

Evaluationsbericht des Fraunhofer-Instituts für Arbeitswirtschaft und Organisation



Institut
Arbeitswirtschaft und
Organisation

ANWENDERTEST PERSONAL TRANSLATOR (PT)

1. PT Usability Test

- 1.1 Testdesign
- 1.2 Versuchsbedingungen
- 1.3 Probanden
- 1.4 Testablauf

2. Präaktionale Befragung

3. Zeitersparnis

4. Produktivitätssteigerung

5. Effektivität und Effizienz

- 5.1 Effektivität
 - 5.1.1 Qualität der Texte
 - 5.1.2 Übersetzungsqualität in Abhängigkeit vom Vorwissen
- 5.2 Effizienz
- 5.3 Effizienzwerte nach Gruppen

6. Subjektive Variablen

7. Postaktionale Befragung

8. Abschließende Bewertung

1. PT Usability Test

1.1 Testdesign

Die vier Versuchsgruppen setzten sich folgendermaßen zusammen:

	Übersetzen E-D			Übersetzen D-E		
	Gruppe 1			Gruppe 3		
Sachtext	A/PT	B/W	2	A/PT	B/W	2
	A/W	B/PT	4	A/W	B/PT	3
	B/W	A/PT	2	B/W	A/PT	2
	B/PT	A/W	3	B/PT	A/W	3
			n=11			n=9

	Übersetzen E-D			Erstellen		
	Gruppe 2			Gruppe 4		
Geschäftsbrief	A/PT	B/W	3	A/PT	B/W	3
	A/W	B/PT	3	A/W	B/PT	3
	B/W	A/PT	2	B/W	A/PT	2
	B/PT	A/W	2	B/PT	A/W	2
			n=10			n=10
				Gesamt: n=40		

Legende:

PT: Nur mit Personal Translator

W: nur mit Print-Wörterbuch

A, B: unterschiedliche Texte/Instruktionen für eine Aufgabenart

Tabelle 1: Testdesign

Insgesamt wurden die folgenden **vier Aufgabentypen** gestellt:

1.	Übersetzen eines englischen Sachtexts ins Deutsche	(Gruppe 1)
2.	Übersetzen eines englischen Geschäftsbriefs ins Deutsche	(Gruppe 2)

3.	Übersetzen eines deutschen Sachtexts ins Englische	(Gruppe 3)
4.	Erstellen eines englischen Geschäftsbriefs auf der Grundlage stichpunktartig dargestellter Szenarios	(Gruppe 4)

Zu jedem Aufgabentyp wurden zwei verschiedene, gleich schwere und gleich umfangreiche Texte bzw. Instruktionen ausgewählt. Es handelte sich dabei um:

zu 1:	A) und B) Abituraufgaben Grundkurs zur Übersetzung von englisch nach deutsch (etwa 200 Wörter).
zu 2:	Englische Geschäftsbriefe Möbelhaus (A) und Softwarehersteller (B) an Kunden (etwa 200 Wörter).
zu 3:	Übersetzung eines deutschen Sachtexts ins Englische. A) Text aus einer Wirtschaftszeitung B) deutsche Musterlösung aus dem Grundkursabitur zur Übersetzung ins Englische.
zu 4:	Erstellen eines englischen Texts. A) Reklamation an eine englische Fluggesellschaft wegen beschädigten Gepäcks. B) Bewerbungsschreiben auf eine vorgelegte Stellenausschreibung.

Die ausgewählten Aufgaben waren bewusst relativ **anspruchsvoll** gewählt, so dass bei angenommener breiter Streuung der Englischkenntnisse auch eine hohe Varianz in den gemessenen Variablen zu erwarten war. Bei zu einfachen oder zu schweren Aufgaben war aufgrund von Decken- bzw. Bodeneffekten zu befürchten, dass eventuelle Effekte der Unterstützung des PT nicht nachgewiesen werden können.

1.2 Versuchsbedingungen

Die 40 Probanden wurden zufällig einer der vier Aufgabentypen zugewiesen. Durch Versuchspersonenausfall (von den ursprünglich 41 geplanten Probanden erschienen nur 40) ergaben sich folgende Gruppengrößen: Gruppe 1 ist die größte Gruppe mit 11 Probanden, Gruppe 2 und 4 verfügen über je 10 Probanden und Gruppe 3 umfasst 9 Probanden. Jede Testperson bearbeitete nur einen Aufgabentyp, dazu jedoch beide Texte bzw. Instruktionen (A und B). Ein Text wurde mit PT, der andere nur mit Hilfe eines Wörterbuchs bearbeitet. Sowohl die Reihenfolge der beiden unterschiedlichen Texte bzw. Instruktionen als auch die Art der Übersetzungshilfe (PT= mit Personal Translator, W= nur mit Wörterbuch) wurden systematisch permutiert, so dass jede mögliche Kombination von zwei Personen bearbeitet wurde.

1.3 Probanden

Die Probanden wurden zufällig aus der **Versuchspersonendatenbank** des MT Usability Engineering ausgewählt. Einzige Auswahlkriterien waren, dass sie grundlegende PC-Kenntnisse sowie Kenntnisse der englischen Sprache besaßen. Vier Probanden von der Uni Regensburg wurden von der Firma Linguatec gestellt. Insgesamt nahmen 16 weibliche und 24 männliche Probanden mit folgender Altersverteilung an der Untersuchung teil:

Alter	Anzahl
18-20 Jahre	3
21-30 Jahre	20
31-40 Jahre	12
Über 40 Jahre	5

Abbildung 1: Geschlecht und Alter der Teilnehmer

Bei der Testdurchführung sollten die Probanden bereits mit dem PT umgehen können, um Verzerrungen (insbesondere der Bearbeitungszeit) durch Unerfahrenheit in der Bedienung zu vermeiden. Deshalb erhielten sie eine Woche vor dem Test den PT zu Übungszwecken. Um sicherzustellen, dass die Versuchspersonen tatsächlich mit dem PT arbeiteten, mussten fünf Übungsaufgaben zur Funktionsweise des Programms beantwortet werden. Die Lösungen wurden vor dem Anwendertest per E-Mail oder Fax eingereicht.

1.4 Testablauf

1.	Empfang und Begrüßung der Testpersonen, Einführung in den Usability Test.
2.	Präaktionale Befragung zu Vorerfahrungen.
3.	Einschätzung der subjektiven Bedeutsamkeit folgender vier Faktoren für die jeweilige Übersetzungsaufgabe: sprachliche Perfektion, geringer Zeitaufwand, geringe Anstrengung, Freude am Arbeiten.
4.	Aufgabenbearbeitung: Szenario A/B mit/ohne PT.
5.	Einschätzung der subjektiv wahrgenommenen Anstrengung, Zeitaufwand, Spaß.
6.	5 Minuten Pause.

7.	Aufgabenbearbeitung: Szenario A/B mit/ohne PT.
8.	Einschätzung der subjektiv wahrgenommenen Anstrengung, Zeitaufwand, Spaß.
9.	Postaktionale Befragung: Interview zur Akzeptanz, Erfragung von Verbesserungsvorschlägen.

2. Präaktionale Befragung

Vor Beginn der Bearbeitung beantworteten die Probanden die drei folgenden Fragen:

1. Hatten Sie vorher schon Erfahrung mit anderen Übersetzungshilfen? Wenn ja, mit welchen? (Mehrere Antworten möglich)

32,5% der Probanden hatten keine Erfahrung mit Übersetzungshilfen.

66,5 % der Testpersonen hatten bereits mindestens eine andere Übersetzungshilfe kennen gelernt.

29,6% der erfahrenen Probanden hatten bereits mehrere Übersetzungshilfen benutzt. Dabei handelte es sich meist um gebührenfreie Internet Services. Am bekanntesten ist Altavista, gefolgt von Leo, Babelfisch und Babylon. Lediglich eine Testperson hatte im Vorfeld mit einem käuflich erwerbbar Übersetzungstool gearbeitet (Langenscheidt's T1).

2. Wofür würden Sie eine Übersetzungshilfe benutzen? Bei dieser Frage waren mehrere Antworten möglich.
Wie aus Tabelle 2 hervorgeht, benötigte die Mehrzahl der Probanden eine Übersetzungshilfe im geschäftlichen Umfeld.

	In Prozent
Geschäftlich: Erstellung von Geschäftsbrief, Vortrag, Bericht, Bewerbung	41 %
Sachtext übersetzen: Zeitungsartikel, Internet-Webseite	22 %
Privater Briefverkehr	18 %
Lernhilfe: Hausaufgabe, Literaturübersetzen	10 %
Technische Texte übersetzen, z.B. Bedienungsanleitungen	9 %

3. Tabelle 2: Wofür würden Sie eine Übersetzungshilfe benutzen?
4. Wie lange haben Sie sich im Vorfeld mit dem PT beschäftigt?

Stunden	In Prozent
bis 30 Min	10 %
bis 1 Std.	13 %
bis 2 Std.	32 %
bis 3 Std.	29 %
bis 4 Std.	8 %
mehr als 4 Std.	8 %

5. Tabelle 3: Dauer der Übung mit dem PT

Die Übungszeit der Probanden variierte in hohem Maße. Dennoch schien sich jede Testperson ausreichend mit der Software beschäftigt zu haben, da im Test nur sehr wenige Bedienprobleme auftraten.

3. Zeitersparnis

Es wurde berechnet ob die Verwendung des PT zu einer Zeitersparnis bei der Bearbeitung führt.

$$\text{Zeitersparnis} = \frac{\text{BearbeitungsdauerPT} - \text{BearbeitungsdauerWörterbuch}}{\text{BearbeitungsdauerWörterbuch}}$$

In den folgenden Tabellen sind die gemittelten Bearbeitungszeiten jeweils für die Bearbeitung mit PT und Wörterbuch sowie die daraus resultierende Zeitersparnis angegeben.

Über alle Gruppen:

	MW Zeitaufwand	Zeitersparnis
Mit PT	21,1 Min	41,4 %
Mit Wörterbuch	35,45 Min	

Tabelle 4: Bearbeitungszeiten und Zeitersparnis über alle Gruppen

Gruppe 1 (Übersetzen engl. Sachtexte):

	MW Zeitaufwand	Zeitersparnis
Mit PT	26,91 Min	26,37 %
Mit Wörterbuch	36,55 Min	

Tabelle 5: Bearbeitungszeiten und Zeitersparnis Gruppe 1

Gruppe 2 (Übersetzen engl. Geschäftsbriefe)

	MW Zeitaufwand	Zeitersparnis
Mit PT	12,50 Min	65,47 %
Mit Wörterbuch	36,20 Min	

Tabelle 6: Bearbeitungszeiten und Zeitersparnis Gruppe 2

Gruppe 3 (Übersetzen deut. Sachtexte)

	MW Zeitaufwand	Zeitersparnis
Mit PT	13,78 Min	61,17 %
Mit Wörterbuch	39,56 Min	

Tabelle 7: Bearbeitungszeiten und Zeitersparnis Gruppe 3

Gruppe 4 (Erstellen engl. Geschäftsbriefe)

	MW Zeitaufwand	Zeitersparnis
Mit PT	29,90 Min	-0,34 %
Mit Wörterbuch	29,80 Min	

Tabelle 8: Bearbeitungszeiten und Zeitersparnis Gruppe 4

Besonders bei den Gruppen 2 und 3 lässt sich eine überaus deutliche Zeitersparnis nachweisen: Die Probanden konnten über 60% schneller arbeiten, wenn sie den PT benutzten. Bei Gruppe vier jedoch führte die Verwendung des PT zu einer um 0.3% verlangsamten Arbeitsleistung. Dieses Ergebnis ist damit zu erklären, dass die Testpersonen es vorzogen, den Text erst auf Deutsch zu formulieren und ihn dann mit Hilfe des PT zu übersetzen. Dadurch brauchten Sie doppelt Zeit: zuerst für die deutsche Formulierung und dann noch für die Korrektur der PT-Übersetzung.

4. Produktivitätssteigerung

Es wurde ermittelt, ob die Verwendung des PT eine gesteigerte Produktivität bewirkt. Die Produktivität wurde als Wortanzahl pro Bearbeitungsdauer gemessen, d.h. die in einem Text erzielte Wortzahl wurde durch die benötigte Bearbeitungszeit geteilt.

$$\text{Produktivitätssteigerung} = \frac{\text{Produktivität} - \text{ProduktivitätWörterbuch}}{\text{ProduktivitätWörterbuch}}$$

In den folgenden Tabellen sind die gemittelten Produktivitätswerte jeweils für die Bearbeitung mit PT und mit Wörterbuch sowie die daraus resultierende Produktivitätssteigerung angegeben.

Über alle Gruppen

	MW Produktivität	Produktivitätssteigerung
Mit PT	14,10	164,45 %
Mit Wörterbuch	5,30	

Tabelle 9: Mittelwerte der Wortzahlen pro Zeitdauer und Produktivitätssteigerung über alle Gruppen

Gruppe 1 (Übersetzen engl. Sachtexte)

	MW Produktivität	Produktivitätssteigerung
Mit PT	8,09	64,97 %
Mit Wörterbuch	4,90	

Tabelle 10: Mittelwerte der Wortzahlen pro Zeitdauer und Produktivitätssteigerung Gruppe 1

Gruppe 2 (Übersetzen engl. Geschäftsbriefe)

	MW Produktivität	Produktivitätssteigerung
Mit PT	21,77	275,78 %
Mit Wörterbuch	5,79	

Tabelle 11: Mittelwerte der Wortzahlen pro Zeitdauer und Produktivitätssteigerung Gruppe 2

Gruppe 3 (Übersetzen deut. Sachtexte)

	MW Produktivität	Produktivitätssteigerung
Mit PT	22,40	373,97 %

Mit Wörterbuch	4,73
----------------	------

Tabelle 12: Mittelwerte der Wortzahlen pro Zeitdauer und Produktivitätssteigerung Gruppe 3

Gruppe 4 (Erstellen engl. Geschäftsbriefe)

	MW Produktivität	Produktivitätssteigerung
Mit PT	5,76	-3,34 %
Mit Wörterbuch	5,96	

Tabelle 13: Mittelwerte der Wortzahlen pro Zeitdauer und Produktivitätssteigerung Gruppe 4

Die **Produktivität** stieg mit PT durchschnittlich um **164%** an. Gruppe 2 und Gruppe 3 profitierten wiederum besonders von einer Verwendung des PT. Die Produktivitätssteigerung lag in diesen Gruppen bei rund 275% bzw. 374%, d.h. die Probanden erzeugten in derselben Zeit eine drei bis viermal höhere Wortanzahl als mit einem Wörterbuch. Bei Gruppe 4 wiederum wurde eine um rund 3% verminderte Produktivität nachgewiesen.

5. Effektivität und Effizienz

5.1 Effektivität

Mit Effektivität wird die **Güte des Ergebnisses** bezeichnet. In diesem Fall wurde untersucht, ob die Bearbeitung der Texte mit dem PT zu besseren Noten führt. Außerdem war in diesem Zusammenhang von Interesse, bei welchem Grad an Vorwissen der PT die größte Hilfe darstellt.

5.1.1 Qualität der Texte

Es wurde über alle Gruppen und Versuchspersonen berechnet, ob die Noten mit PT oder ohne PT besser ausfallen. Zur Berechnung wurden sowohl die Durchschnittsnoten aller Lehrer als auch die Noten der einzelnen Lehrer herangezogen. Nur bei dem zweiten Lehrer konnten mit dem PT bessere Noten erzielt werden. Bei allen anderen Lehrern und bei den Durchschnittsnoten, die aus den Urteilen aller Lehrer gebildet wurden, ist kein Unterschied in der Notengebung feststellbar.

Gruppenstatistiken

	N	Mittelwert	Standardabweichung
MW Note			
Mit PT	46	6,54	3,20
Mit Wörterbuch	40	5,33	4,26
Note 1			
Mit PT	46	3,20	3,40
Mit Wörterbuch	40	2,90	3,89
Note 2			
Mit PT	46	8,72	3,69
Mit Wörterbuch	40	6,29	4,99
Note 3			
Mit PT	46	5,22	3,96
Mit Wörterbuch	40	5,18	4,95
Note 4			
Mit PT	46	8,48	5,10
Mit Wörterbuch	40	6,48	5,76

Tabelle14: Mittelwerte und Standardabweichungen. MWNote bezeichnet den mittleren Notendurchschnitt über alle Lehrer hinweg. Note 1-4 ist der Durchschnitt aller Noten, die der jeweilige Lehrer 1-4 vergab.

T-Test bei unabhängigen Stichproben

	Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit	
	F	Signifikanz	T	Sig. (2-seitig)
MW Note	5,567	0,021	1,483	0,142
Note 1	,101	0,752	0,376	0,708
Note 2	6,929	0,10	2,537	0,013
Note 3	2,72	0,103	0,44	0,965
Note 4	1,219	0,273	1,712	0,091

Tabelle 15: Ergebnisse des T-Tests

Im zweiten Schritt wurde untersucht, ob es innerhalb der Gruppen Unterschiede in den Noten gibt. Hierzu wurde für jede Gruppe einzeln geprüft, ob die Noten mit PT besser ausfallen als ohne. Dahinter steht die Vermutung, dass der PT möglicherweise bei unterschiedlichen Aufgabenstellungen unterschiedlich viel Nutzen zeigt. Dieser Annahme liegen die Ergebnisse für Zeitersparnis und Produktivitätsgewinn zugrunde, die zeigen, dass die verschiedenen Gruppen unterschiedlich stark von der Nutzung des PT profitieren.

Bei den Gruppen 2 (Übersetzung eines englischen Geschäftsbriefs) und 3 (Übersetzung eines deutschen Sachtexts) war ein signifikanter Mittelwertsunterschied festzustellen, welcher sich aber auf die Notengebung von Lehrer 2 beschränkte. Dies bedeutet, dass Lehrer 2 in den beiden Gruppen signifikant bessere Noten für die Textversionen vergab, die mit PT erstellt worden waren. Die Darstellung in den folgenden Tabellen beschränkt sich auf die signifikanten Ergebnisse. Es soll jedoch darauf hingewiesen werden, dass bei Gruppe 2 erst eine mittlere Notendifferenz von 5,40 für ein signifikantes Ergebnis ausreichte. Selbst mittlere Notenunterschiede von 3,70 oder 4,70 in den anderen Gruppen wurden aufgrund der hohen Streuungen nicht statistisch signifikant.

Gruppenstatistiken (englischen Geschäftsbrief übersetzen)

	N	Mittelwert	Standardabweichung
Note 2			
Mit PT	10	12,10	1,79
Mit Wörterbuch	10	6,70	6,72

Tabelle 16: Mittelwerte und Standardabweichungen Gruppe 2, Lehrer 2

T-Test bei unabhängigen Stichproben (englischen Geschäftsbrief übersetzen)

Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit		
	F	Signifikanz	T	Sig. (2-seitig)
Note 2	47,193	0,000	2,456	0,033

Tabelle 17: Ergebnisse des T-Tests, Gruppe 3, Lehrer 2

Gruppenstatistiken (deutschen Sachtext übersetzen)

	N	Mittelwert	Standardabweichung
Note 2			
Mit PT	9	8,89	2,57
Mit Wörterbuch	9	5,11	4,20

Tabelle 18: Mittelwerte und Standardabweichungen Gruppe 3, Lehrer 2

T-Test bei unabhängigen Stichproben (deutschen Sachtext übersetzen)

Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit		
	F	Signifikanz	T	Sig. (2-seitig)
Note 2	2,707	0,119	2,303	0,038

Tabelle 19: Ergebnisse des T-Tests, Gruppe 3, Lehrer 2

5.1.2 Übersetzungsqualität in Abhängigkeit vom Vorwissen

Die folgenden Diagramme zeigen, welchen Einfluss die Verwendung des PT bei entsprechendem Vorwissen der Probanden hatte. Im unteren und mittleren Notenbereich, d.h. wenn die Versuchspersonen über schlechte bis mittelmäßige Englischkenntnisse verfügten, trug der PT zu einer Verbesserung der Note bei. Die Probanden erzielten demnach bessere Noten wenn sie mit PT arbeiteten. Bei guten bis sehr guten Englischkenntnissen der Probanden schien der PT eher kontraproduktiv zu wirken: die Probanden produzierten unter Zuhilfenahme des PT schlechtere Texte als mit Wörterbuch.

In der folgenden Abbildung stellen die Punkte die jeweilige Veränderung in der Note dar, die ein Proband im Vergleich zu einer Bearbeitung mit einem Wörterbuch (Werte der x-Achse) erfuhr. Die Trendlinie zeigt, dass der Benefit des PT bei zunehmend besseren Noten geringer wird und schließlich sogar ins Negative fällt. Die Gerade erhält man durch Regressionsrechnung mit Hilfe der Methode der kleinsten Quadrate. Die Regression dient dazu, den Zusammenhang zwischen der Prädiktorvariable (Note ohne PT) und der Kriteriumsvariable (Verbesserung bzw. Verschlechterung der Note mit PT) darzustellen. Die Beziehung zwischen den beiden Variablen wird graphisch durch die Regressionsgerade repräsentiert, welche rechnerisch durch die Gleichung $y = -0,517x + 4,437$ beschrieben ist. Der Korrelationskoeffizient nach Pearson r zeigt die Enge des linearen Zusammenhang zwischen den beiden Variablen auf, das Bestimmtheitsmaß r^2 ist das Quadrat des Korrelationskoeffizienten.

Der Korrelationskoeffizient zeigt mit $-0,65$ einen relativ engen negativen Zusammenhang zwischen den erzielten Leistungen mit Wörterbuch und der Verbesserung bzw. Verschlechterung mit PT.

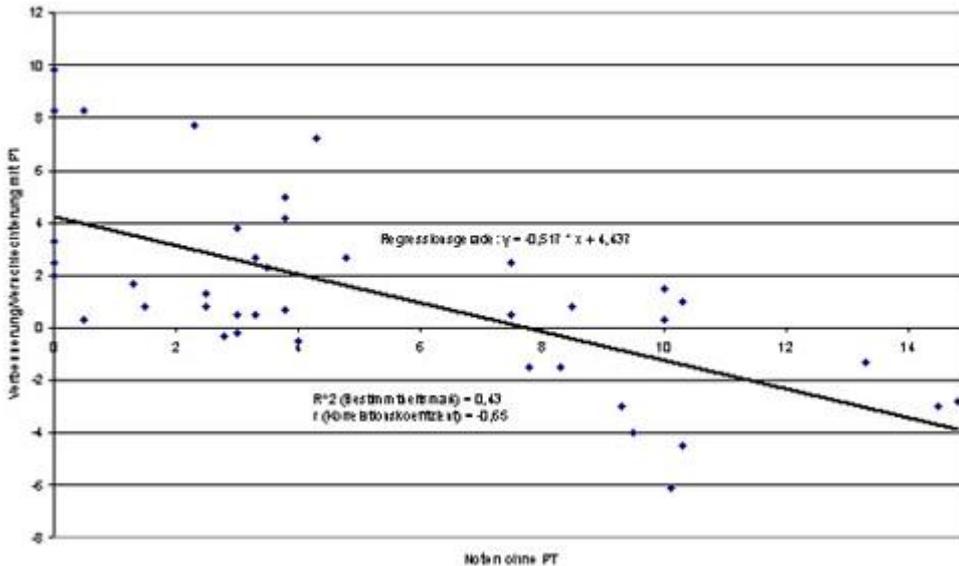


Abbildung 2: lineare Regression

In der folgenden Abbildung wurde die Verbesserung bzw. Verschlechterung durch den PT für jede ganze Notenstufe ermittelt. Die Verbindungslinie zwischen den Notenstufen zeigt den Verlauf der Veränderung der Übersetzungsqualität mit zunehmenden sprachlichen Fähigkeiten der Testpersonen.

Die Abbildung ist folgendermaßen zu interpretieren: bei einer Note von 0 Punkten ohne PT haben die Testpersonen eine durchschnittliche Verbesserung von über 5 Punkten mit PT. Hat eine Versuchsperson ohne PT jedoch 10 Punkte erreicht, so erreicht sie mit PT eine durchschnittliche Verschlechterung von -2, also nur 8 Punkte.

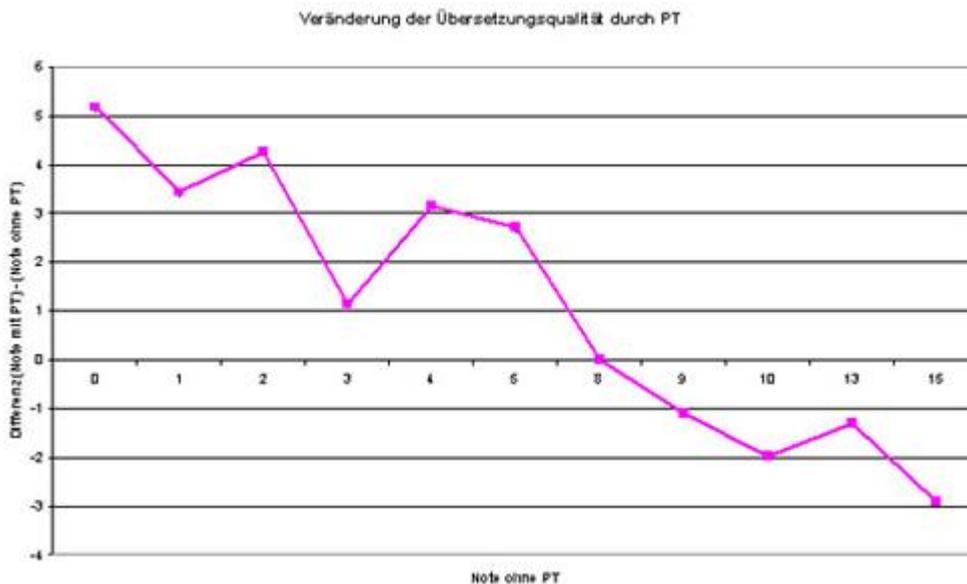


Abbildung 3: Veränderung der Übersetzungsqualität für die ganzen Notenstufen

5.2 Effizienz

Effizienz ist die **Güte des Ergebnisses in Relation zum benötigten Aufwand**. In dieser Untersuchung wurden für die Effizienz vier verschiedene Maße gebildet, um alle rechnerischen Möglichkeiten auszuschöpfen. Alle Variablen (Noten, subjektive Gewichtung vor dem Test, subjektive Bewertung nach der Aufgabebearbeitung, objektive Zeit) wurden zu diesem Zweck normiert.

Das erste Maß ist eine **gewichtete Summe** (abgekürzt mit GEWICHTE). In diese Summe geht die Gewichtung der Versuchspersonen ein, mit der sie die subjektive Bedeutsamkeit verschiedener Faktoren (Qualität, Schnelligkeit, Spaß, Einfachheit) einschätzten. Die Gewichtung wird dann mit dem im Test erhaltenen Wert für die jeweilige Variable (bzw. mit ihrem Komplement) multipliziert.

$$\text{GEWICHTE} = (\text{Gew. Übers. qualität} \cdot \text{Note}) + (\text{Gew. Spaß} \cdot \text{subj. Spaß}) +$$

$$[\text{Gew. Zeit} \cdot (1 - \text{subj. Zeitaufwand})] + [\text{Gew. leichte Übers.} \cdot (1 - \text{subj. Anstr.})]$$

Das zweite Effizienzmaß (EFFIZIEN) bezieht die subjektiven Gewichtungen nicht mit ein, sondern beschränkt sich lediglich auf die **objektive Ergebnisqualität** (Note) und den dazu im Verhältnis stehenden Aufwand (Anstrengung und subjektiver Zeitaufwand).

$$\text{EFFIZIEN} = \frac{\frac{\text{Note}}{\text{Anstrengung}} + \frac{\text{Note}}{\text{subj. Zeitaufwand}}}{2}$$

Das dritte Effizienzmaß (EFFOBJZE) ist genauso aufgebaut wie der zweite Term, außer dass statt dem subjektiv eingeschätzten Zeitaufwand die **objektive Bearbeitungsdauer** herangezogen wird.

$$\text{EFFOBJZE} = \frac{\frac{\text{Note}}{\text{Anstrengung}} + \frac{\text{Note}}{\text{obj. Bearb. dauer}}}{2}$$

Das vierte Effizienzmaß (EFFDAUER) ist die **Note** dividiert durch die **objektive Bearbeitungsdauer**.

$$\text{EFFDAUER} = \frac{\text{Note}}{\text{obj. Bearb. dauer}}$$

Für alle Maße wurde ein T-Test gerechnet, um festzustellen, ob die Bearbeitung mit oder ohne PT effizienter vonstatten geht.

Gruppenstatistiken

Übersetzungsweise	N	Mittelwert	Standardabweichung
EFFDAUER			
Mit PT	40	2,49	3,24
Mit Wörterbuch	40	0,94	0,76
EFFOBJZE			
Mit PT	40	3,93	8,14
Mit Wörterbuch	40	0,86	0,76
GEWICHTE			
Mit PT	40	5594,28	1792,48
Mit Wörterbuch	40	3789,00	1551,30
EFFIZIEN			
Mit PT	40	4,95	12,39
Mit Wörterbuch	40	0,74	0,64

Tabelle 10: Mittelwerte und Standardabweichungen

T-Test bei unabhängigen Stichproben

	Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit	
	F	Signifikanz	T	Sig. (2-seitig)
EFFDAUER	22,186	0,000	2,949	0,005
EFFOBJZE	13,137	0,001	2,376	0,022
GEWICHTE	1,017	0,316	4,816	0,000
EFFIZIEN	11,046	0,001	2,146	0,038

Tabelle 11: Ergebnisse des T-Tests

Bei alle vier Indizes ist eine signifikanter Mittelwertsunterschied erkennbar. Die Effizienz wird also durch die Verwendung des PT signifikant gesteigert.

5.3 Effizienzwerte nach Gruppen

Es wurde für jede Gruppe einzeln untersucht, wie sich die vier Effizienzwerte für die Bearbeitung mit oder ohne PT unterscheiden. In der Gruppe 1 (Übersetzung englischer Sachtexte) sind keine Mittelwertsunterschiede statistisch signifikant. Für Gruppe 2 (Übersetzung englischer Geschäftsbriefe) werden die Mittelwertsunterschiede signifikant, sowohl für gewichtete Summen als auch für EFFDAUER.

Gruppenstatistiken

Übersetzungsweise	N	Mittelwert	Standardabweichung
EFFDAUER			
Mit PT	10	4,99	4,38
Mit Wörterbuch	10	0,87	0,66
EFFOBJZE			
Mit PT	10	9,99	14,43
Mit Wörterbuch	10	0,74	0,54
GEWICHTE			
Mit PT	10	6768,88	1926,27
Mit Wörterbuch	10	3557,75	1530,62

EFFIZIEN			
Mit PT	10	13,77	22,78
Mit Wörterbuch	10	0,61	0,38

Tabelle 12: Mittelwerte und Standardabweichungen, Gruppe 2

T-Test bei unabhängigen Stichproben

Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit		
	F	Signifikanz	T	Sig. (2-seitig)
EFFDAUER	16,688	0,001	2,948	0,016
EFFOBJZE	11,419	0,003	2,026	0,073
GEWICHTE	1,625	0,219	4,127	0,001
EFFIZIEN	8,214	0,010	1,827	0,101

Tabelle 13: Ergebnisse der T-Tests, Gruppe 2

In der Gruppe 3 (Übersetzung eines deutschen Sachtexts) sind die gewichteten Summen signifikant höher bei der Bearbeitung mit PT als bei der Bearbeitung mit Wörterbuch.

Gruppenstatistiken

Übersetzungsweise	N	Mittelwert	Standardabweichung
EFFDAUER			
Mit PT	9	2,65	3,65
Mit Wörterbuch	9	1,08	0,86
EFFOBJZE			
Mit PT	9	3,09	4,56
Mit Wörterbuch	9	0,91	0,66
GEWICHTE			
Mit PT	9	5956,67	2058,26
Mit Wörterbuch	9	3396,53	1821,08
EFFIZIEN			
Mit PT	9	2,94	4,88
Mit Wörterbuch	9	0,74	0,53

Tabelle 14: Mittelwerte und Standardabweichungen, Gruppe 3

T-Test bei unabhängigen Stichproben

Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit		
	F	Signifikanz	T	Sig. (2-seitig)
EFFDAUER	3,592	0,076	1,259	0,240
EFFOBJZE	10,050	0,006	1,417	0,193
GEWICHTE	0,250	0,624	2,795	0,013
EFFIZIEN	10,389	0,005	1,348	0,214

Tabelle 15: Ergebnisse der T-Tests, Gruppe 3

Bei Gruppe 4 (Erstellen eines englischen Geschäftsbriefs) zeigen die Ergebnisse in dieselbe Richtung: die gewichteten Summen sind mit PT signifikant größer als ohne PT.

Gruppenstatistiken

Übersetzungsweise	N	Mittelwert	Standardabweichung
EFFDAUER			
Mit PT	10	0,79	0,51
Mit Wörterbuch	10	0,87	0,83
EFFOBJZE			
Mit PT	10	1,24	1,09
Mit Wörterbuch	10	0,80	0,76
GEWICHTE			
Mit PT	10	4958,25	1003,19
Mit Wörterbuch	10	3866,38	1275,32

EFFIZIEN			
Mit PT	10	1,51	1,65
Mit Wörterbuch	10	0,66	0,52

Tabelle 16: Mittelwerte und Standardabweichungen, Gruppe 4

T-Test bei unabhängigen Stichproben

	Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit	
	F	Signifikanz	T	Sig. (2-seitig)
EFFDAUER	3,592	0,076	1,259	0,240
EFFOBJZE	10,050	0,006	1,417	0,193
GEWICHTE	0,250	0,624	2,795	0,013
EFFIZIEN	10,389	0,005	1,348	0,214

Tabelle 17: Ergebnisse der T-Tests, Gruppe 4

6. Subjektive Variablen

Hier wurde mit Hilfe von T-Tests ermittelt, ob die Bearbeitung mit oder ohne PT anstrengender bzw. zeitaufwendiger war und welche Art der Bearbeitung mehr Spaß gemacht hat. Die Ergebnisse werden hier über alle Gruppen hinweg dargestellt.

Gruppenstatistiken

Übersetzungsweise	N	Mittelwert	Standard-abweichung
ANSTRENG			
Mit PT	40	35,75	27,77
Mit Wörterbuch	40	62,43	24,22
ZEITAUFW			
Mit PT	40	33,08	25,82
Mit Wörterbuch	40	62,75	23,19
FREUDE			
Mit PT	40	52,73	28,76
Mit Wörterbuch	40	39,55	29,06

Tabelle 18: Mittelwerte und Standardabweichungen, subjektive Variablen

T-Test bei unabhängigen Stichproben

	Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit	
	F	Signifikanz	T	Sig. (2-seitig)
ANSTRENG	1,372	0,245	-4,579	0,000
ZEITAUFW	0,290	0,592	-5,408	0,000
FREUDE	0,381	0,539	2,038	0,045

Tabelle 19: Ergebnisse des T-Tests, subj. Variablen

Alle drei subjektiven Variablen zeigen vorteilhafte Ergebnisse für den PT. Die Arbeit mit dem PT wurde als signifikant weniger anstrengend und zeitaufwendig gefunden. Außerdem machte es den Probanden mehr Spaß, mit dem PT zu arbeiten als mit dem Wörterbuch.

7. Postaktionale Befragung

Am Ende der Aufgabenbearbeitung waren im Rahmen eines Interviews folgende Fragen zu beantworten:

1.	<p>Sie haben die Aufgabe nun einmal mit/einmal ohne dem Personal Translator erarbeitet. Welche Art der Bearbeitung hat Ihnen besser gefallen? Warum?</p> <p>80% der Probanden bevorzugten die Bearbeitung mit PT, da sie ihrer Meinung nach schneller, einfacher und effektiver ist. 20% der Probanden arbeiteten lieber mit einem Wörterbuch, da die Nachbearbeitung sehr viel Zeit kostet.</p>
2.	<p>Wie sind mit dem Personal Translator zurecht bekommen?</p> <p>87.5% der Probanden gaben an, dass sie ohne Schwierigkeiten zurecht gekommen seien. 12.5% der Teilnehmer sagten aus, dass sie im Vorfeld des Tests Probleme mit der Installation hatten oder Schwierigkeiten hatten, das Manual zu verstehen.</p>

3. Wie zufrieden sind Sie mit der Übersetzungsqualität, die vom Personal Translator geliefert wird? Die Probanden wurden gebeten, Schulnoten für die Übersetzungsqualität zu vergeben.

Note	Anzahl
1	2
2	19
3	7
4	6
5	6

Tabelle 20: Schulnoten

Es zeichnet sich eine überwiegend gute Bewertung der Übersetzungsqualität ab. 21 Probanden (52,5%) fanden die Übersetzungsqualität sehr gut bzw. gut.

4. Gibt es etwas, was Ihnen bei der Arbeit mit dem Personal Translator besonders gefallen hat? (mehrere Antworten möglich)

Besonderheit	Anzahl
Einfache Bedienung, einfache Übersicht und Funktionalität	20
Wörterbuch und verschiedene Sachgebiete	12
Schnelligkeit	11
Integration mit Word und mit Internet	6
Sprachausgabe	3
Nichts besonders Positives aufgefallen	4

Tabelle 21: Besonders positiv empfundene Aspekte

Drei Punkte wurden von den Benutzern besonders positiv hervorgehoben. Die Bedienung des PT ist überwiegend einfach, mit einem Button kann man mühelos den ganzen Text übersetzen. Das Wörterbuch wurde ebenfalls als sehr nützlich empfunden, da es zusätzliche Funktionalitäten bietet, die über die bloße Übersetzung hinausgehen (automatische Erkennung der Grundform, Angaben zur Syntax, Idiomatik, Übersetzungsbedingungen und Verwendung). Die Schnelligkeit des PT wurde ebenfalls von vielen Teilnehmern lobend erwähnt.

5. Gibt es etwas, was Ihnen gar nicht gefallen hat? (Mehrere Antworten möglich)

Besonderheit	Anzahl
Nichts besonders Negatives aufgefallen	20
Integration: PT mit Word, PT mit BE, PT mit Internet	6
komplizierte Menüführung, komplexe Oberfläche, schlechte Hilfsfunktion	3
schwere Installation	3
Instabilität	3
grobe Unterteilung der Sachgebiete	2
Bedienung anfangs schwierig	1
Sonstiges: Zeitüberschreitung, zu ausführliches Handbuch	2

Tabelle 22: Besonders negativ empfundene Aspekte

Die Hälfte der Teilnehmer hatte nichts zu bemängeln. Als verbesserungsbedürftig wurden vor allem die Integration des PT mit Word, Business English und dem Internet empfunden. Ein paar Teilnehmer merkten an, dass das Interface und die Menüführung zu kompliziert

seien und die Hilfe nicht ausreiche. Folgende Bedienprobleme traten in diesem Zusammenhang auf: Der Begriff "Übersetzungsweise" wurde von vielen Testpersonen anfangs nicht mit Sachgebieten in Verbindung gebracht; es kam zu Verwirrung angesichts vieler unbekannter Zeichen in der Symbolleiste; die Option "unbekannte Wörter suchen" blieb für die Nutzer unklar; einige Benutzer hatten in der Einarbeitungsphase Probleme mit dem Einstellen der Übersetzungsrichtung.

6. Haben Sie Vorschläge, Wünsche, Anregungen, wie man den Personal Translator verbessern könnte?

	Anzahl
Bessere Integration von PT mit Word bzw. Web-Browser und bessere Integration von PT mit BE; einheitliche Oberfläche von PT & BE, Wörter direkt im Web-Browser übersetzen	11
Unbekannte Wörter sofort anzeigen z.B. durch Markierung im Text, Live-Update des Wörterbuches via Internet, Umgangssprachewörterbuch, einfacheres Eintragen unbekannter Wörter in das Wörterbuch	11
Bessere (bzw. einfache) Menüführung und wenig komplexere Oberfläche	6
Automatische Erkennung der Übersetzungsrichtung	3
Anzeigen von Übersetzungsvorschlägen	2
Automatische Erkennung von Sachgebieten	2
Bessere Hilfefunktion	2
Rechtschreibkontrolle im Quelltext	1
Sonstiges: Mehr Sachgebiete, dasselbe Programm für Macintosh, verschiedene Stufen zu übersetzen, keine Zeitbegrenzung, besseres Manual	6

Tabelle 23: Verbesserungsvorschläge

Die meisten Benutzerwünsche bezogen sich auf die Integration des Personal Translator mit anderen Komponenten (Webbrowser, Word, Business English) bzw. auf das Wörterbuch und den Umgang mit unbekanntem Wörtern.

Des Weiteren wurden ein automatisches Update aus dem Internet gewünscht sowie ein leichteres Eintragen neuer Wörter in das Wörterbuch. Das Interface sollte nach Meinung einiger Benutzer auf wenige essentielle Funktionen reduziert werden. Die Vielzahl an unbekanntem Icons wirkte verwirrend, zumal nur sehr wenige Funktionen tatsächlich benutzt werden. Außerdem wurde eine Rechtschreibkontrolle im Quelltext vorgeschlagen. Da Fehler im Quelltext automatisch zu Übersetzungsfehlern im Zieltext führen, wäre die Rechtschreibkontrolle ein sinnvoller Schritt zu weniger Übersetzungsfehlern und zu noch mehr Zeitersparnis bei der automatisierten Übersetzung.

8. Abschließende Bewertung

Durch seine **hohe Aufgabenangemessenheit** und die beachtenswerte Schnelligkeit konnte der Personal Translator im Usability Test einen positiven Eindruck vermitteln. Vor allem die **Effizienz** konnte mit dem Personal Translator **deutlich gesteigert** werden. Auch die subjektiven Bewertungen der Probanden zeigen ein überwiegend positives Bild. Die Arbeit mit dem Personal Translator wurde als weniger anstrengend und zeitaufwendig empfunden und machte mehr Spaß als das Übersetzen oder Texte erstellen mit einem normalen Wörterbuch. Vor allem für Kunden mit mäßigen Englischkenntnissen stellt der Personal Translator eine sehr sinnvolle Unterstützung dar, die nicht nur zu erhöhter Produktivität und geringerem Zeitaufwand verhilft sondern auch zu besseren Qualitätsbewertungen führt.

Das Auffinden der wichtigsten Befehle auf der Symbolleiste könnte durch eine auffälligere Platzierung und den Verzicht auf einige Funktionen vereinfacht werden. Bei der Funktionalität sind ebenfalls noch Ergänzungen oder Veränderungen vorstellbar (vgl. hierzu auch Vorschläge der Probanden, Tabelle 23), die die Bedienung noch weiter vereinfachen könnten.

Es wird in Zukunft verstärkt darum gehen, solche Nutzergruppen anzusprechen, die bislang vom Personal Translator in geringerem Maße profitieren. Personen mit guten bis sehr guten Englischkenntnissen nutzen verstärkt andere Funktionen des Personal Translator und weniger die reine Übersetzungsfunktion. Eine automatische Übersetzung kann für diese Probanden niemals so gut sein wie ihre eigene und selbst mit Nachkorrektur erzielen sie schlechtere Ergebnisse als mit einer selbständigen Übersetzung (vergleiche Ergebnisdarstellung 5.1.2). Probanden mit guten Vorkenntnissen brauchen eher das Wörterbuch für gezieltes Nachschlagen sowie die Idiomatikfunktion für bestimmte Wendungen und Ausdrücke, die man mit Vokabeln allein schwer erschließen kann. In einer so hoch idiomatischen Sprache wie dem Englischen stellt die Idiomatikfunktion ein unerlässliches Feature für Kunden dar, die nicht nur einen korrekten, sondern auch einen sprachlich geschliffenen Text erstellen möchten. Werden Funktionalität und Usability dieser beiden Funktionen weiter optimiert, können auch Nutzer mit hohem Ausgangsniveau in der englischen Sprache noch Effektivitäts- und Effizienzsteigerungen erreichen.

