde *hinzugeben* als Alternative genannt. Schließlich listete je eine Person die *Ränder* von der Chorizo und die *Schnittfläche* der Tomate als nicht gebräuchliche Begriffe auf (siehe Tabelle 7). In weiterer Folge merkte eine Person hinsichtlich der Verwendung von *Toast* im dritten Zubereitungsschritt an, dass sie mit Toast grundsätzlich Weißbrot verbindet, ihr aber durchaus klar ist, dass auch dunkles Brot getoastet werden kann. Eine Person schrieb letztlich in den Kommentaren, dass der Titel gut klingt und zum Nachkochen einlädt. Dies spiegelt auch die Beantwortung der diesbezüglichen Frage wider: Zwölf Teilnehmer*innen (86%) würden das Rezept nachkochen beziehungsweise fanden den Titel ansprechend, wobei lediglich zwei Personen (14%) die Frage mit Nein beantworteten. Die zuvor genannte Person führte im Kommentar weiter aus, dass der spätere Text ab der Zutatenliste an der Qualität der Übersetzung zweifeln ließ.

Unter Hinzunahme der in Tabelle 8 dargestellten Bewertung der Übersetzungen 3 und 4 kann festgestellt werden, dass Google Translate auch in diesem Set bessere Ergebnisse erzielte als DeepL. Hierbei ist jedoch anzumerken, dass Google Translate nicht so deutlich wie in Set 1 der Favorit war. So schnitt DeepL etwa bei den ersten vier Aussagen besser ab. Zwar weist Übersetzung 4 bei der Bewertung der Aussagen zu Akzeptabilität und Bekanntheitsgrad der Zutaten ein niedrigeres arithmetisches Mittel auf, die Werte 1 und 2 auf der Skala wurden jedoch weniger häufig ausgewählt als bei Übersetzung 3 und die Standardabweichung ist dementsprechend niedriger. Während Übersetzung 4 bei den allgemeineren Aussagen bessere Ergebnisse erzielte, fiel die Bewertung von Übersetzung 3 bei den spezifischeren Aussagen zu sprachlichen, idiomatischen beziehungsweise grammatikalischen Fehlern, zur Verwendung korrekter Fachbegriffe und unüblicher Begriffe besser aus (siehe Tabelle 8).

Set 3 mit den Rezepten für Beef Wellington lieferte im Gesamteindruck ebenfalls eher positive Ergebnisse. Wie aus Tabelle 9 ersichtlich, wurden beide Übersetzungen von den Teilnehmer*innen als akzeptabel angesehen, wobei in beiden Fällen der Wert 1 gar nicht angekreuzt wurde; der Wert 2 wurde bei Übersetzung 5 lediglich zweimal ausgewählt. Aus Sicht der Umfrageteilnehmer*innen kann man die Gerichte mit beiden Rezepten zubereiten und

die Kochanleitung gilt als verständlich und umsetzbar. Gemäß Tabelle 9 ergab die Bewertung beider Aussagen Mittelwerte größer als 3. Im Vergleich zu den anderen Sets schnitt das dritte Set jedoch hier aufgrund der höheren Komplexität der Rezepte schlechter ab. Die enthaltenen Zutaten waren den Teilnehmer*innen verhältnismäßig gut bekannt, wobei bei Übersetzung 5 der Wert 5 von vier Personen und bei Übersetzung 6 von sieben Personen angekreuzt wurde. Dementsprechend wurden bei Übersetzung 5 zehn Teilnehmer*innen und bei Übersetzung 6 sieben Teilnehmer*innen zu unbekannten Zutaten befragt. Bei Übersetzung 5 führten, wie in Tabelle 7 aufgelistet, je vier Personen *Madeirawein* und *Worcestershire-Soße* als ihnen nicht vertraute Zutaten an. Zwei Personen kannten den Ausdruck *pariert* als Zusatz zu Rinderfilet nicht und gaben dies entsprechend als unbekannte Zutat an. Weiters listete je eine Person folgende Zutaten auf: *Blätterteigblock*, *englischer Senf*, *Thymian* (siehe Tabelle 7). Bei Übersetzung 6 gaben je zwei von den sieben dazu befragten Teilnehmer*innen *Schalotten* und *Rindfleischabschnitte* als nicht bekannte Zutaten an (siehe Tabelle 7). Zu den *Rindfleischabschnitten* merkte eine Person in diesem Zusammenhang an, dass sie nicht oft mit Fleisch kocht und dies wohl dieser Tatsache geschuldet ist. Ebenfalls zwei Personen beantworteten die Frage mit „keine“. Schließlich führte eine Person, die lieber asiatisch beziehungsweise japanisch kocht, die Gewürze *Lorbeerblatt* und *Thymianzweig* an (siehe Tabelle 7). Die Mengenangaben beziehungsweise Maßeinheiten wurden in diesem Set mit aus der Bewertung ermittelten Mittelwerten größer als 4 (siehe Tabelle 9) als verständlich und nachvollziehbar beurteilt. Bei der Frage zu unüblichen Begriffen merkte eine Person an, dass die Mengenangabe *Stück* im Kontext von Butter in Übersetzung 5 zu ungenau und nicht ganz nachvollziehbar ist (siehe Tabelle 7). Außerdem kommentierte eine Teilnehmerin, dass die Angabe zur Backofentemperatur *Gas 6* in Übersetzung 6 nicht so gebräuchlich und verständlich ist wie etwa *Gasherdstufe 6*. Die Mengenangabe *1 x 750ml Flasche Rotwein* in Übersetzung 6 wurde von zwei Personen thematisiert. Hierzu wurde angemerkt, dass eine solche Angabe eher unüblich ist und vereinfachte Alternativen wie *1 Flasche Rotwein (750ml)* oder *750ml Rotwein* adäquater wären.

Tabelle 9: Bewertung von Set 3 mit DeepL und Google Translate (GT) nach dem Durchschnittswert und Modalwert pro zu bewertender Aussage durch 14 Personen

Aussage	**Übersetzung 5 (GT)**		**Übersetzung 6 (DeepL)**	
	Ø (σ)	**Modal-wert**	**Ø (σ)**	**Modal-wert**
Die Übersetzung ist akzeptabel.	3,5 (0,91)	3 & 4	3,93 (0,7)	4

Mit diesem Rezept kann man dieses Gericht zubereiten.	3,79 (0,94)	3	3,64 (1,11)	4 & 5
Die Kochanleitung ist verständlich und könnte leicht umgesetzt werden.	3,5 (0,98)	4	3,43 (1,12)	2 & 4
Alle Zutaten sind Ihnen bekannt.	3,57 (1,29)	4	4,21 (1,01)	5
Die verwendeten Mengenangaben bzw. Maßeinheiten sind verständlich und nachvollziehbar.	4,14 (1,25)	5	4,21 (1,08)	5
Die richtigen Fachbegriffe werden verwendet.	3,57 (1,24)	4 & 5	3,5 (1,12)	3
Die Übersetzung enthält sprachliche bzw. idiomatische Fehler.	3,29 (1,28)	4	2,5 (1,3)	1 & 3
Die Übersetzung enthält grammatikalische Fehler.	2,79 (1,37)	2	1,71 (0,8)	1
Bestimmte Begriffe sind für Sie im Deutschen nicht üblich.	3,29 (1,39)	4	2,36 (1,34)	1

Bei der Bewertung der Aussage zu Grammatikfehlern wurden, wie aus Tabelle 9 ersichtlich, Mittelwerte von 2,79 und 1,71 berechnet. Daraus ergibt sich, dass in beiden Rezepten eher weniger bis keine Fehler dieser Natur erkannt wurden. In diesem Kontext wurde von zwei Personen die Verwendung von *die Lebern* im Zubereitungsteil von Übersetzung 5 als grammatikalischer Fehler angestrichen (siehe Tabelle 7). Der Plural sei in diesem Fall durch die Grammangabe in der Zutatenliste nicht passend. Ebenso ist aus Tabelle 9 erschließbar, dass in beiden Übersetzungen eher richtige Fachbegriffe verwendet wurden. Die Bewertung zu sprachlichen beziehungsweise idiomatischen Fehlern und der Verwendung unüblicher Begriffe fiel bei den Übersetzungen unterschiedlich aus. Während in Übersetzung 5 wohl mehr Fehler dieser Natur ermittelt werden konnten, bewerteten die Teilnehmer*innen Übersetzung 6 in dieser Hinsicht positiver (siehe Tabelle 9).

Hinsichtlich Übersetzung 5 gaben insgesamt elf Teilnehmer*innen unübliche Begriffe an. Ein großer Teil der Antworten betraf in diesem Zusammenhang die Zutaten. So nannten, wie Tabelle 7 zu entnehmen ist, drei Personen *Blätterteigblock* als unüblichen Ausdruck. Eine Teilnehmerin gab hier den *Madeirawein* an. Eine weitere Person erwähnte den *englischen Senf* und führte in weiterer Folge aus, dass die Bezeichnung *ungesalzene Butter* im deutschsprachigen Raum nicht gebräuchlich ist, da Butter in diesem Gebiet immer ungesalzen ist. Dieser Person war ebenso nicht klar, was mit der Zutat *gemischte Champignons* gemeint ist. Sie fragte sich, ob es sich dabei um eine Mischung aus weißen und braunen Champignons

handeln sollte. Darüber hinaus wurden die Beschreibungen *geputzt* für Leber und *pariert* für Rinderfilet in der Zutatenliste von je einer Person als ungewöhnlich bewertet (siehe Tabelle 7). Der Zusatz *zum Bestäuben* bei der Angabe von Mehl in der Zutatenliste wurde ebenfalls von einer Person als nicht gebräuchlich erachtet; *bestreuen* sei laut einem Kommentar durchaus üblicher. Der Rest der Antworten zu unüblichen Begriffen betraf Ausdrücke beziehungsweise Formulierungen im Zubereitungsteil. In diesem Zusammenhang führte eine Person *Teignaht* in Schritt 10 als ungewöhnlich an (siehe Tabelle 7); die Beschreibung *kühlschrankkalt* soll aus Sicht dieser Person in deutschen Rezepten ebenso wenig üblich sein. Die Angabe, die Pfanne *auf hoher Flamme* zu erhitzen, schien einer Person nicht passend. Im Kontext der Rosmarinblätter gab gemäß Tabelle 7 je eine Person an, dass die Verwendung der Verben *abstreifen* und *kochen* nicht adäquat ist. In ersterem Fall wurde in einem Kommentar für ein besseres Verständnis der Zusatz *von den Zweigen* vorgeschlagen. In letzterem Fall wurde vermutet, dass die Rosmarinblätter *gedünstet* werden sollen. Zwei Personen waren mit der Formulierung *den Inhalt auf ein großes Brett kippen* nicht zufrieden, wobei eine Person die Verwendung von *kippen* störte, während die andere Person anmerkte, dass dieser Vorgang nicht passend scheint und der Inhalt wahrscheinlich eher auf einen Teller geleert werden soll. Die Beschreibung *rustikal* für die zu erreichende Konsistenz der Pilzmasse in Schritt 6 wurde von insgesamt vier Personen als in deutschsprachigen Rezepten ungewöhnlich und nicht verständlich beurteilt (siehe Tabelle 7). Im selben Schritt wurde der Teil mit dem Auslegen des Teigs mit Pfannkuchen als unüblicher Arbeitsablauf angegeben. Diesen Vorgang bezeichneten drei Personen in den weiteren Kommentaren als nicht verständlich beziehungsweise ungewöhnlich (siehe Tabelle 7); eine dieser Personen fragte sich, ob ein Pfannkuchen den Saft tatsächlich problemlos aufsaugen kann; eine andere meinte, dass der Satz falsch übersetzt wirkt und nicht viel Sinn ergibt. Auch das Erhitzen des Blechs auf dem Herd in Schritt 11 wurde von einer Person als unüblich angesehen. In diesem Zusammenhang wurde in den Kommentaren ausgeführt, dass das Blech eigentlich im Ofen warm gemacht werden sollte. Außerdem beurteilte, wie in Tabelle 7 aufgelistet, eine Person die Verwendung des Verbs *bekommen* in der Phrase *bis Sie ein rötliches, saftiges Rindfleisch bekommen* als nicht gebräuchlich beziehungsweise nicht idiomatisch. Die Anmerkung *die beiden Endstücke sind dann stärker gegart, aber manche Leute bevorzugen das normalerweise* wurde in zweierlei Hinsicht thematisiert. Einerseits merkte eine Person an, dass *gegart* nicht diesem Kontext entspricht und etwa die Variante *gebräunt* korrekter wäre. Andererseits wurde der zweite Teil der Anmerkung als ungewöhnlich und für ein Rezept unpassend befunden (siehe Tabelle 7). Die Formulierung der Anweisung, die Sauce köcheln zu lassen, *bis sie glänzt und ziemlich dunkel ist* wurde ebenso als nicht

gebräuchlich erachtet. Darüber hinaus wurde die Anweisung in Schritt 13, die Sauce mit einem Streichholz anzuzünden, von zwei Personen als merkwürdig angesehen. Auch die Formulierung *durch ein Sieb passieren* wurde unter den unüblichen Begriffen aufgelistet (siehe Tabelle 7).

Schließlich haben zu Übersetzung 5 neun Teilnehmer*innen weitere Kommentare hinterlassen. Darin wurde von einer Person geschrieben, dass das Rezept im Allgemeinen ziemlich lang und ausführlich ist. Eine weitere Person führte in diesem Zusammenhang aus, dass die Übersetzung zum Teil falsch ist und gewisse Inhalte seltsam klingen, wodurch sich die Befolgung eines so langen Rezepts durchaus schwierig gestaltet. Außerdem wurden gewisse sprachliche und inhaltliche Aspekte kommentiert. So wurde, wie in Tabelle 7 dargestellt, etwa die Schreibweise von *Soße* thematisiert. Während eine Person anmerkte, dass dieses Wort in der Zutatenliste in Blockbuchstaben fälschlicherweise mit *ß* geschrieben wurde, schrieb eine andere Person, dass die Schreibweise *Sauce* üblicher wäre. Wie auch in anderen Sets, wurde hier von einer Person angemerkt, dass die Verwendung von *plus* als Zusatz in der Zutatenliste eher nicht gebräuchlich ist und Alternativen wie *dazu* oder *mit* gängiger sind. Eine Person fand die Anweisung *perfekt würzen* ohne konkrete Vorstellung nicht nachvollziehbar (siehe Tabelle 7). Weiters hob eine Person hervor, dass bei der Zubereitung der Sauce in Schritt 13 im Teil *nach und nach die Brühe* ein Verb ausgelassen wurde (siehe Tabelle 7); an dieser Stelle wurde das Verb *hinzufügen* vorgeschlagen. Inhaltlich wurde angemerkt, dass man unter Umständen nicht wissen könnte, ob der zu verwendende Wein rot oder weiß ist. In weiterer Folge erkannte eine Teilnehmerin, dass mit den *gemischten Champignons* in der Zutatenliste eigentlich eine *Pilzmischung* gemeint ist. Außerdem ist einer anderen Person aufgefallen, dass in der Zubereitung *Meersalz* verwendet wird, dies aber nicht in den Zutaten vorkommt. Diese Person hob ebenso den Ausdruck *Semmelbrösel wie diese* (siehe Tabelle 7) in der zusätzlichen Anmerkung von Schritt 6 hervor und fragte sich, welche Semmelbrösel damit gemeint sein sollen. Schließlich merkte eine Person an, dass der Gargrad des Filets im Rezept fehlt. Generell stellte laut den Umfrageteilnehmer*innen die Verständlichkeit der Anweisungen aufgrund der Komplexität des Rezepts eine große Herausforderung dar. In diesem Zusammenhang beschrieb eine Person die Schritte 8 bis 11 als nicht verständlich und zu kompliziert. Gemäß Tabelle 7 kommentierten zwei weitere Personen hierzu, dass die Verwendung von *Kante* in Schritt 8 nicht passend ist und Alternativen wie *Rand* angemessener wären. Darüber hinaus sollte laut einer Person in Schritt 4 ergänzt werden, dass der Rosmarin in die Pilzmasse hinzugefügt werden soll. In Schritt 5 hingegen war nicht klar, ob *Leber* und *Worcestershire-Sauce* zu der Pilzmasse

oder in eine eigene Pfanne kommen. Abschließend gaben sechs der vierzehn Teilnehmer*innen (43%) an, dass sie das Rezept nachkochen würden.

Bei Übersetzung 6 wurden acht Teilnehmer*innen zu unüblichen Begriffen befragt. Davon nannten gemäß Tabelle 7 zwei Personen den Titel *Rindfleisch-Wellington* und meinten, dass die englische Bezeichnung *Beef Wellington* im deutschsprachigen Raum gebräuchlicher ist. Außerdem führte eine Person die Zutat *Waldpilze* und zwei Personen *Rindfleischabschnitte* als ungewöhnliche Begriffe an (siehe Tabelle 7). Wie auch im Fall von Übersetzung 1, wurde hierbei von einer Person angemerkt, dass die Überschrift *Methode* für den Zubereitungsteil eher unüblich ist und Begriffe wie *Zubereitung* passender wären. Zwei Personen fanden darüber hinaus, wie Tabelle 7 zu entnehmen ist, die Formulierung *dreifache Lage Frischhaltefolie* ungewöhnlich. Eine weitere Person hob die Formulierung *ein Blatt Frischhaltefolie* zur Beschreibung eines Stücks der Folie als nicht gängigen Ausdruck hervor. Eine Person listete die Verwendung des Begriffs *Duxelle* im Deutschen als unüblich auf (Tabelle 7), wobei eine andere Teilnehmerin dazu in den weiteren Kommentaren anmerkte, dass diese Schreibweise im Deutschen falsch ist und durch *Duxelles* ersetzt werden sollte. In weiterer Folge merkte eine Person an, dass die Verwendung von *Teig* im Zubereitungsteil nicht passend ist, da damit der Blätterteig gemeint ist. Die Formulierung *unter häufigem Rühren* wurde von zwei Personen als unüblich eingestuft (siehe Tabelle 7). Bei Schritt 9 zur Zubereitung der Sauce wurde zunächst *Essig dazugießen* als nicht gebräuchlicher Ausdruck thematisiert (siehe Tabelle 7); weiters wurden von je einer Person die Formulierungen *sprudelnd kochen lassen* und *bis er fast trocken ist* als in deutschsprachigen Rezepten nicht gängig befunden. Schließlich beantwortete eine Person die Frage zu unüblichen Begriffen mit „keine“.

Neun Teilnehmer*innen kommentierten noch gewisse bei Übersetzung 6 aufgefallene Aspekte. So schrieb eine Person, dass die Übersetzung sehr gut ist. Eine weitere Person merkte in diesem Zusammenhang an, dass die Übersetzung trotz mancher „unglücklicher Formulierungen“ sehr gut verständlich ist. Im Gegensatz dazu kommentierte eine Teilnehmerin, dass das Rezept sehr lang und verwirrend ist; dazu schrieb ein weiterer Teilnehmer, dass die Übersetzung etwas künstlich und das Rezept lang ist, wodurch Ungenauigkeiten schwerer auffallen. In stilistischer Hinsicht wurde von einer Person angemerkt, dass die Anweisungen abwechselnd im Infinitiv und der Sie-Form geschrieben wurden. Dieselbe Person führte weiters aus, dass die Schritte 4 bis 7 zur Beschreibung des Einwickelns des Rinderfilets zu kompliziert sind. Zwei andere Personen wiederum fanden die Zubereitungsschritte zur Rotweinsauce verwirrend. In diesem Zusammenhang war einerseits nicht klar, ob man dafür tatsächlich die Rinderfilets

benötigt, und andererseits, was nach dem Sieben der Sauce mit den Fleischstücken aus der Pfanne passieren sollte. Auch terminologische Konsistenz wurde bei Übersetzung 6 thematisiert. So schrieb eine Person, dass in diesem Rezept stark auffällt, wie wichtig terminologische Konsistenz ist, da sowohl Rinderfilets als auch Rindfleischabschnitte verwendet wurden. In der Übersetzung wurden die Begriffe *Rinderfilet* und *Rindfleisch* durcheinandergemischt (siehe Tabelle 7), wodurch nicht ganz klar herauszulesen war, welches Fleisch an welcher Stelle zum Einsatz kommen sollte. Eine Teilnehmerin fand die abwechselnde Verwendung der Schreibweisen *Soße* und *Sauce* in der Übersetzung irritierend und setzte sich für die einheitliche Verwendung einer dieser Schreibweisen ein (siehe Tabelle 7). Im Kontext der unbekannten Zutaten merkte eine Person an, dass die Formulierung der Anmerkung zu Rindfleischabschnitten in der Zutatenliste merkwürdig war, man sich aber die Bedeutung beziehungsweise den Sinn dahinter denken kann. Weiters schlug eine Teilnehmerin vor, anfängerfreundlichere Alternativen für das *Musselin-Tuch* wie etwa *Käse- oder Baumwolltücher* anzugeben. Schließlich merkte dieselbe Teilnehmerin an, dass die Verwendung der Pluralform *Wellingtons* nicht gängig beziehungsweise korrekt ist (siehe Tabelle 7). Die Hälfte der Umfrageteilnehmer*innen, sprich sieben Personen, würden dieses Rezept nachkochen beziehungsweise finden den Titel ansprechend.

Aus der in Tabelle 9 dargestellten Bewertung der Übersetzungen in Set 3 ist zu schließen, dass Übersetzung 6 durch DeepL im Allgemeinen bessere Ergebnisse erzielte. Übersetzung 5 durch Google Translate schnitt lediglich bei der Bewertung der zweiten und dritten Aussage etwas besser als Übersetzung 6 ab. Hervorzuheben ist dabei, dass Übersetzung 6 hinsichtlich Bekanntheit der Zutaten, sprachlichen beziehungsweise idiomatischen Fehlern, Grammatikfehlern und Verwendung von unüblichen Begriffen deutlich besser gelungen ist (siehe Tabelle 9). Die von DeepL erstellte Übersetzung enthält gemäß Tabelle 9 mit einem arithmetischen Mittel von 4,21 mehr bekannte Zutaten als jene von Google Translate, wobei der Wert 5 von sieben Teilnehmer*innen und somit am häufigsten angekreuzt wurde; darauf folgt der Wert 4 mit fünf und Wert 2 mit zwei Teilnehmer*innen, die Werte 1 und 3 wurden hierbei nicht ausgewählt. Übersetzung 5 wurde in dieser Hinsicht von insgesamt vier Personen mit einem Wert kleiner als 3 bewertet, wobei der Wert 4 mit fünf Teilnehmer*innen den Modalwert darstellte (siehe Tabelle 9). Ebenso enthält Übersetzung 6 der Bewertung nach mit einem Mittelwert von 2,5 weniger Fehler sprachlicher beziehungsweise idiomatischer Natur; hierbei kreuzten je fünf Teilnehmer*innen die Werte 1 und 3 an, was zugleich den Modalwert darstellt (siehe Tabelle 9). Übersetzung 5 hingegen wurde mit einem Mittelwert von 3,29

deutlich schlechter bewertet, wobei insgesamt sechs Personen den Wert 4 (= Modalwert) und zwei Personen den Wert 5 ankreuzten. Ähnlich verhält es sich mit der Verwendung unüblicher Begriffe. Wie in Tabelle 9 dargestellt, wurde bei Übersetzung 5 ein Mittelwert von 3,29 und bei Übersetzung 6 ein Mittelwert von 2,36 ermittelt. Hierbei wurde im Fall von Übersetzung 5 der Wert 4 und im Fall von Übersetzung 6 der Wert 1 am häufigsten angekreuzt (siehe Tabelle 9). Hinsichtlich der Grammatik konnte Übersetzung 6 im Vergleich zu Übersetzung 5 bessere Ergebnisse erzielen. So ergab die Bewertung zu dieser Aussage, wie aus Tabelle 9 ersichtlich, bei Übersetzung 6 einen Mittelwert von 1,71, wobei die Werte 4 und 5 in diesem Fall überhaupt nicht angekreuzt wurden; der Wert 1 wurde von sieben Teilnehmer*innen ausgewählt und stellt somit den Modalwert dar. Bei Übersetzung 6 waren sich die Teilnehmer*innen eher nicht einig. Zwar ergab die Bewertung mit einem Mittelwert von 2,79 kein schlechtes Ergebnis, die Antworten der Teilnehmer*innen waren jedoch auf der Skala ziemlich breit verteilt, was nicht zuletzt die in Tabelle 9 angegebene hohe Standardabweichung zeigt. Schließlich ist zu erwähnen, dass Übersetzung 6 bei der Bewertung der Verwendung richtiger Fachbegriffe einen minimal kleineren Mittelwert erzielte (siehe Tabelle 9); die Werte 1 und 2 wurden in diesem Zusammenhang jedoch weniger oft angekreuzt als bei Übersetzung 5, wodurch die Verteilung der Antworten auf der Skala vergleichsweise nicht so breit ist. Daraus lässt sich ableiten, dass die Bewertung von Übersetzung 6 trotz des niedrigeren Mittelwerts in diesem Fall erfolgreicher war.

Die Bewertung von Set 4 fiel ebenfalls tendenziell positiv aus. So wurden die Übersetzungen 7 und 8 durchaus als akzeptabel angesehen, was nicht zuletzt durch die in Tabelle 10 berechneten, im Vergleich zu den anderen drei Sets höheren Mittelwerte bestätigt wird. Im Allgemeinen waren sich die Umfrageteilnehmer*innen einig, dass man mit diesen Rezepten die beschriebenen Desserts zubereiten kann. Dabei erreichten beide Rezepte sehr ähnliche Ergebnisse. Gemäß Tabelle 10 konnten bei der Bewertung der zweiten Aussage Mittelwerte von 4 und 4,07 ermittelt werden, wobei bei beiden Übersetzungen der Wert 5 am häufigsten angekreuzt wurde, und zwar bei Übersetzung 7 sechsmal und bei Übersetzung 8 achtmal. Die Backanleitung galt generell als verständlich und in der Praxis umsetzbar. In diesem Zusammenhang wurde von den Teilnehmer*innen in beiden Fällen der Wert 4 am häufigsten ausgewählt (siehe Tabelle 10). Die Bekanntheit der Zutaten wurde in diesem Set sehr unterschiedlich beurteilt. Während Übersetzung 7 mit einem Mittelwert von 4,86 von allen Übersetzungen das beste Ergebnis lieferte, schnitt Übersetzung 8 hier mit einem Mittelwert von 3,29 deutlich schlechter ab (siehe Tabelle 10). Bei Übersetzung 7 wurde nur eine Teilnehmerin

aufgrund der Auswahl von Wert 3 auf der Skala zu unbekannten Zutaten befragt. Diese gab an dieser Stelle *getrocknete Kokosnüsse* und *kandierte Kirschen* an (siehe Tabelle 7). Im Fall von Übersetzung 8 listeten insgesamt zehn Teilnehmer*innen unbekannte Zutaten auf. In diesem Kontext gaben gemäß Tabelle 7 neun davon an, *Muscovado-Zucker* nicht zu kennen; zwei Personen führten *Dessert-Äpfel* und je eine Person *Gewürzmischung*, *Natriumbikarbonat* und *Natron* an. Die verwendeten Mengenangaben beziehungsweise Maßeinheiten sind der Bewertung nach in beiden Übersetzungen überwiegend nachvollziehbar (siehe Tabelle 10). Dabei kommentierte nur je eine Person, dass die für die Zubereitung des Kuchens wesentlichen Gradangaben nicht verständlich sind (siehe Tabelle 7). In diesem Kontext schlug eine Person bei beiden Übersetzungen statt der Temperaturangabe *Gas 4* mit *Gasherdstufe 4* eine verständlichere Alternative vor. Während eine Teilnehmerin bei Übersetzung 7 die Angabe *Dose […] (à 425 g)* bemängelte (siehe Tabelle 7) und eher eine Variante mit *je* bevorzugte, wurde diese Angabe von einer anderen Person positiv hervorgehoben und als dem Fachjargon entsprechend bezeichnet. Drei Personen fanden, wie in Tabelle 7 aufgelistet, die Anweisung *die größte Bratpfanne […] erhitzen* seltsam, da die erforderliche Größe der Pfanne nicht klar hervorging. Hierzu schlug eine Person eine allgemeinere Variante *eine große Pfanne* und eine etwas genauere Variante *die größte Pfanne, die Sie haben* als verständlichere Optionen vor. Generell wurde empfohlen, genaue Angaben zur Größe der Pfanne zu machen. Darüber hinaus war die Größenangabe der Kuchenform in der Anweisung *Boden einer 23 cm großen, quadratischen Kuchenform einfetten* in Übersetzung 8 für eine Teilnehmerin nicht ganz nachvollziehbar (siehe Tabelle 7). Die Bewertung der Aussage zur Verwendung richtiger Fachbegriffe erzielte im Vergleich zu den anderen Sets die besten Ergebnisse. Wie Tabelle 10 zu entnehmen ist, wurde bei Übersetzung 7 der Wert 4 am häufigsten ausgewählt, während sich bei Übersetzung 8 jeweils gleiche Anteile auf die Werte 3 und 5 verteilten. Aus den in Tabelle 10 ermittelten arithmetischen Mitteln der Grammatikbewertung lässt sich schließen, dass die Übersetzungen 7 und 8 so gut wie keine grammatikalischen Fehler enthalten. Im Fall von Übersetzung 7 kreuzten nur fünf Personen und im Fall von Übersetzung 8 nur zwei Personen den Wert 4 an, wobei der Wert 5 in diesem Kontext von niemandem ausgewählt wurde. So merkte eine Person in den weiteren Kommentaren an, dass in Übersetzung 7 an der Stelle *Kokosraspeln darüber streuen* die Vorsilbe *darüber* und das Verb *streuen* fälschlicherweise getrennt geschrieben wurden (siehe Tabelle 7). Eine weitere Person gab an, dass in der Anweisung *Den Teig darübergießen, die Oberfläche glattstreichen und 30 Minuten backen, oder bis sie schön goldbraun sind.* ein falscher Bezug zwischen dem ersten und letzten Teil des Satzes hergestellt wurde. Da es in diesem Fall um den Teig geht, sollte der letzte Teil

in Singular wie folgt lauten: *bis er schön goldbraun ist*. Zwei Personen thematisierten zu Übersetzung 8 im grammatikalischen Kontext, dass im Rezept zwar zwei Äpfel gerieben werden, in weiterer Folge jedoch irrtümlicherweise der geriebene Apfel in Singular erwähnt wird. Bei der Aussage zu sprachlichen beziehungsweise idiomatischen Fehlern schnitt Set 4 im Vergleich zu den anderen Sets am besten ab. Übersetzung 7 enthält laut Umfrageteilnehmer*innen mit einem Mittelwert von 3,07 (siehe Tabelle 10) noch eher Fehler dieser Natur als Übersetzung 8 mit einem Mittelwert von 2,79. Dem steht die Bewertung zur Verwendung unüblicher Begriffe entgegen. Hierbei erzielten, wie in Tabelle 10 dargestellt, beide Übersetzungen eher positive Ergebnisse. Im Gegensatz zu Set 1 und 2 fiel die Bewertung zu dieser Aussage hier jedenfalls besser aus.

Tabelle 10: Bewertung von Set 4 mit DeepL und Google Translate (GT) nach dem Durchschnittswert und Modalwert pro zu bewertender Aussage durch 14 Personen

Aussage	**Übersetzung 7 (GT)**		**Übersetzung 8 (DeepL)**	
	Ø (σ)	**Modal-wert**	**Ø (σ)**	**Modal-wert**
Die Übersetzung ist akzeptabel.	4 (1,13)	5	3,86 (1,25)	5
Mit diesem Rezept kann man dieses Gericht zubereiten.	4 (1,2)	5	4,07 (1,33)	5
Die Backanleitung ist verständlich und könnte leicht umgesetzt werden.	3,79 (1,15)	4	3,93 (1,16)	4
Alle Zutaten sind Ihnen bekannt.	4,86 (0,52)	5	3,29 (1,39)	2
Die verwendeten Mengenangaben bzw. Maßeinheiten sind verständlich und nachvollziehbar.	4,79 (0,77)	5	4,5 (1,05)	5
Die richtigen Fachbegriffe werden verwendet.	4 (1,07)	4	3,64 (1,23)	3 & 5
Die Übersetzung enthält sprachliche bzw. idiomatische Fehler.	3,07 (1,44)	4	2,79 (1,32)	1; 3 & 4
Die Übersetzung enthält grammatikalische Fehler.	2,43 (1,35)	1	2,07 (1,16)	1
Bestimmte Begriffe sind für Sie im Deutschen nicht üblich.	2,5 (1,45)	1	2,86 (1,55)	1 & 4

Von acht Teilnehmer*innen, die zu unüblichen Begriffen bei Übersetzung 7 geantwortet haben, listeten gemäß Tabelle 7 drei (38%) *Bratensaft* im Kontext einer Süßspeise als nicht geläufig auf. Zwei Personen gaben *Backfolie* als im deutschsprachigen Raum nicht gebräuchlichen

Ausdruck an. Ebenfalls zwei Personen fanden die Bezeichnung *antihaftbeschichtete Pfanne* ungewöhnlich und führten *beschichtete Pfanne* als gängigere Variante an. Die Phrase *bis der Honig fast verschwunden ist* klingt laut zwei Personen merkwürdig (siehe Tabelle 7); eine Person interpretierte die Anweisung so, dass es sich bei diesem Vorgang um das Verkochen des Honigs handelte. Eine Person nannte darüber hinaus die Verwendung einer *Bratpfanne* zum Backen des Kuchens (siehe Tabelle 7). Hierbei sollte wahrscheinlich eher eine *Backform* zum Einsatz kommen. Zwei Personen fanden den Ausdruck *getrocknete Kokosnüsse* und eine Person *Ananaskuchen mit Deckel* unüblich. Eine Person führte in weiterer Folge die Ausdrucksweisen *in einer Lage verteilen*, *Eier auflockern* und *Teig darübergießen* als nicht gebräuchlich an (siehe Tabelle 7). Schließlich beantwortete eine Person die Frage zu unüblichen Begriffen mit „keine". Darüber hinaus kommentierten zehn Teilnehmer*innen weitere in Übersetzung 7 aufgefallene Aspekte. Darin wurde im Allgemeinen angemerkt, dass das Rezept gut beschrieben wurde. In einer Aussage wurde festgestellt, dass die Zubereitung nicht so gut verständlich ist, da zu viele Schritte auf einmal beschrieben werden, was vermutlich nicht nur der Übersetzung, sondern auch dem Original geschuldet ist. Dies wurde durch einen anderen Kommentar abgeschwächt, in dem erläutert wurde, dass die Übersetzung gut verständlich, jedoch teilweise holprig ist. Dabei soll vor allem der Titel seltsam klingen. Die Formulierung des Titels wurde in diesem Zusammenhang, wie Tabelle 7 zu entnehmen ist, von vier Personen thematisiert. Varianten wie *Ananas-Upside-Down-Kuchen*, *umgestürzter Ananaskuchen* oder *gedeckter Ananaskuchen* seien im Deutschen üblicher. Dementsprechend finden weniger als die Hälfte (43%) der Teilnehmer*innen den Titel ansprechend und würden das Rezept nachbacken. Die inkonsistente Verwendung von *getrocknete Kokosnüsse* in der Zutatenliste und *Kokosraspeln* im Zubereitungsteil wurde in fünf Kommentaren bemängelt (siehe Tabelle 7). Dabei bevorzugten die Teilnehmer*innen die Option *Kokosraspeln*, da diese im Deutschen gebräuchlicher ist und man unter der anderen Variante ganze Kokosnüsse verstehen könnte. Außerdem merkte eine Person an, dass das Rezept im Normalfall mit einem anderen Schritt, nämlich dem Abtropfen der Ananas beginnen müsste. Schließlich wurde die Angabe der genauen Pfannengröße im dritten Zubereitungsschritt von einer Person als ungewöhnlich, aber dennoch sinnvoll angesehen.

Die Antworten auf die Frage zu unüblichen Begriffen in Übersetzung 8 bezogen sich primär auf Zutaten. Von den neun dazu befragten Teilnehmer*innen listeten gemäß Tabelle 7 drei *Natriumbikarbonat* als im Kontext eines Rezepts ungewöhnlich auf; zwei gaben *normales Mehl* an, wobei hierzu Alternativen wie *Universalmehl*, *Weizenmehl* und *Backmehl* als

gängigere Varianten angesehen wurden. Je eine Person führte *Dessert-Äpfel* und *feuchte Zutaten* bei dieser Frage an. Beim Untertitel des Zubereitungsteils wurde die Beschreibung des Kuchens *klebrig-süß* im Deutschen von einer Person als unüblich erachtet (siehe Tabelle 7), wobei angemerkt wurde, dass dies im Englischen, zum Beispiel *Sticky Toffee Pudding*, typischer ist. Darüber hinaus wurde die Formulierung *backen, bis ein Spieß sauber herauskommt* von zwei Personen als nicht gebräuchlich beanstandet (siehe Tabelle 7). In deutschsprachigen Rezepten ist vielmehr von der *Stäbchen-* oder *Zahnstocherprobe* die Rede. Eine Person bezeichnete die Verwendung des Verbs *schmelzen* in der Anweisung *Zucker und Margarine in einer Pfanne schmelzen* als ungewöhnlich (siehe Tabelle 7). Je eine Person beantwortete die Frage zu unüblichen Begriffen zuletzt mit „Walnüsse gießen“ und „keine“. Schließlich hinterließen acht Teilnehmer*innen weitere Kommentare zu Übersetzung 8. Generell wurde die Übersetzung von einer Person als gut und verständlich beschrieben. Im Gegensatz dazu sind aus Sicht einer anderen Teilnehmerin die Zutaten nicht ganz klar, da etwa bei der Zutatenliste andere aufgelistet werden als jene, die in der Zubereitung vorkommen. In diesem Zusammenhang wurde gemäß Tabelle 7 von einer weiteren Person die inkonsistente Verwendung von *Natriumbikarbonat* und *Natron* angesprochen. Hier war ebenso nicht ganz klar, um welche Zutat es sich konkret handelt; man fragte sich, ob eventuell *Backpulver* damit gemeint sein könnte. Der Hälfte der Teilnehmer*innen, die einen Kommentar hinterlassen haben, war nicht klar, um welche Gewürzmischung es sich im Rezept handelt (siehe Tabelle 7). Hierbei hätte man sich genauere Angaben zu den benötigten Gewürzen gewünscht. Eine Teilnehmer*in schlug darüber hinaus vor, *Muscovado-Zucker* mit einer im deutschsprachigen Raum bekannten, anfängerfreundlicheren Variante, zum Beispiel *Vollrohrzucker*, zu ersetzen. Die Formulierung *plus extra zum Einfetten* als Zusatz zur Margarine wurde ebenfalls von einer Person als ungewöhnlich erachtet (siehe Tabelle 7). Wie auch im Fall von Übersetzung 1 und 6, wurde bei Übersetzung 8 angemerkt, dass die Überschrift *Methode* für den Zubereitungsteil im Deutschen nicht üblich ist (siehe Tabelle 7). Beim Untertitel des Zubereitungsteils sollte aus Sicht einer Teilnehmerin *Nachmittagstee* an die Zielkultur angepasst werden. Darüber hinaus wurde angemerkt, dass der Kuchen in diesem Untertitel falsch beschrieben wurde, da er nicht mit Äpfeln bestreut wird. Im Zubereitungsteil wurde die Formulierung der Anweisung *Die trockenen Zutaten schnell, aber gründlich mit den feuchten Zutaten vermischen.* als merkwürdig beurteilt (siehe Tabelle 7). Im fünften Zubereitungsschritt sollte laut einem Kommentar klargestellt werden, dass die Walnüsse in die Teigmischung gehören, vermengt werden und der Teig schließlich über die Äpfel gegossen werden soll. Der Titel der Übersetzung 8 wurde als natürlich klingend empfunden, als wäre er von einem Menschen übersetzt worden. Dem-

entsprechend finden neun Teilnehmer*innen (64%) den Titel ansprechend und würden das Rezept nachbacken.

Aus der in Tabelle 10 dargestellten Bewertung beider Übersetzungen ist zu schließen, dass die von Google Translate erstellte Übersetzung 7 bei fünf von neun Aussagen bessere Ergebnisse erzielte. Wie bereits oben ausgeführt, erzielten die Übersetzungen 7 und 8 bei der zweiten Aussage im Allgemeinen sehr ähnliche Ergebnisse, wodurch sich kein klarer Favorit abzeichnete. So kreuzte bei Übersetzung 7 je eine Person auf der Skala die Werte 1 und 2 an, während der Wert 3 zweimal, der Wert 4 fünfmal und der Wert 5 sechsmal angekreuzt wurde. Übersetzung 8 wurde hierbei einmal mit dem Wert 1, zweimal mit dem Wert 2, dreimal mit dem Wert 4 und achtmal mit dem Wert 5 beurteilt, wobei der Wert 3 von keiner Person gewählt wurde. Bei der Aussage zu den Zutaten schnitt gemäß Tabelle 10 Übersetzung 7 mit einem Mittelwert von 4,86 deutlich besser ab als Übersetzung 8. In diesem Kontext haben bei Übersetzung 7 13 Teilnehmer*innen den Wert 5 und nur eine Teilnehmerin den Wert 3 ausgewählt, während die Bewertung bei Übersetzung 8 auf der Skala breiter verteilt war und zum Beispiel fünf Personen den Wert 2 (= Modalwert) beziehungsweise vier Personen den Wert 5 angekreuzt haben. Übersetzung 8 durch DeepL erreichte, wie in Tabelle 10 dargestellt, bei der Aussage zur Verständlichkeit der Backanleitung und Grammatik bessere Ergebnisse. Ebenso schnitt Übersetzung 8 bezüglich sprachlicher beziehungsweise idiomatischer Fehler besser ab. Während Übersetzung 7 mit einem Mittelwert von 3,07 mehr Fehler dieser Natur enthalten soll, erzielte die Bewertung von Übersetzung 8 mit einem Mittelwert von 2,79 ein eindeutig besseres Ergebnis. Hierzu ist jedoch anzumerken, dass die Bewertung von Übersetzung 8 bei dieser Aussage auf der Skala breiter verteilt ist, da je vier Personen die Werte 1, 3 und 4 auswählten.

Bei Betrachtung der Bewertungsergebnisse nach den MÜ-Systemen und unter Berücksichtigung der vorstehenden Ausführungen kann abgeleitet werden, dass Google Translate im Gesamteindruck besser abgeschnitten hat als DeepL. So wurden die Übersetzungen von Google Translate in drei von vier Sets besser bewertet. In Set 1 und 4 erreichte Google Translate bei sechs von neun Aussagen ein besseres arithmetisches Mittel; in Set 2 bei fünf Aussagen. In Set 3 hingegen schnitt DeepL bei sieben Aussagen besser ab. Hierbei ist anzumerken, dass Google Translate zwar in mehr Fällen bessere Ergebnisse erzielte, sich aber nicht als klarer Favorit abzeichnen konnte, was durch eine nähere Betrachtung der Bewertungsergebnisse auf Aussagenebene ersichtlich wird. So schnitt Google Translate bei den drei Aussagen Zubereitung der Rezepte (Aussage 2), Verständlichkeit und Nachvollziehbarkeit der Mengenangaben beziehungsweise Maßeinheiten (Aussage 5), Verwendung unüblicher Begriffe

(Aussage 9) in drei von vier Sets besser ab. Hervorzuheben ist jedoch, dass Google Translate bei Aussage 2 in Set 2 aufgrund einer Differenz von 0,07 beim arithmetischen Mittel nicht eindeutig vorne lag. Im Fall von Aussage 2 wurden die Übersetzungen von DeepL lediglich in Set 2 und im Fall von Aussage 5 und 9 in Set 3 besser bewertet. Während DeepL in diesem Zusammenhang bei Aussage 2 und 9 eindeutiger Favorit war und in letzterem Fall sogar das beste Ergebnis von allen Übersetzungen erzielte, fiel die Bewertung von Aussage 5 in Set 3 nur knapp zugunsten von DeepL aus. Hinsichtlich der Akzeptabilität der Übersetzung (Aussage 1) und Grammatik (Aussage 8) erreichte DeepL bei der Bewertung in drei Sets bessere Ergebnisse. Zu Aussage 1 lag DeepL in den Sets 1, 2 und 3 eher knapp vorne und konnte sich nicht so eindeutig abheben, da die arithmetischen Mittel mit Werten zwischen 3 und 4 eher niedrig ausfielen. Aus Sicht der Umfrageteilnehmer*innen enthielten die Übersetzungen von DeepL in den Sets 1, 3 und 4 deutlich weniger Grammatikfehler als jene von Google Translate. Lediglich in Set 2 erreichte Google Translate hierzu ein besseres Ergebnis. Bezüglich der Verständlichkeit beziehungsweise Umsetzung der Koch-/Backanleitung (Aussage 3), Bekanntheit der Zutaten (Aussage 4), Verwendung richtiger Fachbegriffe (Aussage 6) und Sprache beziehungsweise Idiomatik (Aussage 7) fiel die Bewertung der Übersetzungen mit je zwei besseren Mittelwerten pro MÜ-System unentschieden aus. So lag DeepL hinsichtlich Aussage 3 in Set 4 einmal knapp und in Set 2 einmal eindeutig vorne. Google Translate hingegen erzielte in Set 1 ein eindeutig besseres Ergebnis, während sich das System in Set 3 nur knapp von DeepL abhob. Ähnlich verhielt es sich bei der Bewertung von Aussage 4: Während die Bewertung der Übersetzungen von DeepL und Google Translate in den Sets 1 und 2 ausgeglichen war und ein System nur jeweils knapp vorne lag, schnitt in Set 3 DeepL und in Set 4 Google Translate deutlich besser ab. Bei Aussage 6 erzielten die MÜ-Systeme relativ ähnliche Ergebnisse. DeepL konnte sich hierbei in den Sets 1 und 3 nicht eindeutig hervorheben, erzielte jedoch durch eine günstigere Verteilung der Antworten auf der Skala eine bessere Bewertung. Google Translate hingegen lag in Set 2 und insbesondere in Set 4 klar vorne. Ähnlich verhielt es sich bei der Bewertung von Aussage 7. Während Google Translate in den ersten zwei Sets relativ knapp besser abschnitt, erzielte DeepL in den letzten zwei Sets deutlich bessere Ergebnisse. Aus diesen Ausführungen ist zu erkennen, dass die Übersetzungen von Google Translate zwar im Allgemeinen besser bewertet wurden, sich jedoch kein MÜ-System mit seiner Leistung besonders behaupten konnte. In vielen Fällen fiel die Bewertung sehr ausgeglichen aus, wodurch sich kein klarer Favorit abzeichnete. Google Translate ist durch seine Leistung bei der Bewertung der Aussagen 5 und 9 herausgestochen, während DeepL bei den Aussagen 7 und 8 besonders punkten konnte.

Darüber hinaus ist zu berücksichtigen, inwiefern die Frage, von welchem Koch die jeweiligen Rezepte stammten, in die Bewertungsergebnisse hineinspielte. So schnitten in den Sets 1 und 3, in denen sowohl ein Rezept von Gordon Ramsay als auch eines von Jamie Oliver enthalten waren, jeweils die Übersetzungen der Rezepte von Gordon Ramsay (Übersetzungen 2 und 6) besser ab. Hierbei spielte das verwendete MÜ-System keine Rolle. Während Gordon Ramsays Rezept in Set 1 mit Google Translate übersetzt wurde, wurde dafür in Set 3 DeepL verwendet. Nichtsdestotrotz würde weniger als die Hälfte der Teilnehmer*innen die Suppe von Gordon Ramsay nachkochen, wohingegen jene von Jamie Oliver von neun Teilnehmer*innen ausprobiert werden würde. In Set 3 konnte sich bei dieser Frage das Rezept von Gordon Ramsay eher behaupten. Genau die Hälfte der Befragten würde das Rezept nachkochen, während Jamie Olivers Rezept von einer Person weniger nachgemacht werden würde. Im zweiten Set mit zwei Rezepten von Gordon Ramsay konnte Übersetzung 3 von Google Translate bei der Mehrzahl der Aussagen überzeugen. Hierbei fanden jedoch weniger Teilnehmer*innen das Rezept ansprechend; zwölf Teilnehmer*innen würden das Rezept aus Übersetzung 4 nachmachen, wohingegen es beim Rezept aus Übersetzung 3 zwei Personen weniger sind. Letztlich erzielte in Set 4 mit zwei Rezepten von Jamie Oliver Übersetzung 7 von Google Translate bessere Bewertungsergebnisse. Entgegen dieser Tendenz würden auch in diesem Fall weniger als die Hälfte der Teilnehmer*innen dieses Rezept nachbacken, während bei Übersetzung 8 neun Personen angaben, das Rezept nachmachen zu wollen. Aus diesen Ausführungen ergibt sich, dass im direkten Vergleich die Übersetzungen der Rezepte von Gordon Ramsay besser bewertet wurden als jene von Jamie Oliver. Google Translate erzielte hierbei im Allgemeinen die besseren Ergebnisse, wobei in Set 3 mit den komplexesten Rezepten DeepL vorne lag. Darüber hinaus kamen die Rezepte von Gordon Ramsay bei den Teilnehmer*innen etwas besser an, wodurch insgesamt ein höherer Anteil der Befragten diese nachkochen würde.

Die Bewertungsergebnisse wurden weiters unter dem Aspekt der Koch- beziehungsweise Backerfahrung betrachtet. In diesem Kontext wurde für die Sets 1 bis 3 die Bewertung der zehn Teilnehmer*innen, die bei der Frage zur Vertrautheit mit Kochtechniken auf der Skala die Werte 4 und 5 (Wert 4: neun Personen, Wert 5: eine Person) angekreuzt hatten, herausgefiltert. In Tabelle 11 sind die arithmetischen Mittel der Übersetzungen der ersten drei Sets dargestellt. Zum einen findet sich darin der Bewertungsdurchschnitt jener Teilnehmer*innen mit einer Kochbegeisterung größer als 3, zum anderen wird auch der Mittelwert jener Teilnehmer*innen mit einer Kochbegeisterung kleiner gleich 3 zum Vergleich aufgelistet. Gemäß Tabelle 11 konnte bei den Befragten mit einer höheren Vertrautheit mit Kochtechniken

im Vergleich zu den Bewertungsergebnissen der Gesamtgruppe keine klare Tendenz in eine bestimmte Richtung festgestellt werden. Während in den meisten Fällen die arithmetischen Mittel relativ ähnliche Werte aufweisen, schwanken jene bei einer Kochbegeisterung größer als 3 teilweise sowohl in die negative als auch in die positive Richtung; sie weichen jedoch nicht mit einem erheblichen Unterschied von der Gesamtgruppe ab. Die Bewertung jener Teilnehmer*innen mit einer Kochbegeisterung kleiner gleich 3 fiel hingegen tendenziell besser aus. In den Sets 1 und 2 wurden eher bessere arithmetische Mittel ermittelt, wobei es in einigen Fällen auch zu Ausnahmen kam (siehe Tabelle 11). In Set 3 spiegelte die Bewertung von Übersetzung 5 die Ergebnisse der Gesamtgruppe wider, wobei die Bewertung bei dieser Teilgruppe teilweise besser ausfiel. Übersetzung 6 wurde in diesem Zusammenhang eher schlechter bewertet, wobei in einigen Fällen die Tendenz auch in eine positivere Richtung ging. Betrachtet man die Bewertung der ersten Aussage zur Akzeptabilität, ergibt die Bewertung der Teilnehmer*innen mit einer höheren Kochbegeisterung relativ ähnliche Mittelwerte wie jene der Gesamtgruppe (siehe Tabelle 11). Lediglich Übersetzung 3 wurde hierbei tendenziell besser und Übersetzung 4 tendenziell schlechter bewertet. Die Bewertung der Teilnehmer*innen mit einer niedrigeren Kochbegeisterung hingegen erzielte in den meisten Fällen bessere Ergebnisse, wobei nur im Fall von Übersetzung 3 ein zur Gesamtgruppe relativ ähnliches Ergebnis erreicht wurde.

Tabelle 11: Bewertung der ersten drei Sets nach Kochbegeisterung

Aussage	**Ü1 (>3)**	**Ü1 (≤3)**	**Ü2 (>3)**	**Ü2 (≤3)**	**Ü3 (>3)**	**Ü3 (≤3)**	**Ü4 (>3)**	**Ü4 (≤3)**	**Ü5 (>3)**	**Ü5 (≤3)**	**Ü6 (>3)**	**Ü6 (≤3)**
1	3,8	4,25	3,8	4	3,6	3,5	3,2	4	3,4	3,75	3,9	4
2	4,2	4,5	4,4	4,5	3,4	4,25	3,8	4,5	3,8	3,75	3,6	3,75
3	3,5	4	4,3	3,75	3,4	3,25	3,7	4	3,5	3,5	3,5	3,25
4	3	3,75	3,3	3,25	2,8	3,25	2,9	2,75	3,7	3,25	4,3	4
5	4,2	4,25	4,5	4,5	2,6	3,25	2	3,25	3,9	4,75	4	4,75
6	3,4	3,75	3,5	3,75	3,5	3,5	3	3,75	3,4	4	3,4	3,75
7	3,3	2,25	3,1	2,5	3,2	3,25	3,7	3	3,3	3,25	2,3	3
8	2,2	2,5	2,9	2,75	1,8	2	2,1	2,5	2,9	2,5	1,7	1,75
9	3,8	2,75	3,2	2,5	3,7	4	3,9	4,25	3,3	3,25	2,1	3

Ähnlich verhielt es sich bei der Gebräuchlichkeit der Rezepte. Wie aus Tabelle 11 herauszulesen ist, fiel die Bewertung der zweiten Aussage durch die Teilnehmer*innen mit

höherer Kochbegeisterung in den Sets 1 und 3 relativ ähnlich im Vergleich zur Gesamtgruppe aus, während sie bei den Übersetzungen aus Set 2 etwas schlechtere Mittelwerte erreichte. Die Teilnehmer*innen mit niedrigerer Kochbegeisterung bewerteten die Übersetzungen aus den Sets 1 und 2 sowie Übersetzung 6 aus Set 3 besser, während die Bewertung von Übersetzung 5 nicht sehr vom Mittelwert der Gesamtgruppe abwich.

Auch im Fall von Set 4 wurden die Bewertungsergebnisse der zehn Teilnehmer*innen, die bei der Frage zur Vertrautheit mit Backtechniken auf der Skala die Werte 4 und 5 ausgewählt hatten, herausgefiltert. In Tabelle 12 sind die arithmetischen Mittel der zwei Übersetzungen dargestellt. Darin findet sich wie auch in Tabelle 11 der Bewertungsdurchschnitt jener Teilnehmer*innen mit einer Vertrautheit größer als 3 sowie der Mittelwert jener Teilnehmer*innen mit einer Backbegeisterung kleiner gleich 3. Da im Gegensatz zur Kochbegeisterung bei der Backbegeisterung mehrere Personen – genauer gesagt drei – auf der Skala den Wert 5 angekreuzt hatten, wurden hierbei die arithmetischen Mittel der Bewertung dieser Teilgruppe zum Vergleich hinzugefügt. In diesem Zusammenhang konnte, wie in Tabelle 12 ersichtlich, im Vergleich zur Bewertung der Gesamtgruppe bei den Teilnehmer*innen mit einer Backbegeisterung größer als 3 eher eine Tendenz in die negative Richtung festgestellt werden. Sowohl bei Übersetzung 7 als auch bei Übersetzung 8 waren die Mittelwerte in den meisten Fällen schlechter, wobei bei einzelnen Aussagen die Bewertung zu jener der Gesamtgruppe relativ ähnlich ausfiel. Die Bewertung jener Teilnehmer*innen mit einer Backbegeisterung größer als 4 erzielte gemäß Tabelle 12 bei Übersetzung 7 tendenziell bessere Ergebnisse als bei jenen, die einen Wert größer als 3 auswählten, wobei auch hier Ausnahmen festzustellen sind. Im Vergleich zur Gesamtgruppe wurde die Übersetzung bei dieser Gruppe eher schlechter bewertet. Übersetzung 8 hingegen wurde von den Teilnehmer*innen mit einer Backbegeisterung größer als 4 positiver beurteilt als von jenen mit einer Backbegeisterung größer als 3 beziehungsweise der Gesamtgruppe (siehe Tabelle 12). Auch hier kam es bei gewissen Aussagen zu Ausnahmen, da die Ergebnisse eine Tendenz in die andere Richtung aufwiesen. Die Teilnehmer*innen mit einer niedrigeren Backbegeisterung bewerteten die Übersetzungen in diesem Set, ähnlich wie bei den anderen Sets, besser als die anderen Teilnehmer*innengruppen (siehe Tabelle 12). Während die Bewertung von Übersetzung 7 in diesem Zusammenhang ausnahmslos besser als bei der Gesamtgruppe beziehungsweise den Teilnehmer*innen mit einer höheren Backbegeisterung ausfiel, erzielte Übersetzung 8 generell bessere Ergebnisse als bei der Gesamtgruppe und den Teilnehmer*innen mit einer Backbegeisterung größer als 3. Im Vergleich zur Gruppe mit einer

Backbegeisterung größer als 4 fiel die Bewertung auch besser aus, jedoch nicht so eindeutig; bei vier Aussagen wurden, wie in Tabelle 12 dargestellt, schlechtere arithmetische Mittel berechnet. Betrachtet man die Bewertung der Akzeptabilität, erreichte Übersetzung 7 bei den Teilnehmer*innen mit einer Backbegeisterung größer als 3 einen leicht schlechteren Mittelwert als bei der Gesamtgruppe und bei jenen mit einer Backbegeisterung größer als 4 einen besseren. Das errechnete arithmetische Mittel der Teilnehmer*innen mit niedrigerer Backbegeisterung stellt in diesem Zusammenhang das beste Ergebnis dar (siehe Tabelle 12). Ebenso erzielte die Bewertung der Akzeptabilität von Übersetzung 8 bei dieser Gruppe einen leicht schlechteren Mittelwert. Wie Tabelle 12 zu entnehmen ist, fiel die Bewertung der Teilnehmer*innen mit niedriger Backbegeisterung hierzu zwar besser aus, das beste arithmetische Mittel wurde jedoch in der Gruppe der Teilnehmer*innen mit einer Backbegeisterung größer als 4 erreicht. Ähnlich verhielt es sich bei der Gebräuchlichkeit der Rezepte. Während die Bewertung von Übersetzung 7 in diesem Zusammenhang in den jeweiligen Gruppen gleich verlief wie bei der Akzeptabilität, erzielte Übersetzung 8 bei den Teilnehmer*innen mit einer Backbegeisterung größer als 3 einen relativ ähnlichen Mittelwert wie bei der Gesamtgruppe.

Tabelle 12: Bewertung des vierten Sets nach Backbegeisterung

Aussage	**Ü7 (>3)**	**Ü7 (>4)**	**Ü7 (≤3)**	**Ü8 (>3)**	**Ü8 (>4)**	**Ü8 (≤3)**
1	3,8	4,33	4,5	3,6	4,67	4,5
2	3,7	4,33	4,75	4	4,67	4,25
3	3,4	3,67	4,75	3,6	4,33	4,75
4	4,8	5	5	3	3	4
5	4,7	4	5	4,3	4,67	5
6	3,9	4	4,25	3,6	4	3,75
7	3,3	4	2,5	2,9	2,33	2,5
8	2,7	2,67	1,75	2,2	2,33	1,75
9	2,5	3	2,5	3	3	2,5

Im Schlussteil der Umfrage wurden die Teilnehmer*innen gefragt, wie leicht ihnen die Bewertung im Allgemeinen gefallen ist. In diesem Zusammenhang empfanden, wie in Abb. 8 dargestellt, die Befragten die Bewertung weder als sehr leicht noch als sehr schwer. Acht Personen ist die Aufgabe weder leicht noch schwer gefallen. Vier Personen empfanden die Bewertungsaufgabe als leicht und zwei Personen als schwer. Hierbei spielte die Vertrautheit mit Koch- beziehungsweise Backprozessen keine gravierende Rolle, da jede Antworten-

Gruppe Teilnehmer*innen enthielt, die unterschiedliche Werte auf der Skala ankreuzten. Darüber hinaus fanden elf Teilnehmer*innen (79%), dass ihre Koch- und Backkenntnisse für die Bewertung ausreichend waren. Dementsprechend kreuzten drei Teilnehmer*innen (21%) bei dieser Frage Nein an. In diesem Kontext begründete eine Teilnehmerin ihre Antwort damit, dass sie kein Fleisch konsumiert und nicht mit den Zutaten vertraut ist. Die Fleischgerichte waren für sie dementsprechend schwer zu beurteilen. Ein Teilnehmer schrieb hierzu, dass er einige Fachbegriffe und Kocharten nicht kannte und die Richtigkeit der Inhalte der Übersetzungen daher schwer zu beurteilen war. Darüber hinaus meinte er, dass manche Rezepte so lang waren, dass sich die genaue Befolgung derselben schwierig gestaltete. Schließlich begründete die dritte Teilnehmerin ihre Antwort zu dieser Frage damit, dass sie weniger oft mit schriftlichen Rezepten arbeitet und eher deutsche YouTube-Videos ansieht, in denen alle Schritte einzeln gezeigt beziehungsweise erklärt und kaum Fachbegriffe verwendet werden. Dementsprechend müsste sie Fachbegriffe wie *pariert*, *gehäufter Löffel* oder *tranchieren* nachschlagen. Sie führte weiters aus, dass die Gerichte, die sie zubereitet, einfacher sind und sie zum Beispiel Beef Wellington nie zubereitet hatte.

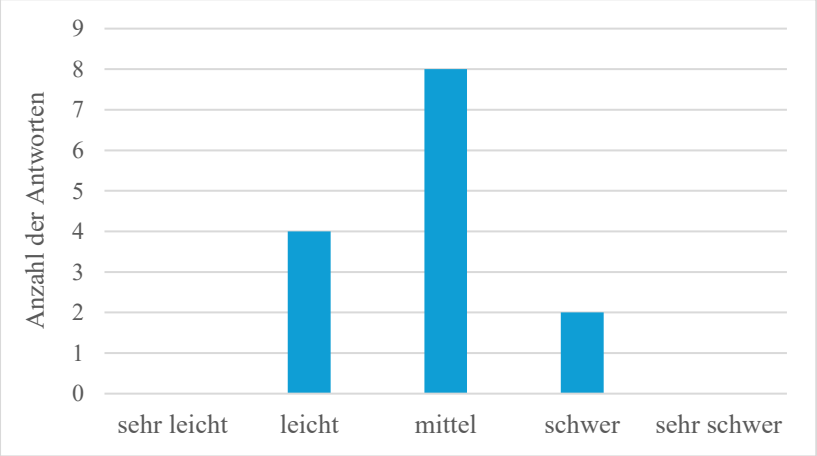

Abbildung 8: Einschätzung der Umfrageteilnehmer*innen zur Leichtigkeit der Bewertung

5.5. Kommentare zur Umfrage und MÜ

Zum Schluss der Umfrage nahmen sich fünf Teilnehmer*innen noch die Zeit und teilten ihre Gedanken zur Umfrage, Thematik und MÜ im Allgemeinen mit. Die Anmerkungen beinhalteten dabei sowohl allgemeine Kommentare zur Thematik als auch detailliertes Feedback zur Umfrage und eine Stellungnahme zur Verwendung von MÜ in diesem Bereich. So wurde von drei Teilnehmer*innen hervorgehoben, dass sie das Thema sehr interessant

finden. Auch die Rezeptauswahl wurde gelobt. In diesem Zusammenhang fand eine Teilnehmerin die Rezepte sehr spannend und versicherte, dass sie einige davon ausprobieren wird. Eine andere Teilnehmerin bezeichnete die Rezeptauswahl als gelungen und erläuterte, dass die Rezepte trotz mancher unbekannter Zutaten und Sprach- beziehungsweise Grammatikfehler zum Nachkochen respektive Nachbacken einladen. Außerdem führte ein Teilnehmer aus, dass er sich das Rezept für den veganen Toffee-Apfel-Upside-Down-Kuchen abgespeichert hat. In weiterer Folge merkte dieser Teilnehmer an, dass vor allem die Bewertung etwas ermüdend war und es unter anderem aufgrund der Länge der Rezepte schwierig war, den Inhalten genau zu folgen. Hierzu führte er Folgendes aus: „Man hätte zwar eine Pause machen können, aber es ist fraglich, wie sich das auf die Bewertung auswirkt oder wie diese Umfrage das vorsieht. Vielleicht wären Erinnerungen, dass man Pause machen soll, hilfreich für die Motivation […].“ Darüber hinaus fand er es schwierig einzuschätzen, wie man die fünf Punkte auf der Skala bei der Bewertung interpretieren soll. Es stellte sich dabei auch die Frage, ob man wieder nach oben scrollen soll, um zu prüfen, was aufgefallen ist beziehungsweise ob man etwas vergessen hat.

Schließlich hinterließen noch zwei professionelle Translatorinnen Kommentare zur MÜ-Qualität beziehungsweise zum Einsatz von MÜ in diesem Bereich. In diesem Kontext schrieb eine Translatorin, dass die Ergebnisse beziehungsweise die Qualität der MÜ sehr unterschiedlich waren. Während manche etwas holprig waren, waren andere sehr gut verständlich und regten zum Nachkochen/-backen an. Eine andere Translatorin teilte mit, dass man als Übersetzer*in oft über die MÜ anderer Textsorten, wie zum Beispiel Verträge, Pressetexte, Social-Media-Texte oder Website-Texte, nachdenkt, aber selten, zumindest in ihrem Fall, über die Übersetzung von Rezepten. Dabei erläuterte sie, dass Koch- beziehungsweise Backrezepte ein hohes Maß an Kulturkompetenz verlangen, da in bestimmten Ländern zahlreiche Zutaten existieren und vielleicht sogar sehr üblich sind, während man diese in anderen Teilen der Welt überhaupt nicht kennt. Zur Frage, ob der Einsatz von MÜ in diesem Bereich sinnvoll ist, zeigte sich die Teilnehmerin eher skeptisch und hob die Tatsache hervor, dass gerade bei Texten wie Rezepten wenig Zeit zum Nachdenken beziehungsweise Hineininterpretieren bleibt. Hierzu teilte die Translatorin ihre Gedanken wie folgt mit: „Rezepte sind meiner Meinung nach Texte, die sich gut lesen lassen sollten und keine allzu hohe Konzentrationsfähigkeit verlangen sollten (oft muss es beim Kochen / Backen schnell gehen). Daher finde ich die menschliche, gut überdachte Übersetzung – im Moment (noch) – viel sinnvoller. Gerade bei dieser Textsorte sollte nicht bei der Übersetzung gespart werden.“

6. Diskussion

In Kapitel 3 wurden bereits durchgeführte Studien zur Qualitätsbewertung maschinell übersetzter Texte vorgestellt. Da bisher keine Studien in der Art der vorliegenden Masterarbeit existieren, wurden hierfür Studien im Bereich der MÜ herangezogen, die sich auf andere Textsorten fokussierten und deren Bewertungsteile eher aus der manuellen Identifizierung und Klassifizierung von Fehlern als der Kriterienbewertung auf einer Likert-Skala bestanden (vgl. Wiesmann 2019, Heinisch & Lušicky 2019, Larassati et al. 2019). Darüber hinaus handelte es sich bei den Teilnehmer*innen dieser Studien um professionelle Translator*innen. So wurden bei Wiesmann (2019) und Larassati et al. (2019) die Übersetzungen von den Autor*innen selbst bewertet beziehungsweise auf Fehler untersucht. Heinisch und Lušicky (2019) hingegen ließen Studierende der Translationswissenschaft die Fehler in den maschinell erstellten Übersetzungen identifizieren und klassifizieren. In der vorliegenden Masterarbeit wurde die Zielgruppe für die Umfrage beziehungsweise Bewertung auf ein sachgebietsspezifisches Publikum, das nicht zwingend nur aus Translator*innen bestehen musste, ausgeweitet.

Die Studie von Wiesmann (2019) behandelte die Gebräuchlichkeit der MÜ von Rechtstexten. Dabei wurden repräsentative Textsorten dieses Fachgebiets herangezogen. Auch in der vorliegenden Masterarbeit geht es um die Qualität von MÜ in einem bestimmten Fachgebiet, nämlich jenem der Kulinarik. Dafür wurden ebenfalls repräsentative Texte, genauer gesagt Koch- und Backrezepte von renommierten Köchen mit einer Vielfalt an Gerichten und unterschiedlichen Schwierigkeitsgraden herangezogen. Wiesmann (2019) kam zu dem Schluss, dass die maschinell erstellten Übersetzungen zwar verständlich, jedoch nicht immer korrekt waren, was sich vor allem in der inkonsistenten Terminologieverwendung äußerte. In diesem Zusammenhang fanden die Umfrageteilnehmer*innen der vorliegenden Masterarbeit die Übersetzungen der Rezepte tendenziell gebräuchlich und die Anweisungen verständlich beziehungsweise umsetzbar. Hierbei verwiesen die Teilnehmer*innen auf bestimmte Passagen, in denen unübliche Begriffe verwendet wurden, deren Sinn beziehungsweise Bedeutung sich jedoch problemlos ableiten ließ. Im Gegensatz dazu führte die inkonsistente Verwendung gewisser Termini beziehungsweise die teilweise fehlerhafte Verwendung gewisser Ausdrücke an einigen Stellen zu Missverständnissen. Schließlich stellte Wiesmann (2019) fest, dass die Übersetzungsqualität auch von Faktoren wie Textlänge und Komplexität abhängt. In diesem Kontext konnte im Rahmen der vorliegenden Masterarbeit festgestellt werden, dass die Übersetzungen der Rezepte von Gordon Ramsay im direkten Vergleich besser bewertet wurden. Dies ist vermutlich auf seine Ausdrucksweise, seinen Stil und die vergleichsweise kürzere Syntax

zurückzuführen. Darüber hinaus wurden die Übersetzungen von den komplexesten Rezepten am schlechtesten bewertet. Auch die Tatsache, aus welcher Küche die Teilnehmer*innen besonders gerne Gerichte zubereiten, spielte bei der Bewertung der Übersetzungsqualität zumindest hinsichtlich Kenntnis der Zutaten eine Rolle. So führten Personen, die regelmäßig asiatisch kochen, Zutaten asiatischen Ursprungs weniger oft als unbekannt an.

Heinisch und Lušicky (2019) untersuchten, wie bereits in Kapitel 3 beschrieben, die Qualität von MÜ unter Berücksichtigung der Erwartungen der Nutzer*innen. In der vorliegenden Masterarbeit wurden zwar nicht die Erwartungen der Teilnehmer*innen an die MÜ-Qualität erfragt, die Umfrage enthielt jedoch wie auch der Fragebogen von Heinisch und Lušicky (2019) Teile zum Beruf und der Erfahrung mit MÜ. Da die Zielgruppe sowohl aus professionellen Translator*innen als auch Lai*innen bestand, wurden Translator*innen in diesem Kontext zu ihren Arbeitssprachen und allgemeinen Sprachkenntnissen und Lai*innen zu ihrer Branche und Erstsprache befragt. Dabei gaben alle befragten professionellen Translatorinnen an, bereits einmal Koch- und Backrezepte übersetzt zu haben. Daraus ließ sich ableiten, dass sie sich der Herausforderungen dieser Textsorte im Übersetzungsprozess bewusst sind. Im Umfrageabschnitt zur Erfahrung mit MÜ-Systemen gaben zwölf von 14 Teilnehmer*innen an, MÜ-Systeme zu nutzen. Dabei führten elf Personen Google Translate und acht DeepL als regelmäßig genutzte MÜ-Systeme an. Da die Teilnehmer*innen folglich mit den Ergebnissen von Google Translate vertrauter waren, ist es wenig überraschend, dass dieses System bei der Bewertung im Allgemeinen besser abgeschnitten hat. Darüber hinaus sollten die Teilnehmer*innen dem Kontext der Masterarbeit entsprechend Angaben zu ihrer Koch- und Backerfahrung machen. Die Umfrageteilnehmer*innen wurden auch gefragt, ob sie bereits in der Vergangenheit MÜ-Systeme für Rezepte verwendet hatten. Dabei bejahten vier Personen die Frage, wobei diese eher nicht mit der Qualität der MÜ zufrieden waren. Weiters konnte festgestellt werden, dass Teilnehmer*innen mit einer höheren Koch- beziehungsweise Backerfahrung die maschinell erstellten Übersetzungen schlechter bewerteten und als weniger akzeptabel beziehungsweise gebräuchlich empfanden als jene Teilnehmer*innen mit einer niedrigeren Koch- beziehungsweise Backerfahrung. Dementsprechend konnte die Schlussfolgerung von Heinisch und Lušicky (2019), dass die Erfahrungen der Nutzer*innen die MÜ-Bewertung beeinflussen, in der vorliegenden Masterarbeit im weitesten Sinne bestätigt werden.

Die Studie von Larassati et al. (2019) stellt die einzige in dieser Arbeit beschriebene Studie zu prozeduralen Textsorten dar. Motivation für die Textsortenauswahl war dabei, wie auch im Fall der vorliegenden Masterarbeit, dass die Leser*innen zur Erreichung eines Ziels

gewisse Anweisungen befolgen müssen. Bei Larassati et al. (2019) wurden wie auch in dieser Masterarbeit die Ergebnisse zweier Systeme, nämlich Google Translate und Instagram, herangezogen und deren Bewertungen gegenübergestellt. Während bei Larassati et al. (2019) die Fehler manuell anhand einer Fehlertypologie analysiert wurden, wurden die MÜ der vorliegenden Masterarbeit auf einer Likert-Skala nach bestimmten Aussagen bewertet. Google Translate erzielte in der Studie von Larassati et al. (2019) bessere Ergebnisse als die Übersetzungsfunktion von Instagram. Im Vergleich dazu schnitt in der vorliegenden Arbeit Google Translate im Allgemeinen besser ab als DeepL, wobei sich kein MÜ-System mit seiner Leistung deutlich abheben konnte. Larassati et al. (2019) dokumentierten am häufigsten Fehler der Kategorien Terminologie, Syntax und Wörtlichkeit; auch wurden zum Beispiel Verwendungsfehler identifiziert. Die Umfrageteilnehmer*innen der vorliegenden Masterarbeit bewerteten hingegen die Aussage zur Verwendung richtiger Fachbegriffe tendenziell positiv, thematisierten aber in einigen Fällen eine fehlerhafte beziehungsweise missverständliche Terminologie. In der Bewertung und den Kommentaren wurden in diesem Kontext zahlreiche unübliche Begriffe und Ausdrücke angesprochen. Darüber hinaus kamen Larassati et al. (2019) zum Schluss, dass sich die Übersetzungsqualität von Google Translate beachtlich erhöht hat, dieses System dementsprechend Fachbegriffe beziehungsweise Abkürzungen identifiziert und Grammatik, Wortstellung sowie Wortwahl berücksichtigt. Google Translate konnte bei der Bewertungsaufgabe in der vorliegenden Masterarbeit vor allem hinsichtlich Verständlichkeit der verwendeten Mengenangaben und Maßeinheiten sowie Verwendung unüblicher Begriffe mit seiner Leistung herausstechen. DeepL hingegen punktete besonders hinsichtlich Grammatik, Sprache und Idiomatik. Letztlich konnten Larassati et al. (2019) bei Google Translate keine textsortenspezifischen Fehler feststellen und schlossen daraus, dass das MÜ-System die entsprechenden Konventionen befolgt. Die Umfrageteilnehmer*innen dieser Masterarbeit hingegen thematisierten in den Anmerkungen die Verwendung unüblicher Ausdrücke bei den Überschriften der Rezepte. Auch wurde mehrfach auf den inkonsistenten Stil bei den Anweisungen durch die abwechselnde Verwendung des Infinitivs und der Sie-Form hingewiesen. Daraus lässt sich ableiten, dass die MÜ-Systeme nicht alle Textsortenkonventionen berücksichtigten.

Alle drei beschriebenen Studien gelangten zum Ergebnis, dass die Qualität der MÜ noch unzureichend ist und nicht an professionelle Humanübersetzungen herankommt. Dies spiegelt ebenfalls die Bewertungsergebnisse der vorliegenden Masterarbeit wider. So lassen sowohl die errechneten Mittelwerte als auch die Kommentare zur Umfrage darauf schließen. In diesem

Zusammenhang wurden etwa in einem Kommentar die Tatsachen, dass gerade bei der Textsorte der Rezepte wenig Zeit zum Überlegen bleibt, bei der Übersetzung von Rezepten ein hohes Maß an Kulturkompetenz erforderlich ist und die MÜ noch nicht die nötige Qualität erreicht hat, hervorgehoben.

Ziel der vorliegenden Masterarbeit war es herauszufinden, wie Kochbegeisterte die Übersetzungen von Rezepten durch zwei NMÜ-Systeme bewerten. In diesem Zusammenhang wurde zur Untersuchung der Forschungsfragen und Überprüfung der Annahmen eine Online-Umfrage erstellt, in der die tatsächliche Zielgruppe der Textsorte der Koch- und Backrezepte in den Bewertungsprozess miteinbezogen wurde. Insgesamt füllten 14 Personen die Umfrage vollständig aus. Diese erfüllten allesamt mit einer ausreichend hohen Einschätzung der Koch- und Backexpertise sowie guten Deutschkenntnissen die Teilnahmevoraussetzungen. Von den 14 Befragten waren neun professionelle Translatorinnen und fünf Personen aus anderen Branchen vertreten, wodurch eine differenziertere Betrachtung der MÜ-Qualität gewährleistet werden konnte. Alle befragten professionellen Translatorinnen verfügten bereits über Erfahrungen mit der Übersetzung von Rezepten und waren sich folglich der in diesem Kontext auftretenden Herausforderungen bewusst. Im Allgemeinen nutzen zwölf Teilnehmer*innen – alle Translatorinnen und drei von fünf Lai*innen – relativ regelmäßig MÜ-Systeme, vor allem für private, berufliche und Ausbildungszwecke. Vier Personen nutzten bereits einmal MÜ-Systeme für Koch- beziehungsweise Backrezepte, waren dabei jedoch eher nicht mit der Qualität zufrieden.

Die Bewertung der maschinell übersetzten Koch- und Backrezepte ergab insgesamt positive Ergebnisse. Mit steigender Komplexität der Rezepte wurden die Übersetzungen schlechter bewertet, wobei die Beurteilung nie erheblich in den negativen Bereich ging. Alle übersetzten Rezepte galten tendenziell als akzeptabel und in der Praxis umsetzbar. Die Verständlichkeit der Koch- beziehungsweise Backanleitungen stellte mit steigender Komplexität eine zunehmende Herausforderung dar, wobei die arithmetischen Mittel nie unter den Wert 3 fielen. Die Zutaten und deren Bekanntheit stellten eines der Hauptprobleme dar. In diesem Zusammenhang wurden bei allen Rezepten mehrere Zutaten aufgelistet, mit denen eine große Zahl der Teilnehmer*innen nichts anfangen konnte. Ohne einschlägige Recherche würde dies den Zubereitungsprozess erheblich erschweren. Daraus ist zu schließen, dass eine entsprechende Anpassung der Zutaten an den deutschsprachigen Raum oder eine nähere Erklärung mit Auflistung von alternativen Zutaten durch professionelle Translator*innen unabdingbar ist. Die Tatsache, aus welcher Küche die Teilnehmer*innen besonders gerne

Gerichte zubereiten, hatte zwar einen Einfluss auf die Thematisierung von unbekannten Zutaten, diese fiel jedoch nicht massiv aus. So gab einerseits eine Person, die eher asiatisch kocht, bei den Beef-Wellington Rezepten Gewürze wie *Thymian* als nicht bekannt an. Andererseits bezeichneten Personen, die gerne asiatisch kochen, mit wenigen Ausnahmen asiatische Zutaten als unbekannt. Die Verständlichkeit und Nachvollziehbarkeit der Mengenangaben und Maßeinheiten wurde in den meisten Fällen sehr gut bewertet. Dies liegt sicherlich vor allem daran, dass in den Originalrezepten hauptsächlich die im deutschsprachigen Raum bekannten Maßeinheiten *Gramm* und *Milliliter* verwendet wurden. Bei den Rezepten in zweiten Set fiel die Bewertung aufgrund der Verwendung der Einheiten *ounces*, *pounds* und *cups* eher negativ aus. Die wörtliche Übersetzung dieser Einheiten ins Deutsche mit *Unzen*, *Pfund* und *Tassen* wurde von den Teilnehmer*innen als nicht verständlich und ungewöhnlich angesehen. Ebenso kam es bei den Teilnehmer*innen generell durch ungenaue Mengenangaben wie *Klecks* oder *Stück* zu Verständnisschwierigkeiten. Laut den Umfrageteilnehmer*innen wurden in den Übersetzungen tendenziell die richtigen Fachbegriffe verwendet. Nichtsdestotrotz sind den MÜ-Systemen in bestimmten Fällen Fehler beziehungsweise Unstimmigkeiten unterlaufen. Beispielsweise wurde in einem Rezept des zweiten Sets irrtümlich eine Überschrift im Zubereitungsteil mit *Anweisungen zum Garen* übersetzt, obwohl darin gar nicht gegart werden sollte. Eine größere Herausforderung für die MÜ-Systeme stellten jedoch mehrdeutige Termini dar. In diesem Zusammenhang wurde etwa das englische Verb *cook* in den meisten Sets mit *kochen* übersetzt, wobei damit Vorgänge wie *anbraten* oder *andünsten* gemeint waren. Hierzu ist anzumerken, dass die Teilnehmer*innen in diesen Fällen den eigentlichen Sinn aus den Anweisungen erschließen konnten. Schwieriger gestaltete es sich bei der fälschlichen Übersetzung von *pepper* mit *Pfeffer* statt *Paprika* in Übersetzung 1. Hier waren viele Teilnehmer*innen irritiert. Auch die irrtümliche Übersetzung von *mushrooms* mit *Champignons* statt *Pilze* wurde nicht von allen Befragten erkannt. In diesem Kontext führte auch die inkonsistente Übersetzung gewisser Termini zu Irritationen. Da für die gleichen Elemente oder Zutaten teilweise unterschiedliche Ausdrücke verwendet wurden, waren sich die Teilnehmer*innen an gewissen Stellen nicht sicher, wann welche zum Einsatz kommen, oder haben teilweise nicht verstanden, was im Rezept benötigt wurde. Ein Beispiel hierfür ist das *Paprika-Topping* in Übersetzung 1.

Die Bewertung der Sprache und Idiomatik in der MÜ fiel unterschiedlich aus. In manchen Fällen wurden eher keine diesbezüglichen Mängel identifiziert, während in anderen Fällen durchaus Fehler dieser Natur erkannt wurden. So wurde sowohl im Kontext der

unüblichen Begriffe als auch der zusätzlichen Anmerkungen auf zahlreiche nicht idiomatische Ausdrücke beziehungsweise Phrasen verwiesen. Diese wurden zwar von den Teilnehmer*innen hervorgehoben und als nachbearbeitungsbedürftig angesehen, sie wurden dennoch im Wesentlichen verstanden. In sprachlicher Hinsicht spielten auch regionale Unterschiede zwischen Österreich und Deutschland bei der Bewertung eine Rolle. So wurden etwa die Begriffe *Blumenkohl* und *Tomate* thematisiert. Außerdem wurde die Verwendung des femininen Artikels für *Paprika* angesprochen. Hinsichtlich der Grammatik traten insgesamt wenig Fehler auf. Dennoch wurde in mehreren Fällen durch die Verwendung des falschen Pronomens ein falscher Bezug hergestellt. In Übersetzung 2 kam es zur falschen Verwendung des femininen Artikels für *Blumenkohl* und zu einer fehlerhaften Satzstellung. Darüber hinaus verzeichneten die Umfrageteilnehmer*innen in allen Übersetzungen mehrere unübliche Begriffe. Während diese in manchen Fällen dennoch verständlich waren, sorgten sie in anderen wiederum für Irritationen beziehungsweise Missverständnisse. Derartige Fehler und auch die nicht idiomatischen Ausdrücke beziehungsweise Phrasen sind vermutlich auf die in Larassati et al. (2019) erwähnte Wörtlichkeit der MÜ-Systeme zurückzuführen. Beispiele dafür sind etwa die *schwere Bratpfanne* oder die Phrase *bis der Spinat verwelkt ist*. In diesem Zusammenhang ist der Kommentar einer Translatorin zu nennen. Sie thematisierte darin, dass sich Rezepte gut lesen lassen und keine allzu hohe Konzentrationsfähigkeit erfordern sollten.

Durch die Verwendung von unüblichen Begriffen beziehungsweise nicht idiomatischen Ausdrücken müssen sich die Leser*innen dieser maschinell übersetzten Rezepte durchaus mehr konzentrieren. Obwohl der Großteil dieser Begriffe und Ausdrücke von den Teilnehmer*innen verstanden wurde, erfordern diese dennoch mehr Konzentration, Mühe und Zeit, die beim Zubereiten der Gerichte nicht immer vorhanden ist. Aufgrund dieser Ausführungen lässt sich die Annahme, dass durch die kulturellen Besonderheiten von Rezepten als Textsorte und die terminologische Inkonsistenz von MÜ-Systemen die Qualität von MÜ nicht zur problemlosen Befolgung der Anleitungen aus den Rezepten ausreicht, bestätigen. Die im Zuge der MÜ entstandenen Fehler spielten bei der Frage, ob die Teilnehmer*innen diese Gerichte zubereiten würden, keine maßgebliche Rolle. Trotz gewisser unbekannter Zutaten und Sprach- beziehungsweise Grammatikfehler lud, wie auch in einem Abschlusskommentar zur Umfrage angemerkt, der Großteil der Rezepte zum Nachkochen beziehungsweise -backen ein.

Hinsichtlich der Frage, welches MÜ-System bei der Übersetzung von Rezepten aus dem Englischen ins Deutsche bessere Ergebnisse erzielt, konnte festgestellt werden, dass Google Translate im Gesamtbild besser bewertet wurde. Dies kann unter anderem darauf

zurückzuführen sein, dass insgesamt elf der zwölf MÜ-Nutzer*innen der gegenständlichen Umfrage dieses MÜ-System nutzen und dementsprechend mit dessen Ergebnissen vertrauter sind. Hierbei ist anzumerken, dass DeepL bei den komplexesten Rezepten besser abschnitt. Es konnte sich in diesem Zusammenhang jedoch kein MÜ-System mit seiner Leistung besonders vom anderen abheben. In vielen Fällen fiel die Bewertung sehr ausgeglichen aus, wodurch sich kein klarer Favorit abzeichnen konnte. Google Translate stach im Vergleich zu DeepL durch seine Leistung bei der Bewertung der Aussagen zur Verständlichkeit und Nachvollziehbarkeit der Mengenangaben beziehungsweise Maßeinheiten und zur Verwendung von unüblichen Begriffen heraus. DeepL hingegen brillierte besonders im Kontext der Aussagen zu Sprache beziehungsweise Idiomatik und Grammatik. Daraus ist zu schließen, dass von den Umfrageteilnehmer*innen kein bestimmtes MÜ-System für die Übersetzung von Koch- beziehungsweise Backrezepten empfohlen werden konnte. Sowohl Google Translate als auch DeepL liefern in dieser Hinsicht ähnliche Ergebnisse und stehen vor denselben Herausforderungen.

Weiters spielte bei den Bewertungsergebnissen die Frage, von welchem Koch die jeweiligen Rezepte stammten, eine nicht unerhebliche Rolle. So schnitten in den Sets mit Rezepten beider Köche die Rezepte von Gordon Ramsay besser ab, wobei in diesen Fällen für seine Rezepte unterschiedliche MÜ-Systeme verwendet wurden. Daraus lässt sich ableiten, dass die MÜ-Systeme generell besser mit der Ausdrucksweise, dem Stil und der Syntax von Gordon Ramsay umgehen können. In den Sets mit je zwei Rezepten vom gleichen Koch hingegen konnten die Übersetzungen von Google Translate mehr überzeugen. Die bessere Bewertung bestimmter Rezepte hing jedoch nicht mit einem höheren Anteil an Teilnehmer*innen, die den Titel ansprechend fanden und das Rezept nachkochen/-backen würden, zusammen. In drei von vier Sets würden im direkten Vergleich weniger Umfrageteilnehmer*innen die Gerichte der besser bewerteten Übersetzungen zubereiten wollen. Insgesamt würde jedoch ein höherer Anteil der Befragten die Gerichte von Gordon Ramsay nachkochen. Somit beeinflusste die Attraktivität der Gerichte die Bewertungsergebnisse nicht.

Darüber hinaus bewerteten Teilnehmer*innen mit einer niedrigeren Koch- beziehungsweise Backbegeisterung die Übersetzungen tendenziell besser. Damit konnte die Annahme, dass die Gesamtbewertung der MÜ bei Personen mit einem höheren Grad des prozeduralen Wissens schlechter ausfällt, anhand der Ergebnisse tendenziell positiv bewertet werden. Entgegen den Erwartungen erzielten die übersetzten Rezepte bei Personen mit einer höheren Koch- beziehungsweise Backbegeisterung bezüglich der Akzeptabilität und Gebräuchlichkeit der Übersetzungen niedrigere Mittelwerte. Dementsprechend scheinen Personen mit

ausgeprägteren Kenntnissen im Bereich der Kulinarik maschinelle Übersetzungen weniger zu akzeptieren und finden diese in der Praxis schlechter umsetzbar als Personen mit einem niedrigeren Grad an prozeduralem Wissen.

In methodischer Hinsicht kam es zu gewissen Herausforderungen. So gestaltete sich die Rezeptauswahl schwieriger als erwartet, da es vergleichbare Rezepte zu suchen galt, die sich vom Gericht, Schwierigkeitsgrad und der Länge her ähnlich waren. Darüber hinaus wurde bei der Rezeptauswahl nicht berücksichtigt, dass Personen, die kein Fleisch essen, bei der Bewertung von fleischhaltigen Rezepten Schwierigkeiten und nicht ausreichend Kenntnisse haben könnten. Zurückblickend hätte man hier entweder die Zielgruppe weiter spezifizieren oder die Textauswahl fairer gestalten sollen. Die größte Herausforderung stellte jedoch die Länge der Umfrage dar. Von mehreren Teilnehmer*innen wurde mündlich oder schriftlich berichtet, dass die Umfrage und vor allem die Bewertungsaufgabe ziemlich langwierig waren. In diesem Zusammenhang kommentierte ein Teilnehmer zum Schluss der Umfrage, dass es auch aufgrund der Rezeptlängen schwierig war, diese problemlos zu befolgen. Hierbei merkte er weiters an, dass Erinnerungen, eine Pause zu machen, die Motivation fördern könnten. Retrospektiv hätte man in diesem Kontext die Bewertungsaufgabe entweder um ein weiteres Set kürzen oder eine Komponente zum Abbrechen und späteren Fortsetzen der Umfrage einbauen können. Dabei wäre jedoch fraglich gewesen, wie viele Teilnehmer*innen die Umfrage zu einem späteren Zeitpunkt tatsächlich fortgesetzt hätten. Dies hätte eventuell auch dazu geführt, dass eine höhere Zahl an Teilnehmer*innen die Umfrage vollständig ausgefüllt hätte. Letztlich wurde vom oben genannten Teilnehmer angemerkt, dass auch die Interpretation der fünf Punkte auf der Bewertungsskala nicht einfach war. In diesem Zusammenhang hätte man zum besseren Verständnis und der Gewährleistung eines konsistenten Bewertungsverhaltens die einzelnen Punkte noch näher definieren können.

7. Fazit

Die vorliegende Masterarbeit beschäftigte sich mit der Frage, wie Kochbegeisterte die Übersetzung von Koch- beziehungsweise Backrezepten aus dem Englischen ins Deutsche durch zwei NMÜ-Systeme bewerten. Dabei wurde untersucht, welches NMÜ-System die besseren Ergebnisse erzielt, inwiefern die maschinell übersetzten Rezepte eine erfolgreiche Zubereitung der Gerichte gewährleisten und welche Einschränkungen dabei festgestellt werden können sowie inwiefern das kulinarische Hintergrundwissen der Befragten die Bewertung beeinflusst. Zur Untersuchung dieser Fragestellungen wurde eine Online-Umfrage erstellt, in der ein sachgebietsspezifisches Publikum in den Bewertungsprozess miteinbezogen wurde.

Die maschinell erstellten Übersetzungen der Koch- und Backrezepte wurden insgesamt als akzeptabel und in der Praxis umsetzbar angesehen, wobei die Verständlichkeit der Anweisungen mit steigender Komplexität der Rezepte eine zunehmende Herausforderung darstellte. Die Bewertung der Bekanntheit der Zutaten ergab, dass die Arbeit von professionellen Translator*innen in diesem Bereich zur Gewährleistung eines erfolgreichen Zubereitungsprozesses unerlässlich ist. Aus Sicht der Umfrageteilnehmer*innen leisteten die NMÜ-Systeme bei der Übersetzung von im deutschsprachigen Raum bekannten Maßeinheiten sehr gute Arbeit. Die MÜ-Resultate zeigen jedoch in Bezug auf die Übersetzung von im britischen und amerikanischen Raum verwendeten Maßeinheiten sowie von ungenauen Mengenangaben Herausforderungen auf. Darüber hinaus wurden in den Übersetzungen tendenziell richtige Fachbegriffe verwendet. Die fehlerhafte Übersetzung von mehrdeutigen Termini und auch die inkonsistente Übersetzung gewisser Termini führten dennoch zu Verständnisschwierigkeiten und Irritationen. Bei der Bewertung der Sprache und Idiomatik wurden bei den jeweiligen Sets unterschiedliche Ergebnisse erzielt. Die Befragten listeten in diesem Kontext zahlreiche nicht idiomatische Ausdrücke und Phrasen auf, die allerdings größtenteils keine gravierenden Verständnisprobleme verursachten. Insgesamt sind im Zuge der MÜ nur wenige Grammatikfehler entstanden. Hauptfehlerquelle stellten hierbei Übereinstimmungsfehler durch die Verwendung falscher Pronomen dar. Schließlich wurden bei der Bewertung zahlreiche unübliche Begriffe in den maschinellen Übersetzungen wahrgenommen, die vermutlich durch die wörtliche Übersetzung der MÜ-Systeme verursacht wurden. Diese sorgten im Allgemeinen für keine gravierenden Verständnisschwierigkeiten, führten in einigen Fällen jedoch durchaus zu Irritationen und Missverständnissen. Trotz gewisser im Zuge der MÜ entstandener Fehler würden die Befragten einen Großteil der Rezepte nachkochen beziehungsweise -backen.

Aus den Ergebnissen der Bewertung lässt sich weiters schließen, dass sowohl Google Translate als auch DeepL bei der Übersetzung von Koch- und Backrezepten aus dem Englischen ins Deutsche ähnliche Ergebnisse liefern und vor denselben Herausforderungen stehen. Im Allgemeinen wurden die Übersetzungen von Google Translate besser bewertet, es zeichnete sich jedoch kein klarer Favorit ab. Während Google Translate aus Sicht der Umfrageteilnehmer*innen bei der Bewertung der Verständlichkeit und Nachvollziehbarkeit der Mengenangaben beziehungsweise Maßeinheiten und der Verwendung von unüblichen Begriffen erfolgreicher war, unterliefen DeepL weniger Fehler sprachlicher, idiomatischer und grammatikalischer Natur. Weiters ergab die Bewertung, dass die MÜ-Systeme mit den Rezepten von Gordon Ramsay besser umgehen konnten. Darüber hinaus bewerteten Teilnehmer*innen mit einer höheren Koch- beziehungsweise Backbegeisterung die Übersetzungen tendenziell schlechter. Entgegen der Annahme scheinen Personen mit einem höheren Grad an prozeduralem Wissen im Bereich der Kulinarik maschinelle Übersetzungen weniger zu akzeptieren und finden diese weniger gebräuchlich als Personen mit weniger ausgeprägten Kenntnissen. Um hierzu fundiertere Aussagen treffen zu können, ist weitere Forschungsarbeit erforderlich.

Somit gelangte die vorliegende Masterarbeit zu dem Ergebnis, dass die MÜ-Qualität bei der Übersetzung von Koch- und Backrezepten noch nicht ausreicht. Durch die kulturellen Besonderheiten von Rezepten als Textsorte und vor allem die terminologische Inkonsistenz kann keine problemlose Befolgung der Anleitungen gewährleistet werden. Zwar ist der Großteil der übersetzten Rezepte aus Sicht der Umfrageteilnehmer*innen gebräuchlich und verständlich, allerdings ist ein höheres Maß an Zeit und Konzentration erforderlich.

Zusammenfassend lässt sich aus den Ergebnissen der durchgeführten Umfrage schließen, dass die Miteinbeziehung der tatsächlichen Textsortenzielgruppe im Bewertungsprozess sinnvoll war und aufschlussreiche Ergebnisse lieferte. Die Zugehörigkeit der befragten Teilnehmer*innen sowohl zum Tätigkeitsbereich der Translation als auch zu anderen Branchen ermöglichte dabei eine differenzierte Betrachtung der Übersetzungsqualität. Im Hinblick auf zukünftige Forschung würde die Miteinbeziehung von weniger erfahrenen Köch*innen im Bewertungsprozess sowie die Gegenüberstellung der Bewertung von Koch- beziehungsweise Backbegeisterten und Lai*innen interessante und aufschlussreiche Ergebnisse liefern. Darüber hinaus könnte eine manuelle Analyse anhand einer adaptierten Fehlertypologie weitere interessante Aufschlüsse zu konkreten Problemen der MÜ bei der Übersetzung von Koch- und Backrezepten und zum genaueren Ausmaß gewisser Fehlerkategorien ergeben.

8. Literaturverzeichnis

Ahrend, Klaus (2006). Kriterien für die Bewertung von Fachübersetzungen. In: Schippel, Larisa (Hg.) *Übersetzungsqualität: Kritik – Kriterien – Bewertungshandeln*. Berlin: Frank & Timme, 31-42.

ALPAC (1966). *Language and Machines: Computers in Translation and Linguistics*. Washington: National Academy of Sciences, National Research Council.

Banerjee, Satanjeev & Lavie, Alon (2005). METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments. In: Goldstein, Jade; Lavie, Alon; Lin, Chin-Yeq & Voss, Clare (Hg.) *Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*, 65-72.

Bar-Hillel, Yehoshua (1960). The Present Status of Automatic Translation of Languages. In: Alt, Franz L. (Hg.) *Advances in Computers*. Band 1. New York, London: Academic Press, 91-163.

Burchardt, Aljoscha & Porsiel, Jörg (2017). Vorwort: Was kann maschinelle Übersetzung und was nicht? In: Porsiel, Jörg (Hg.) *Maschinelle Übersetzung. Grundlagen für den professionellen Einsatz*. Berlin: BDÜ Weiterbildungs- und Fachverlagsgesellschaft mbH, 11-17.

Castilho, Sheila; Doherty, Stephen; Gaspari, Federico & Moorkens, Joss (2018). Approaches to Human and Machine Translation Quality Assessment. In: Moorkens, Joss; Castilho, Sheila; Gaspari, Federico & Doherty, Stephen (Hg.) *Translation Quality Assessment – From Principles to Practice*. Dublin: Springer, 9-38.

Cölfen, Hermann (2007). Vom Kochrezept zur Kochanleitung: Sprachliche und mediale Aspekte einer verständlichen Vermittlung von Kochkenntnissen. *UNIKATE: Berichte aus Forschung und Lehre* 30, 84-93.

DeepL (2021). *Unternehmensprofil*. https://static.deepl.com/files/press/companyProfile_DE.pdf (Stand: 04.11.2024).

DeepL (2024a). *DeepL Übersetzer: Der präziseste Übersetzer der Welt*. https://www.deepl.com/de/translator (Stand: 04.11.2024).

DeepL (2024b). *DeepL Pro – schnelle, präzise und sichere Übersetzungen*. https://www.deepl.com/de/pro (Stand: 04.11.2024).

DeepL (2024c). *Warum DeepL?* https://www.deepl.com/de/why-deepl-pro (Stand: 04.11.2024).

Der Standard (2005). *„Österreichisches Deutsch“ existiert in der EU nur kulinarisch*. https://www.derstandard.at/story/1947362/oesterreichisches-deutsch-existiert-in-der-eu-nur-kulinarisch (Stand: 02.12.2024).

Devlin, Jacob; Chang, Ming-Wei; Lee, Kenton & Toutanova, Kristina (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: Burstein, Jill; Doran, Christy & Solorio, Thamar (Hg.) *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 4171-4186.

Donalies, Elke (2012). Man nehme ... Verbformen in Kochrezept oder Warum das Prototypische nicht immer das Typische ist. *Sprachreport* 2, 25-31.

Doss Mohan, Krishna & Skotdal, Jann (2021). *Microsoft Translator: Now translating 100 languages and counting!* https://www.microsoft.com/en-us/research/blog/microsoft-translator-now-translating-100-languages-and-counting/ (Stand: 14.11.2024).

Duden (2025). *Tranchieren*. https://www.duden.de/rechtschreibung/tranchieren (Stand: 31.01.2025).

EU (1995). *Amtsblatt der Europäischen Gemeinschaften, C 241, 29. August 1994. Dokumente betreffend den Beitritt des Königreichs Norwegen, der Republik Österreich, der Republik Finnland und des Königreichs Schweden zur Europäischen Union.* https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=OJ:C:1994:241:TOC (Stand: 02.12.2024).

Europäische Kommission (2025). *eTranslation*. https://commission.europa.eu/resources-partners/etranslation_de (Stand: 10.02.2025).

Forcada, Mikel L. (2017). Making sense of neural machine translation. *Translation Spaces* 6 (2), 291-309.

GER (2025). *Gemeinsamer Europäischer Referenzrahmen für Sprachen*. https://www.europaeischer-referenzrahmen.de/ (Stand: 24.04.2025).

Google (2024). *110 new languages are coming to Google Translate*. https://blog.google/products/translate/google-translate-new-languages-2024/ (Stand: 04.11.2024).

Google Übersetzer (2024a). *Google Übersetzer*. https://translate.google.com/ (Stand: 04.11.2024).

Google Übersetzer (2024b). *Über Google Übersetzer*. https://translate.google.com/about/ (Stand: 04.11.2024).

Hassan, Hany; Aue, Anthony; Chen, Chang; Chowdhary, Vishal; Clark, Jonathan; Federmann, Christian; Huang, Xuedong; Junczys-Dowmunt, Marcin; Lewis, William; Li, Mu; Liu, Shujie; Liu, Tie-Yan; Luo, Renqian; Menezes, Arul; Qin, Tao; Seide, Frank; Tan, Xu; Tian, Fei; Wu, Lijun; Wu, Shuangzhi; Xia, Yingce; Zhang, Dongdong; Zhang, Zhirui & Zhou, Ming (2018). Achieving Human Parity on Automatic Chinese to English News Translation. arXiv:1803.05567.

Heinisch, Barbara & Lušicky, Vesna (2019). User expectations towards machine translation: A case study. In: Forcada, Mikel; Way, Andy; Tinsley, John; Shterionov, Dimitar; Rico, Celia & Gaspari, Federico (Hg.) *Proceedings of the MT Summit XVII. Volume 2: Translator, Project and User Tracks*, 42-48.

Hiebl, Bettina & Gromann, Dagmar (2024). Comparative Quality Assessment of Human and Machine Translation with Best-Worst-Scaling. In: Scarton, Carolina; Prescott, Charlotte; Bayliss, Chris; Oakley, Chris; Wright, Joanna; Wrigley, Stuart; Song, Xingyi; Gow-Smith, Edward; Bawden, Rachel; Sánchez-Cartagena, Víctor M.; Cadwell, Patrick; Lapshinova-Koltunski, Ekaterina; Cabarrão, Vera; Chatzitheodorou, Konstantinos; Nurminen, Mary; Kanojia, Diptesh & Moniz, Helena (Hg.) *Proceedings of the 25th Annual Conference of the European Association for Machine Translation (Volume 1)*, 507-536.

Hutchins, William John (2007). Machine translation: a concise history. *Computer-Aided Translation: Theory and Practice* 13 (29-70), 1-21.

Hutchins, William John & Somers, Harold L. (1992). *An introduction to machine translation*. London: Academic Press.

Kenny, Dorothy (2018). Machine translation. In: Rawling, J. Piers & Wilson, Philip (Hg.) *The Routledge Handbook of Translation and Philosophy*. London: Routledge, 428-445.

Kenny, Dorothy (2022). Human and machine translation. In: Kenny, Dorothy (Hg.) *Machine translation for everyone: Empowering users in the age of artificial intelligence*. Berlin: Language Science Press, 23-49.

Koby, Geoffrey S.; Fields, Paul; Hague, Daryl R.; Lommel, Arle & Melby, Alan (2014). Defining Translation Quality. *Tradumàtica* 12, 413-420.

Kochwiki (2016). *Mixed Spice*. https://www.kochwiki.org/wiki/Mixed_Spice (Stand: 24.01.2025).

Koehn, Philipp (2020). *Neural Machine Translation*. Cambridge: Cambridge University Press.

LamaPoll (2025a). *Online Umfrage erstellen*. https://www.lamapoll.de/ (Stand: 24.04.2025).

LamaPoll (2025b). *Modelle & Preise*. https://www.lamapoll.de/Lizenz (Stand: 24.04.2025).

Langeder, Laura (o. J.). *Grammel und Powidl im EU-Beitrittsvertrag*. https://hdgoe.at/erdaepfelsalat_bleibt_erdaepfelsalat (Stand: 02.12.2024).

Larassati, Anisa; Setyaningsih, Nina; Nugroho, Raden Arief; Suryaningtyas, Valentina Widya; Cahyono, Setyo Prasiyanto & Pamelasari, Stephani Diah (2019). Google vs. Instagram Machine Translation: Multilingual Application Program Interface Errors in Translating Procedure Text Genre. In: *2019 International Seminar on Application for Technology of Information and Communication (iSemantic)*, 554-558.

Läubli, Samuel; Sennrich, Rico & Volk, Martin (2018). Has Machine Translation Achieved Human Parity? A Case for Document-level Evaluation. In: Riloff, Ellen; Chiang, David; Hockenmaier, Julia & Tsuji, Jun'chi (Hg.) *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 4791-4796.

LimeSurvey (2025). *LimeSurvey*. https://www.limesurvey.org/de (Stand: 24.04.2025).

Lommel, Arle; Uszkoreit, Hans & Burchardt, Aljoscha (2014). Multidimensional Quality Metrics (MQM): A Framework for Declaring and Describing Translation Quality Metrics. *Tradumàtica* 12, 455-463.

Lommel, Arle (2018). Metrics for Translation Quality Assessment: A Case for Standardising Error Typologies. In: Moorkens, Joss; Castilho, Sheila; Gaspari, Federico & Doherty, Stephen (Hg.) *Translation Quality Assessment – From Principles to Practice*. Dublin: Springer, 109-127.

Microsoft (2024a). *Microsoft Bing*. https://www.bing.com/translator/ (Stand: 14.11.2024).

Microsoft (2024b). *Microsoft Translator – Sprachen*. https://www.microsoft.com/de-de/translator/languages/ (Stand: 14.11.2024).

Microsoft (2024c). *Microsoft Translator*. https://www.microsoft.com/de-de/translator/ (Stand: 14.11.2024).

Nord, Christiane (2006). Translationsqualität aus funktionaler Sicht. In: Schippel, Larisa (Hg.) *Übersetzungsqualität: Kritik – Kriterien – Bewertungshandeln*. Berlin: Frank & Timme, 11-30.

Papineni, Kishore; Roukos, Salim; Ward, Todd & Zhu, Wie-Jing (2002). BLEU: a Method for Automatic Evaluation of Machine Translation. In: Isabelle, Pierre; Charniak, Eugene & Lin, Dekang (Hg.) *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, 311-318.

Poibeau, Thierry (2017). *Machine translation*. Cambridge, London: The MIT Press.

PONS (2025). *Poppadoms*. https://de.pons.com/%C3%BCbersetzung-2/englisch-deutsch/poppadoms (Stand: 16.01.2025).

Popel, Martin; Tomkova, Marketa; Tomek, Jakub; Kaiser, Łukasz; Uszkoreit, Jakob; Bojar, Ondřej & Žabokrtský, Zdeněk (2020). Transforming machine translation: a deep learning system reaches news translation quality comparable to human professionals. *Nature Communications* 11 (1), 1-15.

Reiß, Katharina & Vermeer, Hans J. (1984). *Grundlegung einer allgemeinen Translationstheorie*. Tübingen: Max Niemeyer Verlag.

Rivera-Trigueros, Irene (2022). Machine translation systems and quality assessment: a systematic review. *Language Resources and Evaluation* 56 (2), 593-619.

Rossi, Caroline & Carré, Alice (2022). How to choose a suitable neural machine translation solution: Evaluation of MT quality. In: Kenny, Dorothy (Hg.) *Machine translation for everyone: Empowering users in the age of artificial intelligence*. Berlin: Language Science Press, 51-79.

Schmalz, Antonia (2019). Maschinelle Übersetzung. In: Wittenpahl, Volker (Hg.) *Künstliche Intelligenz*. Berlin: Springer, 194-208.

Snover, Matthew; Dorr, Bonnie; Schwartz, Richard; Micciulla, Linnea & Makhoul, John (2006). A Study of Translation Edit Rate with Targeted Human Annotation. In:

Proceedings of the 7th Conference of the Association for Machine Translation in the Americas, 223-231.

Specia, Lucia; Raj, Dhwaj & Turchi, Marco (2010). Machine translation evaluation versus quality estimation. *Machine Translation* 24 (1), 39-50.

Specia, Lucia & Shah, Kashif (2018). Machine Translation Quality Estimation: Applications and Future Perspectives. In: Moorkens, Joss; Castilho, Sheila; Gaspari, Federico & Doherty, Stephen (Hg.) *Translation Quality Assessment – From Principles to Practice*. Dublin: Springer, 201-235.

Stein, Daniel (2009). Maschinelle Übersetzung – ein Überblick. *JLCL – Journal for Language Technology and Computational Linguistics* 24 (3), 5-18.

Stein, Daniel (2018). Machine Translation: Past, present and future. In: Rehm, Georg; Sasaki, Felix; Stein, Daniel & Witt, Andreas (Hg). *Language Technologies for a Multilingual Europe*. Berlin: Language Science Press, 5-17.

Survio (2025). *Survio*. http://survio.com/l-de-7-k3-2-umfrage-erstellen?campaignid=176334660&keywordid=kwd-43200881820&keyword=survio&matchtype=e&adgroupid=12965367900&adposition=&device=c&trc_cp=AT-DE&utm_source=google&utm_medium=cpc&utm_campaign=S-AT-DE-SEA&utm_term=survio&utm_campaign=S-AT-DE-SEA&utm_source=adwords&utm_medium=ppc&hsa_acc=8232418607&hsa_cam=176334660&hsa_grp=12965367900&hsa_ad=630693401725&hsa_src=g&hsa_tgt=kwd-43200881820&hsa_kw=survio&hsa_mt=e&hsa_net=adwords&hsa_ver=3&gad_source=1&gbraid=0AAAADzqj7yPCVlBzQmXXycHZIPEUpdU6&gclid=CjwKCAjwwqfABhBcEiwAZJjC3kwPxbo71tVZCuU3K02ZB5FsFbTFh6fzK_idQMkL9hNkHITAvTX-xBoCVbgQAvD_BwE (Stand: 24.04.2025).

SYSTRAN (2024a). *Mit SYSTRAN translate unten können Sie Ihren Text übersetzen*. https://www.systransoft.com/de/ubersetzen/?_gl=1*ocqpq2*_up*MQ (Stand: 07.11.2024).

SYSTRAN (2024b). *50 Jahre MT Innovation*. https://www.systransoft.com/de/systran/ (Stand: 07.11.2024).

SYSTRAN (2024c). *Unsere Geschichte – Mehr als 50 Jahre Erfahrung in der Entwicklung von maschinellen Übersetzungstechnologien*. https://www.systransoft.com/de/systran/ (Stand: 07.11.2024).

SYSTRAN (2024d). *Produkte – Eine Übersetzungslösung für jeden Bedarf*. https://www.systransoft.com/de/ubersetzungsprodukt/ (Stand: 07.11.2024).

The MQM Council (2024a). *The MQM Error Typology*. https://themqm.org/error-types-2/typology/ (Stand: 22.11.2024).

The MQM Council (2024b). *The MQM FULL Typology*. https://themqm.org/the-mqm-full-typology/ (Stand: 22.11.2024).

UmfrageOnline (2025). *UmfrageOnline*. https://www.umfrageonline.com/?cid=21908153696&aid=169134352934&kwd=umfrage%20erstellen&gad_source=1&gbraid=0AAAADshgI_OPapKY5eljzNLeJ02TBisD&gclid=CjwKCAjwwqfABhBcEiwAZJjC3rUnfm4GxzWdLwz0XiR7NS_aevOg1zaYg_xV78jqm_KA7F0gbrbkwhoCZVsQAvD_BwE (Stand: 24.04.2025).

Van Slype, Georges (1979). *Critical study of methods for evaluating the quality of machine translation. Final Report*. Brüssel: Bureau Marcel van Dijk.

Werthmann, Antonina & Witt, Andreas (2014). Maschinelle Übersetzung – Gegenwart und Perspektiven. In: Stickel, Gerhard (Hg.) *Translation and interpretation in Europe: Contributions to the Annual Conference 2013 of EFNIL in Vilnius*. Frankfurt am Main/Berlin/Bern/Brüssel/New York/Oxford/Wien: Lang, 79-103.

Wiesmann, Eva (2019). Machine Translation in the Field of Law: A Study of the Translation of Italian Legal Texts Into German. *Comparative Legilinguistics* 37 (1), 117-153.

Wolańska-Köller, Anna (2010). *Funktionaler Textaufbau und sprachliche Mittel in Kochrezepten des 19. und frühen 20. Jahrhunderts*. Stuttgart: Ibidem-Verlag.

Wu, Yonghui; Schuster, Mike; Chen, Zhifeng; Le, Quoc V.; Norouzi, Mohammad; Macherey, Wolfgang; Krikun, Maxim; Cao, Yuan; Gao, Qin; Macherey, Klaus; Klinger, Jeff; Shah, Apurva; Johnson, Melvin; Liu, Xiaobing; Kaiser, Łukasz; Gouws, Stephan; Kato, Yoshikiyo; Kudo, Taku; Kazawa, Hideto; Stevens, Keith; Kurian, George; Patil, Nishant; Wang, Wei; Young, Cliff; Smith, Jason; Riesa, Jason; Rudnick, Alex; Vinyals, Oriol; Corrado, Greg; Hughes, Macduff & Dean, Jeffrey (2016). Google's Neural

Machine Translation System: Bridging the Gap between Human and Machine Translation. *CORR* abs/1609.08144, 1-23.

Zenjob (2022). *Gastro-Lexikon (grundlegende Begriffe)*. https://www.zenjob.com/de/wp-content/uploads/sites/2/2022/07/Gastro-LexikonServiceTraining210623022.pdf (Stand: 06.12.2024).

9. Anhang A

Text 1:

Sweet potato, coconut & cardamom soup

With a red pepper & spinach topping

Ingredients

3 green cardamom pods
1 onion
600g sweet potato
4cm piece of ginger
2 cloves of garlic
2 tablespoons groundnut oil
1 x 400ml tin of low-fat coconut milk
800ml organic vegetable or chicken stock
1 lemon

RED PEPPER TOPPING

1 tablespoon groundnut oil
1 teaspoon coriander seeds
1 pinch of dried chilli flakes
optional: 30g coconut flakes
2 jarred roasted peppers
100g baby spinach
optional: 4 large or 12 mini poppadoms

Method

1. Bruise the cardamom pods and shake out the seeds, then roughly crush them with a pestle and mortar.
2. Peel and chop the onion and sweet potato, peel and finely grate the ginger, then peel and crush the garlic.
3. Heat the oil in a large pan over a low heat. Add the onion and a small pinch of sea salt and cook for 10 minutes, or until soft and sweet, stirring often.
4. Stir in the sweet potato, ginger, garlic and crushed cardamom seeds. Cook for 2 minutes, before adding the coconut milk.
5. Let it simmer for 1 to 2 minutes, then stir in the stock. Cover with a lid, and leave to simmer gently for 15 minutes.
6. Liquidise the soup with a stick blender until smooth, then season to taste with a pinch of salt and black pepper and a squeeze of lemon juice.
7. To make the red pepper topping, heat the oil in a frying pan, then roughly crush and add the coriander seeds and add the chilli flakes. Cook for 1 minute, or until lovely and fragrant.
8. In a dry pan, toast the coconut flakes (if using).
9. Rinse and finely slice the peppers, then add to the spices with the spinach. Continue cooking until the spinach has wilted down, then season and stir in the toasted coconut flakes.
10. Ladle the soup into bowls, finishing with a dollop of red pepper topping, then serve with broken up pieces of poppadoms on the side, if you like.

Quelle: https://www.jamieoliver.com/recipes/vegetables/sweet-potato-coconut-cardamom-soup/

Übersetzung 1 durch DeepL:

Suppe aus Süßkartoffeln, Kokosnuss und Kardamom

Mit roter Paprika und Spinat überbacken

Zutaten

3 grüne Kardamomkapseln
1 Zwiebel
600g Süßkartoffel
4cm Stück Ingwer
2 Knoblauchzehen
2 Esslöffel Erdnussöl
1 x 400ml Dose fettarme Kokosnussmilch
800 ml Bio-Gemüse- oder Hühnerbrühe
1 Zitrone

ROTER PFEFFER ALS BELAG

1 Esslöffel Erdnussöl
1 Teelöffel Koriandersamen
1 Prise getrocknete Chiliflocken
optional: 30 g Kokosnussflocken
2 gebratene Paprikaschoten aus dem Glas
100 g Babyspinat
wahlweise: 4 große oder 12 Mini-Poppadoms

Methode

1. Die Kardamomkapseln zerdrücken und die Samen herausschütteln, dann mit einem Mörser grob zerstoßen.
2. Die Zwiebel und die Süßkartoffel schälen und würfeln, den Ingwer schälen und fein reiben, den Knoblauch schälen und zerdrücken.
3. Das Öl in einer großen Pfanne auf kleiner Flamme erhitzen. Die Zwiebel und eine kleine Prise Meersalz hinzugeben und unter häufigem Rühren 10 Minuten kochen, bis sie weich und süß sind.
4. Die Süßkartoffel, den Ingwer, den Knoblauch und die zerstoßenen Kardamomsamen unterrühren. 2 Minuten kochen lassen, dann die Kokosmilch hinzufügen.
5. 1 bis 2 Minuten köcheln lassen, dann die Brühe einrühren. Mit einem Deckel abdecken und 15 Minuten sanft köcheln lassen.
6. Die Suppe mit einem Stabmixer glatt pürieren und mit einer Prise Salz und schwarzem Pfeffer sowie einem Spritzer Zitronensaft abschmecken.
7. Für den Paprikaaufstrich das Öl in einer Pfanne erhitzen, die Koriandersamen grob zerstoßen und die Chiliflocken hinzufügen. 1 Minute lang kochen, bis sie schön duften.
8. In einer trockenen Pfanne die Kokosflocken (falls verwendet) rösten.
9. Die Paprikaschoten waschen und in feine Scheiben schneiden, dann mit dem Spinat zu den Gewürzen geben. Weiter kochen, bis der Spinat verwelkt ist, dann würzen und die gerösteten Kokosflocken unterrühren.
10. Die Suppe in Schüsseln füllen, mit einem Klecks Paprika-Topping abschließen und nach Belieben mit zerkleinerten Poppadoms servieren.

Text 2:

Cauliflower Soup And Ham Cheese Toastie

Ingredients

Serves 4

For the soup

- 1 cauliflower
- 650ml milk
- 350ml chicken stock
- 1 sprig thyme
- Salt and pepper to taste

For the toastie

- Butter, for spreading and frying
- 8 slices white bread
- 12 slices Bayonne or Parma ham
- 40g Montgomery cheddar or other strong traditional cheddar cheese

For the garnish

- 75g-90g ceps, or 4 chestnut mushrooms
- Butter, for frying

Cooking instructions

To make the soup

1. Remove almost all the leaves from the cauliflower and cut it into even-sized pieces, including the stem
2. Put it in a large saucepan and cover with the milk and stock
3. Add the thyme and some seasoning and bring to a boil over a medium-high heat
4. Simmer steadily until the cauliflower is soft
5. Strain the cauliflower, reserving the cooking liquid
6. Tip the cauliflower into a blender or food processor and blend until smooth
7. Carefully add the cooking liquid, adjusting the consistency to your preference
8. Taste and add more salt and pepper as necessary, then pass the soup through a sieve into a saucepan and keep warm until ready to serve

To make the toastie

1. Butter each slice of bread on both sides
2. Lay three slices of the ham on each of four pieces of bread
3. Grate the cheese and sprinkle it evenly over the ham
4. Sandwich with the remaining bread
5. Heat some butter in a heavy frying pan and, once melted, lay one sandwich in the pan
6. Cook until golden on each side, pressing gently to help the layers stick together

7. Remove the toastie from the pan, press gently together once more and keep warm in a low oven while you cook the remaining sandwiches, using extra butter as necessary

To make the garnish

1. Cut the ceps into bite-sized pieces
2. Melt a knob of butter in a small frying pan and cook the ceps, stirring occasionally, for a few minutes, until browned

To serve

1. Divide the soup between four bowls
2. Top each with a spoonful of the ceps
3. Cut up the toasties as preferred then put them on base plates or side plates and serve

Quelle: https://www.gordonramsayrestaurants.com/recipes/cauliflower-soup-and-ham-cheese-toastie/

Übersetzung 2 durch Google Translate:

Blumenkohlsuppe und Schinken-Käse-Toastie

Zutaten

Für 4 Personen

Für die Suppe

- 1 Blumenkohl
- 650 ml Milch
- 350 ml Hühnerbrühe
- 1 Zweig Thymian
- Salz und Pfeffer nach Geschmack

Für das Toastie

- Butter zum Bestreichen und Braten
- 8 Scheiben Weißbrot
- 12 Scheiben Bayonner- oder Parmaschinken
- 40 g Montgomery Cheddar oder anderer kräftiger traditioneller Cheddar-Käse

Für die Garnitur

- 75–90 g Steinpilze oder 4 Kastanienpilze
- Butter zum Braten

Kochanleitung

Zur Zubereitung der Suppe

1. Entfernen Sie fast alle Blätter vom Blumenkohl und schneiden Sie ihn in gleich große Stücke, einschließlich des Strunks
2. Geben Sie ihn in einen großen Topf und bedecken Sie ihn mit Milch und Brühe
3. Fügen Sie den Thymian und etwas Gewürz hinzu und bringen Sie alles bei mittlerer bis hoher Hitze zum Kochen
4. Lassen Sie es köcheln, bis die Blumenkohl ist weich
5. Blumenkohl abgießen und Kochflüssigkeit aufbewahren
6. Blumenkohl in einen Mixer oder eine Küchenmaschine geben und glatt pürieren
7. Kochflüssigkeit vorsichtig hinzufügen und Konsistenz nach Belieben anpassen

8. Abschmecken und bei Bedarf mehr Salz und Pfeffer hinzufügen, dann die Suppe durch ein Sieb in einen Topf passieren und warm halten, bis sie serviert wird

Zur Zubereitung des Toasties

1. Jede Brotscheibe auf beiden Seiten mit Butter bestreichen
2. Auf jede der vier Brotscheiben drei Scheiben Schinken legen
3. Käse reiben und gleichmäßig über den Schinken streuen
4. Mit dem restlichen Brot belegen
5. Etwas Butter in einer schweren Bratpfanne erhitzen und, sobald sie geschmolzen ist, ein Sandwich in die Pfanne legen
6. Auf jeder Seite goldbraun braten, dabei leicht andrücken, damit die Schichten zusammenkleben
7. Toastie aus der Pfanne nehmen, noch einmal leicht zusammendrücken und in einem schwach geheizten Ofen warm halten, während Sie die restlichen Sandwiches braten, bei Bedarf zusätzliche Butter verwenden

Zur Garnierung

1. Steinpilze in mundgerechte Stücke schneiden
2. Ein Stück Butter in einer kleinen Pfanne schmelzen und die Steinpilze unter gelegentlichem Rühren einige Minuten braten, bis sie gebräunt sind

Zum Servieren

1. Die Suppe auf vier Schüsseln verteilen
2. Auf jede Schüssel einen Löffel Steinpilze geben
3. Die Toasties nach Belieben in Stücke schneiden, auf Unterteller oder Beilagenteller legen und servieren

Text 3:

Harissa Tomato Jam Toast with Poached Egg, Feta, Preserved Lemon

Ingredients

Harissa Tomato Jam:

- 2 pounds tomatoes, cored and diced
- 1 yellow onion, finely diced
- ¾ cup brown sugar, packed
- 3 tablespoons lemon juice
- 2 tablespoons harissa paste (more if you want an extra kick)
- 1 teaspoon ground coriander
- 1 teaspoon ground cumin
- 1 teaspoon kosher salt

Assembly:

- Hearty Artisan Bread, sliced and toasted
- Arugula
- Poached eggs
- Feta cheese, crumbled
- Preserved lemons, chopped
- Thinly shaved red onions, soaked in ice water for 5 minutes then patted dry
- Extra virgin olive oil
- Flaky salt
- Freshly cracked black pepper

Cooking instructions

1. Combine all ingredients in a medium pot and bring to a boil over medium-high heat. Lower heat and simmer rapidly for 55 minutes - 1 hour and 15 minutes. Once done, the jam should be thick and most of the moisture should have evaporated. The jam will also thicken slightly as it cools. Leftover cooled jam can be stored in an airtight container in the refrigerator for up to a week.
2. To assemble, spread a generous amount of tomato jam on toasted bread and top with dots of feta cheese and arugula, followed by preserved lemon and red onion. Nestle a poached egg on top and sprinkle with flaky salt and pepper. Drizzle with olive oil and dollop with more tomato jam, if desired.

Quelle: https://www.gordonramsay.com/gr/recipes/harissatoast/

Übersetzung 3 durch Google Translate:

Toast mit Harissa-Tomatenmarmelade, pochiertem Ei, Feta, eingelegter Zitrone

Zutaten

Harissa-Tomatenmarmelade:

- 2 Pfund Tomaten, entkernt und gewürfelt
- 1 gelbe Zwiebel, fein gewürfelt
- ¾ Tasse brauner Zucker, gepresst
- 3 Esslöffel Zitronensaft
- 2 Esslöffel Harissa-Paste (mehr, wenn Sie es besonders pikant mögen)
- 1 Teelöffel gemahlener Koriander
- 1 Teelöffel gemahlener Kreuzkümmel
- 1 Teelöffel koscheres Salz

Zusammenstellung:

- Herzhaftes Artisan-Brot, in Scheiben geschnitten und geröstet
- Rucola
- Pochierte Eier
- Fetakäse, zerbröselt
- Eingelegte Zitronen, gehackt
- Dünn gehobelte rote Zwiebeln, 5 Minuten in Eiswasser eingeweicht und dann trocken getupft
- Natives Olivenöl extra
- Flockensalz
- Frisch gemahlener schwarzer Pfeffer

Kochanleitung

1. Alle Zutaten in einem mittelgroßen Topf vermischen und bei mittlerer bis hoher Hitze zum Kochen bringen. Reduzieren Sie die Hitze und lassen Sie alles 55 Minuten bis 1 Stunde und 15 Minuten sprudelnd köcheln. Wenn dies erledigt ist, sollte die Marmelade dick sein und die meiste Feuchtigkeit verdunstet sein. Die Marmelade wird beim Abkühlen auch etwas dicker. Übrig gebliebene abgekühlte Marmelade kann bis zu einer Woche in einem luftdichten Behälter im Kühlschrank aufbewahrt werden.
2. Zum Anrichten eine großzügige Menge Tomatenmarmelade auf geröstetes Brot streichen und mit Fetakäse und Rucola belegen, gefolgt von eingelegter Zitrone und roten Zwiebeln. Ein pochiertes Ei darauf legen und mit Salz- und Pfefferflocken bestreuen. Mit Olivenöl beträufeln und nach Belieben noch mehr Tomatenmarmelade darauf geben.

Text 4:

SPANISH-STYLE AVOCADO TOAST WITH CHORIZO AND TOMATOES

Ingredients

Chorizo & Tomatoes:

- Olive oil
- 6 ounces Spanish hard chorizo, cut into small cubes
- ½ small yellow onion, finely diced (about ⅓ cup)
- 2 cloves garlic, minced
- 1 cup cherry tomatoes, halved
- Chicken stock
- Pedro Ximenez or sherry vinegar
- Parsley, chopped

To Serve:

- 1 ripe avocado, sliced thin, plus salt and lemon juice
- 2 thick slices sourdough or country bread, toasted in a pan with butter
- Frisée lettuce, tossed in olive oil, salt and Pedro Ximenez or sherry vinegar
- Aged Manchego cheese, shaved

Cooking instructions

1. Heat a large skillet over medium-high heat and add a splash of olive oil. Add chorizo and cook, stirring occasionally, until the edges are crispy. Lower the heat to medium and add onions. Cook until translucent, then add garlic and cook until fragrant, about 30 seconds.
2. Push all ingredients to the edges of the pan, then add the tomatoes, cut-side down if possible, to the empty space in the center of the pan and cook until they blister. Add a splash of chicken stock to deglaze, stirring to scrape up any browned bits on the pan. Simmer until the liquid has reduced a bit, check for seasoning and adjust if needed before removing from heat and stirring in a splash of Pedro Ximenez and most of the parsley.
3. Spread the avocado slices out between toasts, then sprinkle with salt and a squeeze of lemon juice. Divide chorizo mixture between the two toasts, then top with a few leaves of dressed frisée. Generously shave Manchego cheese all over, sprinkle with remaining parsley and serve.

Quelle: https://www.gordonramsay.com/gr/recipes/avocadotoastwithchorizo/

Übersetzung 4 durch DeepL:

AVOCADO-TOAST NACH SPANISCHER ART MIT CHORIZO UND TOMATEN

Zutaten

Chorizo und Tomaten:

- Olivenöl
- 6 Unzen spanische harte Chorizo, in kleine Würfel geschnitten
- ½ kleine gelbe Zwiebel, fein gewürfelt (etwa ⅓ Tasse)
- 2 Knoblauchzehen, gehackt
- 1 Tasse Kirschtomaten, halbiert
- Hühnerbrühe
- Pedro Ximenez oder Sherry-Essig
- Petersilie, gehackt

Zum Servieren:

- 1 reife Avocado, in dünne Scheiben geschnitten, plus Salz und Zitronensaft
- 2 dicke Scheiben Sauerteig- oder Landbrot, in einer Pfanne mit Butter geröstet
- Frisée-Salat, in Olivenöl, Salz und Pedro Ximenez- oder Sherry-Essig geschwenkt
- Gereifter Manchego-Käse, gehobelt

Anweisungen zum Garen

1. Eine große Pfanne auf mittlerer bis hoher Stufe erhitzen und einen Spritzer Olivenöl hineingeben. Chorizo hinzufügen und unter gelegentlichem Rühren braten, bis die Ränder knusprig sind. Die Hitze auf mittlere Stufe reduzieren und die Zwiebeln hinzufügen. Braten, bis sie glasig sind, dann den Knoblauch hinzufügen und etwa 30 Sekunden lang braten, bis er duftet.
2. Alle Zutaten an den Rand der Pfanne schieben, dann die Tomaten, wenn möglich mit der Schnittfläche nach unten, in die Mitte der Pfanne geben und kochen, bis sie Blasen werfen. Mit einem Spritzer Hühnerbrühe ablöschen, dabei umrühren, um alle gebräunten Stücke aus der Pfanne zu kratzen. Köcheln lassen, bis die Flüssigkeit etwas eingekocht ist, die Gewürze überprüfen und bei Bedarf anpassen, dann vom Herd nehmen und einen Spritzer Pedro Ximenez und den größten Teil der Petersilie unterrühren.
3. Die Avocadoscheiben auf den Toasts verteilen, mit Salz und einem Spritzer Zitronensaft bestreuen. Die Chorizo-Mischung auf die beiden Toasts verteilen und mit ein paar Blättern Frisée belegen. Manchego-Käse großzügig darüber raspeln, mit der restlichen Petersilie bestreuen und servieren.

Text 5:

Beef Wellington

With Madeira gravy

Ingredients

1kg centre fillet of beef, trimmed (the timings below work perfectly for a fillet of roughly 10cm in diameter)
olive oil
2 knobs of unsalted butter
3 sprigs of fresh rosemary
1 red onion
2 cloves of garlic
600g mixed mushrooms
100g free-range chicken livers (cleaned)
1 tablespoon Worcestershire sauce
optional: ½ teaspoon truffle oil
50g fresh breadcrumbs
1 x 500g block of puff pastry
1 large free-range egg

GRAVY

2 onions
4 sprigs of fresh thyme
1 heaped teaspoon blackcurrant jam
100ml Madeira wine
1 heaped teaspoon English mustard
2 heaped tablespoons plain flour, plus extra for dusting
600ml organic beef stock (hot)

Method

1. Preheat a large frying pan on a high heat. Rub the beef all over with sea salt and black pepper. Pour a good lug of oil into the pan, then add the beef, 1 knob of butter and 1 sprig of rosemary.
2. Sear the beef for 4 minutes in total, turning regularly with tongs, then remove to a plate.
3. Wipe out the pan and return to a medium heat. Peel the onion and garlic, then very finely chop with the mushrooms and put into the pan with the remaining knob of butter and another lug of oil.
4. Strip in the rest of the rosemary leaves and cook for 15 minutes, or until soft and starting to caramelise, stirring regularly.
5. Toss the livers and Worcestershire sauce into the pan and cook for another few minutes, then tip the contents onto a large board and drizzle with the truffle oil (if using).
6. Finely chop it all by hand with a big knife, to a rustic, spreadable consistency. Taste and season to perfection, then stir in the breadcrumbs (you can use pancakes to line the pastry and absorb the juices, but I prefer using breadcrumbs like this).
7. Preheat the oven to 210°C/425°F/gas 7.
8. On a flour-dusted surface, roll out the pastry to 30cm x 40cm. With one of the longer edges in front of you, spread the mushroom pâté over the pastry, leaving a 5cm gap at either end and at the edge furthest away from you – eggwash these edges.
9. Sit the beef on the pâté then, starting with the edge nearest to you, snugly wrap the pastry around the beef, pinching the ends to seal.

10. Transfer the Wellington to a large baking tray lined with greaseproof paper, with the pastry seal at the base, and brush all over with eggwash (you can prep to this stage, then refrigerate until needed – just get it out 1½ hours before cooking so it's not fridge-cold).
11. When you're ready to cook, heat the tray on the hob for a couple of minutes to start crisping up the base, then transfer to the oven and cook for 40 minutes for blushing, juicy beef – the two end portions will be more cooked, but usually some people prefer that.
12. For the gravy, peel and roughly chop the onions and put into a large pan on a medium heat with a lug of oil and the thyme leaves. Cook for 20 minutes, stirring occasionally, then stir in the jam and simmer until shiny and quite dark.
13. Add the Madeira, flame with a match, cook away, then stir in the mustard and flour, gradually followed by the stock. Simmer to the consistency you like, then blend with a stick blender and pass through a sieve, or leave chunky.
14. Once cooked, rest the Wellington for 5 minutes, then serve in 2cm-thick slices with the gravy and steamed greens.

Quelle: https://www.jamieoliver.com/recipes/beef/beef-wellington/

Übersetzung 5 durch Google Translate:

Beef Wellington

Mit Madeira-Soße

Zutaten

1 kg Rinderfilet, pariert (die unten angegebenen Zeitangaben sind perfekt für ein Filet mit etwa 10 cm Durchmesser)
Olivenöl
2 Stücke ungesalzene Butter
3 Zweige frischer Rosmarin
1 rote Zwiebel
2 Knoblauchzehen
600 g gemischte Champignons
100 g Leber von Hühnern aus Freilandhaltung (geputzt)
1 Esslöffel Worcestershire-Soße
optional: ½ Teelöffel Trüffelöl
50 g frische Semmelbrösel
1 x 500 g Blätterteigblock
1 großes Ei aus Freilandhaltung

SOßE

2 Zwiebeln
4 Zweige frischer Thymian
1 gehäufter Teelöffel schwarze Johannisbeermarmelade
100 ml Madeirawein
1 gehäufter Teelöffel englischer Senf
2 gehäufte Esslöffel Mehl, plus etwas mehr zum Bestäuben
600 ml Bio-Rinderbrühe (heiß)

Zubereitung

1. Eine große Bratpfanne auf hoher Flamme vorheizen. Das Rindfleisch rundherum mit Meersalz und schwarzem Pfeffer einreiben. Einen guten Schuss Öl in die Pfanne geben, dann das Rindfleisch, 1 Stück Butter und 1 Zweig Rosmarin hinzufügen.

2. Das Rindfleisch insgesamt 4 Minuten anbraten, dabei regelmäßig mit einer Zange wenden, dann auf einen Teller legen.
3. Die Pfanne auswischen und wieder auf mittlere Hitze stellen. Zwiebel und Knoblauch schälen, dann mit den Pilzen sehr fein hacken und mit dem restlichen Stück Butter und einem weiteren Schuss Öl in die Pfanne geben.
4. Den Rest der Rosmarinblätter abstreifen und 15 Minuten kochen, oder bis sie weich sind und anfangen zu karamellisieren, dabei regelmäßig umrühren.
5. Die Lebern und die Worcestershire-Sauce in die Pfanne geben und noch ein paar Minuten kochen, dann den Inhalt auf ein großes Brett kippen und mit dem Trüffelöl (falls verwendet) beträufeln.
6. Alles mit einem großen Messer von Hand fein hacken, bis eine rustikale, streichfähige Konsistenz entsteht. Abschmecken und perfekt würzen, dann die Semmelbrösel unterrühren (Sie können den Teig auch mit Pfannkuchen auslegen und den Saft aufsaugen, aber ich verwende lieber Semmelbrösel wie diese).
7. Den Backofen auf 210 °C/425 °F/Gasherdstufe 7 vorheizen.
8. Den Teig auf einer mit Mehl bestäubten Fläche auf 30 cm x 40 cm ausrollen. Mit einer der längeren Kanten vor Ihnen die Pilzpastete darauf verteilen und dabei an beiden Enden und an der Kante, die am weitesten von Ihnen entfernt ist, einen Abstand von 5 cm lassen – diese Kanten mit Ei bestreichen.
9. Das Rindfleisch auf die Pastete legen und dann, beginnend mit der Kante, die Ihnen am nächsten ist, den Teig eng um das Rindfleisch wickeln und die Enden zusammendrücken, um es abzudichten.
10. Legen Sie das Wellington auf ein großes, mit Backpapier ausgelegtes Backblech, mit der Teignaht nach unten, und bestreichen Sie es rundherum mit Eigelb (Sie können es bis zu diesem Schritt vorbereiten und dann bis zur Verwendung im Kühlschrank aufbewahren – nehmen Sie es einfach 1½ Stunden vor dem Kochen heraus, damit es nicht kühlschrankkalt ist).
11. Wenn Sie bereit zum Kochen sind, erhitzen Sie das Blech ein paar Minuten auf dem Herd, damit der Boden knusprig wird, geben Sie es dann in den Ofen und lassen Sie es 40 Minuten lang backen, bis Sie ein rötliches, saftiges Rindfleisch bekommen – die beiden Endstücke sind dann stärker gegart, aber manche Leute bevorzugen das normalerweise.
12. Für die Soße die Zwiebeln schälen, grob hacken und in eine große Pfanne bei mittlerer Hitze mit einem Schuss Öl und den Thymianblättern geben. 20 Minuten unter gelegentlichem Umrühren kochen, dann die Marmelade einrühren und köcheln lassen, bis sie glänzt und ziemlich dunkel ist.
13. Den Madeira hinzufügen, mit einem Streichholz anzünden, aufkochen lassen, dann Senf und Mehl einrühren, nach und nach die Brühe. Bis zur gewünschten Konsistenz köcheln lassen, dann mit einem Stabmixer pürieren und durch ein Sieb passieren oder stückig lassen.
14. Nach dem Garen das Wellington 5 Minuten ruhen lassen, dann in 2 cm dicken Scheiben mit der Soße und gedünstetem Gemüse servieren.

Text 6:

Beef Wellington

Ingredients

- 2 x 400g beef fillets
- Olive oil, for frying
- 500g mixture of wild mushrooms, cleaned
- 1 thyme sprig, leaves only
- 500g puff pastry
- 8 slices of Parma ham
- 2 egg yolks, beaten with 1 tbsp water and a pinch of salt
- Sea salt and freshly ground black pepper

For the red wine sauce

- 2 tbsp olive oil
- 200g beef trimmings (ask the butcher to reserve these when trimming the fillet)
- 4 large shallots, peeled and sliced
- 12 black peppercorns
- 1 bay leaf
- 1 thyme sprig
- Splash of red wine vinegar
- 1 x 750ml bottle red wine
- 750ml beef stock

Method

Serves 4

1. Wrap each piece of beef tightly in a triple layer of cling film to set its shape, then chill overnight.
2. Remove the cling film, then quickly sear the beef fillets in a hot pan with a little olive oil for 30-60 seconds until browned all over and rare in the middle. Remove from the pan and leave to cool.
3. Finely chop the mushrooms and fry in a hot pan with a little olive oil, the thyme leaves and some seasoning. When the mushrooms begin to release their juices, continue to cook over a high heat for about 10 minutes until all the excess moisture has evaporated and you are left with a mushroom paste (known as a duxelle). Remove the duxelle from the pan and leave to cool.
4. Cut the pastry in half, place on a lightly floured surface and roll each piece into a rectangle large enough to envelop one of the beef fillets. Chill in the refrigerator.
5. Lay a large sheet of cling film on a work surface and place 4 slices of Parma ham in the middle, overlapping them slightly, to create a square. Spread half the duxelle evenly over the ham.

6. Season the beef fillets, then place them on top of the mushroom-covered ham. Using the cling film, roll the Parma ham over the beef, then roll and tie the cling film to get a nice, evenly thick log. Repeat this step with the other beef fillet, then chill for at least 30 minutes.
7. Brush the pastry with the egg wash. Remove the cling film from the beef, then wrap the pastry around each ham-wrapped fillet. Trim the pastry and brush all over with the egg wash. Cover with cling film and chill for at least 30 minutes.
8. Meanwhile, make the red wine sauce. Heat the oil in a large pan, then fry the beef trimmings for a few minutes until browned on all sides. Stir in the shallots with the peppercorns, bay and thyme and continue to cook for about 5 minutes, stirring frequently, until the shallots turn golden brown.
9. Pour in the vinegar and let it bubble for a few minutes until almost dry. Now add the wine and boil until almost completely reduced. Add the stock and bring to the boil again. Lower the heat and simmer gently for 1 hour, removing any scum from the surface of the sauce, until you have the desired consistency. Strain the liquid through a fine sieve lined with muslin. Check for seasoning and set aside.
10. When you are ready to cook the beef wellingtons, score the pastry lightly and brush with the egg wash again, then bake at 200°C/Gas 6 for 15-20 minutes until the pastry is golden brown and cooked. Rest for 10 minutes before carving.
11. Meanwhile, reheat the sauce. Serve the beef wellingtons sliced, with the sauce as an accompaniment.

Quelle: https://www.gordonramsay.com/gr/recipes/beef-wellington/

Übersetzung 6 durch DeepL:

Rindfleisch-Wellington

Zutaten

- 2 x 400g Rinderfilets
- Olivenöl, zum Braten
- 500 g gemischte Waldpilze, geputzt
- 1 Thymianzweig, nur die Blätter
- 500 g Blätterteig
- 8 Scheiben Parmaschinken
- 2 Eigelb, verquirlt mit 1 Esslöffel Wasser und einer Prise Salz
- Meersalz und frisch gemahlener schwarzer Pfeffer

Für die Rotweinsauce

- 2 Esslöffel Olivenöl
- 200 g Rindfleischabschnitte (bitten Sie den Metzger, diese beim Abschneiden des Filets zu reservieren)
- 4 große Schalotten, geschält und in Scheiben geschnitten
- 12 schwarze Pfefferkörner
- 1 Lorbeerblatt
- 1 Thymianzweig
- Ein Spritzer Rotweinessig
- 1 x 750ml Flasche Rotwein
- 750 ml Rinderbrühe

Methode

Für 4 Personen

1. Jedes Stück Rindfleisch fest in eine dreifache Lage Frischhaltefolie einwickeln, damit es seine Form behält, und dann über Nacht kühl stellen.
2. Die Frischhaltefolie entfernen und die Rinderfilets in einer heißen Pfanne mit etwas Olivenöl 30-60 Sekunden lang anbraten, bis sie von allen Seiten gebräunt und in der Mitte blutig sind. Aus der Pfanne nehmen und abkühlen lassen.
3. Die Champignons fein hacken und in einer heißen Pfanne mit etwas Olivenöl, den Thymianblättern und etwas Gewürz anbraten. Wenn die Pilze anfangen, ihren Saft abzugeben, etwa 10 Minuten bei starker Hitze weiterbraten, bis die gesamte überschüssige Feuchtigkeit verdampft ist und eine Pilzpaste (Duxelle) entsteht. Die Duxelle aus der Pfanne nehmen und abkühlen lassen.
4. Den Teig halbieren, auf eine leicht bemehlte Fläche legen und jedes Stück zu einem Rechteck ausrollen, das groß genug ist, um eines der Rinderfilets zu umhüllen. Im Kühlschrank abkühlen lassen.
5. Ein großes Blatt Frischhaltefolie auf eine Arbeitsfläche legen und 4 Scheiben Parmaschinken leicht überlappend in die Mitte legen, so dass ein Quadrat entsteht. Die Hälfte der Duxelle gleichmäßig auf dem Schinken verteilen.
6. Die Rinderfilets würzen und auf den mit Pilzen bedeckten Schinken legen. Rollen Sie den Parmaschinken mit Hilfe der Frischhaltefolie über das Rindfleisch, rollen Sie die Folie zusammen und binden Sie sie fest, so dass ein schöner, gleichmäßig dicker Block entsteht. Diesen Schritt mit dem anderen Rinderfilet wiederholen, dann mindestens 30 Minuten lang kühlen.
7. Den Teig mit dem Ei bestreichen. Die Frischhaltefolie vom Rindfleisch entfernen, dann den Teig um jedes mit Schinken umwickelte Filet wickeln. Den Teig zuschneiden und rundherum mit Ei bestreichen. Mit Frischhaltefolie abdecken und mindestens 30 Minuten kühl stellen.
8. In der Zwischenzeit die Rotweinsoße zubereiten. Das Öl in einer großen Pfanne erhitzen und die Rinderfilets einige Minuten lang anbraten, bis sie von allen Seiten gebräunt sind. Die Schalotten mit den Pfefferkörnern, dem Lorbeer und dem Thymian dazugeben und unter häufigem Rühren etwa 5 Minuten weiterbraten, bis die Schalotten goldbraun werden.
9. Den Essig dazugießen und einige Minuten sprudelnd kochen lassen, bis er fast trocken ist. Nun den Wein hinzufügen und einkochen lassen, bis er fast vollständig reduziert ist. Die Brühe hinzugeben und erneut zum Kochen bringen. Die Hitze reduzieren und die Sauce 1 Stunde lang sanft köcheln lassen, dabei den Schaum von der Oberfläche der Sauce entfernen, bis sie die gewünschte Konsistenz hat. Die Flüssigkeit durch ein feines, mit Musselin ausgelegtes Sieb abseihen. Nachwürzen und beiseite stellen.
10. Wenn Sie bereit sind, die Wellingtons zuzubereiten, ritzen Sie den Teig leicht ein und bepinseln Sie ihn erneut mit der Eimasse, dann backen Sie ihn bei 200°C/Gas 6 für 15-20 Minuten, bis der Teig goldbraun und gar ist. Vor dem Tranchieren 10 Minuten ruhen lassen.
11. In der Zwischenzeit die Soße erneut erhitzen. Servieren Sie die in Scheiben geschnittenen Wellingtons mit der Sauce dazu.

Text 7:

Pineapple upside-down cake

Glacé cherries & desiccated coconut

Ingredients

200ml runny honey
2 x 425g tins of pineapple rings in juice
200ml vegetable oil
200g self-raising flour
500ml natural yoghurt
200g desiccated coconut
2 large eggs
12 glacé cherries (60g)

Method

1. Preheat the oven to 180°C/350°F/gas 4. Place your largest non-stick frying pan on a medium-high heat with 1 tablespoon of the honey, then arrange 12 pineapple rings on top in a single layer, reserving the juice for later. Cook for 4 to 5 minutes on just one side, or until the honey has almost disappeared and the pineapple is starting to caramelise.
2. Place the remaining honey into a large mixing bowl with the vegetable oil, flour, 200ml of yoghurt and most of the desiccated coconut. Crack in the eggs and use a spatula to mix together, then loosen with 2 tablespoons of the reserved pineapple juice.
3. Line a 30cm x 20cm roasting tray with greaseproof paper. Arrange the pineapple rings in the tray caramelised-side down and so they're slightly overlapping, drizzling over any remaining pan juices, then place a glacé cherry in the centre of each ring. Pour over the batter, level off the top, then bake for 30 minutes, or until beautifully golden.
4. Carefully turn the cake out onto a board and peel off the greaseproof. Top with a dollop of yoghurt and sprinkle over the reserved desiccated coconut, to serve.

Quelle: https://www.jamieoliver.com/recipes/desserts/pineapple-upside-down-cake/

Übersetzung 7 durch Google Translate:

Ananaskuchen mit Deckel

Kandierte Kirschen und getrocknete Kokosnüsse

Zutaten

200 ml flüssiger Honig
2 Dosen Ananasringe in Saft (à 425 g)
200 ml Pflanzenöl
200 g Mehl
500 ml Naturjoghurt
200 g getrocknete Kokosnüsse
2 große Eier
12 kandierte Kirschen (60 g)

Zubereitung

1. Den Backofen auf 180 °C/350 °F/Gas 4 vorheizen. Die größte antihaftbeschichtete Bratpfanne mit 1 Esslöffel Honig auf mittlerer bis hoher Hitze erhitzen und dann 12 Ananasringe in einer Lage darauf verteilen. Den Saft für später aufbewahren. 4 bis 5 Minuten auf nur einer Seite braten, oder bis der Honig fast verschwunden ist und die Ananas anfängt zu karamellisieren.
2. Den restlichen Honig mit dem Pflanzenöl, Mehl, 200 ml Joghurt und dem Großteil der Kokosraspeln in eine große Rührschüssel geben. Die Eier aufschlagen und mit einem Spatel verrühren, dann mit 2 Esslöffeln des beiseite gestellten Ananassafts auflockern.
3. Eine 30 x 20 cm große Bratpfanne mit Backpapier auslegen. Die Ananasringe mit der karamellisierten Seite nach unten und so, dass sie sich leicht überlappen, in die Pfanne legen, den restlichen Bratensaft darüberträufeln und in die Mitte jedes Rings eine kandierte Kirsche legen. Den Teig darübergießen, die Oberfläche glattstreichen und 30 Minuten backen, oder bis sie schön goldbraun sind.
4. Den Kuchen vorsichtig auf ein Brett stürzen und die Backfolie abziehen. Zum Servieren einen Klecks Joghurt darauf geben und die beiseite gelegten Kokosraspeln darüber streuen.

Text 8:

Vegan toffee apple upside-down cake

Mixed spice, lemon & walnuts

Ingredients

25g vegan margarine, plus extra for greasing
3 dessert apples
195g muscovado sugar
180g plain flour
1 teaspoon bicarbonate of soda
1½ teaspoons mixed spice
80ml sunflower oil
1 teaspoon vinegar
1 lemon
85g shelled walnuts

Method

Sticky-sweet and topped with lightly spiced apples, this makes the perfect afternoon tea treat!

1. Preheat the oven to 180ºC/350ºF/gas 4, and grease and line the base of a 23cm square cake tin.
2. Core and coarsely grate 2 of the apples and finely slice the remaining apple.
3. Melt 85g of the sugar and the margarine in a pan, then pour into the prepared tin. Top with the sliced apple in a single layer.
4. Combine the flour, 110g of sugar, the bicarbonate of soda and mixed spice in a bowl. In a separate bowl, combine the oil, 180ml water, the vinegar, grated apple and lemon zest. Mix the dry ingredients with the wet, quickly but thoroughly.
5. Roughly chop and stir in the walnuts, then pour over the layer of apples in the cake tin.
6. Bake for 30 minutes, or until a skewer comes out clean. Leave the cake to cool for 5 minutes before turning out.

Quelle: https://www.jamieoliver.com/recipes/fruit/vegan-toffee-apple-upside-down-cake/

Übersetzung 8 durch DeepL:

Veganer Toffee-Apfel-Upside-Down-Kuchen

Gewürzmischung, Zitrone & Walnüsse

Zutaten

25 g vegane Margarine, plus extra zum Einfetten
3 Dessert-Äpfel
195 g Muscovado-Zucker
180 g normales Mehl
1 Teelöffel Natriumbikarbonat
1½ Teelöffel Gewürzmischung
80 ml Sonnenblumenöl
1 Teelöffel Essig
1 Zitrone
85 g geschälte Walnüsse

Methode

Klebrig-süß und mit leicht gewürzten Äpfeln bestreut, ist dies der perfekte Genuss für den Nachmittagstee!

1. Den Ofen auf 180ºC/350ºF/Gas 4 vorheizen und den Boden einer 23 cm großen, quadratischen Kuchenform einfetten und auslegen.
2. 2 der Äpfel entkernen und grob raspeln, den restlichen Apfel in feine Scheiben schneiden.
3. 85 g Zucker und Margarine in einer Pfanne schmelzen und in die vorbereitete Form geben. Die Apfelscheiben in einer einzigen Schicht darauf verteilen.
4. Das Mehl, 110 g Zucker, das Natron und die Gewürzmischung in einer Schüssel mischen. In einer anderen Schüssel das Öl, 180 ml Wasser, den Essig, den geriebenen Apfel und die Zitronenschale vermischen. Die trockenen Zutaten schnell, aber gründlich mit den feuchten vermischen.
5. Die Walnüsse grob hacken und unterrühren, dann über die Apfelschicht in der Kuchenform gießen.
6. 30 Minuten backen, oder bis ein Spieß sauber herauskommt. Lassen Sie den Kuchen 5 Minuten abkühlen, bevor Sie ihn stürzen.

10. Anhang B

Wahrnehmung der Übersetzungsqualität von Koch- bzw. Backrezepten verschiedener Systeme der neuronalen maschinellen Übersetzung durch Kochbegeisterte

Liebe Teilnehmer*innen,
herzlich willkommen zur Umfrage! Mein Name ist Melanie Leko und ich schreibe gerade meine Masterarbeit im Studium Translation am Zentrum für Translationswissenschaft der Universität Wien. Diese Umfrage ist Teil der Masterarbeit.
Im Rahmen der Umfrage möchte ich herausfinden, wie verständlich und praktisch Kochbegeisterte maschinell erstellte Übersetzungen von Koch- bzw. Backrezepten finden. Dafür werden Sie zunächst in einem allgemeinen Teil zu Ihrer Person, Koch-/Backerfahrung und Erfahrung mit maschineller Übersetzung befragt. Danach kommen Sie zum Bewertungsteil: Hierzu wurden insgesamt acht Rezepte von den berühmten britischen Köchen Gordon Ramsay und Jamie Oliver herangezogen. Diese Rezepte wurden maschinell aus dem Englischen ins Deutsche übersetzt - Sie bekommen dabei nur die deutsche Version zu sehen. Sie werden gebeten, die Übersetzungen anhand von allgemeineren Fragen bzw. Aussagen zu bewerten und zu kommentieren. Es geht vor allem um die Frage, ob die übersetzten Rezepte aus Ihrer Sicht verständlich und in der Praxis umsetzbar sind. Zum Schluss können Sie noch Ihre Gedanken zur Umfrage sowie zum Thema im Allgemeinen mitteilen.
Voraussetzung für die Teilnahme: Die Umfrage richtet sich an Personen, die Deutsch gut beherrschen und mit grundlegenden Koch- bzw. Backtechniken vertraut sind.
Die Umfrage sollte etwa 40 Minuten Ihrer Zeit beanspruchen.
Vielen Dank für Ihre Unterstützung - ich freue mich auf Ihre Rückmeldungen!

Allgemeine Fragen

Wie alt sind Sie? *

Bitte wählen Sie die zutreffende Antwort aus.

- ◯ unter 18
- ◯ 18-24
- ◯ 25-34
- ◯ 35-44
- ◯ 45-54
- ◯ 55-64
- ◯ 65 und älter

Welchem Geschlecht fühlen Sie sich zugehörig? *

Bitte wählen Sie die zutreffende Antwort aus.

- ◯ Männlich
- ◯ Weiblich
- ◯ Divers

Was ist Ihr höchster Bildungsabschluss? *

Bitte wählen Sie die zutreffende Antwort aus.

- ◯ Kein Schulabschluss
- ◯ Pflichtschule
- ◯ Lehrabschluss
- ◯ Matura
- ◯ Universität bzw. Hochschule (Bachelor)
- ◯ Universität bzw. Hochschule (Master)
- ◯ Universität bzw. Hochschule (Magister)
- ◯ Universität bzw. Hochschule (Doktorat)
- ◯ Anderer Abschluss nach Matura

Sind Sie professionelle*r Translator*in oder studieren Sie im Bereich der Transkulturellen Kommunikation bzw. Translation? *

- ◯ Ja
- ◯ Nein

Was sind Ihre Arbeitssprachen? *

Bitte füllen Sie die Felder entsprechend aus. Sollten Sie zum Beispiel mehrere Zweitsprachen haben, geben Sie bitte alle an.

Erst- bzw. Bildungssprache (A-Sprache):

Zweitsprache (B-Sprache):

Drittsprache (C-Sprache):

Welche weiteren Sprachen beherrschen Sie und auf welchem Niveau? *

Bitte geben Sie alle weiteren Sprachkenntnisse, die Sie nicht zwingend im Beruf verwenden, mit dem entsprechenden Niveau (A1, A2, B1, B2, C1, C2) an. Eine detaillierte Erklärung der Sprachniveaus finden Sie unter: https://www.europaeischer-referenzrahmen.de/
Wenn Sie keine weiteren Sprachen beherrschen, schreiben Sie bitte "keine" ins Textfeld.

In welcher Branche sind Sie tätig? *

Bitte wählen Sie die zutreffende Antwort aus.

- ◯ Tourismus, Gastgewerbe
- ◯ Bildung
- ◯ Forschung
- ◯ Umwelt
- ◯ Soziales, Gesundheit
- ◯ Medien, Grafik, Design, Kunst
- ◯ IT
- ◯ Handel, Logistik
- ◯ Marketing
- ◯ Psychologie
- ◯ Finanz, Wirtschaft
- ◯ Recht
- ◯ Lebensmittel
- ◯ Baugewerbe
- ◯ Sprachen, Kultur
- ◯ Sonstiges:

Ist Deutsch Ihre Erstsprache?

- ◯ Ja
- ◯ Nein

Was ist Ihre Erstsprache? *

Bitte geben Sie Ihre Erstsprache an.

Wie hoch schätzen Sie Ihr Sprachniveau auf Deutsch ein?

Bitte wählen Sie die zutreffende Antwort aus. Eine detaillierte Erklärung der Sprachniveaus finden Sie unter: https://www.europaeischer-referenzrahmen.de/

- ◯ A1 - Anfänger
- ◯ A2 - Grundlegende Kenntnisse
- ◯ B1 - Fortgeschrittene Sprachverwendung
- ◯ B2 - Selbständige Sprachverwendung
- ◯ C1 - Fachkundige Sprachkenntnisse
- ◯ C2 - Annähernd muttersprachliche Kenntnisse

Koch- bzw. Backerfahrung

Wie oft kochen Sie nach Rezept? *

Bitte wählen Sie die zutreffende Antwort aus.

- ◯ täglich
- ◯ mehrmals in der Woche
- ◯ mehrmals im Monat
- ◯ mehrmals im Jahr

Wie oft backen Sie nach Rezept? *

Bitte wählen Sie die zutreffende Antwort aus.

- ◯ täglich
- ◯ mehrmals in der Woche
- ◯ mehrmals im Monat
- ◯ mehrmals im Jahr

Wie vertraut sind Sie mit Kochtechniken? *

Auf einer Skala von 1 (= nicht vertraut) bis 5 (= sehr vertraut)

nicht vertraut		mittelmäßig vertraut		sehr vertraut
1	2	3	4	5
◯	◯	◯	◯	◯

Wie vertraut sind Sie mit Backtechniken? *

Auf einer Skala von 1 (= nicht vertraut) bis 5 (= sehr vertraut)

nicht vertraut		mittelmäßig vertraut		sehr vertraut
1	2	3	4	5
◯	◯	◯	◯	◯

Aus welcher Küche bereiten Sie besonders gerne Gerichte zu? *

Mehrfachauswahl möglich.

- ☐ Österreichisch
- ☐ Italienisch
- ☐ Asiatisch
- ☐ Amerikanisch
- ☐ Balkan
- ☐ Orientalisch
- ☐ Sonstiges:

Haben Sie schon einmal ein Koch- bzw. Backrezept übersetzt? *

- ◯ Ja
- ◯ Nein

Erfahrung mit maschineller Übersetzung

Nutzen Sie maschinelle Übersetzungssysteme? *

Anmerkung: Dazu zählen zum Beispiel Google Translate, DeepL, Microsoft Translator und SYSTRAN.

- ◯ Ja
- ◯ Nein

Welche maschinellen Übersetzungssysteme nutzen Sie? *

Bitte geben Sie alle Systeme maschineller Übersetzung an, die sie benutzen.

Wie häufig nutzen Sie maschinelle Übersetzungssysteme? *

Bitte wählen Sie die zutreffende Antwort aus.

- ◯ täglich
- ◯ mehrmals in der Woche
- ◯ mehrmals im Monat
- ◯ mehrmals im Jahr

Für welche(n) Zweck(e) nutzen Sie maschinelle Übersetzungssysteme? *

Mehrfachauswahl möglich.

- ☐ private Zwecke
- ☐ berufliche Zwecke
- ☐ Ausbildung
- ☐ Kommunikation
- ☐ Sonstiges:

Wie zufrieden sind Sie mit den Ergebnissen maschineller Übersetzungssysteme? *

Bewerten Sie auf einer Skala von 1 (= nicht zufrieden) bis 5 (= sehr zufrieden).

nicht zufrieden		zufrieden		sehr zufrieden
1	2	3	4	5
◯	◯	◯	◯	◯

Haben Sie bereits maschinelle Übersetzungssysteme für Koch- bzw. Backrezepte verwendet? *

- ◯ Ja
- ◯ Nein

Wie zufrieden waren Sie mit der Qualität der maschinellen Übersetzung von Koch- bzw. Backrezepten? *

Bewerten Sie auf einer Skala von 1 (= nicht zufrieden) bis 5 (= sehr zufrieden).

nicht zufrieden		zufrieden		sehr zufrieden
1	2	3	4	5
◯	◯	◯	◯	◯

Bewertungsaufgabe

Nun haben Sie bereits den Bewertungsteil erreicht. Dafür wurden insgesamt acht Rezepte von den berühmten britischen Köchen Gordon Ramsay und Jamie Oliver herangezogen. Diese wurden maschinell aus dem Englischen ins Deutsche übersetzt.
Sie bekommen dabei nur die deutschen Versionen zu sehen. Sie werden gebeten, die Übersetzungen anhand von allgemeineren Fragen bzw. Aussagen auf einer Skala von 1 (= stimme überhaupt nicht zu) bis 5 (= stimme völlig zu) zu bewerten und zu kommentieren.
Es geht hier vor allem um die Frage, ob die übersetzten Rezepte aus Ihrer Sicht verständlich und in der Praxis umsetzbar sind.
Viel Spaß bei der gedanklichen Zubereitung der Gerichte! :)

Set 1 - Suppenrezepte

Übersetzung 1

Bitte bewerten Sie die maschinelle Übersetzung anhand der folgenden Aussagen. *

Bewerten Sie auf einer Skala von 1 (= stimme überhaupt nicht zu) bis 5 (= stimme völlig zu).

Die Übersetzung ist akzeptabel.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Mit diesem Rezept kann man dieses Gericht zubereiten.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Die Kochanleitung ist verständlich und könnte leicht umgesetzt werden.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Alle Zutaten sind Ihnen bekannt.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Die verwendeten Mengenangaben bzw. Maßeinheiten sind verständlich und nachvollziehbar.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die richtigen Fachbegriffe werden verwendet.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die Übersetzung enthält sprachliche bzw. idiomatische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die Übersetzung enthält grammatikalische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Bestimmte Begriffe sind für Sie im Deutschen nicht üblich.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Welche Zutaten kannten Sie bisher nicht? *

Welche Begriffe sind für Sie im Deutschen nicht üblich? *

Bitte fügen Sie hier weitere Kommentare zu dieser Übersetzung hinzu.

Was ist Ihnen sonst noch aufgefallen (sprachlich, inhaltlich etc.)? Sie können Ihr Kommentar gerne auch in Stichworten schreiben.

Würde Sie ein solches Rezept zum Nachkochen einladen? Ist der Titel ansprechend? *

- ◯ Ja
- ◯ Nein

Übersetzung 2

Bitte bewerten Sie die maschinelle Übersetzung anhand der folgenden Aussagen. *

Bewerten Sie auf einer Skala von 1 (= stimme überhaupt nicht zu) bis 5 (= stimme völlig zu).

Die Übersetzung ist akzeptabel.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Mit diesem Rezept kann man dieses Gericht zubereiten.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Die Kochanleitung ist verständlich und könnte leicht umgesetzt werden.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Alle Zutaten sind Ihnen bekannt.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Die verwendeten Mengenangaben bzw. Maßeinheiten sind verständlich und nachvollziehbar.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Die richtigen Fachbegriffe werden verwendet.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Die Übersetzung enthält sprachliche bzw. idiomatische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Die Übersetzung enthält grammatikalische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Bestimmte Begriffe sind für Sie im Deutschen nicht üblich.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Welche Zutaten kannten Sie bisher nicht? *

Welche Begriffe sind für Sie im Deutschen nicht üblich? *

Bitte fügen Sie hier weitere Kommentare zu dieser Übersetzung hinzu.

Was ist Ihnen sonst noch aufgefallen (sprachlich, inhaltlich etc.)? Sie können Ihr Kommentar gerne auch in Stichworten schreiben.

Würde Sie ein solches Rezept zum Nachkochen einladen? Ist der Titel ansprechend? *

- ◯ Ja
- ◯ Nein

Set 2 - Toastrezepte

Übersetzung 3

Bitte bewerten Sie die maschinelle Übersetzung anhand der folgenden Aussagen. *

Bewerten Sie auf einer Skala von 1 (= stimme überhaupt nicht zu) bis 5 (= stimme völlig zu).

Die Übersetzung ist akzeptabel.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Mit diesem Rezept kann man dieses Gericht zubereiten.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die Kochanleitung ist verständlich und könnte leicht umgesetzt werden.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Alle Zutaten sind Ihnen bekannt.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die verwendeten Mengenangaben bzw. Maßeinheiten sind verständlich und nachvollziehbar.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die richtigen Fachbegriffe werden verwendet.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die Übersetzung enthält sprachliche bzw. idiomatische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die Übersetzung enthält grammatikalische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Bestimmte Begriffe sind für Sie im Deutschen nicht üblich.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Welche Zutaten kannten Sie bisher nicht? *

Welche Begriffe sind für Sie im Deutschen nicht üblich? *

Bitte fügen Sie hier weitere Kommentare zu dieser Übersetzung hinzu.

Was ist Ihnen sonst noch aufgefallen (sprachlich, inhaltlich etc.)? Sie können Ihr Kommentar gerne auch in Stichworten schreiben.

Würde Sie ein solches Rezept zum Nachkochen einladen? Ist der Titel ansprechend? *

- ◯ Ja
- ◯ Nein

Übersetzung 4

Bitte bewerten Sie die maschinelle Übersetzung anhand der folgenden Aussagen. *

Bewerten Sie auf einer Skala von 1 (= stimme überhaupt nicht zu) bis 5 (= stimme völlig zu).

Die Übersetzung ist akzeptabel.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Mit diesem Rezept kann man dieses Gericht zubereiten.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die Kochanleitung ist verständlich und könnte leicht umgesetzt werden.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Alle Zutaten sind Ihnen bekannt.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die verwendeten Mengenangaben bzw. Maßeinheiten sind verständlich und nachvollziehbar.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die richtigen Fachbegriffe werden verwendet.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die Übersetzung enthält sprachliche bzw. idiomatische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Die Übersetzung enthält grammatikalische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Bestimmte Begriffe sind für Sie im Deutschen nicht üblich.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Welche Zutaten kannten Sie bisher nicht? *

Welche Begriffe sind für Sie im Deutschen nicht üblich? *

Bitte fügen Sie hier weitere Kommentare zu dieser Übersetzung hinzu.

Was ist Ihnen sonst noch aufgefallen (sprachlich, inhaltlich etc.)? Sie können Ihr Kommentar gerne auch in Stichworten schreiben.

Würde Sie ein solches Rezept zum Nachkochen einladen? Ist der Titel ansprechend? *

◯ Ja

◯ Nein

Set 3 - Rezepte für Beef Wellington

Übersetzung 5

Bitte bewerten Sie die maschinelle Übersetzung anhand der folgenden Aussagen. *

Bewerten Sie auf einer Skala von 1 (= stimme überhaupt nicht zu) bis 5 (= stimme völlig zu).

Die Übersetzung ist akzeptabel.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Mit diesem Rezept kann man dieses Gericht zubereiten.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Die Kochanleitung ist verständlich und könnte leicht umgesetzt werden.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Alle Zutaten sind Ihnen bekannt.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Die verwendeten Mengenangaben bzw. Maßeinheiten sind verständlich und nachvollziehbar.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Die richtigen Fachbegriffe werden verwendet.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Die Übersetzung enthält sprachliche bzw. idiomatische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Die Übersetzung enthält grammatikalische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Bestimmte Begriffe sind für Sie im Deutschen nicht üblich.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Welche Zutaten kannten Sie bisher nicht?*

Welche Begriffe sind für Sie im Deutschen nicht üblich?*

Bitte fügen Sie hier weitere Kommentare zu dieser Übersetzung hinzu.

Was ist Ihnen sonst noch aufgefallen (sprachlich, inhaltlich etc.)? Sie können Ihr Kommentar gerne auch in Stichworten schreiben.

Würde Sie ein solches Rezept zum Nachkochen einladen? Ist der Titel ansprechend? *

◯ Ja

◯ Nein

Übersetzung 6

Bitte bewerten Sie die maschinelle Übersetzung anhand der folgenden Aussagen. *

Bewerten Sie auf einer Skala von 1 (= stimme überhaupt nicht zu) bis 5 (= stimme völlig zu).

Die Übersetzung ist akzeptabel.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Mit diesem Rezept kann man dieses Gericht zubereiten.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Die Kochanleitung ist verständlich und könnte leicht umgesetzt werden.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Alle Zutaten sind Ihnen bekannt.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die verwendeten Mengenangaben bzw. Maßeinheiten sind verständlich und nachvollziehbar.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die richtigen Fachbegriffe werden verwendet.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die Übersetzung enthält sprachliche bzw. idiomatische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die Übersetzung enthält grammatikalische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Bestimmte Begriffe sind für Sie im Deutschen nicht üblich.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Welche Zutaten kannten Sie bisher nicht? *

Welche Begriffe sind für Sie im Deutschen nicht üblich? *

Bitte fügen Sie hier weitere Kommentare zu dieser Übersetzung hinzu.

Was ist Ihnen sonst noch aufgefallen (sprachlich, inhaltlich etc.)? Sie können Ihr Kommentar gerne auch in Stichworten schreiben.

Würde Sie ein solches Rezept zum Nachkochen einladen? Ist der Titel ansprechend? *

- ◯ Ja
- ◯ Nein

Set 4 - Dessertrezepte

Übersetzung 7

Bitte bewerten Sie die maschinelle Übersetzung anhand der folgenden Aussagen. *

Bewerten Sie auf einer Skala von 1 (= stimme überhaupt nicht zu) bis 5 (= stimme völlig zu).

Die Übersetzung ist akzeptabel.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Mit diesem Rezept kann man dieses Gericht zubereiten.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die Backanleitung ist verständlich und könnte leicht umgesetzt werden.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Alle Zutaten sind Ihnen bekannt.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die verwendeten Mengenangaben bzw. Maßeinheiten sind verständlich und nachvollziehbar.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die richtigen Fachbegriffe werden verwendet.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die Übersetzung enthält sprachliche bzw. idiomatische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die Übersetzung enthält grammatikalische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Bestimmte Begriffe sind für Sie im Deutschen nicht üblich.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Welche Zutaten kannten Sie bisher nicht? *

Welche Begriffe sind für Sie im Deutschen nicht üblich? *

Bitte fügen Sie hier weitere Kommentare zu dieser Übersetzung hinzu.

Was ist Ihnen sonst noch aufgefallen (sprachlich, inhaltlich etc.)? Sie können Ihr Kommentar gerne auch in Stichworten schreiben.

Würde Sie ein solches Rezept zum Nachkochen einladen? Ist der Titel ansprechend? *

- ◯ Ja
- ◯ Nein

Übersetzung 8

Bitte bewerten Sie die maschinelle Übersetzung anhand der folgenden Aussagen. *

Bewerten Sie auf einer Skala von 1 (= stimme überhaupt nicht zu) bis 5 (= stimme völlig zu).

Die Übersetzung ist akzeptabel.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Mit diesem Rezept kann man dieses Gericht zubereiten.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die Backanleitung ist verständlich und könnte leicht umgesetzt werden.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Alle Zutaten sind Ihnen bekannt.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die verwendeten Mengenangaben bzw. Maßeinheiten sind verständlich und nachvollziehbar.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die richtigen Fachbegriffe werden verwendet.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
○	○	○	○	○

Die Übersetzung enthält sprachliche bzw. idiomatische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Die Übersetzung enthält grammatikalische Fehler.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Bestimmte Begriffe sind für Sie im Deutschen nicht üblich.

stimme überhaupt nicht zu		neutral		stimme völlig zu
1	2	3	4	5
◯	◯	◯	◯	◯

Welche Zutaten kannten Sie bisher nicht? *

Welche Begriffe sind für Sie im Deutschen nicht üblich? *

Bitte fügen Sie hier weitere Kommentare zu dieser Übersetzung hinzu.

Was ist Ihnen sonst noch aufgefallen (sprachlich, inhaltlich etc.)? Sie können Ihr Kommentar gerne auch in Stichworten schreiben.

Würde Sie ein solches Rezept zum Nachkochen einladen? Ist der Titel ansprechend? *

- ◯ Ja
- ◯ Nein

Schlussteil

Wie leicht ist Ihnen die Bewertung gefallen? *

Bitte wählen Sie die zutreffende Antwort aus.

- ◯ sehr leicht
- ◯ leicht
- ◯ mittel
- ◯ schwer
- ◯ sehr schwer

Finden Sie, dass Ihre Koch- bzw. Backkenntnisse für die Bewertung ausreichend waren? *

- ◯ Ja
- ◯ Nein

Haben Sie noch Anmerkungen zum Thema bzw. zur Umfrage?

Abschluss

Sie haben die Umfrage erfolgreich abgeschlossen. Vielen Dank, dass Sie sich die Zeit genommen haben! Falls Sie Fragen haben oder an den Umfrageergebnissen interessiert sind, können Sie mich gerne unter folgender E-Mail-Adresse kontaktieren: a11911756@unet.univie.ac.at

Klicken Sie bitte auf "Beenden", um die Umfrage abzuschließen und die Ergebnisse für die Auswertung abzuspeichern.

Ich wünsche Ihnen weiterhin viel Freude beim Kochen und Backen!

Melanie Leko, BA

Abstract (Deutsch)

Die vorliegende Masterarbeit untersuchte die Wahrnehmung der Übersetzungsqualität von Koch- beziehungsweise Backrezepten verschiedener Systeme der neuronalen maschinellen Übersetzung durch Kochbegeisterte. Hierbei wurde folgende Forschungsfrage behandelt: Wie bewerten Kochbegeisterte die Übersetzung von Rezepten aus dem Englischen ins Deutsche durch zwei NMÜ-Systeme? In diesem Zusammenhang sollte herausgefunden werden, welches NMÜ-System die besseren Ergebnisse erzielt, inwiefern die maschinell übersetzten Rezepte eine erfolgreiche Zubereitung der Gerichte gewährleisten und welche Einschränkungen dabei festgestellt werden können sowie inwiefern das kulinarische Hintergrundwissen der Befragten die Bewertung beeinflusst.

Zur Beantwortung der Forschungsfrage wurde eine Online-Umfrage erstellt, in der die Teilnehmer*innen zunächst zu ihrer Person, Koch- und Backerfahrung sowie Erfahrung mit MÜ befragt wurden. Für die Bewertungsaufgabe wurden insgesamt acht Rezepte von den britischen Köchen Gordon Ramsay und Jamie Oliver ausgewählt und vier Sets mit je zwei vergleichbaren Rezepten erstellt. Pro Set wurde jeweils ein Rezept mit DeepL und Google Translate aus dem Englischen ins Deutsche übersetzt. Den Befragten wurden die maschinellen Übersetzungen ohne die Ausgangstexte vorgelegt, welche sie anhand von allgemeineren Aussagen auf einer Likert-Skala von 1 (= stimme überhaupt nicht zu) bis 5 (= stimme völlig zu) bewerten sollten. Abschließend konnten die Befragten ihre Gedanken zur Umfrage sowie zur MÜ im Allgemeinen mitteilen.

Die MÜ erzielte insgesamt positive Ergebnisse und wurde als akzeptabel und gebräuchlich beurteilt. Aus Sicht der Befragten stellten die Übersetzung von Zutaten, im deutschsprachigen Raum ungewöhnlichen Maßeinheiten und mehrdeutigen Termini sowie die inkonsistente Terminologieverwendung die größten Herausforderungen für die MÜ-Systeme dar. Die Übersetzungen enthielten darüber hinaus zahlreiche nicht idiomatische Ausdrücke und unübliche Begriffe, die jedoch überwiegend verständlich waren. Google Translate schnitt im Gesamteindruck besser als DeepL ab, wobei beide MÜ-Systeme ähnliche Ergebnisse erzielten und mit denselben Problemen konfrontiert waren. Die Auswertungsergebnisse zeigten weiters, dass Personen mit einem höheren Grad des prozeduralen Wissens die MÜ schlechter beurteilen und als weniger akzeptabel beziehungsweise gebräuchlich ansehen als Personen mit einer niedrigeren Koch- und Backbegeisterung. Aufgrund der kulturellen Merkmale von Rezepten und der terminologischen Inkonsistenz der Übersetzungen herrscht hinsichtlich der MÜ-Qualität jedoch weiterhin Verbesserungsbedarf.

Abstract (Englisch)

This master's thesis researched the perception of the translation quality of recipes from different neural machine translation systems by cooking enthusiasts. The following research question was discussed: How do cooking enthusiasts evaluate the translation of recipes from English into German by two NMT systems? In this context, the aim was to find out which NMT system achieves the better results, to what extent the machine-translated recipes ensure a successful preparation of the dishes and which limits can be identified in this area, as well as to what extent the culinary background knowledge of the respondents influences the evaluation.

To answer the research question, an online survey was created in which the participants were first asked about their person, their cooking and baking experience and their experience with MT. For the evaluation task, a total of eight recipes from the British chefs Gordon Ramsay and Jamie Oliver were selected and four sets of two comparable recipes were created. For each set, one recipe was translated from English into German using DeepL and one using Google Translate. The respondents were presented with the machine translations without the source texts, which they were asked to rate on a Likert scale from 1 (= strongly disagree) to 5 (= strongly agree) based on rather general statements. Finally, the respondents were given the opportunity to reflect on the survey and on MT in general.

Overall, the MT achieved positive results and was assessed as acceptable and useable. From the respondents' point of view, the translation of ingredients, unusual measurement units in the German-speaking region and ambiguous terms as well as the inconsistent use of terminology represented major challenges for the MT systems. The translations also contained numerous unidiomatic expressions and unusual terms, which were, however, largely intelligible. In general, Google Translate performed better than DeepL, with both MT systems achieving similar results and facing the same problems. The evaluation results also showed that people with a higher level of procedural knowledge rated the MT less favourably and considered it less acceptable or useable than people with a lower level of enthusiasm for cooking and baking. However, taking into account the cultural characteristics of recipes and the terminological inconsistency of the translations, there is still room for improvement in terms of MT quality.