

BACHELORARBEIT

Entwicklung eines chinesischen Schriftzeichen-Lernprogramms in Virtual Reality mit Auswertung über Optical Character Recognition

vorgelegt am 05 Dezember 2023
Kevin Kaiwen Yip

[REDACTED]
Media Systems

Erstprüferin: Prof. Dr. Roland Greule,
Zweitprüfer: Mirco Hülsemann

**HOCHSCHULE FÜR ANGEWANDTE
WISSENSCHAFTEN HAMBURG**

Department Medientechnik
Finkenau 35
20081 Hamburg

Inhaltsverzeichnis

Abkürzungsverzeichnis	5
Abbildungsverzeichnis	6
Eigenständigkeitserklärung	8
1 Einleitung.....	1
1.1 Fragestellung der Bachelor Arbeit.....	3
1.2 Aufbau der Bachelor Arbeit	3
2 Definition von CALL	4
2.1 Behavioristischer CALL.....	5
2.2 Kommunikativer CALL	6
2.3 Integrativer CALL.....	8
2.4 Multimedialer CALL.....	8
2.5 Web-basierter CALL.....	9
2.6 Mobil assistiertes Sprachlernen (MALL).....	11
3 CALL und die Implementierung von VR.....	11
3.1 Beispielprogramme für VR Implementierung.....	14
3.1.1 NounTown.....	14
3.1.2 Hanzi	16
4 Die chinesische Schrift und Sprache	18
4.1 Sprachgeschichte Mandarin.....	18
4.2 Geschichte der chinesischen Schrift.....	20
4.3 Vereinfachung der chinesischen Zeichen.....	23
4.4 Die Herausforderung von Mandarin-Chinesisch.....	24
4.4.1 Tonalität und Aussprache.....	24
4.4.2 Zusammenhang zwischen Wort, Silbe und Morphem.....	27
4.4.3 Anzahl der Zeichen.....	28
4.4.4 Kategorien der chinesischen Zeichen.....	28
4.4.5 Piktogramme - Xiàngxíng (象形) „Bildzeichen“	29

4.4.6	Einfache Ideogramme - Zhǐshì (指事) „auf Tatbestände deuten“	29
4.4.7	Zusammengesetzte Ideogramme - Huìyì (会意) „Vereinigte Bedeutungen“	29
4.4.8	Schriftbild: Schreibrichtung und Strichfolge.....	32
4.4.9	Die Erlernbarkeit von Chinesisch als L2.....	36
5	Virtuelle Realität – Was ist VR?	37
5.1	Historie und Meilensteine der Virtual Reality.....	38
5.2	Motivation und Einsatzmöglichkeiten von VR außerhalb Gaming.....	39
5.2.1	Visualisierung.....	40
5.2.2	Virtuelles Training.....	40
5.3	Wie wird der Effekt einer virtuellen Realität erreicht?	41
5.3.1	Wahrnehmung in VR und menschliches Informationsverarbeitungssystem.....	42
5.3.2	Technische Anforderungen	44
5.3.2.1	Echtzeitfähigkeit.....	44
5.3.2.2	Interaktivität	44
5.4	Eingabesysteme	44
5.4.1	Optisches Tracking.....	46
5.4.2	Finger Tracking	47
5.4.3	Laser Tracking.....	48
5.4.4	Eye Tracking	49
5.5	Ausgabegeräte	52
5.5.1	Valve Index	53
5.5.2	Meta Quest 2	54
6	Zukunft und Ausblick der VR	55
6.1	Bruch der Immersion/Probleme/Nachteile von VR.....	56
7	OCR Erkennung mittels Tesseract.....	56
7.1	Meilensteine mit Tesseract.....	57
7.2	Funktion.....	58
7.3	Herausforderungen mit Tesseract.....	60

8	Methodik.....	60
8.1	Auswahl der Methodik	60
8.2	Begründung der Auswahl technischer Elemente.....	60
8.2.1	Oculus/Meta Quest 2	60
8.2.2	Unity	61
8.2.3	Integration von Tesseract 4.0 als OCR.....	63
8.2.4	Grundlegendes Tesseract API Setup	63
8.2.5	Darstellung des Textes	67
8.2.6	CharacterVideoScript.cs.....	70
8.2.7	SoundMananagerScript.cs.....	71
8.3	Vorstellung der Software aus Nutzersicht.....	71
8.3.1	Experiment mit Versuchspersonen.....	76
8.3.1.1	Versuchsablauf:	76
8.3.1.2	Versuchsbeobachtungen:.....	76
8.3.1.3	Versuchsergebnisse:	77
8.3.1.4	Interpretation der Ergebnisse.....	77
8.3.1.5	Einschätzung eines Zeitrahmens bei Softwarenutzung des Chinese Writing Game 78	
8.4	Vorstellung der Software.....	78
8.5	Ablauf des Programms	79
9	Fazit	81
	Literaturverzeichnis.....	83

Abkürzungsverzeichnis

VR	Virtual Reality
AR	Augmented Reality
CALL	Computer Assisted Language Learning

Bei der Gestaltung des Abkürzungsverzeichnisses gilt es zu beachten, dass in der Regel nur solche Abkürzungen angegeben werden, die nicht Teil des üblichen Sprachgebrauchs sind. Abkürzungen wie z. B., u. a. oder bzw. gehören nicht in ein Abkürzungsverzeichnis. Ob ein Abkürzungs-, Abbildungs-, Tabellen- oder Formelverzeichnis oder andere Verzeichnisse erforderlich sind, hängt von der Anzahl der aufzulistenden Elemente ab.

Abbildungsverzeichnis

<i>Abbildung 1: Youtube Video von Metaquest</i>	15
<i>Abbildung 2: Schrullige klischeebehaftete spanische Dame</i>	15
Abbildung 3: Vergleich der Lernweise mit Writing Paper und in der Software	16
Abbildung 4: Zeichenaufbau und Ablauffolge in Hanzi	16
Abbildung 5 Bild aus der Entwicklerpräsentation mit Quellenangabe der Entwickler	18
Abbildung 6: Chinesische Schriftzeichen und deren Entwicklung	24
Abbildung 7: Die vier betonten Töne in Hochchinesisch	25
Abbildung 8: Die fünf Töne in Hochchinesisch	26
Abbildung 9: Pinyin Tabelle von Mandarin-Chinesisch	26
Abbildung 10: Venn-Diagramm von Homophone, Homographen und Synonymen	27
Abbildung 11: Darstellung des Begriffs Pferd	29
Abbildung 12: Abstrakta	29
Abbildung 13: Zusammengesetzte Ideogramme	30
Abbildung 14: Signifikum Phonetikum Kombination	31
Abbildung 15: Striche, Grundstriche und Kombinationsstriche	33
Abbildung 16: die acht Teile der fünf Strichzüge am Zeichen "Yong	34
Abbildung 17: Strichfolge Bild 1	34
Abbildung 18: Strichfolge Bild 2	34
Abbildung 19 Strichfolge Bild 3	34
Abbildung 20 Strichfolge Bild 4	34
Abbildung 21 Strichfolge Bild 5	36
Abbildung 22 Strichfolge Bild 6	36
Abbildung 23 Strichfolge Bild 6	36
Abbildung 24: Virtual Continuum	38
Abbildung 25 Abbildung das Modul zum Training der aseptischen Wundreinigung in Virtual Reality	41
Abbildung 26 Modell der Menschlichen Informationsverarbeitung. Nach Card et al. (1986)	43
Abbildung 27: Die sechs Freiheitsgrade	45
Abbildung 28: Inside-Out und Outside-In Verfahren	47
Abbildung 29: Sticker zur Kalibrierung	47
Abbildung 30: Fingertracking	48
Abbildung 31:; Prinzipieller Aufbau eines Eye-Tracking-HMDs	49
Abbildung 32: Erkennung mittels Anpassung durch die Eye Tracking Software	51
Abbildung 33: Valve Index Gesamtsystem bestehend aus Brille, Controllern und Basisstationen	53

Abbildung 34: Meta Quest 2 VR.....	55
Abbildung 35; OCR Grundschemata.....	57
Abbildung 36: Tesseract Parameter, Sicht aus dem Programm auf ein Zeichen.....	59
Abbildung 37: Startbild Chinese Writing Game	71
Abbildung 38: Virtueller Classroom aus dem Chinese Writing Game	72
Abbildung 39: Blick auf die Malwand	73
Abbildung 40: Zeichnungen auf der Malwand.....	74
Abbildung 41: Auswertung eines korrekt gezeichneten Schriftzeichens	75

Eigenständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Bachelorarbeit mit dem Titel:

Entwicklung eines chinesischen Schriftzeichen- Lernprogramms in Virtual Reality mit Auswertung über Optical Character Recognition

selbständig und nur mit den angegebenen Hilfsmitteln verfasst habe. Alle Passagen, die ich wörtlich aus der Literatur oder aus anderen Quellen wie z. B. Internetseiten übernommen habe, habe ich deutlich als Zitat mit Angabe der Quelle kenntlich gemacht.

Hamburg, 05.12.2023

Datum

Unterschrift

1 Einleitung

Mandarin-Chinesisch ist die meistgesprochene Muttersprache weltweit mit über 918 Millionen Muttersprachlern. Zählt man Muttersprachler und Nicht-Muttersprachler zusammen, liegt Mandarin-Chinesisch mit über 1117 Millionen SprecherInnen als zweitmeistgesprochene Sprache nur knapp hinter Englisch mit über 1132 Millionen SprecherInnen (2). Durch das sehr unterschiedliche Sprachprofil im Gegensatz zu lateinalphabetischen Sprachen wird Mandarin-Chinesisch als eine der herausforderndsten Sprachen angesehen, insbesondere für Lernende, die es als sogenannte L2, also als Fremdsprache lernen.

Durch die Schwierigkeit der Erlernbarkeit von Mandarin-Chinesisch, hat das Interesse an innovativen Lehrmethoden stark zugenommen. In diesem Zusammenhang hat sich Computer Assisted Language Learning (CALL) als wegweisende Herangehensweise herauskristallisiert. Der Begriff CALL bezieht sich allgemein auf computerunterstützte Anwendungen, die den Sprachlernprozess unterstützen. Diese reichen von traditionellen Programmen, die bereits seit 1960 und 1970 Anwendung finden, bis hin zur Integration in modernen Lösungen mit Virtual Reality und Augmented Reality.

Bei CALL liegt der Schwerpunkt in der Förderung der Eigeninitiative der Lernenden, um selbstständig und im eigenen Tempo zu arbeiten. Dabei werden allgemeine Prinzipien der Sprachmethodik und -pädagogik verschiedener Lerntheorien berücksichtigt. Die steigende Verfügbarkeit der CALL-Plattformen, insbesondere im Bereich VR, ermöglicht es Lernenden in immersive Umgebungen einzutauchen und Sprachen auf eine interaktive und individualisierte Weise zu erlernen. Da die VR-Systeme ihre Nutzer in künstliche Welten versetzen können, wird ihnen durch physische und immersive Elemente eine besonders realistische Lernerfahrung angeboten. Diese technologischen Fortschritte werfen nicht nur ein neues Licht auf die Möglichkeiten des Sprachlernens, sondern bieten auch eine vielversprechende Perspektive für die Bewältigung der Herausforderungen, die mit dem Erwerb von Sprachen wie Mandarin-Chinesisch verbunden sind.

Des Weiteren benötigt ein Lernprogramm, egal ob CALL oder Lernspiel eine Möglichkeit dem Nutzer bzw. Lerner Feedback mitzugeben, ob die Lektion absolviert wurde, aber auch ob Wörter oder Zeichen richtig geschrieben wurden. Ohne Feedback zur Lektion und Überprüfung der Nutzereingaben wäre das Lernprogramm letztendlich nur ein großer Notizblock oder eine große Malwand. In computerunterstützten Sprachprogrammen gibt es dafür unterschiedliche Ansätze wie Nutzereingaben auf Korrektheit geprüft werden. Bei einer tonalen zeichenbasierten Sprache wie Mandarin-Chinesisch, wo der Ton nicht einzigartig zu einem Zeichen gehört und das Erlernen der Bedeutung des Zeichens das Ziel ist, sollte die Überprüfung der korrekten Zeichenschreibweise Vorrang bei der Lehre der Zeichen nehmen. Hier würde sich Optical Character Recognition (OCR) als passende Auswertungsmethode

eignen. OCR ist der englische Begriff für Texterkennung, im genauen bezeichnet es die automatisierte Texterkennung bzw. automatische Schrifterkennung innerhalb von Bildern.

Aus diesen Punkten ergibt sich nun eine interessante Fragestellung:

1.1 Fragestellung der Bachelor Arbeit

„Inwiefern kann die Integration von Virtual Reality in Computer Assisted Language Learning Programme den Lernerfolg und die Effektivität beim Erlernen von Mandarin-Chinesisch unter Anwendung von Optical Character Recognition als Überprüfung- und Auswertungsmethode steigern?“

1.2 Aufbau der Bachelor Arbeit

Im Laufe dieser Bachelorarbeit wird zunächst die Historie von Chinesisch als Zeichensprache beleuchtet und anhand Mandarin-Chinesisch die Herausforderungen vom Lernen eine Fremdsprache veranschaulicht. Es wird kurz weiter im Detail auf computerunterstütztes Sprachlernprogrammen (CALL) eingegangen, wo Beispiele und die Vor- und Nachteile dessen erläutert werden.

Danach werden die Grundlagen und Vergangenheit von Virtual Reality angesprochen, es folgt eine Erklärung zur Funktionsweise der Immersion von VR, insbesondere in CALL-Programmen, wie VR in diesem Bereich als Werkzeug für das Erlernen einer Fremdsprache funktioniert und die Effizienz von CALL mit der Integration von VR erörtert.

Darüber hinaus wird betrachtet wie Optical Character Recognition heutzutage in Sprachlernprogrammen integriert werden kann, um den Lerneffekt zu erhöhen. Hier wird sich auf die Anwendung Tesseract 4.0 fokussiert, welches ein leistungsfähiges Open-Source OCR Programm ist, das von Google entwickelt wurde. Es ermöglicht die Umwandlung von gedrucktem oder handgeschriebenem Text auf Bildern in eine maschinelle Textausgabe. Es wird diskutiert wie effektiv Tesseract 4.0 als Evaluierungsmöglichkeit des Lernprozesses im Bildungsbereich ist.

Anschließend wird dies im Zusammenhang mit der Vorstellung Lernprogramms „Chinese Writing Game“ gezeigt, welches mit Unity Version 2021.3.15f1 und in Rahmen dieser Bachelorarbeit entwickelt wurde. Es handelt sich um ein chinesisches Schriftzeichenlernprogramm gerichtet an Anfänger mit keinen chinesischen Sprachkenntnissen oder Lerner mit einem Sprachniveau von A1. Es wird über Entwicklungsprozesse, Designüberlegungen und Funktionalitäten des VR-Lernprogramms gesprochen und es wird auf einen qualitativen Versuch mit Testpersonen eingegangen, wo die Effektivität des Programms mit Lernpersonen mit verschiedenem Vertrautheitslevel hinsichtlich VR und Mandarin-Chinesisch bewertet wird. Ebenfalls wird die Effektivität von Tesseract bei der Integration von VR-basiertem CALL in dieser Sprachlernanwendung untersucht.

Schließlich wird die Bachelorarbeit mit dem Fazit beendet, wo ein Ausblick von CALL-Programmen erfolgt. Anhand von „Chinese Writing Game“ wird gezeigt, wo Verbesserungs- und

Erweiterungsmöglichkeiten sind, um VR-basierte CALL-Programme noch effektiver für das Lernen von Chinesisch zu gestalten.

Das Ziel dieser Arbeit ist es, einen Beitrag zur Verbesserung des Chinesischlernens für Anfänger zu leisten, innovative Lernumgebungen zu nutzen und die Wirksamkeit der verwendeten Technologien

2 Definition von CALL

In der Einleitung wurde kurz auf die Historie als auch auf den Zweck von CALL eingegangen. Wie angesprochen kann CALL definiert werden als die Untersuchung und Nutzung der Anwendungen des Computers beim Lehren und Lernen von Sprachen (Levy, 1997: 1). Es repräsentiert auch eine Seite des E-Learnings (Yuan, 2007:416) und konzentriert sich hauptsächlich auf den Zweitspracherwerb (Chapelle). Dieses Kapitel dient dazu die verschiedenen Einsatzzwecke von CALL zu beleuchten und die technologische Entwicklung zu erörtern und diese in Bezug zum Programm „Chinese Writing Game“ zu setzen.

Hierzu genügt ein elementares Verständnis der Entwicklung innerhalb CALLs als solcher. Es folgt eine grundlegende Verknüpfung mit dem Kontext der Sprachmethodik respektive -pädagogik, da eine inhaltlich detaillierte Ausführung die Abgrenzung unnötig erschweren würde. Nach dieser Entwicklung und Verständnis der Technologie hinter den Applikationen gibt es beispielhafte Erläuterungen aus der Realanwendung – hier PflegeDigital 2.0 und [Feuer VR Testzweck]

CALL gibt es seit den 1960ern. Damals noch in der Einsatzform in Mainframes und Plattenspielern, später auf CDs oder MP3 Spielern, DVD und dann auch über interaktive Whiteboards, heutzutage in Form der Smartphones, Tablets. Das Lernerlebnis erstreckt sich aktuell sogar auf virtuellen Welten, durch VR- und AR (Augmented Reality) Head-Mounted-Displays. Im Prinzip kann jedes technologische Gerät eingesetzt werden, um CALL zu unterstützen. Der Ansatz von CALL entstand durch CAL „computer assisted learning“, dem allgemeinen Oberbegriff für den Einsatz von Computern im Bildungsbereich, wohingegen die Ergänzung des zusätzlichen L für Language also für den Kontext des Sprachlernens steht.

„Die Nutzung von elektronischen Geräten/Computern für die Bereitstellung zu Bildungs- und Lernzwecken.“ the use of electronic devices/computers to provide educational instruction and to learn.”

Nach Warschauer (1996) wurden im Laufe des technologischen Fortschrittes 3 Hauptphasen von CALL mit unterschiedlichen Ansätze entwickelt:

- behavioristischer CALL (1960 bis 1970)
- kommunikativer CALL (1970 bis 1980)

- integrativer CALL (1990 bis heute)

Jeder Ansatz korrespondiert zu bestimmten technologischen Mitteln und zu bestimmten pädagogischen Ansätzen. Sie unterscheiden sich in der Integration und Aufgaben der Computertechnologie im Lernbereich, der Interaktion der Schüler zu den Computern sowie welche Kontrolle die Schüler über ihr Lernverhalten und -tempo der Software haben. Wichtig ist anzumerken, dass die Einführung einer neuen CALL-Phase nicht immer die vergangene Phase ausschließt und ersetzt, viel mehr wird eine alte CALL-Phase in die neuen subsumiert, miteingebunden und weiterentwickelt.

2.1 Behavioristischer CALL

Behavioristischer CALL wurde als erster Ansatz in den 1950ern konzipiert und in den 1960 bis 1970ern implementiert. Das behavioristische CALL basiert auf behavioristische Lerntheorien von B.F. Skinner, welche davon ausgehen, dass Lerneffekte durch Reize von Beobachtungen und den daraus resultierenden Reaktionen in Form einer Belohnung hervorgerufen werden. Dies geschieht durch das Aufbauen einer Angewohnheit in Form von Mimikry und Auswendiglernen, im Sprachlernen z.B. das Auswendiglernen von Dialogen und Satzmustern. Die Konditionierung zu positiven Konsequenzen wie Lob oder Belohnungen sollen als Motivation dienen, sodass weiteres Verlangen nach Erfolg und Anerkennung verspürt wird und die Lernübungen wiederholt und weiterverfolgt werden.

Für den Zweck von CALL wurden Sprachlernübungen mit dem PLATO-System (Programmed Logic for Automatic Teaching Operations) entwickelt, welches ein Mainframe ist, das zu der Zeit für computerbasierte Bildungs-, Lern- und Tutorialzwecke eingesetzt wurde. Die Lernübungen bestanden aus repetitiven, umfangreichen und expliziten Vokabel- und Grammatiktests oder Übersetzungsaufgaben. Einen Fokus gab es auf audio-linguale Lernmethoden, da es nach dem Zweiten Weltkrieg eine große Nachfrage nach Fremdsprachlicher Kompetenz im Hör- und Sprechverstehen gab. Nach Warschauer (2012) ist ein Computer ideal, um wiederholende Drillaufgaben auszuführen, da der Maschine nie langweilig wird und der Computer Lernmaterial stets dem Lerntempo der Lernenden anpassen kann. So wird der Lernprozess in Blöcke geteilt und in jedem Lernblock wird das Wissen eingedrillt, bevor es zum nächsten Lernblock geht. Diese Lernprogramme wurden so designt, dass den Lernenden unmittelbar positives oder negatives Feedback zur formalen Genauigkeit gegeben wurde. Behavioristischer CALL verfolgt strikt den Ansatz, dass der Computer nur wie ein mechanischer Tutor angesehen wird. Das Bearbeiten der Aufgaben im eigenen Tempo und die individuelle Gestaltung des Lernmaterials wurde als revolutionär angesehen. Die Auseinandersetzung der Lernenden zu repetitiven Übungen wurde als nützlich befunden und waren zu der Zeit das Hauptelement von CALL.

Allerdings wurde der behavioristische CALL gegen 1970 viel kritisiert. Kritiker betonten, dass der Schwerpunkt auf Wiederholungen und Genauigkeit letzten Endes nicht ausreichend sind, um die kommunikative Kompetenz der Zielsprache zu erreichen. Durch den behavioristischen Ansatz wird

Lernenden zu stark der Fokus auf das Einprägen von einzelnen Vokabeln, Grammatikregeln und isolierten Strukturen gesetzt. Die Annahme, dass Lernende durch die richtige Aneinanderreihung diskreter Komponenten Sprachen erwerben können, vernachlässigt den kommunikativen Aspekt der Sprache. Nach Noam Chomsky sei Sprache keine Gewohnheitsstruktur. Die gewöhnliche Sprache, Gewohnheiten und Verhalten beinhalten durch Innovation, die Bildung neuer Sätze und Muster gemäß Regeln von großer Abstraktheit und Komplexität, was durch den behavioristischen CALL nicht authentisch vermittelt werden konnte.

Durch die Kritik wurden zwar behavioristische Lernmethoden abgelehnt, die Lehrprogramme, die mit dem behavioristischen CALL entwickelt wurden, dennoch nicht, da mehrere Begründungen immer noch für die Nutzung sprachen. Repetition durch Computer wurde zum Teil immer noch als essenziell fürs Lernen angesehen. Das Bestimmen des eigenen Lerntempos und der verfolgbare Lernfortschritt waren hierfür ausschlaggebend. Generell motivierte der Umgang mit der noch aufkommenden und sich stetig mehr verbreitenden Computertechnologie. Dies wirkte sich auf die Lernintensität und Häufigkeit aus, da die Konzentration der Lernenden deutlich länger aufrecht erhalten konnte als mit traditionellen Lehrmethoden und ein Lehrfortschritt kontrolliert werden konnte. (Simonson and Thompson, 1997: 96).

2.2 Kommunikativer CALL

Die Kritik an den behavioristischen CALL erreichte in den 1970ern und 1980ern den Höhepunkt, als das kommunikative CALL entstand. Dieser korrespondiert zu kognitiven Lerntheorien, welche davon ausgehen, dass Lernen ein kreativer Prozess durch Entdeckungen, Ausdruck und Entwicklung ist. Befürworter der kognitiven Lerntheorien legen Wert auf wie Lehranweisungen strukturiert sind. In der Sprachlehre bedeutet das, dass die Sprache als Ganzes gesehen wird statt in isolierte Einzelteile, zum Ausdruck, um Intentionen klarzustellen, wie wenn man z.B. nach Weganweisungen fragt oder sich über die eigenen Hobbies ausdrücken möchte. Insgesamt soll Sprache als Medium zur Kommunikation gesehen werden. Übungen sind mehr aufgabenbasierend und sollen auch kollaborativ unter Lernenden ausgeführt werden. Grammatik soll eher implizit statt explizit beigebracht werden und es soll mehr Wert auf die Nutzung der Zielsprache im Unterricht gelegt werden (Jones & Foretsue, 1987; Phillips, 1987; Underwood, 1984) und die Rolle des Feedbacks wurde stark überdacht. Lernende sollten fähig sein eigenständig zu lernen und sich selbst intrinsisches Feedback zu geben, anstatt nur auf extrinsisches Feedback von Lernenden oder vom Computer abhängig zu sein. Es entstand ein Wechsel von einer lehrerzentrierten Lehrform in eine lernendenzentrierten Lehrform.

Für die computertechnologische Umsetzung des kommunikativen CALLS wurden Mainframes durch Verfügbarkeit, die sinkenden Preise und Marktstrukturen weitergehend durch personalisierte Computer ersetzt. Die Übungen bestanden deutlich weniger aus einfachen repetitiven Aufgaben, sondern aus Lernspielen, Lese-, Verständnis- oder auch Schreibaufgaben. Die Programme und Lehrende sollten

Lernende viel mehr durch den Lernprozess leiten und unterstützen, anstatt ihnen klare Lernvorgaben und eine kontrollierte Lernumgebung zu geben.

Einen großen Fokus gab es auf Textrekonstruktionprogramme. Diese Programme verschleiern einen kurzen Text auf dem Bildschirm und ersetzen jeden Buchstaben durch ein Zeichen. Durch das Erraten einzelner Wörter muss der Text wiederhergestellt werden. Verschiedene Hilfestufe stehen zur Verfügung, darunter einzelne Buchstaben, Wörter oder der ganze Text. Ein anderer Ansatz waren Textmanipulationsprogramme, bei denen Lernende verschiedene Arten von Manipulation durchführen wie die Neuordnung eines durcheinander geworfenen Textes.

Bei den Lernprogrammen, die mit dem Ansatz des kommunikativen CALLS erstellt wurden, wird nicht vorgesehen Lernende für falsche Antworten zu bewerten und für ihre Leistungen belohnt zu werden, egal ob mit einer sichtbaren oder mit einer auditiven Erfolgsmeldung. Der Computer wird weiterhin als Tutor eingesetzt, allerdings werden die Lernenden nicht mehr gedrillt und es gibt ausgeprägtere Wahl-, Kontroll- und Interaktionsmöglichkeiten um den Aspekt der sozialen Interaktion erweitert. Durch den kommunikativen CALL entwickelte sich der Computer als Stimulus der Diskussionen und kritisches Denken unter Lernenden zu stimulieren und zu initiieren. In Rahmen dessen entstanden Simulationsprogramme, in denen es um Problemlösungen ging. Diese Simulatoren sollen den Lernenden reale Situationen geben und sie aktiv in eine Lernumgebung einschließen, wo sie Entscheidungen treffen können und Lernresultate in der Lernumgebung sehen können. Beispiel hierfür ist „Treasure Hunt“, ein Lernspiel in Form einer Web Quest, wo Hinweise und Rätsel gelöst werden müssen, um den Schatz zu finden.

Zusätzlich wurde anfangen den Computer als Werkzeug zu sehen, um eigenes Schreiben anzuregen. Hier wurden Programme genutzt, die nicht unbedingt spezifisch fürs Sprachlernen entwickelt wurden, aber helfen Sprache besser zu verstehen. Rechtschreib- und Grammatikprüfprogramme stellen effektive und kosteneffiziente Werkzeuge für Lernende dar, um ihren Text zu prüfen. Kollaborative Schreibwerkzeuge wie „Daedalus Integrated Writing“ ermöglichen Lernenden über ein lokales Netzwerk eine gemeinsame Arbeitsumgebung, fördern den Austausch von Ideen und unterstützen die kollektive Entwicklung schriftlicher Kompetenzen und tragen ultimativ zur verstärkten Interaktion und Zusammenarbeit im Sprachlernprozess bei.

Die Aufgaben konnten allein oder in Gruppen bearbeitet werden. Das Lernen der Sprachschemata und Bedeutungen von Sätzen wurde durch die Zusammenarbeit und Diskussionen in Lerngruppen angeregt. Somit lag der Fokus nicht auf dem, was die Lernende mit den Computern interaktiv durchgeführt haben, sondern was die Lernende untereinander erarbeitet haben, als sie mit den Computern gearbeitet haben.

Dennoch gab es Kritiken zum kommunikativen CALL. Es wurde angemerkt, dass der Computer ad hoc genutzt wurde. Die Anwendung des Computers im Lernprozess war unkoordiniert und erfolgte in einer unstrukturierten Weise. Dies wies auf eine fehlende Integration des Computers in den Lehrplan hin.

Zudem war die Vermittlung von Sprachfähigkeiten immer noch zu stark segmentiert, isoliert und demnach unzureichend verknüpft, da verschiedene Sprachfähigkeiten nicht in Zusammenhang gebracht wurden. Das führte dazu, dass Lernende immer noch Schwierigkeiten hatten das Gelernte in umfassenderen sprachlichen Situationen anzuwenden. Durch die Kritik wurde sich von Lehrenden gewünscht eine integrativere und vernetztere Herangehensweise an den Sprachunterricht zu fördern.

2.3 Integrativer CALL

Die dritte Phase von CALL, der integrative CALL entwickelte sich in den 1990ern basierend auf den Kritiken des kommunikativen Ansatzes. Lehrende wandten sich von dem kognitiv kommunikativen Ansatz ab und verlagerten ihren Fokus auf eine soziokognitive und soziokulturelle Ansicht, die beide den authentischen sozialen Kontext betonen. Während die soziokognitive Ansicht vertritt, dass Sprachlernentwicklung nicht nur auf individuellen kognitiven Fähigkeiten beruht, sondern auch stark von der Interaktion des sozialen Umfelds beeinflusst wird, hebt die soziokulturelle Ansicht zusätzlich die Wichtigkeit der kulturellen Umgebung hervor - in der in einem sozialen Kontext die zu erlernende Sprache gemeinsam genutzt wird. Beispielsweise mit einem Tandempartner, der die Sprache besser beherrscht als die lernende Person (Schieffelin & Ochs, 1986; cited in Warschauer and Meskill, 2000). Soziokulturelle Theorien beschreiben nicht explizit, wie Lernenden die kulturellen Werte eine Fremdsprache erwerben können, aber sie erklären, wie das Wissen einer Fremdsprache durch soziokulturelle Erfahrungen verinnerlicht wird. Damit den Lernenden authentische soziale Interaktionen und damit verständlicher Input (Krashen, 1983) mit möglichst vielen Sprachdaten gegeben werden kann, muss eine Umgebung erschaffen werden, in der Aufgaben und Projekte kollaborativ durchgeführt werden können, während sie durch Lehrende unterstützt werden.

Auf der technischen Seite vom integrativen CALL war das Ziel die technologische Integration von verschiedenen Sprachfähigkeiten und somit sowohl das Hör- und Leseverständnis als auch die Sprach- und Schreibfähigkeit zu verbinden und diese in den Sprachlernprozess einzubeziehen. Übungen sollen mehr aufgaben-, projekt- und inhaltsbasierte Ansätze verfolgen, damit Sprachlehre im authentischen und sozialen Kontext verwendet werden kann. In diesem Zusammenhang haben sich der multimediale CALL und der Web-basierte CALL sich als die zwei wichtigsten Ansätze herauskristallisiert, später auch das mobil assistierte Sprachlernen (MALL) durch Smartphone-Technologien, was revolutionär für den computertechnologischen Zweitspracherwerb war und derzeitig der verbreitetste Stand ist.

2.4 Multimedialer CALL

Der multimediale CALL bezieht sich auf Lernprogramme und -übungen, die verschiedene Arten von Medien und Technologien integrieren, um ein umfassenderes Sprachlernumfeld zu erschaffen. Computern in den 90ern gelang es multimediale Elemente wie Text, Audio, Video, Grafiken und

interaktive Elemente zu verbinden. Beispielsweise konnten Texte mit Audioaufnahmen verknüpft werden, um das Leseverständnis zu fördern oder interaktive Grafiken genutzt werden, um Sprache spielerisch zu erkunden. Zur gleichen Zeit wurden CD-ROM erstmals veröffentlicht und Mitte 1990 gab es bereits eine große Nachfrage an CD-ROM Programmen für Sprachlernende, wie zum Beispiel an „*Who is Oscar Lake? An Interactive Learning Adventure*“. Die Oscar Lake-Reihe sind Point-and-Click Videospiele in einem Mystery-Abenteuer-Setting. Dialoge von Charakteren sind von Muttersprachlern vertont und unbekannte Wörter und Sätze können durch Tools übersetzt werden. Dazu kann der Spieler das Tempo des Spiel und seinen Lernverlaufs kontrollieren. Durch die Interaktion des Spielers mit einer dynamischen Umgebung werden nicht mehr Vokabeln oder Grammatik passiv wiederholt und Lernenden wird eine effektivere Möglichkeit des Zweitspracherwerbs geboten und durch Gamification wird deutlich länger im Spracherwerb gearbeitet.

Mit der Veröffentlichung von DVDs am Ende der 1990er konnten auch Videoaufnahmen mit einer höheren Qualität erzeugt werden – durch den Zuwachs an unmittelbarer Speichermenge pro Medium. Ein beliebtes DVD Lernprogramm war Eurotalk DVD-ROM. Eurotalk spezialisiert sich auf Sprachlernsoftware vom Anfänger bis hin zum Fortgeschrittenen für verschiedene Sprachen und fokussiert sich auf interaktive Möglichkeiten des Vokabel-, Grammatik, Aussprache-Erlernens.

2.5 Web-basierter CALL

Gegen 1993 befanden sich die Grundzüge des heutigen World Wide Web in rapider Entwicklung. Zu der Zeit war es erstmals möglich Zugang zu digitalen Lernressourcen zu eröffnen und Lehrmaterial über das Internet zu verbreiten. Lernende konnten in einer Vielzahl von Lernmaterial recherchieren, dass zu Ihren Bedürfnissen passt (z.B. in virtuellen Bibliotheken) und dieses in wenigen Minuten ausfindig machen. Auf Texte, Audio- und Videomaterial konnte beliebig zugegriffen werden und interaktive Tests absolviert werden.

Mit dem Aufkommen des Web 2.0 ab 2003 waren noch mehr fortgeschrittenere und interaktive Aktivitäten möglich. Die mittlerweile schnelleren Internetzugänge erlaubten es mit dem sogenannten Web 2.0 Informationen in Form von aufwendigeren Dateiinhalten und -größen zu verbreiten, sondern auch selbst Inhalte zu kreieren und diese mit jedermann zu teilen. Möglich war das mit Videoportalen wie YouTube und Vimeo oder mit Online-Enzyklopädien wie Wikipedia. Zum anderen gab es durch computervermittelte Kommunikation (CMC) mehr Möglichkeiten soziokulturelle Beziehungen herzustellen. CMC meint den kommunikativen Austausch von Nachrichten und digitalen Medien durchs Internet. Es wird unterschieden zwischen der synchronen CMC in Echtzeitdiskussionsumgebungen (Livechat) oder Onlinekonferenzen und der asynchronen CMC mit E-Mails, Blogs, Diskussionsforen oder sozialen Medienplattformen wie Facebook. In beiden Fällen ist eine Kommunikation mit anderen

Lernenden der Zielsprache oder Muttersprachlern möglich und die zeitlichen und örtlichen Limitierungen können überbrückt werden.

2.6 Mobil assistiertes Sprachlernen (MALL)

Mobil assistiertes Sprachlernen lässt Lernende noch flexibler ihren Lernort und Lernzeit bestimmen, da mit der Portabilität keine örtlichen und zeitlichen Grenzen mehr ausgesetzt waren. Durch die Einführung von Mobiltelefonen und Smartphones und der Kombination zu Web-Technologien ist der Kontext vom multimedialen CALL und web-basierten CALL jederzeit verfügbar. Zudem entstanden neue Lernmöglichkeiten, die explizit auf die Nutzung von Smartphones ausgelegt ist, wie zum Beispiel mobile Sprachlernplattformen wie Duolingo oder Babbel. Obwohl die Nutzung eines Computers vorteilhafter als ein Smartphone ist bei der Darstellung von Visuellem, Sound und Textinformationen ist die Portabilität der Mobiltelefone dem Computer überlegen (Yamaguchi, 2005).

3 CALL und die Implementierung von VR

Im integrativen CALL werden Lernende mehr angeregt die Zielsprache in realistischen Situationen anzuwenden. Es bietet die bisher multimedial interaktivsten Tools mit den meisten Ressourcen zum Sprachlernen an. Die Möglichkeit zur internationalen Vernetzung zu Menschen verschiedener Kulturen oder zur Kontenterstellung und -verbreitung regt die Kreativität von Lernenden an. Durch das textgesteuerte Internet können Zeit und Ort keine Barriere, sodass schüchtern, zurückhaltende Menschen ebenfalls in Ruhe an sprachlichen Diskussionen und Informationsaustausch beteiligt sein können, mit einem möglichen Resultat für bessere sprachliche und kommunikative Fähigkeiten, verbesserte Einstellung gegenüber des Lernprozesses und auch einen Aufbau im Selbstbewusstsein.

Hindernisse im integrativen CALL schienen erstmal zu existieren. Zunächst betraf dies finanzielle Limitierung zum Erwerb der Computertechnologie und technische Probleme, die allerdings im Laufe der Zeit weniger zur Beschränkung wurden. Zum einen wird Computertechnologie, gerade solche, die nicht kürzlich auf dem Markt erschien, immer erwerbbarer für privat Konsumierende, dies gilt sogar für Virtual Reality HMD. Als verhältnismäßig kostengünstiges Produkt erwies sich die Meta Quest 2 als einer der meist gekauften VR Headsets bisher. Die Integration von VR in CALL wird im Kapitel 5.2 näher angesprochen. Technische Probleme mögen zwar noch in der Aufrechterhaltung und der Datenbandbreite einer kabellosen Internetverbindung liegen, dies wird jedoch mit der Einführung von Glasfaserverbindungen ausgebaut. Auch in der künstlichen Intelligenz gab es in den letzten Jahren revolutionäre Entwicklungen und Veröffentlichungen wie ChatGPT als Chatbot mit einem KI-gesteuerten Sprachmodell, welches von OpenAI entwickelt wurde. Mit ChatGPT können komplexe, detaillierte Eingaben verarbeitet werden und sowohl sprachlich als auch menschlich realistisch klingende Ausgaben zurückgegeben werden können, dessen Schreibstil und Informationsmuster sogar beeinflusst werden kann.

Auch im Sprachlernbereich hat künstliche Intelligenz eine große Rolle gespielt. Der Onlinedienst DeepL für maschinelle Übersetzung, der erstmals 2017 veröffentlicht wurde, ist in den letzten Jahren zunehmend bekannt geworden qualitativ hochwertige, kontexthaftete und sprachlich flüssig und natürlich klingende Übersetzungen liefern zu können. Zugrunde liegt hier der Deep Learning Ansatz, womit neuronale Netzwerke des Systems mit großen Mengen von Textdaten trainiert wurden. DeepL verbleibt damit bisher ein fortgeschrittenes und konkurrenzfähiges Sprachlernwerkzeug. Wobei dennoch der Einsatz von Computertechnologie im Sprachlernprozess weiterhin auf menschliche Unterstützung angewiesen ist beim Umgang mit unerwarteten Situationen. Es wird ein Experte im Umgang mit grundlegender Informations- und Kommunikationstechnik (ICT) benötigt, bevorzugterweise die trainierte Lehrperson, die die Lernende in die Nutzung der benötigten Computertechnik einweisen kann. Ohne eine Einweisung könnte der Umgang mit der Technik sich als herausfordernd erweisen und auf Seite der Lernenden Frustration statt Lernmotivation aufrufen. Zudem kann durch die Verfügbarkeit einer immer größer werdende Sprachdatenmenge im Internet eine Entscheidungsblockade entstehen, weil es herausfordernd sein kann das passendste Lernmaterial auszuwählen wird. Auf Seiten der Lehrenden könnte eine Ablehnung der Computertechnologie erfolgen, falls ihnen die Nutzung nicht gelingt, weshalb die Nutzerfreundlichkeit und ein Training zur Verinnerlichung der Nutzung der CALL-Software bzw. Technologie gerade für Laien essenziell ist. Die Wichtigkeit der Lehrenden ist maßgeblich zu betonen, da ihnen die Entscheidungskompetenz inne wohnt wie sie die Internetressourcen von CALL in den Sprachunterricht einbinden. Warschauer (1997) erstellte Richtlinien dazu, viele handeln von einer guten Beziehung und Einstellung zwischen Lehrenden, Lernenden und der Computertechnologie. Der Computer sollte als Werkzeug möglichst so in den Unterricht eingebunden werden, dass solches Lernmaterial gezeigt wird, das durch Bücher oder andere Übungen aufgrund seiner Art, Aufbau oder Struktur nicht möglich wäre. So erhält die Computertechnologie einen qualitativen und innovativen Stellenwert, der das Sprachlernen motiviert und die Teilnahme der Lernenden steigert.

Der Einsatz vom integrativen CALL reicht bis heute. CALL-Forschungen und Methoden haben eine große Änderung vollzogen, da sie sich von den anfänglichen Ansichten den Computer ausschließlich als Tutor für repetitive Drillübungen zu nutzen stetig zur Ansicht weiterentwickelten den Computer als grenzenloses Medium fürs Sprachlernen zu sehen. Mittlerweile besteht die Möglichkeit ein Leben lang das Lernen zu transferieren. (Warschauer & Healey, 1998). Durch die Geschichte von CALL hat sich herausgestellt, dass ungeachtet des technologischen Fortschritts die Rolle der Lehrenden in der Lehrumgebung immer wichtig bleibt und nicht vollständig ersetzt werden kann. Computertechnologie bleibt ein wertvolles Werkzeug zum Sprachlernprozess, aber nur durch Lehrende wird wertvolles Feedback mitgeteilt, wie, ob und welche Internetressourcen ausgewählt werden können. Lehrende nehmen die Rolle des Mentors ein, die dann die soziale Entwicklung und intellektuelles Wachstum der

Lernenden verfolgen und bestimmen indem Sie die Integration der Computertechnologie in den Sprachunterricht ermöglichen.

3.1 Beispielprogramme für VR Implementierung

Programme wie „You, Calligrapher“, „Noun Town“ und „HanZi“ sind Produkte, die im Rahmen dessen kreiert wurden. „You, Calligrapher“, „Noun Town“ sind zwar Beispiele für Lernprogramme, fallen durch ihren hohen Gamification-Anteil aber fast in die Kategorie Lernspiel, welches einen gleichen Ansatz wie CALL verfolgt, aber viel mehr einem Videospiel oder einem Lernspiel gleicht statt einem Lernprogramm mit spielerischem Ansatz.

3.1.1 NounTown

Noun Town ist ein Kickstarter VR Sprachlerngame, das sich zur Aufgabe gemacht hat über 1000 Wörter in den verschiedenen Sprachen beizubringen. Die Gestaltung entspricht einer kleinen Stadt auf einer Insel mit knuffigen klischeebehafteten Figuren im Visuellen Style von Animal Crossing – siehe Screenshots. Das Ziel des Spiels im Bereich Gamification ist die Insel zu retten, indem Worte erlernt und die farbenlose Welt wieder bunt und vielfältig wird – was durch korrektes Bezeichnen und der Aussprache der Vokabeln abgefragt wird. Das Spiel stellt dann den Gesamt und Levelfortschritt entsprechend dar indem die Welt zunehmend bunter wird. Es gibt in der Stadt zahlreiche Alltagsszenarien, die gelöst werden müssen und in einfach gehaltenen Architekturen wie einem Krankenhaus, einer Farm oder einem Supermarkt stattfinden.

Das Lernen wird dadurch begünstigt, dass die jeweiligen Aufgabenelemente (also Gegenstände bzw. Nomen) jeweils interaktiv hochgehoben werden und durch eine kleine Animation das Belohnungssystem triggern und sich dadurch der Lern- und Behaltens-Effekt verstärkt. Während des Hochhebens von Gegenständen wird der Begriff ausgesprochen und muss vom Lernenden dann korrekt nachgesagt werden. Auch dieser richtige Abschluss wird durch optisches Feedback und Sound belohnt.

Dieses Spiel gibt es mittlerweile in verschiedenen Fassungen darunter Deutsch, Chinesisch, Spanisch, Koreanisch, Italienisch, Japanisch und Französisch

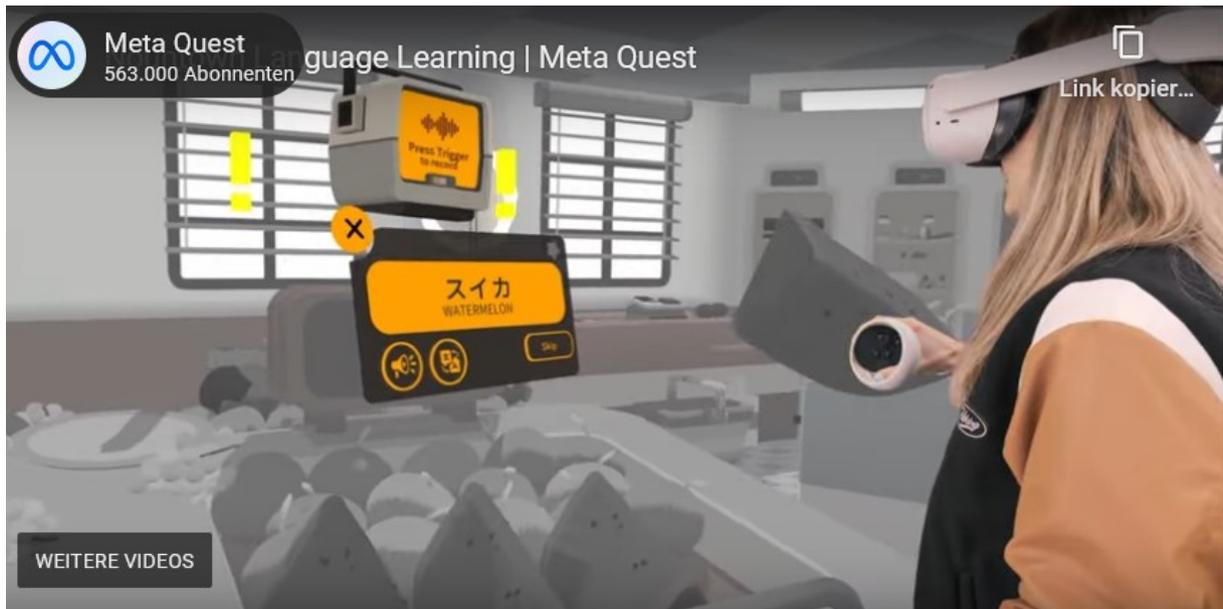


Abbildung 1: Youtube Video von Metaquest

Quelle: <https://youtu.be/CdYxKuMpufc> Stand 02.12.2023 19:55 Uhr



Abbildung 2: Schrullige klischeebehaftete spanische Dame

Quelle: <https://youtu.be/CdYxKuMpufc> Stand 02.12.2023 19:55 Uhr

3.1.2 Hanzi

Hanzi ist ein 3-dimensionales Game in der VR Welt welches sich zum Ziel gesetzt hat die Schreibweise der chinesischen Zeichen zu erlernen. Der User nutzt den VR Controller, der wie ein Pinsel modelliert ist und muss die Zeichen in der korrekten Strichreihenfolge und Strichführung vervollständigen, wobei diese sich je nach ausgewähltem Skill Level in Geschwindigkeit auf Ihn zubewegen und verschiedene Optionen in der Gesamtzeichendarstellung angeboten werden – was immens den Schwierigkeitsgrad bezeichnet, da ein Abruf des Gesamtzeichens ohne Visualisierung aus dem Erlernten erfolgen muss. Das Spiel ist durchgehend animiert und setzt auf den Faktor der Wiederholung und Herausforderung Dies steht der klassischen Lernweise von purer Wiederholung in Papierform entgegen



Abbildung 3: Vergleich der Lernweise mit Writing Paper und in der Software
Quelle: <https://dl.acm.org/doi/fullHtml/10.1145/3550472.3558415>

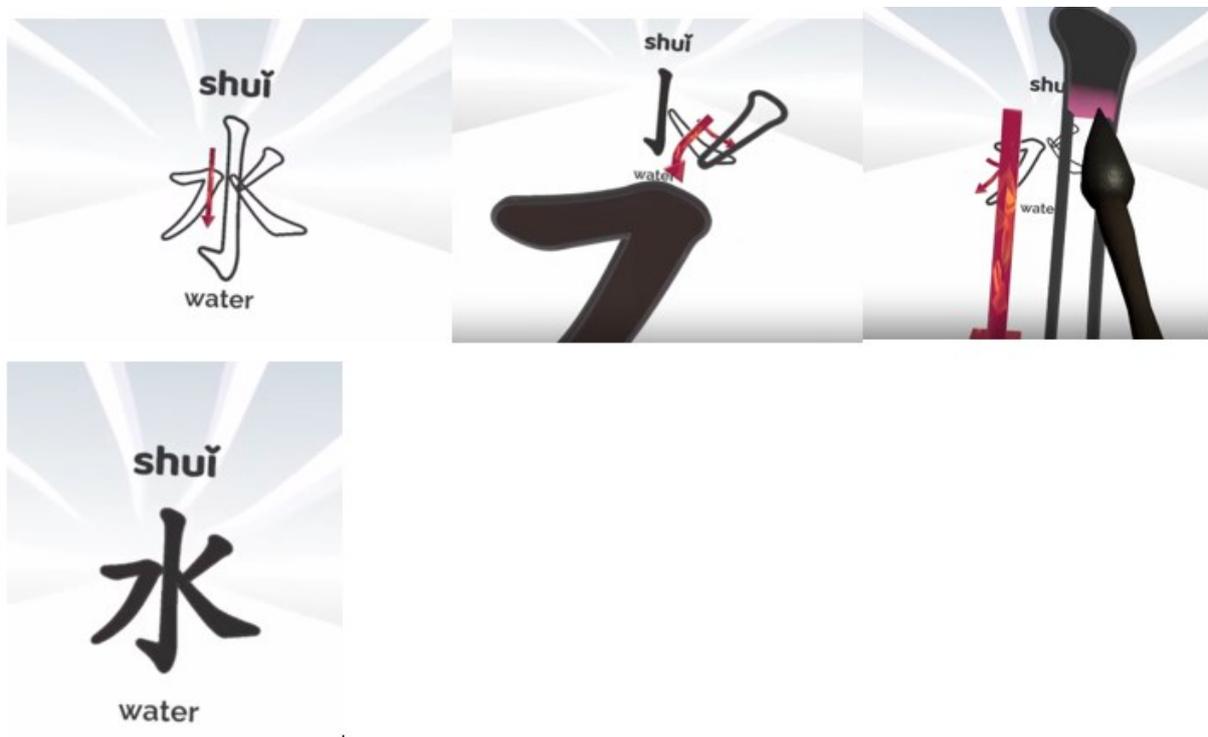


Abbildung 4: Zeichenaufbau und Ablauffolge in Hanzi

Anzumerken ist, dass die Entwickler selbst davon sprechen das ursprüngliche und als (CHDR Chinese Character handwriting und remedial instruction) Tool bekannte System als Grundlage genommen zu haben und dieses um den Aspekt der Gamification erweitert zu haben. Es gibt verschiedene Schwierigkeitsstufen und Ansichten. Neben der Ansicht von jedes Zeichen von vorn auf den User zu fliegt gibt es auch eine Ansicht in der die Zeichen von links nach rechts in das Bild einfliegen und damit eine Version, die einfacher im Erlernen ist, da die Verdeckung des Ursprungszeichen über die Mitte minimiert werden konnte.



CHDRI – Chinese character handwriting diagnosis and remedial instruction

Hsiao et al. (2013) ^[1]

[1] Hsien-Sheng Hsiao, Cheng-Sian Chang, Chiao-Jia Chen, Chia-Hou Wu, and Chien-Yu Lin. 2013. The influence of Chinese character handwriting diagnosis and remedial instruction system on learners of Chinese as a foreign language. *Computer Assisted Language Learning* 28, 4 (jul 2013), 306–324. <https://doi.org/10.1080/09588221.2013.818562>

Abbildung 5 Bild aus der Entwicklerpräsentation mit Quellenangabe der Entwickler

4 Die chinesische Schrift und Sprache

4.1 Sprachgeschichte Mandarin

Mandarin ist nicht nur die chinesische Sprache mit den meisten Sprechern weltweit. Sie ist ebenfalls die am meisten gesprochene sinitische Sprache unter allen chinesischen Sprachen und Dialekten. Das Wort „Mandarin“ hat seinen Ursprung interessanterweise nicht im Chinesischen, sondern entstammt dem malaiisch-portugiesischen Wort „mantari“, was Verwalter oder Minister heißt. Mandarin war die Sprache der Zivilbeamten im Kaiserreich China, besonders in der Ming-Dynastie (1368-1644) und Qing-Dynastie (1644-1911) und wurde als Guānhuà (chinesisch: 官话) bezeichnet, was übersetzt Beamtensprache bedeutet. Portugiesische Seefahrer, die im 16. Jahrhundert China erreichten, übertrugen somit den Begriff auf die chinesische Verwaltung.

Mandarin wird heutzutage in seiner Bedeutung gleichgesetzt wie Hochchinesisch. Sprachhistorisch betrachtet allerdings ist Mandarin eine Gruppierung verschiedener Dialektgruppen, unter anderem dem Pekinger Dialekt, auf dem das standardisierte Hochchinesisch basiert. Zur Ming-Dynastie und Qing-

Dynastie wurde die Hauptstadt nach Peking verlagert, weshalb nördliche chinesische Sprachen Anerkennung gewannen. Daher basiert die Beamtensprache maßgeblich auf nördliche Regionalsprachen. Gegen Ende der Qing-Dynastie und den verlorenen Opiumkrieg (1840, 1856), sowie des Japankrieges (1894) wurden die Forderungen des Volkes nach einer gemeinsamen Sprache lauter. Dies mündete darin, dass erstmals 1913 auf der Nationalen Konferenz zur Vereinheitlichung der Aussprache in der Beiyang-Regierung der Aussprachestandard festgelegt wurde. 1924 wurde entschieden, dass die Pekinger Aussprache die Standardaussprache werden sollte. Man sprach seitdem von Guóyǔ (国语), was sinnhaftig „Nationalsprache“ bedeutet. Eine Änderung der Benennung erfolgte 1955 auf der „Nationalen Konferenz über Schriftreform“, somit unmittelbar nach Gründung der Volksrepublik China. Statt Guóyǔ sollte nun die Bezeichnung „Pǔtōnghuà“ (普通话, wörtlich für „normale Verkehrssprache“ oder „allgemeine Sprache“) zu verwendet werden, um nationale Minderheiten zu integrieren. So kam es final zu dem Begriff Pǔtōnghuà, der sich als Standardchinesisch, Hochchinesisch oder als Mandarin durchsetzte und die gesprochene Amtssprache Volksrepublik Chinas, aber auch in Hongkong, Macau, Taiwan und Singapur ist. Taiwan verliert aber bei der Bezeichnung „Guóyǔ“.

Hochchinesisch vereinfachte die heterogene Sprachsituation zwischen verschiedenen Provinzen, deren Dialekt und Akzente das gegenseitige Verständnis ursprünglich massiv erschwerte, die geschah parallel zu der Vereinfachung der chinesischen Schrift in die vereinfachte Regelschrift, welche im 4.3.8 erwähnt wird. Trotz der Einigung auf eine gemeinsame chinesische Sprache und Durchsetzung von Hochchinesisch ist weiterhin das Sprechen der orts- und provinziellen Dialekte präsent, insbesondere bei älteren Generationen (mit Ausnahme von Kantonesisch). So sind regionale Akzente vom Hochchinesisch in mehreren Provinzen und Ländern außerhalb der Volksrepublik China bemerkbar. Stärkere und konsistente Abweichungen vom gesprochenen Standardchinesisch werden vor allem bei Menschen wahrgenommen, die Hochchinesisch als Zweitsprache lernen.

Mandarin wird im Ausland mit Hochchinesisch gleichgesetzt, was nicht vollständig korrekt ist, da Mandarin auch alle Dialekte im nordchinesischen Raum einschließt. Weitere Bezeichnungen für die chinesische Sprache sind Zhōngwén (中文), ein eher allgemeiner Begriff für die gesprochene als auch geschriebene Sprache oder aber, Hànyǔ (汉语), was umgangssprachlich für Hochchinesisch, aber zudem alle gesprochenen Dialekte der Han-Chinesen inkludiert oder Huáyǔ (华语), das eher von Auslandschinesen in der Diaspora außerhalb Chinas benutzt wird.

4.2 Geschichte der chinesischen Schrift

Neben der Einführung vom gesprochenen Hochchinesisch dient die chinesische Schrift teils auch als eine dialektübergreifende Kommunikation zwischen den verschiedenen Provinzen und Ländern.

Die chinesische Schriftsprache ist als Zhōngwénzì (中文字) oder als Hànzì (汉字) bekannt. Allerdings schreibt die chinesische Schrift seit über 3000 Jahren Geschichte und ist trotz ihrer Tradition im permanenten Wandel und durchlief unterschiedliche Entwicklungsstufen, bis es zum heutigen vereinfachten Chinesisch kam.

Die ältesten Funde von chinesischen Schriftzeichen fanden sich auf **Orakelknochen aus der Shang-Dynastie (17. bis 11. Jahrh. v.Chr.) bzw. den** eingeritzten **Bildzeichen**. Gewöhnlich waren dies Tierknochen und Schildkrötenpanzer, auf denen chinesische Schrift für Worte und Gedanken in Form von **Weissagungen** verwendet wurden. Ab dem 11. Jahrh. V.Chr. wurden während der Shang und Zhou Dynastie (770–221 v. Chr.) Bronzeinschriften auf Bronzegefäßen verfasst. Die niedergeschriebene Sprache wird Altchinesisch oder archaisches Chinesisch genannt und fand Anwendung in literarischen Werken wie Gedichtsammlungen. Altchinesisch beschreibt im Allgemeinen auch die chinesische Sprache der Zhou-Dynastie vor 256 v.Chr., sie durchlief mehrere Entwicklungen, darunter die große Siegelschrift Dàzhuàn (大篆) in der frühen Zhou-Dynastie und die kleine Siegelschrift Xiǎozhuàn (小篆) zur späteren Zhou-Dynastie bis zur Qin-Dynastie. Diesen Schriften wurden auf Steinen oder Bronzekesseln gefunden. Die kleine Siegelschrift war eine Entwicklung der Großen und konnte sich zu einer verkehrsfähigen Schrift entwickeln, die Sprachsyntax und -semantik nach Anforderungen einer Sprache abbildet.

Die letzte altchinesische Phase bildet das klassische Chinesisch Wényán (文言), welches „Literatursprache“ bedeutet. Es bezeichnet sowohl die gesprochene als auch die geschriebene Sprache und bildet den Vorläufer der modernen sinitischen Sprachen.

Unter Chinas erstem Kaiser zur Qin-Dynastie (221-207 v.Chr.) fand die **erste große Vereinheitlichung der chinesischen Schrift** statt. Die Kanzleischrift Lishū (隶书) entstand durch die Abwandlung der Siegelschrift. Diese Zeichen sind geradlinig, eher breit als hoch und haben geschwungenen Einzelstriche. Diese schnell ausführbare Schrift sparte Zeit, was eine essenzielle Rolle in der für offizielle Dokumentverwaltung in Kanzleien war und durch die Förderung und Standardisierung der Kanzleischrift konnte eine einheitliche Schreibweise in der Qin-Dynastie aufgebaut werden. Sie ähnelt damit dem heutigen Stenographieren was in der heutigen Jurisdiktion und Gerichtbarkeit weiterhin Verwendung findet – wenn auch zumeist durch Spezialtastaturen unterstützt und dadurch weitaus weniger abhängig vom Stil des Stenographen.

Obwohl das gesprochene klassische Chinesisch nach der Qin-Dynastie nicht mehr verwendet wurde, so wird es selbst in heutiger Zeit immer noch als Literatursprache konserviert und beigebracht. Die Kompetenz bezieht sich meistens nur aufs Leseverständnis und der kulturellen Erhaltung statt des Verfassens von Texten. Spuren davon überdauern aber in Form von Zitaten oder den Anekdotenwörter Chéngyǔ (成语) merklich. Diese Anekdotenwörter bestehen aus vier Zeichen bzw. Morphemen und finden immer noch regelmäßige Anwendungen in Schrift und Sprache.

Entgegen aller Standardisierung der Kanzleischrift war die Struktur und Stilistik nicht ideal für den alltäglichen Gebrauch. Man sehnte sich weiter nach einer noch mehr effizienten Schrift, weshalb die Kanzleischrift zum Ende der Han-Dynastie (206 v. Chr. bis 220 n. Chr.) durch die Regelschrift Kǎishū (楷书) abgelöst wurde, die als eine Art Kursivschrift begann. Die Kursivschrift Xíngshū (行书) und Grasschrift (Cǎoshū 草书) entwickelten sich als informellere Schreibstile, die mehr Freiheit in der Strichführung gaben. Diese Schriftstile wurde insbesondere in der Kalligrafie geschätzt und war eine weitere Ausführung der künstlerischen Dimension der chinesischen Schrift. Neben der Siegelschrift und Kanzleischrift bilden sie mit der Kursivschrift, Grasschrift und Regelschrift die fünf Schriftarten der chinesischen Kalligrafie.

Die Regelschrift wurde zwischen der Han- und Wei-Dynastie (220–265) initiiert, fand aber keine Anwendung bis zum 7. Jahrh. Den Höhepunkt der Anwendung fand die Regelschrift zur Tang-Dynastie (618–907), in der die Schrift sich zur heutigen Standard-Schreibschrift Zhèngkǎi (正楷) entwickelte.

Die Etablierung dieser offiziellen Schreibweise diente als Grundlage zur Entwicklung zum heutigen traditionellen und vereinfachten Chinesisch. Zur Song-Dynastie (960–1279 n. Chr.) wurden viele klassische Texte mit der Regelschrift veröffentlicht, was zur Festigung dieser trug. Die Regelschrift wurde zur Norm für die schriftliche Kommunikation und bildete die Grundlage für die Entwicklung der Drucktechnologie in der Zeit.

Bis zum Ende der Kaiserzeit wurde die chinesische Schrift vorwiegend zum Schreiben der klassischen Schriftsprache verwendet, die nur für die gebildete Elite verständlich war. Seit dem 20. Jahrhundert wird die chinesische Schrift vorwiegend zum Schreiben der Standardschriftsprache Báihuàwén (白话文) eingesetzt, die grammatisch den modernen nördlichen Dialekten ähnelt und von südlichen Chinesen leichter als Wényán erlernt werden kann. Schon in den späten 1920er Jahren aber wurden sehr viele chinesischen Zeitungen, Bücher und offizielle Dokumente in Báihuà verfasst. Báihuàwén dient als überregionales Medium zur Verständigung trotz der heterogenen Sprachsituation. Aus diesem Grund sind im chinesischen Fernsehen chinesische Untertitel zu sehen, sodass unabhängig von Aussprache und Dialekten der Menschen der Inhalt verstanden werden kann. Dies soll neben einer gemeinsamen

Verständigung die Anwendung und das durch verknüpfte Lese- und Hörverstehen und damit das Lernen vom gesprochenen Hochchinesisch unterstützen.

Mit der Zeit entstanden zwei Hauptvarianten der Regelschrift: das traditionelle Chinesisch und das vereinfachte Chinesisch. Das traditionelle Chinesisch basiert stark auf die Regelschrift und verwendet die sogenannten Langzeichen Fántǐzì (繁体字). Es wird in Taiwan, Hongkong und Macau verwendet.

Das vereinfachte Chinesisch hingegen nutzt die Kurzzeichen Jiǎntǐzì (简体字). Die Kurzzeichen im das vereinfachten Chinesisch entstanden durch die Schriftreform 1958, in der ca. 2300 vereinfacht wurden. Seitdem wird es dort und in Singapur verwendet. Ziel der Vereinfachung war die Effizienz bei Schriftstücken als auch der elektronischen Verarbeitung und synergetisch den Alphabetisierungsprozess der Bevölkerung zu beschleunigen und dadurch die chinesische Schrift international zugänglicher zu gestalten.

4.3 Vereinfachung der chinesischen Zeichen

Wie genau gestaltete sich nun diese Vereinfachung? Die Vereinfachung der Langzeichen in Kurzzeichen bestand aus verschiedenen Schritten. Zunächst wurde grundlegend entschieden, dass die Varianz von Wort und Bedeutung reduziert wird. Dies half in der Ausprägung die Zahl der nötigen Striche zu verringern. Die große Zahl vereinfachter Schriftzeichen (mehrere Tausend) kommt vor allem dadurch zustande, dass auch Komponenten vereinfacht wurden, die in vielen Schriftzeichen vorkommen, die sogenannten Radikale. Offiziell wurden die Schriftzeichen nach zehn Methodiken vereinfacht Um es greifbar zu machen folgen zwei Beispiele.

Bsp. 1: Vereinfachung des Zeichens 飛 → 飞 (dt. „fliegen“)

Hier wurden Teile eines ursprünglichen Zeichens erhalten bzw. auf diese Teile reduziert

Bsp. 2: Vereinfachung des Zeichens 燈 → 灯 (dt. „Lampe“)

Hier erfolgte ein Ersatz von Radikalen, z.B. beim chinesischen Zeichen für das Wort Lampe wurde das Phonetikum ersetzt (hier rechtes Radikal).

Die Vereinfachung brachte grundsätzlich Schreibeffizienz und -schnelligkeit der chinesischen Zeichen, hatte aber zur Folge, dass bedingt dadurch noch mehrheitlich aus Langzeichen bestehende Texte mit einem ausschließlichen Verständnis über Kurzzeichen massive Herausforderungen im Leseverständnis mit sich brachten. Teils konnten Wörter durch den Wegfall von Varianz und Analogien nur durch Kontext erschlossen werden. In der Praxis erscheint die jeweils andere Schriftzeichenform erstmal ungewöhnlich, nach und nach ist eine Gewöhnung mit Zeit und Erfahrung möglich, was sich aber eher auf das passive Leseverständnis bezieht. Diese Hürde wird heutzutage durch die EDV-technischen Möglichkeiten gebrochen, wodurch eine Umwandlung in die andere Schriftform besonders für digitales Schreiben schnell möglich gemacht werden kann.

Die Standardisierung und Vereinfachung der chinesischen Schrift gaben einen enormen Einfluss auf die chinesische Kultur und Gesellschaft, aber auch in anderen asiatischen Ländern wie Korea und Japan gelangte im **7. Jahrhundert als Kulturexport dorthin, wo sie sich ebenfalls** etablierte, weiterentwickelte und als Teil des nationalen Schriftsystems genutzt wird. Heutzutage bleibt die vereinfachte Regelschrift eine fundamentale Komponente in der Ausbildung und Anwendung der chinesischen Schrift. Traditionelle Schrift wird allerdings informell in Handschrift oder für Grafiken verwendet.

Bedeutung	Sonne	Regen	Vogel	Berg
Orakelknochen				
Siegelschrift				
Kanzleischrift				
Kursivschrift				
Grasschrift				
Regelschrift				

Abbildung 6: Chinesische Schriftzeichen und deren Entwicklung

4.4 Die Herausforderung von Mandarin-Chinesisch

Um zu erklären, warum Chinesisch als L2 besonders herausfordernd sein kann, ist es sinnvoll, einen genaueren Blick auf die Charakteristik, den Aufbau und Anwendung der chinesischen Schrift zu werfen. Die Untersuchung konzentriert sich im Rahmen der Arbeit speziell auf den Aufbau der chinesischen Schriftzeichen, das Schrift, und Tonalität konzentriert. Es wird darauf hingewiesen, dass die Analyse auf die spezifischen Merkmale der chinesischen Schrift innerhalb der Volksrepublik China fokussiert ist.

4.4.1 Tonalität und Aussprache

Chinesisch ist eine tonal gesprochene Sprache. Das bedeutet alle einzelnen Zeichensilben werden mit bestimmter Intonation angesprochen, an die sich die Bedeutung bindet. Obwohl die chinesische Schrift bis zu 100.000 Zeichen verfügt, gibt es nur 450 einzigartige Silben. Eine Änderung im Ton, Tonhöhenverlauf oder die falsche Betonung verändert die Bedeutung des Zeichens sorgt folgerichtig dazu, dass man nicht mehr verstanden wird. Chinesische Silben bestehen meistens aus einem konsonantischen Anlaut und einem nasalten Auslaut. Dazu gibt es vier Töne und einen zusätzlichen unbetonten Ton.

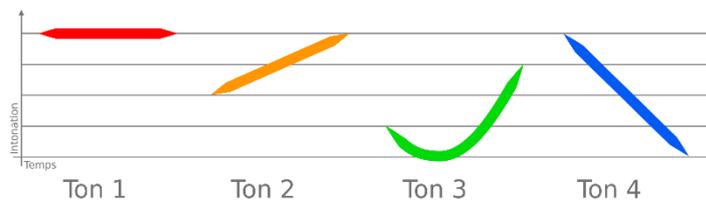


Abbildung 7: Die vier betonten Töne in Hochchinesisch

Als Hilfe für die Aussprache der Zeichen wurde 1956 das Hanyu Pinyin (汉语拼音) eingeführt, welches die chinesische Romanisierung von Hochchinesisch in Volksrepublik China ist. Das Pinyin ist eine phonetische Umschrift und nutzt das lateinische Alphabet, um die chinesischen Zeichensilben alternativ darzustellen. Im Folgenden wird die Silbe „ma“ als Beispiel genommen, um die fünf Töne zu veranschaulichen:

Ton	traditioneller Name & Bedeutung	Charakteristik des Tons	Silbe „ma“ unter Ton	mögliche Zeichen aus Silbe & Bedeutung
Ton 1	yīn píng (阴平) „Yin Pegel“	hoher Ton und konstante Tonhöhe	mā	妈: Mutter 抹: wischen
Ton 2	yáng píng (阳平) „Yang Pegel“	steigender Ton von unterer bis zur mittleren Tonhöhe (wie eine Frage im Deutschen)	má	麻: rau, Hanf
Ton 3	shǎng shēng (上声) „ansteigender Ton“	zunächst sinkender Ton von einer mittleren Tonhöhe, dann aufsteigend in mittlere Tonhöhe (Hatschek)	mǎ	马: Pferd 码: Nummer 蚂: (in Komposita蚂蚁 mǎyǐ): Ameise
Ton 4	qù shēng (去声) „verlassender Ton“	scharf von oben nach unten fallender Ton, vergleichbar mit deutscher Betonung eines Befehls „Sitz!“	mà	骂: schimpfen
Neutraler Ton	qīng shēng (轻声) „leichter Ton“	kurz und leichter, unbetonter Ton	ma	吗: Partikel am Ende zur Bildung von Fragesätzen

Abbildung 8: Die fünf Töne in Hochchinesisch

Abbildung 13 zeigt die Pinyin Tabelle und alle möglichen Kombinationen aus An- und Auslauten. Bei der Aussprachelehre werden zwar oft An- und Auslaute als einzelne Komponenten beigebracht, doch versteht sich die Kombination beide als eine ausgesprochene Silbe.

	a	ao	ai	an	ang	i	ia	ie	iu	iao	ian	iang	iong	in	ing	e	ei	en	eng	o	ou	ong	u	ua	uai	uo	ui	uan	uang	un	ü	üe	ün	üan	ê				
g	ga	gao	gai	gan	gang											ge	gei	gen	geng	gou	gong	gu	gua	guai	guo	gui	guan	guang	gun										
k	ka	kao	kai	kan	kang											ke	kei	ken	keng	kou	kong	ku	kua	kuai	kuo	kui	kuan	kuang	kun										
h	ha	hao	hai	han	hang											he	hei	hen	heng	hou	hong	hu	hua	huai	huo	hui	huan	huang	hun										
f	fa			fan	fang											fei	fen	feng	fo			fu																	
w	wa			wai	wan	wang										wei	wen	weng	wo			wu																	
z	za	zao	zai	zan	zang	zi										ze	zen	zeng		zou	zong	zu			zuo	zui	zuan		zun										
c	ca	cao	cai	can	cang	ci										ce	cen	ceng		cou	cong	cu			cuo	cui	cuan		cun										
s	sa	sao	sai	san	sang	si										se	sen	seng		sou	song	su			suo	sui	suan		sun										
zh	zha	zhao	zhai	zhan	zhang	zhi										zhe	zhei	zhen	zheng	zhou	zhong	zhu	zhua	zhuai	zhuo	zhui	zhuan	zhuang	zhun										
ch	cha	chao	chai	chan	chang	chi										che	chen	cheng		chou	chong	chu	chua	chuai	chuo	chui	chuan	chuang	chun										
sh	sha	shao	shai	shan	shang	shi										she	shei	shen	sheng	shou		shu	shua	shuai	shuo	shui	shuan	shuang	shun										
r	rao			ran	rang	ri										re	ren	reng		rou	rong	ru			ruo	rui	ruan		run										
j						ji	jia	jie	jiu	jiao	jian	jiang	jiong	jin	jing																					ju	jue	jun	juan
q						qi	qia	qie	qiu	qiao	qian	qiang	qiong	qin	qing																					qu	que	qun	quan
x						xi	xia	xie	xiu	xiao	xian	xiang	xiong	xin	xing																					xu	xue	xun	xuan
b	ba	bao	bai	ban	bang	bi	bie		biao	bian				bin	bing	bei	ben	beng	bo				bu																
p	pa	pao	pai	pan	pang	pi	pie		piao	pian				pin	ping	pei	pen	peng	po				pu																
m	ma	mao	mai	man	mang	mi	mie		miao	mian				min	ming	mei	men	meng	mo				mu																
d	da	dao	dai	dan	dang	di	die		diu	diao	dian					ding	de	deng				du			duo	dui	duan		dun										
t	ta	tao	tai	tan	tang	ti	tie		tiao	tian					ting	te		teng				tu			tuo	tui	tuan		tun										
n	na	nao	nai	nan	nang	ni	nie		niu	niao	nian			nin	ning	nei	ne	neng			nong	nu			nuo		nuan				nü	nüe							
l	la	lao	lai	lan	lang	li	lie		liu	liao	lian	liang		lin	ling	lei		leng	lo		lou	long	lu			luo		luan		lun	lü	lüe							
y	ya	yao	yan	yang	yi									yin	ying						you	yong									yu	yue	yun	yuan	ye				

Abbildung 9: Pinyin Tabelle von Mandarin-Chinesisch

Aus der niedrigen Anzahl an einzigartigen Silben und der verglichen hohen Anzahl an Zeichen ist abzuleiten, dass selbst eine betonte Silbe ein Pinyin mit Intonation fast keine einzigartige Bedeutung hat und immer noch Interpretierungsspielraum liefert, welches Zeichen gemeint ist. Dies ist auch in der letzten Spalte der Abbildung 13 erkennbar, dass z.B. die Silbe mā mehrere Bedeutungen annehmen kann. Andersrum hat ein Zeichen immer eine oder mehrere fest zugewiesenen Silben und Intonation. Daher ist es essenziell beim Lernen von chinesischen Schriftzeichen sich die Bedeutung, die Aussprache und die Schreibweise der Zeichen zeitgleich zu lernen.

Dieses sprachliche Phänomen, dass ein gesprochenes Wort verschiedene Bedeutungen zugewiesen werden kann, nennt man Homophonie. Durch die geringe Silbenanzahl hat Chinesisch einen hohen Anteil homophoner Morpheme. Hochchinesisch besitzt ebenso sogenannten Homographen, diese werden Duōyīnzì (多音字, dt: „Doppellautzeichen“) genannt und beschreiben dieselben Zeichen mit mindestens zwei verschiedenen Lauten und semantischen Bedeutungen. Ein Beispiel wäre das Zeichen 还, welches entweder als „hái“ ausgesprochen werden kann und „noch, doch“ heißt oder mit „huán“ das Verb „zurückgeben“ bedeutet. Durch Homographie erschwert sich der Lesefluss eines Satzes, da Verwechslungsgefahr in der Aussprache kommen entstehen könnten. So ist wichtig in einem Satz den Kontext der Morpheme zu erkennen. Da viele Morpheme als Zeichenpaar auftauchen, besteht die Möglichkeit daran sich durch die kontextuelle Bedeutung die Aussprache zu erschließen. Im Falle von

还 (hái) kann das Wort 还有 (hái yǒu) gebildet werden, was „außerdem, es gibt noch“ bedeutet und damit eindeutig identifizierbar ist, da das Wort 还有 mit der Alternativaussprache „huán yǒu“ nicht existiert. Trotzdem gibt es Ausnahmen und Fälle, in denen sich dies nicht gilt und somit weiter auf den Kontext des gesamten Satzes geachtet werden muss.

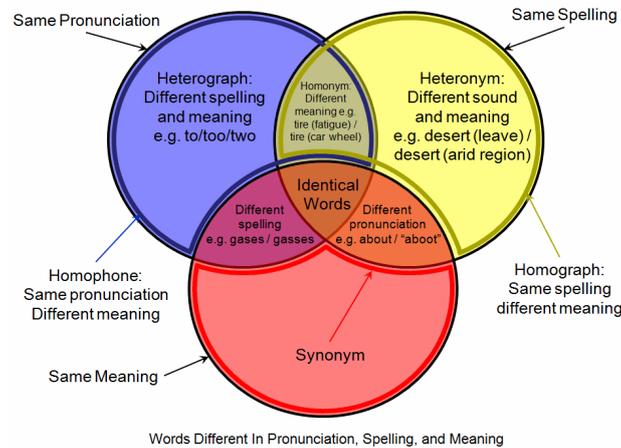


Abbildung 10: Venn-Diagramm von Homophone, Homographen und Synonymen

https://upload.wikimedia.org/wikipedia/commons/e/e7/Homograph_homophone_venn_diagram.png

4.4.2 Zusammenhang zwischen Wort, Silbe und Morphem

Die grundlegenden Unterschiede zwischen dem lateinischen Alphabet und den chinesischen Zeichen kann neben dem logographischen Aussehen und dem Zusammenhang zwischen Wort, Morphem (kleinste bedeutungstragende Einheit) und Silbe erklärt werden.

In der chinesischen Sprache repräsentiert eine Silbe oft sowohl ein Wort als auch ein Morphem, da einzelne chinesische Zeichen oft einsilbig sind, aber eine oder mehrere feste Bedeutungen haben können. Letzteres bezeichnet man als Homographen.

Wie aus dem vorherigen Abschnitt bekannt wird das Zeichen „妈“ wird mit der Silbe „mā“ ausgesprochen und bedeutet Mutter. Ein besonderes Merkmal ist, dass alle Morpheme in der chinesischen Sprache für sich alleinstehen können. Morpheme sind die kleinste bedeutungstragende Einheit in Wörtern einer Sprache. Die Charakteristik, dass chinesischen Zeichen sowohl Silbe, Wort und Morphem zugleich sein können unterscheidet sich vom Deutschen, wo Morpheme nicht immer eigenständige Wörter sind. Z.B. Das Wort „Häuser“ beinhaltet die Silben „Häus-“ und „-er“, welche sich aus der Morphemkombination aus den Morphemen „Haus“ und „-er“, der Pluralendung bildet.

Im Chinesischen sind Morphemkombination Zusammensetzungen von mehreren einzelnen Zeichen, die dadurch neue semantische Bedeutungen bekommen können. Beispiel ist das Wort 中国 (zhōng guó),

bestehend aus den Morphemen 中 (zhōng) für „Mitte“ und „国“ (guó) für Land oder Nation, was zusammengesetzt dann „Mittelland“ bzw. „Mittleres Reich“ bedeutet und aus der Sicht der Chinesen, sowohl kulturell als auch geopolitisch eine maßgebliche Bezeichnung für China ist. Auch können Morphemkombination aus Synonymen bestehen, um die Bedeutung zu verstärken. So bedeutet „饥饿“ (jī' è) beides jeweils Hunger.

4.4.3 Anzahl der Zeichen

Ein oftmals angeführter Umstand für die Herausforderung beim Erlernen von Chinesisch ist die umfangreiche Anzahl individuellen Schriftzeichen, die scheinbar endlos wirken. Verschiedene Zeichenlexika listen eine unterschiedliche Anzahl von chinesischen Zeichen auf. So listet das Variantenwörterbuch der chinesischen Schriftzeichen des Bildungsministeriums Taiwans in der sechsten Auflage von 2017 (Dictionary of Chinese Character Variants *yìtǐ zìdiǎn* 異體字字典) über 100.000 Schriftzeichen, was im ersten Moment für L2 Lernende mehr als erschlagend wirkt. Real gesehen finden viele Zeichen selten bis keinen Gebrauch mehr in der heutigen Gesellschaft. Ein sehr stark etabliertes Wörterbuch in der VR China ist das *Xīnhuá Zìdiǎn*, in seiner 10. Auflage 2004 listet es insgesamt 11.200 Schriftzeichen.

Tatsächlich werden in der Praxis bei Schrift und der gesprochenen Sprache deutlich weniger Zeichen eingesetzt. In der Volksrepublik China werden 2500 Zeichen für den erforderlichen Alltagsgebrauch benötigt, um z.B. eine chinesische Tageszeitung zu lesen, Chinesen besitzen nach Beendigung der Oberstufe ca. 4500 und nach dem Hochschulabschluss ca. 6000. Als L2 Lernende wird für die effektive Kommunikation im Alltag das Beherrschen von 2000 Zeichen empfohlen. Entscheidend sind aber nicht ausschließlich die Anzahl der bekannten Zeichen, sondern ob diese zum aktiven und passiven Wortschatz geeignet sind. Während der aktive Wortschatz alltägliche Gebrauch bei Sprache und Schrift sieht, handelt es beim passiven Wortschatz um die Erkennung durch Hören oder Lesen.

4.4.4 Kategorien der chinesischen Zeichen

Im Unterschied zur lateinischen oder arabischen Schrift besteht die heutige chinesische Sprache überwiegend aus Logogrammen, welche für Wortbedeutungen, ein Wort oder Morphem stehen. Chinesische Logogramme lassen sich insgesamt in drei Kategorien unterteilen: Piktogramme, Ideogramme und Phonogramme. Diese Unterklassen unterscheiden sich vorallem in ihrem Aufbau und in der semantischen Assoziation und Bedeutung.

Im Allgemeinen ist einem Zeichen eine Silbe zugeordnet, mit Ausnahme der Homophone.

4.4.5 Piktogramme - Xiàngxíng (象形) „Bildzeichen“

Mit Piktogrammen sind Bildzeichen gemeint sind, die einen ganzen Gedanken repräsentieren, zu denen keine Wiedergabe der sprachlichen Form existiert. In der chinesischen Schriftentwicklung sind Piktogramme durch Bilder von Objekten entstanden und versuchten die Erscheinungsform mit Symbolen wiederzugeben. Es bildet die erste chinesische Schriftzeichenart seit der Anwendung durch die Orakel-Schrift und durchging mehrere Veränderungen in Form von stilistischen Vereinfachungen. Beim Übergang in die fortgeschrittenen Schriftarten verloren Piktogramme den bildlichen Charakter. Chinesische Piktogramme machen einen geringen Anteil der chinesischen Schriftzeichen aus.

Entsprechend anschaulich wird es, wenn man sich die Zeichen und Ihre Entwicklung anschaut und hier insbesondere die Darstellung des Begriffs Pferd



Abbildung 11: Darstellung des Begriffs Pferd

https://miro.medium.com/v2/resize:fit:640/format:webp/1*gpiEXknM4fCQ2oQEcigo1g.jpeg

4.4.6 Einfache Ideogramme - Zhǐshì (指事) „auf Tatbestände deuten“

Als Ideogramme bezeichnet man die symbolische Abbildung abstrakter Gedanken, deren Bedeutung eher indirekt durch Indikatoren des Zeichens vermittelt wird. Ideogramme können Sachen zeigen, die visuell nicht wahrnehmbar sind. So sind Abstrakta wie Zahlenschriftzeichen oder „oben“ und „unten“ darstellbar.

一	二	三	上	下	本	末
yī	èr	sān	shàng	xià	běn	mò
eins	zwei	drei	oben	unten	Wurzel	Wipfel

Abbildung 12: Abstrakta

4.4.7 Zusammengesetzte Ideogramme - Huìyì (会意) „Vereinigte Bedeutungen“

Wie einfache Ideogramme bilden zusammengesetzte Ideogramme abstrakten Gedanken ab, allerdings indem die Bedeutung von zwei oder mehreren Piktogrammen und Ideogrammen symbolisch

zusammengesetzt werden, die jeweils zur neue Gesamtbedeutung beitragen bzw. ein neues Morphem bilden.

Häufig werden diese Zeichen in Form von Radikalen abgekürzt. Als Radikal wird in der chinesischen Schrift die graphische, phonetische oder semantische Zuordnungskomponente bezeichnet. Durch die Kombination von mehreren Sinnträgern, den Signifikas, entstehen bei zusammengesetzten Ideogrammen Schriftzeichenverbindungen, die neue Zeichen mit neuen semantischen Bedeutungen bilden. Das Schriftzeichenlexikon „Hanyu Da Zidian“, mit der jüngsten Veröffentlichung im Jahre 1993 zählt bis heute 200 verschiedene Radikale. Beispielsweise wird durch die Zusammensetzung der Radikale für Mensch „亻“ (Radikal 9) und Baum 木 (Radikal 75) das Zeichen für „ausruhen“ gebildet mit der symbolischen Bedeutung „Mensch, der sich an einem Baum lehnt“.

Kombination	Bedeutung
木 × 2 = 林 lín	zwei Bäume: „Wald, Forst, Gehölz, -wäldchen, -hain“
木 × 3 = 森 sēn	drei Bäume: „Fülle von Bäumen“
亻 + 木 = 休 xiū	Mensch, an einen Baum gelehnt: „ausruhen“

Abbildung 13: Zusammengesetzte Ideogramme

Phonogramm - Xíngshēng 形声 „Form und Laut“

Die überwiegende Mehrheit der chinesischen Zeichen bilden die Phonogramme mit einem Anteil von über 90%. Im genaueren Sinne sind damit nicht Phonogramme im ursprünglichen Sinne gemeint, da diese Zeichen nicht direkt Ausschluss über die Aussprache geben. Die korrekte Bezeichnung von chinesischen Phonogrammen ist „**Signifikum-Phonetikum-Kombination**“, wo das **Signifikum für die bedeutungstragende Komponente und das Phonetikum für die phonetische Komponente des Schriftzeichens steht. In den meisten Phonogrammen findet sich mindestens jeweils eine Komponente im Zeichen wieder. Während das Signifikum ein Radikal ist, nach dem die Zeichen in Lexika geordnet werden, halten sich chinesische Phonogramme oft an die Aussprache des Phonetikums, daher gibt dieses daher oft Hinweise über die Aussprache.**

Sinngeber	Lautgeber	Zeichen
 Wasser (Radikal 85)	 mù	 mù, „sich die Haare waschen“
 Wasser (Radikal 85)	 lín	 lín, „ausgießen“
 Gras (Radikal 140)	 cǎi	 cài, „Gemüse“

Abbildung 14: Signifikum Phonetikum Kombination

In der Abbildung sehen wir in der ersten Zeile das Wasser-Radikal „氵“ und erneut das Baum-Radikal 木 (Radikal 75), welches zusammen „duschen“ oder „waschen“ bedeuten. Verglichen zu den Ideogrammen ist die Bedeutung von Phonogrammen daher nicht mehr direkt ableitbar aus den Komponenten des Zeichens

Im dritten Beispiel wiederum sehen wir, dass das Gras-Radikal mit dem Phonetikum 采(cài, deutsch: „pflücken“) kombiniert wird und so Zeichen für Gemüse 菜 (cǎi) hergestellt wird. Dies von der Zusammensetzung nachvollziehbarer, in diesem Fall erfolgte zusätzlich ein Tonwechsel vom 3. Ton zum 4. Ton.

Der Unterschied zu traditionellen Phonogrammen, beispielsweise des lateinischen Alphabets, führt zu einer komplett anderen Lernweise. Entgegen dem üblichen Lernprozess, dass Kindern in deutschen Schulen das Alphabet mit einer Fibel beigebracht wird, um die alphabetischen Laute aussprechen zu können, steht im Chinesischen das Erlernen der Silben als unterste Stufe. Die Silben benötigen eine Assoziation zu den Schriftzeichen für eine eindeutige Zuordnung der Silben zur Aussprache und Morpheme.

Schriftzeichen zugehörig zur Gruppe der Entlehnungen wurden durch Ähnlichkeit der Aussprache zu anderen Zeichen von ihrer ursprünglichen Bedeutung entlehnt oder entfremdet, während die ursprüngliche Bedeutung des Zeichens ein neues Zeichen bekam. Als Beispiel wurde 其(jī) ursprünglich für "Schaufel" genutzt, aber klang ähnlich zum Pronomen „er/sie/es“, weshalb diese Bedeutung übernommen wurde und nun das Zeichen „箕“ für die Bedeutung der Schaufel genutzt wird. Mit den zusammengesetzten Ideogrammen, den Entlehnungen und den Synonymen wird mit dem ersten Zeichenlexikon der chinesischen Schrift im 2. Jahrh., dem Shuowen Jiezi sogar in sechs Kategorien eingeteilt. Wovon die letzten beiden allerdings einen sehr geringen Teil der Anzahl ausmachen.

Chinesische Synonyme entstanden aus verschiedenen Dialekten und bezeichnen die Nutzung von zwei Zeichen, die sowohl verschieden geschrieben als auch ausgesprochen werden, für dieselbe semantische Bedeutung.

4.4.8 Schriftbild: Schreibrichtung und Strichfolge

Alle chinesischen Zeichen bestehen aus graphischen Charakterstrichen, welche elementare Bestandteile der chinesischen Schrift und Schriftkunst sind.

Striche können in Grundstriche und komplexe Striche aufgeteilt werden. Während die Grundstriche horizontal, vertikal und diagonale Striche enthalten, bestehen komplexe Striche aus Kombinationen verschiedener Grundstriche oder Richtungsänderungen in einem Strichzug. Die Kombination von verschiedenen Grundstrichen und komplexen Strichen gab der Charakterkreierung endlose Möglichkeiten.

Mit der Neuerung innerhalb der Schriftreform 1958 kam auch eine Anpassung der Schreibrichtung in Anlehnung an die westliche Schreibrichtung von links nach rechts statt wie bis dahin üblich oben nach unten - wie zur vormodernen Zeit, wobei diese Schreibweise noch zu kalligrafischen Zwecken angewandt wird.

Abb. 19 listet die ersten sechs Grundstriche und fünf weitere Kombinationsstriche.

Grundstrich	Bezeichnung	Pinyin	Bedeutung	Schreibrichtung	Beispielzeichen	Bedeutung
丶	点	(diǎn)	dot	↘	不, 六	nein, sechs
一	横	(héng)	horizontal	→	不	nein
丨	竖	(shù)	vertical	↓	不, 你	nein, du
丿	撇	(piě)	throw Away	↙	心	Herz
㇇	捺	(nà)	press Down	↘	八, 大	acht, groß
㇇	提	(tí)	rise	↗	我	ich
㇇	横钩	(hénggōu)	horizontal hook	↗	你, 好	du, gut
丨	竖钩	(shùgōu)	vertical hook	↓	你, 好	du, gut
㇇	斜钩	(xiégōu)	slant hook	↘	我	ich
㇇	横折	(héngzhé)	horizontal bend	→	口	Mund
丨	竖折	(shùzhé)	vertical bend	↓	七	sieben

Abbildung 15: Striche, Grundstriche und Kombinationsstriche

Generell gibt es Abweichungen bei der Klassifizierung und Benennung der Striche. Der Grund für die unterschiedliche Benennung von Autor zu Autor liegt an einer fehlenden Standardisierung der Striche. So wie das letzte Zeichen in Abb. 19 von Schmidt ShuZhe genannt wird, so wird die Drehung nach dem vertikalen Strich auch als 弯 (Wān, englisch: "Curved") bezeichnet.

Neben der Schreibrichtung spielt ebenfalls die Strichreihenfolge eine essenzielle Rolle. Das Schreiben von chinesischen Zeichen richtet sich nach klaren Grundregeln, um einen geordneten Schreibprozess zu gewährleisten. Diese Strichreihenfolgen tragen nicht nur zur Ästhetik der Zeichen bei, sondern verhelfen auch zur Standardisierung der Zeichen. Indem eine bestimmte Schreibreihenfolge für ein Zeichen herrscht und immer wieder geübt wird, trägt dies unmittelbar zur Schreibgeschwindigkeit und -effizienz bei.

Anhand des Zeichens 永 „Yǒng“, was für „ewig“ steht, kann man beispielhaft die acht Teile der kalligrafischen Striche in chinesischen Zeichen darstellen. Hier sind sowohl der sechste und siebte Strich innerhalb der gleichen Strichkategorie, „throw away“. Die folgende Abbildung zeigt die vorhin genannte Strichfolge ebenfalls durch die Nummerierung der Strichzeichen.

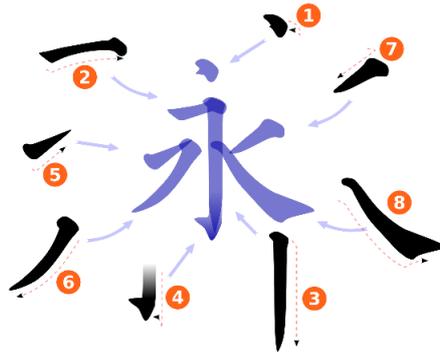


Abbildung 16: die acht Teile der fünf Strichzüge am Zeichen "Yong"

1. Horizontale Striche werden vor vertikalen Strichen geschrieben.



Abbildung 17: Strichfolge Bild 1

2. Linke Striche werden vor rechten Strichen geschrieben. Dies gilt auch für steigende/fallende diagonale Striche.



Abbildung 18: Strichfolge Bild 2

3. Vertikale aufeinanderfolgende Striche werden von oben nach unten geschrieben. Obere Radikale werden vor unteren geschrieben.

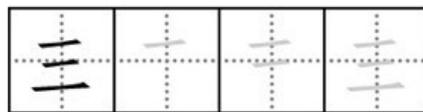


Abbildung 19 Strichfolge Bild 3

4. Horizontal aufeinanderfolgende Striche werden von links nach rechts geschrieben. Linke Radikale werden vor rechten geschrieben



Abbildung 20 Strichfolge Bild 4

5. Umschließende Striche werden vor umschlossenen Strichen geschrieben.



Abbildung 21 Strichfolge Bild 5

6. Umschlossene Komponenten werden geschrieben, bevor die umschließenden Striche geschlossen werden.

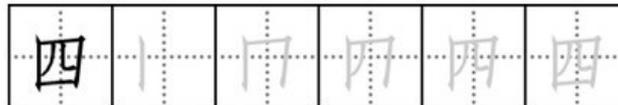


Abbildung 22 Strichfolge Bild 6

7. Bei mehreren horizontal angeordneten vertikalen Grundstrichen wird der mittlere Grundstrich zuerst geschrieben.

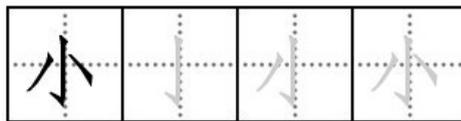


Abbildung 23 Strichfolge Bild 6

4.4.9 Die Erlernbarkeit von Chinesisch als L2

Das Foreign Service Institute (FSI), die Ausbildungseinrichtung der amerikanischen Regierung für die US-Außenpolitik bietet Sprach- und Kulturtraining zur Ausbildung von amerikanischen Diplomaten an. Um die Schwierigkeit von Sprachen zu bewerten hat das FSI eine Liste veröffentlicht, in der Sprachen in vier Kategorien gelistet wurden. Diese gliedern sich nach ihrer Unterschiedlichkeit hinsichtlich der Rechtschreibung und Grammatik im Verhältnis zur Englischen Sprache und darin welchen erwarteter Zeitfaktor von Lernenden im Durchschnitt benötigt wird um das Leistungsniveau „General Professional Proficiency“ zu erreichen - welches eine Leistungsbewertung nach den Kriterien der Sprachlevel des Interagency Language Roundtable Scale (ILR) folgt (3) (4). Dies würde nach der CEFR (Common European Framework of Reference for Languages) in etwa dem Sprachlevel B2 bis C1 entsprechen. (5) Das FSI kategorisiert Mandarin-Chinesisch in die Sprachkategorie IV, Lernende benötigen für Sprachen in dieser Kategorie ca. 88 Wochen bzw. 2200 Klassenstunden, damit sie das Niveau „General Professional Proficiency“ erreichen. Damit unterscheidet sich Mandarin-Chinesisch eindeutig von anderen Sprachen der Sprachkategorie I bis III, die eine Lernzeit von 24 bis 44 Wochen bzw. 600 bis 1100 Klassenstunden benötigen. In der Kategorie IV andere asiatische Sprachen wie Koreanisch,

Japanisch, Kantonesisch und auch Arabisch. Die Einstufung des FSI listet diese Sprachen somit als besonders schwierig zu lernen – mit Bezug auf nativ Englisch sprechende Muttersprachlern.

Das Lernen oder Schreiben chinesischer Schriftzeichen erscheint aufgrund der komplexen Komponenten herausfordernd für Lernende, die Chinesisch als L2 lernen. Die Vielzahl an Zeichen, die große Menge semantischen Logogrammen, welche zum inhaltlichen Verständnis und zur Aussprache einzeln erlernt werden müssen, kombiniert mit dem Charakterstrichfolgen und -richtungen bringt besonders für Lernende mit lateinischen Schriftsystemen als Grundlage die grundlegende Herausforderung.

Im Hinblick auf die Entwicklung der chinesischen Sprach- und Schrifthistorie wurden mehrere Strukturen und Vereinfachungen etabliert. Sowohl die bedeutungsbehafteten Radikalbausteine, Grundstriche und komplexen Strichen sind systematisch aufgebaut und ermöglichen Lernenden, Muster zu erkennen. Auch die Tonalität weist Muster zur Silbenstruktur, Phonologie und den Logogrammen auf. Für ein effizientes und umfassendes Verständnis sollten all diese Komponenten der chinesischen Schrift und Sprache schrittweise, aber gleichzeitig erlernt werden. So ist der Anfang des Chinesisch Lernens deutlich herausfordernder als nicht-asiatische Sprachen. Das Erkennen von Mustern, die Assoziation von Schrift und Sprache und der kontextbezogene Gebrauch können Lernenden helfen die linguistische Herausforderungen von Chinesisch zu bewältigen.

5 Virtuelle Realität – Was ist VR?

Es herrschen mehrere Definitionen zu Virtual Reality. So definiert Riva (2016) VR als beliebige Technologie, die eine simulierte Erfahrung erzeugen kann, bei der man in eine virtuelle Welt eintauchen kann und diese die physische Welt ersetzt. Cruz-Neira (1993) spricht wiederum von einer computergenerierten Welt, die immersiv, interaktiv, multisensorisch, eine nutzerzentrierte Sicht zeigt und in einer dreidimensionalen Umgebung ist. Ähnliche Aussagen werden auch bei anderen Forschern gefunden. Gemeinsame Punkte liegen bei der Immersion in eine computergenerierte, künstliche Welt und in der Interaktivität mit der Umgebung.

Virtuelle Realität ist ein Teil des Milgram-Kontinuums (Milgram et al. 1995), welches ein Spektrum von realen und virtuellen Anteilen in computergenerierten Welten ist. Als übergreifender Begriff wird Mixed Reality (MR) genutzt. Auf der anderen Seite findet sich die Realität, sprich die echte Welt, die von Menschen wahrgenommen wird, sowie Augmented Reality (AR), welche eine Anreicherung von virtueller Information in der Realität ist und Augmented Virtuality (AV), was eine Anreicherung von realer Information in Virtualität zu Grunde legt. Somit bildet VR den letzten Punkt mit rein virtuellen Erfahrungen.

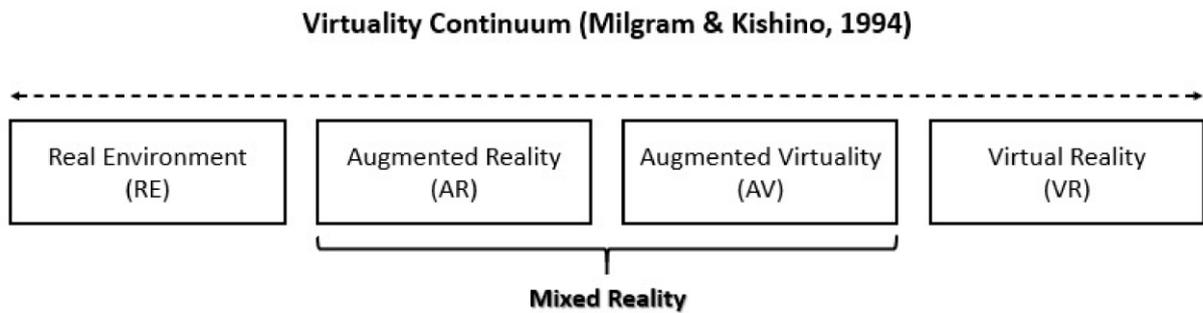


Abbildung 24: Virtual Continuum

Der Realisierungsprozess der virtuellen Realität in Computersystem bildet ein VR-System. Der Inhalt eines VR-Systems wird als virtuelle Welt bezeichnet, was z.B. die Gestaltung der Welt durch Anordnung der Modelle und Objekte umfasst. Diese dargestellte Welt für einen oder mehrere Nutzenden wird somit als virtuelle Umgebung bezeichnet.

5.1 Historie und Meilensteine der Virtual Reality

Im Folgenden werden die wichtigsten Meilensteine der VR-Entwicklung als historischer Überblick im Wesentlichen dargestellt.

Virtual Reality hat seinen Anfang und wichtige Meilensteine bereits Mitte des 20. Jahrhunderts. Ivan Sutherland formulierte 1965 sein Konzept des „ultimativen Displays“, welches grundlegend für die heutige VR ist. Drei Jahre später präsentierte er sein Buch mit dem „Head-Mounted-Three Dimensional Display“ und im gleichen Jahr zeigte er das erste HMD „Sword of Damocles“. Am Rechner wurden Bilder erzeugt, die durch einen augenrandnahen Bildschirm dargestellt wurden. Das HMD war allerdings zu schwer für direktes Tragen und wurde mit einem Gestell unter der Raumdecke angebracht. 1992 wurde erstmal eine Cave Automatic Virtual Environment (CAVE) an der Universität of Illinois als Alternative zum HMD entwickelt, welche aber wegen ihrer hohen Kosten und Rechenaufwands erfolglos blieb. Im Bereich der Videospiele brachte Nintendo 1995 das HMD „Virtual Boy“ raus, welches sich aber in der Nachbetrachtung auf seiner geringen Rechenkapazität und schwacher Bildqualität nicht durchsetzte.

Weitere Projekte scheiterten ebenfalls in den Folgejahren. Seit 2012 konnte sich VR endlich dynamisch entwickeln was durch Fortschritte in Kommunikations- und Informationstechnologien begründet ist. Die Verfügbarkeit von leistungsfähigeren Prozessoren, Grafikkarten und kostengünstigere, hochauflösende Monitore waren hier kausal. Das damalige Startupunternehmen Oculus VR präsentierte 2012 die Oculus Rift welche im Rahmen eines Crowdfundings finanziert wurde. Dieses kostengünstige HMD wurde von Anfang an den Gaming Markt gerichtet. Die Übernahme durch Facebook im Jahr 2014

konnte VR nun die nötige mediale Aufmerksamkeit geben und Tech-Unternehmen wie Google, Samsung, Sony und HTC starten ebenfalls mit Ihrer VR -Entwicklung. Zum gleichen Zeitpunkt machte die Verbreitung der Smartphones einen weiteren Schritt und durch den Umstand das nun hochwertigere Flachbildschirme und Gyroskopie eingebaut wurden profitierte ebenfalls die technische Hardwareverbreitung für VR-Systeme. Der Weg zum leistungsstarken Computer, der Bildschirme mit höherer Auflösung und Bildwiederholrate unterstützte tat sein Übriges. Ein Prototyp der Oculus Rift HMD (Dev Kit 2) wurde 2014 verschickt und fand Einsatz für klinische Forschungszwecke.

- Laut dem Marktforschungsunternehmen SuperData wurden im Jahr 2020 weltweit rund 6,7 Milliarden US-Dollar an VR-Hardware und -Softwareumsätzen erzielt. Es wird erwartet, dass der Umsatz bis Ende 2023 auf über 16 Milliarden US-Dollar steigen wird. Eine andere Studie von Statista schätzte den weltweiten Umsatz mit VR- und AR-Hardware im Jahr 2020 auf rund 6,1 Milliarden US-Dollar. Die Prognose für das Jahr 2023 lag bei etwa 20,9 Milliarden US-Dollar.
- Der Gaming-Sektor ist einer der Haupttreiber des VR-Marktes. Laut dem VR/AR-Industriereport von Greenlight Insights wurde der weltweite Umsatz mit VR-Spielen im Jahr 2020 auf rund 589 Millionen US-Dollar geschätzt. Es wird erwartet, dass dieser Umsatz bis 2025 auf über 4 Milliarden US-Dollar steigen wird. Neben Gaming wird VR auch in anderen Branchen wie Gesundheitswesen, Bildung, Immobilien, Tourismus und Industrie eingesetzt. Diese Anwendungsbereiche tragen ebenfalls zum Gesamtumsatz bei, obwohl genaue Zahlen schwer zu ermitteln sind.
- Es ist wichtig zu beachten, dass der VR-Markt weiterhin relativ jung ist und sich schnell weiterentwickelt. Die Nachfrage nach VR-Technologie steigt, neue Produkte kommen auf den Markt und immer mehr Unternehmen erkennen das Potenzial von VR-Anwendungen. Daher werden sich die Umsatzzahlen voraussichtlich in den kommenden Jahren weiter verändern und wachsen.

5.2 Motivation und Einsatzmöglichkeiten von VR außerhalb Gaming

Die Gaming-Industrie waren die Haupttreiber der VR-Hardware und Spiele und bleiben dies wahrscheinlich auch weiterhin, doch gibt es auch andere Möglichkeiten VR einzusetzen neben Unterhaltungszwecken. Gaming hat aber elementar dazu beigetragen, dass eine VR eine breitere Akzeptanz gefunden hat. Im Bereich des Serious-Gamings findet VR ein großes Anwendungsgebiet als digitaler Zwilling. Das Konzept beschreibt ein Abbild von physisch realen Objekten oder Situationen in einer virtuellen Umgebung simuliert und erlaubt die Simulation, Steuerung und Verbesserung. Der Einsatz von VR ist vielfältig und selbst in den Unterkategorien gibt es unzählige Use Cases, in denen

virtuelle Welten sich als nützlich erweisen. Folgend wird sich daher exemplarisch auf jeweils ein Beispiel und ein Anwendungsfall in den Visualisierung und virtuellem Training beschränkt.

5.2.1 Visualisierung

In der Modellbauplanung können Entwürfe realitätsnah visualisiert werden und Prototypen kostengünstig hergestellt werden als im traditionellen Modellbau. Die Nachsimulation von Prozessabläufen wie z.B. die Planung einer virtuellen Fabrik oder Anlage kann Planungsfehler analysieren und beseitigen. Die Fähigkeit durch virtuelle 3D-Welten die Visualisierung von 2D-Bauplänen herzustellen, dass man Eindrücke von Gebäuden und Objekte in der Realität vor der praktischen Umsetzung bekommt, kann hilfreich bei der Bauplanung sein. Auf Bauplänen ist die Höheninformation der Objekte nicht direkt erkennbar, sondern der Vorstellung der Bauplanenden ausgesetzt. Ebenso können Urteile zu Realisierungsalternativen fundierter getroffen werden.

Als Beispiel entstand das Projekt PaKOMM (Partizipation und Kollaboration multimedial) aus der Motivation Stadtplanung zu digitalisieren und modernisieren, BürgerInnen zu involvieren und es nahbar für sie zu machen. Statt mit traditionellen Bauplänen soll eine interaktive und multimediale Möglichkeit gegeben werden. Durch VR, AR und einem Touchtisch können Bepflanzung-, Bebauung- und Beleuchtungselementen in Form von 3D Objekten platziert werden und bestehende Planung verändert werden. An dem flachen Touchtisch soll Einzelpersonen oder Gruppen ein erster Eindruck über die Umgebung in Form einer Kartenansicht in der Draufsicht gegeben und damit ein Überblick über die Planungen ermöglicht werden. Diese haben dann die Optio diese zu verändern, was noch eher der ursprünglichen Bauplanungsart entspricht. Über die VR HMD können Nutzende immersiv in das Gebiet der Planung aus Sicht der ersten Person eintauchen und die Planung ändern. Darüber hinaus kann über AR ein Tablet vor Ort genutzt werden, um sich die platzierten Objekte anzuschauen und ebenso zu verändern. Ein großer Vorteil ist, dass mehrere Nutzende mit unterschiedlichen Geräten an einem Stadtplanungsentwurf zu synchroner Zeit arbeiten können. Ziel ist es eine Umgebung zu erschaffen, da ein kreativer, kollaborativer Prozess und ein offener Diskussionsaustausch der Planungsstrukturen angeregt wird

5.2.2 Virtuelles Training

Hierbei handelt es sich um eine Trainingsmethode, in der eine virtuelle Umgebung von Auszubildenden zu Schulungszwecken verwendet wird. Im Bereich der Medizin der könnten Ärzte z.B. Operationen in einer virtuellen Umgebung trainieren und planen ohne die Gefährdung des Patientenwohls. Pflegepersonal kann ebenfalls an Patienten den Umgang und die Versorgung üben. Zu Therapiezwecken

könnte es sogar ebenfalls Patienten dienen, z.B. zur Behandlung von Phobien können Patienten mit Angst auslösenden Faktoren konfrontiert werden und Schritt für Schritt eine Toleranz aufbauen.

Ein nennenswertes Projekt hier ist PflegeDigital 2.0, welches eine digital-interaktive Ausbildung zur Vermittlung Handlungs- und Sprachkenntnisse im Bereich Pflege ist. Die Plattform soll zwischen Pflegepersonal mit Migrationshintergrund, sprich Fachkräfte mit Deutsch als L2, und PatientInnen unterstützen, sodass zwischen ihnen eine unmissverständliche Kommunikation erfolgen kann, was zu einer effektiveren Patientenbetreuung beiträgt und Berührungsängste abbaut. Die Realisierung erfolgt über Moodle Boards, in denen angehende PflegerInnen verschiedene Module belegen. Das Spektrum erstreckt sich vom Erlernen elementarer, praxis- und prüfungsrelevanter Handlungsprozesse in der Pflege bis hin zu Trainings in virtueller Realität in der das vermittelte Wissen der Lernmodule verfestigt und vertieft wird. (Abb. 25)



Abbildung 25 Abbildung das Modul zum Training der aseptischen Wundreinigung in Virtual Reality

PflegeDigital 2.0 kann vorbereitend und begleitend zum Präsenzunterricht oder als Selbstlernwerkzeug in der Pflegehelferausbildung genutzt werden. Als Anwendung die die Nutzung von computertechnologischen Geräten und Medien zu Bildungs- und Lernzwecken bereitstellt qualifiziert PflegeDigital 2.0 als CAL. Zusätzlich leistet es durch die Integrität von fachspezifischer Sprachvermittlung einen Beitrag zu CALL in der aktuellen Zeit.

5.3 Wie wird der Effekt einer virtuellen Realität erreicht?

Das Potential von VR den Menschen in eine virtuelle Umgebung zu transportieren ist maßgeblich von zwei Faktoren abhängig. Auf der einen Seite von der menschlichen Wahrnehmung und auf der anderen Seite den technischen Voraussetzungen, die die Raum- und Bewegungswahrnehmung der virtuellen Umgebung begünstigen. Während Wahrnehmungsaspekte die Informationsverarbeitung zur Interpretation der vorgestellten Sinnesreize anregt, beschäftigt sich die Technik von VR mit der Darstellung der Sinnesreize, aber auch mit der Verarbeitung der aufgenommenen Sinnesreize.

5.3.1 Wahrnehmung in VR und menschliches Informationsverarbeitungssystem

Die klassische Einteilung der Wahrnehmung ist bekannt als fünf Sinne: Hören, Sehen, Tasten, Schmecken und Riechen, wobei die Aufnahme von Reizen in der virtuellen Realität sich größtenteils mit den ersten drei Sinnen und dem vestibulären Sinn zusammenhängt. Passend zu den Sinnen spielen daher folgende Wahrnehmungsaspekte bei der Erfahrung mit Virtual Reality eine Rolle.

- Visuelle Wahrnehmung, geteilt in:
 - Stereosehen
 - Raumwahrnehmung
- Auditive Wahrnehmung
- Haptische Wahrnehmung

Zusätzlich gibt es Wahrnehmungsaspekte der Tiefensensibilität

- Tiefenwahrnehmung (Empfindungen, die mit Körperpositionen zu tun haben)
- Bewegungswahrnehmung (Fähigkeit die Bewegungsrichtung zu erkennen und zu bestimmen)

VR ist im Stande mächtige Wahrnehmungsimaginationen hervorzurufen, die dennoch nur eine Illusion sind. So sind Nutzende fähig zu verspüren, dass eine VR Umgebung nicht real ist. Slater (2018) hat festgestellt, dass diese kognitive Beurteilung allerdings erst erfolgt, nachdem die Nutzenden kongruent mit der Umgebung interagiert haben, was suggeriert, dass die Umgebung dennoch als real wahrgenommen wurde.

Eine Erklärung für dieses Phänomen könnte das Modell der menschlichen Informationsverarbeitung von Carter et al. (1986) sein. Nach ihm wird es in drei Phasen verarbeitet. In dem perzeptuellen System werden Reize der äußeren Welt als Eingaben aufgenommen und an das System übergeben. Die Verarbeitung der wahrgenommenen Reize passiert im kognitiven System, in dem auch auf das Arbeits- und Langzeitgedächtnis zugegriffen werden kann. Diese können im ständigen Austausch zueinanderstehen.

Durch das mehrfache Durchlaufen der Verarbeitung der Reize wird bei kognitiven Aufgaben mehr Zeit benötigt, wie bei solchen wo der kognitive Prozess im nur einmal durchlaufen wird oder vom Arbeitsgedächtnis es direkt in die Ausgabe der Reize als motorische Prozesse, welche die in kognitiven Prozess geplante Handlung dann durchführt. So kann der Grad der Präsenz anhand physiologischer Daten oder dem Verhalten des Menschen bestimmt werden. Bei einem hohen Grad von Präsenz werden NutzerInnen in VR simulierte Gefahrensituationen entsprechend mit erhöhter Hautleitfähigkeit oder Herzschlagrate reagieren (Slater et. Al 1994), was als motorischer Prozess und Reaktion verstanden werden kann.

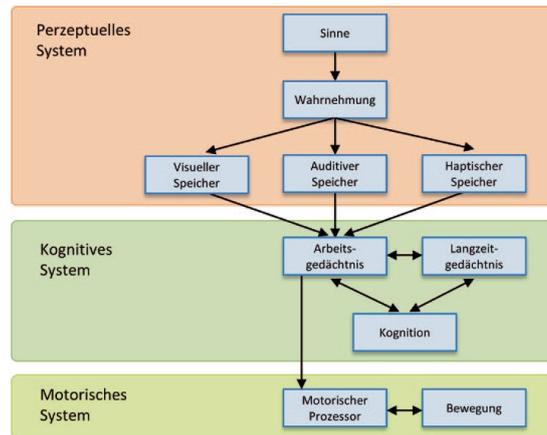


Abbildung 26 Modell der Menschlichen Informationsverarbeitung. Nach Card et al. (1986)

Die multisensorische Wahrnehmung begünstigt das Gefühl vom Eintauchen in eine virtuelle Welt. Dabei gibt es zwei Möglichkeiten dieses Gefühl zu beschreiben: Immersion und Präsenz.

Das Gefühl von Immersion beschreibt den Grad des Eintauchens in einer objektiven Sicht durch quantifizierbare Stimuli bzw. durch die eben beschriebenen multimodalen Stimulationen der menschlichen Wahrnehmung. Allerdings ist die objektive Wahrnehmung allein nicht ausreichend. Denn selbst wenn Stimulationen gegeben sind, passiert eine vollkommene immersive Erfahrung erst dann wenn man durch die subjektive Wahrnehmung in der virtuellen Umgebung Existenz spürt. Wenn Nutzende sich in der VR-Welt befinden, aber agieren und reagieren, wie in der realen Welt und die Reize der Realität ausgeblendet werden können, verspüren sie **Präsenz** ihrer Selbst in der virtuellen Welt. Dies ist besonders bei sozialen Interaktionen möglich. Beispielsweise ist bei der Kommunikation zwischen VR Avataren das Präsenzgefühl besonders glaubhaft und nahbar, sobald eine Möglichkeit zum sozialen Austausch besteht. Demnach kann ein hoher Grad an Immersion durch gute Qualität der Wahrnehmung das Gefühl von Präsenz begünstigen, aber nicht garantieren, wie Slater (2018) festgestellt hat.

Das Präsenzgefühl kann zusätzlich beeinflusst werden durch folgende Illusionen:

- Plausibilitätsillusion (Ereignisse in VR erscheinen plausibel oder wahrheitsgemäß)
- Ortsillusion (Darstellung der Szene)
- Involviertheit (Grad der Aufmerksamkeit und Interesses des Nutzers, Zusammenspiel von Plausibilitätsillusion und Plausibilitätsillusion)

5.3.2 Technische Anforderungen

Damit die in der menschlichen Wahrnehmung aufgenommene Reize zur Immersion und Präsenz in virtuellen Welten ausgeschöpft werden, müssen sie technisch umgesetzt werden. Dabei erfolgt dies wie beim menschlichen Informationssystem durch Ein- und Ausgabesysteme. Technische Eingabesysteme in der VR sind neben der Ausstattung auch Licht- und Audioquellen sowie virtuelle Kameras. In der virtuellen Welt steht die Echtzeitfähigkeit des virtuellen Systems und Interaktivität der Nutzenden im Mittelpunkt. Die folgende zwei Punkten bilden die Grundlage der technischen Anforderungen der virtuelle Welt: Echtzeit und Interaktivität.

5.3.2.1 Echtzeitfähigkeit

Ein System wird als echtzeitfähig bezeichnet, wenn es Resultate einer Eingabe zuverlässig in vorhersagbaren Zeiträumen liefern kann. Um eine virtuelle Welt überzeugend darzustellen sollte diese möglichst verzögerungsfrei aktualisiert werden, um die reale Welt zu simulieren ist eine Bildfrequenz ab 60 FPS ideal und zudem eine Bildwiederholrate von 100 Hz, damit das Bild als flimmerfrei gilt. Zudem ist das Echtzeitgefühl abhängig von der Latenz. Im Sinne von VR beschreibt die Latenz z.B. die Verzögerung zwischen der Bewegung einer VR Headset-tragenden Person und der visuellen Ausgabe der Brille. Nach Brooks (1999) ist eine Latenz unter 50ms für HMDs geeignet, während Abrash (2012) sogar eine Latenz unter 20ms empfiehlt, was einer Bildwiederholrate von 50 Hz entspricht.

5.3.2.2 Interaktivität

Ebenso wichtig wie die Echtzeitfähigkeit ist die Interaktivität eines virtuellen Systems. Interaktivität beschreibt die Manipulierung der 3D Objekte in der virtuellen Umgebung und die Umgebung als solches. Für eine Simulierung einer realen Umgebung werden Kollisionsgeometrien benötigt. Je einfacher diese Geometrien gebaut sind, desto weniger rechnerintensiv ist die Berechnung, was zur weniger Latenz führt und den Echtzeitaspekt vereinfacht. Interaktion mit VR Systemen kann vielfältig sein und passieren durch die Eingaben der Nutzenden (wie werden 3D Objekte gebaut).

5.4 Eingabesysteme

Damit die Eingaben verarbeitet und zu Ausgaben führen können, werden sensorischen Erfassung von Nutzerinteraktionen in Form von Interaktionsgeräten oder bestimmten Körperteile des Nutzenden erkannt. Eingaben vom Nutzenden können z.B. bewusste Handlungen wie Knopf- und/oder Tastendruck

des VR Controllers sein. Diese Eingaben werden diskrete Eingabesysteme bezeichnet, welche einmalige Ereignisse erzeugen. Daneben gibt es natürliche Interaktionen mit dem Controller, der Hand oder des Kopfes in Form von Gesten. Bei diesen kontinuierlichen Ereignisströmen, die diese Eingabegeräte erzeugen, müssen Position und Orientierung der Objekte verfolgt werden, weshalb Bezeichnung *Tracking* geläufig ist. Die Verschiebung des Objekts im Raum wird *Translation* genannt, während die Drehung *Rotation* meint. Somit hat ein virtuelles System, dessen Objekte und die kontinuierlichen Eingabesysteme sechs Freiheitsgrade aufgeteilt auf die Koordinatenachse für Translation (x, y, z-Achse) und Rotation (pitch, roll, yaw) (Abb: 31). Durch Tracking sollen die Freiheitsgrade der verfolgten Körper oder Interaktionen bestimmt werden, wodurch die Interaktion mit einem virtuellen System möglich ist.

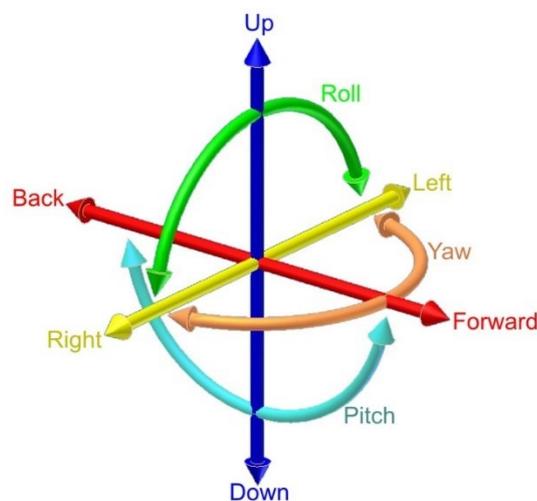


Abbildung 27: Die sechs Freiheitsgrade

https://upload.wikimedia.org/wikipedia/commons/f/fa/6DOF_en.jpg

Folgende Charakteristika sind bei verschiedenen Eingabe- bzw. Trackingsystemen qualitativ unterschiedlich ausgeprägt:

- Anzahl der Freiheitsgrade pro verfolgtem Körper
- Anzahl der gleichzeitig verfolgten Objekte
- Größe der überwachten Fläche bzw. des überwachten Volumens
- Genauigkeit
- Wiederholrate
- Latenz
- Drift (Positionsungenauigkeiten)
- Empfindlichkeit gegenüber äußeren Rahmenbedingungen
- Kalibrierung

- Usability

Im Folgenden werden verschiedene Trackingsysteme vorgestellt. Es wird angemerkt, dass es sich erstens nicht alle existierenden Trackingsysteme vorgestellt werden, sondern nur solche, die von den in Kapitel 5.6 vorgestellten VR HMD Ausgabegeräten genutzt werden.

5.4.1 Optisches Tracking

Bei optischen Trackingsystemen geht es um die Bestimmung der relativen Positionierung und Orientierung der Objekte, die im Videostrom aufgenommen werden, zur Kamera. Es bietet eine hohe Genauigkeit für die Positionsbestimmung an. Durch flexible Einsatzmöglichkeiten wird dieses System für moderne VR HMD attraktiv.

Optisches Tracking nutzt mehrere Verfahren. Zum einen unterscheidet man zwischen *Outside-In Tracking* und *Inside-Out Tracking*. Diese geben Auskunft über den Ort der Kameras, welche die zu trackenden Objekte verfolgt.

Beim Outside-In Tracking werden die Trackingkameras in die externe Umgebung gesetzt, wo die getrackten Geräte sich innerhalb des Sichtfeld der Kameras befinden. Meistens werden mehrere Kameras kombiniert, sodass der Interaktionsbereich bestimmt und vergrößert wird. Outside-In leidet weniger unter Latenzproblemen. Die Anzahl der Kamerasensoren kann jederzeit erhöht und so die Genauigkeit verbessert werden. Ein größeres Interaktionsgebiet geht allerdings mit dem Erwerb zusätzlicher Kameras und damit Kosten einher, je größer das Gebiet sein soll. Dies ist besonders dann nötig, wenn Teilgebiete von der Kamera nicht zu sehen sind oder das Sichtfeld der Kameras (Field of View FOV) nicht breit genug ist).

Beim Inside-Out Tracking wiederum werden die Trackingkameras auf oder an dem HMD platziert. Die Kamera nimmt einen Videostrom auf und baut mehrere Referenzpunkte zu Objekten und Stellen in der Umgebung auf, sodass Position und Orientierung der Kamera zu den Objekten aufgebaut werden kann und bestimmt wird. Nutzende sind mobiler in ihrer Bewegungsfreiheit und im Interaktionsraum als mit einem Outside-In HMD. Allerdings kann das Tragen von Headset, Kamera, Akku sich schwer anfühlen trotz kleinerer Kameramodule. Die Positionsbestimmung zudem weniger präzise wie Outside-In.

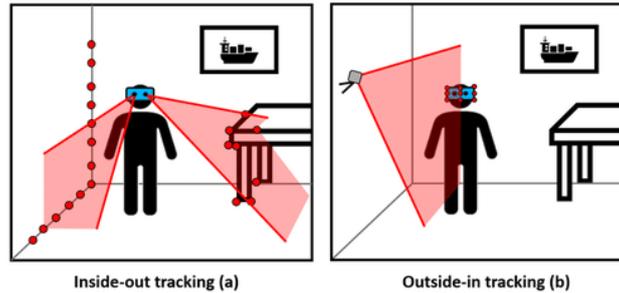


Abbildung 28: Inside-Out und Outside-In Verfahren

Grundsätzlich wird zwischen *markerbasierte Verfahren* und *markerlose Verfahren* unterschieden. Marker sind Objekte für VR-Eingabesysteme die als Referenz- oder Messpunkte dienen sollen. Sie sind im aufgenommen Videostrom leicht erkennbar und werden durch Schwellwertfilter schnell gefunden. Sie sollen die Berechnungskomplexität verringern und die Fehleranfälligkeit gegenüber unterschiedlichen Beleuchtungssituationen vermeiden. Dabei wird mindestens eine Marke gebracht, die der Positionsbestimmung dient als auch mehrere Marker für die Orientierung. Aktive Marker strahlen Licht selbst ab und passive Marker reflektieren Licht. Für eine bessere Erkennung der Marker werden Szenen mit Infrarotlicht beleuchtet, ohne die Nutzenden dabei zu blenden. Heutzutage versteht man unter Marker in VR nicht nur monochrome Sticker wie Abb. 33, sondern auch unter „Lighthouses“ oder „Basisstationen“, welche Infrarotlicht aussenden und durch ihr Outside-In Tracking eine hohe Erfassungsrate der zu trackenden Objekte liefern können.

Markerlose Verfahren nutzen diese Referenzobjekte nicht und arbeiten nur mit dem, was die Kamerasensoren erkennen, um die Freiheitsgrade zu bestimmen.



Abbildung 29: Sticker zur Kalibrierung

5.4.2 Finger Tracking

Die Standardeingabe mit VR Controllerknöpfen mag zwar meistens ausreichend sein, bildet virtuell aber nicht genug echten natürlichen Interaktionen wie Finger- und Handgesten von Menschen ab. Direkte Interaktionen der NutzerInnen mit deren echten Händen und Fingern in der virtuellen Welt sind meistens

intuitiver (Bowman et al. 2004), wobei Standardinteraktionen schneller und einfach sind (Möhring und Fröhlich 2011; Hummel et al. 2012). Unter Finger-Tracking ist meistens umfassenderes Tracking als nur die Erfassung von Position und Orientierung vom ganzen Finger und ganzer Hand gemeint. Jeder Finger hat vier Freiheitsgrade, zwei rotatorische an der Fingerwurzel und je einen rotatorischen Freiheitsgrad für die Gelenke an den Fingergliedern. Um Gesten zu erkennen, müssen also 27 Freiheitsgrade bedient werden. Die dichte Lage von Fingern zueinander und die Verschiedenheit von menschlichen Händen macht Finger- und Handerkennung schwer und verlangt nach Anpassbarkeit.

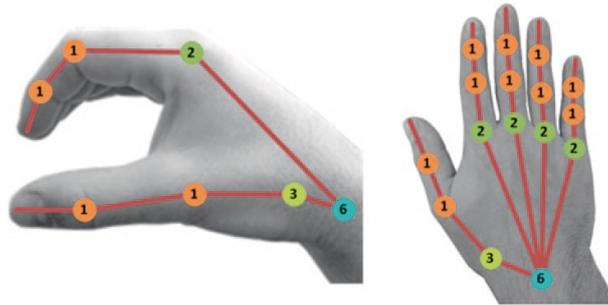


Abbildung 30: Fingertracking

Dörner, R., Broll, W., Grimm, P et. Al (2013)

Gängig für Fingertracking sind mechanische Trackingverfahren wie „Gloves“, die mechanische Trackingverfahren nutzen oder Fingertrackinggeräte, die auch markerbasiertes Tracking nutzen. Allerdings waren optische Finger auch seit 1983 trackbar. Dabei wurden Leuchtdioden genommen, die von externen Kamerasystem ausgenommen werden.

5.4.3 Laser Tracking

Lasertracking nutzt Basisstationen, die zwei rotierende Laserstrahle aussenden und legt Fotosensoren auf die zu trackenden Objekte, die die Laserstrahlen erkennen.

Beim Laser-Tracking werden Fotosensoren, die Laserstrahlen erkennen, auf den zu trackenden Objekten angebracht. Von einer Basis, die zwei rotierende Laser benutzt, werden Laserstrahlen ausgesendet. Es werde mehrere Basisstationen gebraucht da sie oft verdeckt werden. Dazu werden Infrarotsignale verwendet, um Laserstrahlen zu übertragen.

Innerhalb einer Basisstation rotieren zwei Laser horizontal und vertikal um eine Achse, wodurch der Laserstrahl in einem bestimmten Öffnungswinkel gesendet wird. Die Laser werden ca. 100-mal in den Sekunden an den Fotosensoren erkannt. Die Erkennung zu unterschiedlichen Zeiten gibt Ausschluss über die Position und Orientierung des zu trackenden Objekts.

Die Laserstrahlen werden zu unterschiedlichen Zeiten von den Sensoren erkannt, wodurch die Position und Orientierung des Objekts berechnet werden können, ohne Überlappungen zu gefährden.

5.4.4 Eye Tracking

Eye Tracking bezeichnet im Allgemeinen sowohl die Verfolgung der Blickrichtung beim menschlichen Auge als auch die Blickregistrierung und Erfassung sowie die Auswertung des Blickverlaufes.

Es wird zwischen invasiven und non-invasiven Verfahren unterschieden. Ersteres nutzt einen direkten Eingriff am Körper durch Elektroden, während bei letzterem das Auge mittels einer Kamera videografiert wird. Anhand des erfassten Videos werden mittels Bildverarbeitungsalgorithmen unmittelbar die Blickrichtung bestimmt. Generell erfolgt eine Unterscheidung zwischen passiver und aktiver Augenbestrahlung. Bei passiver Bestrahlung wird die Augenszene mit Umgebungslicht bestrahlt und die Kontur zwischen Lederhaut und Iris als Identifikationsmerkmal genutzt. Aktive Bestrahlungsverfahren funktionieren durch Infrarotbestrahlung und unterscheiden zwischen dunkler und heller Pupillentechnik. Die Pupille wird zum dunkelsten Objekt der Augenszene, wenn die Bestrahlungsquelle außerhalb der optischen Achse der Eye-Trackingkamera liegt, sodass die Strahlung durch Iris und Lederhaut reflektiert wird. Zum hellsten Objekt wird sie, wenn Lichtquelle und Kamera in gleicher optischer Achse geordnet werden, sodass eine Reflexion der Strahlung an der Netzhaut im inneren Auge passiert.

Die Verwendung von Eye Tracking in HMDs erfolgt fast immer mit non-invasiven videobasiertem Tracking durch die Kamera am HMD. Dort fokussiert die Kamera das Auge entsprechend und nimmt ein Bild auf, das an einen Computer mit Eye-Trackingalgorithmus übertragen wird.

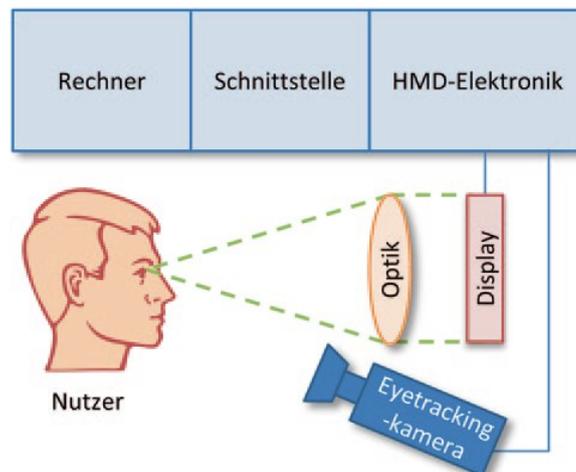


Abbildung 31:: Prinzipieller Aufbau eines Eye-Tracking-HMDs

Dörner, R., Broll, W., Grimm, P

Die Abbildung 36 zeigt den prinzipiellen Ablauf einer Eye-Tracking-Routine mit aktiver Beleuchtung bei heller Pupillentechnik. Eine Eye-Tracking-Kamera, die auf das Auge des Nutzers fokussiert ist, nimmt ein digitales Graustufenbild auf. Dieses Bild wird der Eye-Tracking-Bildverarbeitung übergeben.

In Schritt eins wird eine Anpassung der Grauwerte durchgeführt und das Bild vorgefiltert, um Einflüsse wie beispielsweise ein verrauschtes Bild zu minimieren. Danach wird eine Histogrammspreizung durchgeführt, damit die Objektkonturen des Auges wie die Pupille oder die Iris herausstechend erfasst werden. In der Folge werden die Kanten detektiert und anhand von mathematischen Algorithmen errechnet. In einer aktiven Verfahrenslegung wird zudem die aktive Beleuchtung über eine Reflexion (die grünen Punkte im Schaubild) genutzt. Wichtig zu wissen ist, dass bei einem Head-Mounted Display (HMD) mit integriertem Eye-Tracking oftmals über diese Reflexionen die Referenzpunkte erfasst werden. Hieraus ergibt sich schlussfolgernd die Möglichkeit einer Erkennung des User Fokus, da aus den errechneten Pupillenmittelpunkt-Koordinaten in Verbindung mit den Hornhautreflexionen über Algorithmen dieser exakt bestimmt werden kann. Bedingt durch die beiden Verfahren und die heutige Rechenleistung spricht man von hybriden Verfahren, da je nach Situation die Werte übereinander gelegt und Abweichungen und Fehlinformationen ausgeglichen werden können.

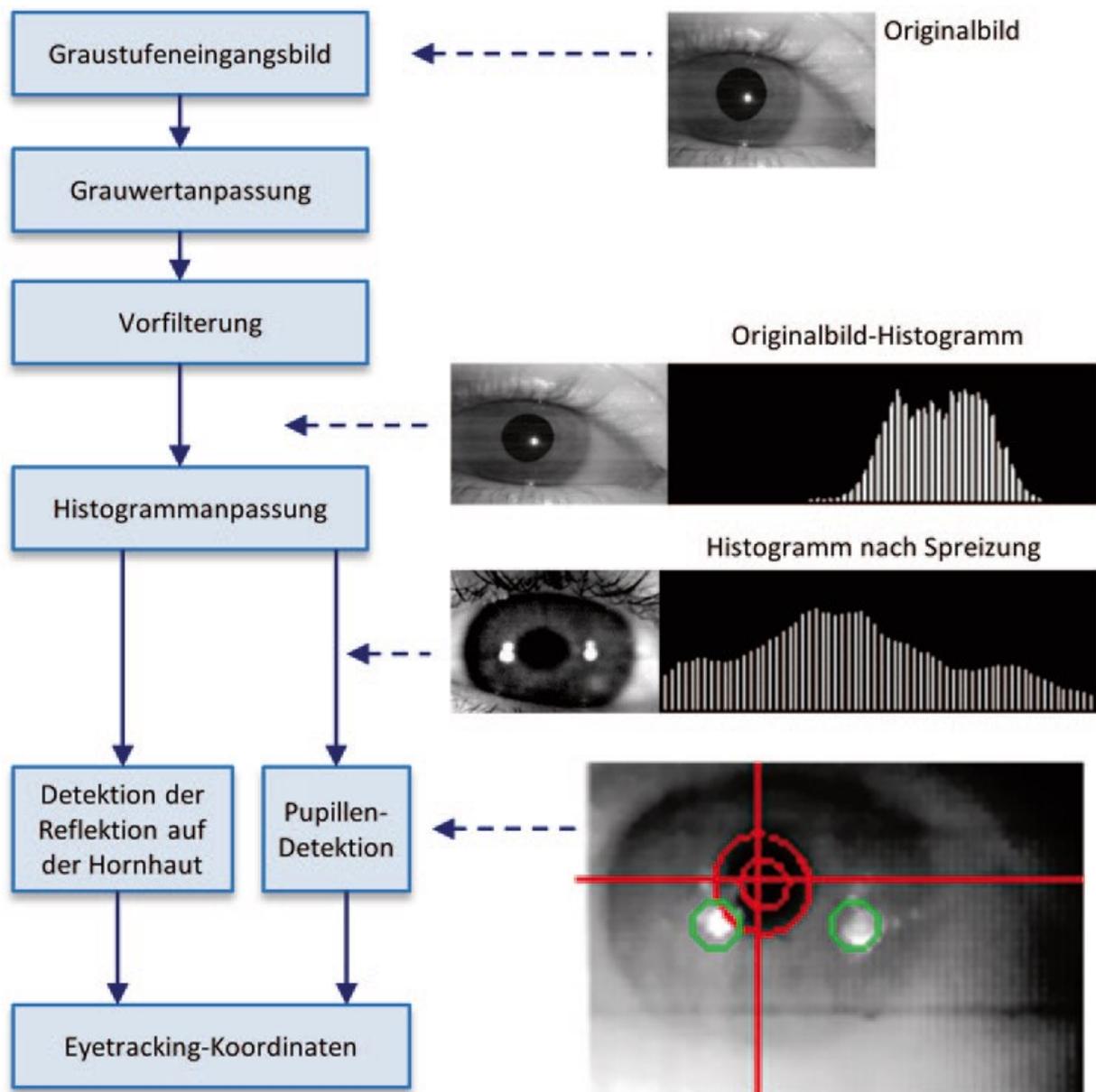


Abbildung 32: Erkennung mittels Anpassung durch die Eye Tracking Software

Dörner, R., Broll, W., Grimm, P

5.5 Ausgabegeräte

In diesem Kapitel vorgestellte Ausgabegeräte dienen alle samt dem Zweck eine Präsenz zu erzielen. Dies bedeutet den Nutzer in einen Status zu versetzen in dem die reale Welt und Virtual Reality nahezu deckungsgleich im Empfinden, dem Umgang und der Aktion und Reaktion als solcher empfunden werden. Hierzu benötigt es verschiedene Parameter die im Zusammenspiel stehen und allesamt für sich gesehen das Potential haben die Präsenz von Testpersonen zu erhöhen oder einzuleiten.

Dazu werden Parameter anhand einer getroffenen Vorauswahl und einem entsprechenden Vergleich der Hardware tiefergehend erläutert, bevor dann auf den Faktor der technischen Umsetzbarkeit geblickt wird.

Maßgeblich für die Lernumgebung der Applikation die Möglichkeit der Bewegung und Durchschreiten einer virtuellen Stadt in Anlehnung an die Schrittvariante analog zum bekannten Dienst Google Street View. Hierzu ist neben dem Feature das alle Brillen durch Ihre Eigenschaft als Brille mitbringen auch ein weiteres Eingabegerät nötig das gleichzeitig Einfachheit in seiner Bedienung bieten muss aber dennoch nicht vom Fokus des Sprachenlernens in einer nachgeahmten Umgebung ablenkt.

Stellen wir uns einmal folgende Situation vor. Wir betreten ein Cafe in der Innenstadt und wollen als Lernziel am Ende des Nutzungserfahrung einen Kaffee bestellen und zumindest einen generellen Smalltalk führen können. Dies kann im Alltag mit bereits erlernten Sprachkenntnissen in einer selbst erschaffenen Drucksituation schwierig sein. In der virtuellen Welt hingegen ist man bewusst damit konfrontiert und möchte anhand dieser Situation wachsen und anhand des Gamification Setups auch Wachstum durch Scheitern und Repetition erzielen. Hierfür ist es wichtig, dass sowohl die eigenen Bewegungen als auch Gestiken verzögerungsfrei in die Lernsoftware übernommen werden und die Situation als real empfunden wird.

5.5.1 Valve Index

Pro

- Hervorragende Controller
- Präzises Tracking
- Gutes Zusammenspiel mit SteamVR
- Gelungene Audio-Lösung

Contra

- Basisstationen erforderlich
- Teuer
- Bildschirm nicht mehr auf dem neuesten Stand



Abbildung 33: Valve Index Gesamtsystem bestehend aus Brille, Controllern und Basisstationen

Derzeitig liegt der Preis für die Gesamtlösung bestehend aus der Basisstation, sowie zwei Kameras und jeweils einem Controller pro Hand bei ca. 1400 EUR (Stand 03.12.2023 in der Preisvergleichssuchmaschine IDEALO)

Das Headset hat zwei Bildschirme, die mit jeweils 1440 x 1600 Pixeln auflösen und eine native Abtastrate von 120 Hz, 90 Hz Abwärtskompatibilität und einem experimentellen Modus von 144 Hz.

Das verwandte Tracking ist ein Laser-Tracking System. Besonderheit ist hier, dass über zwei Basisstationen auch die Outside-In Informationen erfasst werden. Dies ist dem Umstand geschuldet, dass dieses System bereits von Beginn durch den Spieleanbieter Steam und dessen SteamVR mitentwickelt wurde und Anforderungen der Spielehersteller auch in präzisiertem Tracking lagen. Durch die Verwendung von zwei Kameraerfassungssystemen kann der Umstand das es zu Verdeckungen kommt deutlich minimiert werden.

Die Index verwendet ein Laser-Tracking System. Hierfür werden über das HMD selbst noch zwei Basisstationen (2.0, rückwärts kompatibel mit dem Vorgänger Headset von HTC, der HTC Vive Pro) benötigt. Diese senden pro Sekunde 100 mal einen Rotationslaser über das abzudeckende Feld, die von Fotosensoren auf dem Headset und den Controllern eingefangen werden, um die Position und Orientierung in Relation zu den Basisstationen zu berechnen. Die Auflösung liegt in den sub-millimetern und durch das Verwenden von zwei Basisstationen sollte es seltenst zu Verdeckungen kommen.

5.5.2 Meta Quest 2

Die Meta Quest 2 hat VR massentauglich gemacht als sie im Jahr 2020 erschien. Es war das maßgebliche Referenz VR Headset bis es im Oktober 2023 vom Meta Quest 3 abgelöst wurde. Gelauncht wurde es damals noch als Oculus Quest 2. Es entstammte als Nachfolger des ursprünglichen „Oculus“ und bietet ein schlankeres Design und eine auf VR Games ausgelegte Performance. Der verbaute Snapdragon XR2 und die 6 Gigabyte Ram bieten einen enormen Performance Schub und lassen auch komplexere Anwendungen flüssig laufen.

Pro

- Kabellos via Bluetooth und WIFI
- 1832x1920 Pixel hochauflösendes Display
- Präzises Motion Tracking
- 6-fach DOF
- Gute Soundqualität (on Ear Lautsprecher)
- Vielfaches Zubehör von Drittanbietern (Controller, Band etc.)

Contra

- Kurze Batterielaufzeit von nur 3 Stunden
- Controller nicht über USB aufladbar, erfordern AA Batterien



Abbildung 34: Meta Quest 2 VR

Quelle: <https://cdn.mos.cms.futurecdn.net/t5kEThbXSZQThX5RuL3ujL-1200-80.jpg.webp>

6 Zukunft und Ausblick der VR

Die Zukunft von VR ist derzeit maßgeblich durch die Entwicklungsgeschwindigkeit der neu verfügbaren KI bestimmt. Aktuell zeichnet sich ab, dass die technologische Entwicklung und der damit einhergehende Hardware Upscale der VR Brillen dieses Tempo nicht mitzugehen scheint und somit der limitierende Faktor sein kann, wenn es um VR Technologie und deren Anwendung in der Fläche geht. Als derzeitiges massentaugliches Referenzmodell kann die Playstation VR Brille gesehen werden, da der Gamingsektor stets technischen Innovationen als auch die damit einhergehenden Möglichkeiten wie 6 DOF, Adaptive Trigger, Finger Feedback, Touch Detection, Eye Tracking und Tracking Vibration im Headset nutzt, um ein maximum an Immersion zu erzeugen. Sony ist als Gaming Konzern natürlich daran interessiert die Marktführer Rolle einzunehmen und sich als Benchmark der Branche und der neuen Hardware Nische zu etablieren – wie Nokia dies mit seinen Handymodellen gelang. Es sind erste vorsichtige Vorhersagen über die Zukunft der klinischen VR beispielsweise im Bereich der psychischen Gesundheit möglich, obwohl es offensichtliche Lücken in der bisherigen Forschung Forschungsliteratur und Trends gibt, da die Wissenschaft derzeit mit dem technologischen Fortschritt nicht Schritt hält. Es hat sich generell als schwierig erwiesen, technologische Fortschritte vorherzusagen, die wiederum neue therapeutische Anwendungen vorantreiben könnten: Der Fortschritt findet sowohl linear (wie z. B. bei Display-Eigenschaften wie Auflösung und Bildwiederholfrequenz) als auch sprunghaft und unerwartet (wie z. B. die Entwicklung des Inside-Out-Trackings, das 6DOF auch bei mobilen VR HMDs) statt und ist somit schlussendlich eine komplexe Kombination daraus und damit nicht planbar vorhersehbar.

Es ist jedoch wichtig zu beachten, dass der Einsatz von VR im Sprachenlernen immer noch ein aufstrebendes Feld ist, und es gibt Raum für weitere Forschung und Entwicklung. Die technologische

Weiterentwicklung und die Integration von KI könnten zukünftig das Potenzial von VR im Bildungsbereich noch weiter steigern.

Für die aktuellsten Studien und Entwicklungen auf diesem Gebiet empfehle ich, wissenschaftliche Datenbanken zu durchsuchen oder auf dem Laufenden zu bleiben, da die Forschung und Technologie in diesem Bereich ständig voranschreiten.

6.1 Bruch der Immersion/Probleme/Nachteile von VR

Wie bereits von (Rebenitsch and Owen 2016) festgestellt: *“Not surprisingly, the heavy and bulky HMDs of the time, with their low resolutions and framerates, were prone to induce so-called cybersickness symptoms resembling motion sickness believed to be caused primarily by sensory conflict between the visual and vestibular/proprioceptor systems, although display properties in themselves also play a role (Saredakis et al. 2020). Cybersickness was a well recognized phenomenon already in the 1990s (McCauley and Sharkey 1992)“*

Weitere Faktoren, die einen Bruch der VR bewirken können

- abweichender FOV Winkel
- Doppelbilder
- Diskrepanzen in der Raumwahrnehmung (Größe, Distanz)
- Diskrepanzen in der Bewegungswahrnehmung (Distanz, Beschleunigung) 2.4.5 2.4.6
- Cybersickness

Wird in Kombination mit diesen Displays ein Head-Tracker eingesetzt, welcher veranlasst, dass das angezeigte Bild in Echtzeit zu einer Kopfbewegung verändert wird, kann es zu einem Auftreten der Symptome kommen, falls die Bilddarstellung zu spät oder asynchron angepasst

7 OCR Erkennung mittels Tesseract

Um die Grundlagen der Software dieser Bachelorarbeit zu verstehen ist ein Blick auf OCR (optical character recognition) – zu deutsche optische Zeichenerkennung unerlässlich.

Der heutige Stand der OCR Entwicklung setzt maßgeblich auf künstliche Intelligenzen und Maschinenlearning und entwickelt sich mit bahnbrechender Geschwindigkeit. Die Grundlage von OCR und das weltweit erste Patent findet sich in den USA im Jahre 1931 und wurde von IBM beantragt.

Tesseract:

- Open Source
- Komplettlösung „All-in-1“

- Mehr als 100 Sprachen / mehr als 30 Schriften
- Liest Bilder in allen gängigen Formaten (nicht PDF!)
- Erzeugt Text, PDF, hOCR, ALTO, TSV
- Große, weltweite Anwender-Community
- Technologisch aktuell (Texterkennung mit neuronalem Netz)

Erste wirkliche Anwendungsfälle ergaben sich 1981 aus dem Transferieren von rechtlichen Dokumenten in eine Datenbank und deren automatisierter Katalogisierung sowie von Urteilen und Rechtsthemen. Der Wunsch lag in der Schaffung einer Datenbasis zur Recherche in einer Zeit bevor das heutige Internet einen globalen Zugriff auf einen weltweiten Fundus nur einen Mausklick entfernt bereitstellt. Die Firma wurde später veräußert und an Xerox, der seit den Anfängen des Digitaldrucks- und Scannen bis heute einen Stellenwert als Global Player hat.

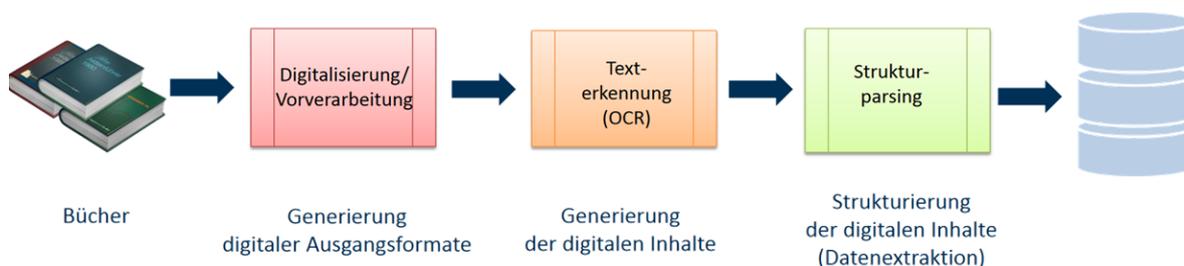


Abbildung 35; OCR Grundschemata

Quelle: <https://madoc.bib.uni-mannheim.de/52213/1/2019-09-19.pdf>

Nennenswerte Programme sind unter anderem Abby Finereader, Acrobat Pro DC aber auch die in der Software verwandte und kostenlose Open Source Software wie tesseract ocr. 1996 von HP ins Leben gerufen ging die Software 2001 in den Bestand von Google über und wird seitdem als Open Source Basisprogramm vertrieben. Dieser Quellcode ist durch Offenlegung und Möglichkeiten der Veränderung grundlegender Bestandteil der Software. Nur dadurch ist es möglich durch Code-Integration bzw. API Schnittstellen eigene Programme zu entwickeln.

7.1 Meilensteine mit Tesseract

Folgende wichtige Meilensteine hat Tesseract OCR seit Entwicklungsbeginn erreicht:

- Open Source OCR Software
- Entwicklungsbeginn 1985 von HP
- 2005 als Open Source Projekt veröffentlicht

- 2006 von Google übernommen
- 2016 Grundlegende Umstellung von Pattern Matching auf neuronale Netzwerke

7.2 Funktion

Herz und Kernstück jeder OCR sind Datenbanken. Dort enthalten sind Millionen möglicher Abbildungen eines jeden verfügbaren Buchstabens, einer Zahl oder chinesischer Zeichen. Diese Bilder werden auf Basis der Pixel analysiert und mit bestehenden elektronischen Abbildungen der Datenbank gematcht.

Anhand eines Beispiels lässt sich dies greifbar machen.

Die Software erkennt anhand der Vektoren und Pixeln und der hinterlegten Datenbasis ob es sich grundsätzlich und mit welcher Wahrscheinlichkeit um dieses Zeichen handelt und gibt diesen Wert oder Übersetzung auf Grundlage der Datenbasis weiter. Die Adaption erfolgt dann durch Feedback und somit stetiger Verbesserung der Algorithmen.

Dies ist für den Großteil der Anwendungsfälle z.B. in der Wirtschaftssektor nur eine Grundlage. Heutzutage als Basis verfügbare Computer Vision Algorithmen sind in der Lage Rechnungsnummern, Beträge, Wörter, oder Mitarbeiternamen zu kennen und beispielsweise Akten nach Öffnen der analogen Post bereits zielgerichtet innerhalb der Unternehmensstruktur zu verteilen und beispielsweise für die Buchhaltung mit allen relevanten Daten anzureichern. Computer Vision erfasst die Elemente eines Bildes, Video, die Text enthalten, die zuvor in seiner Bibliothek definiert wurden und den vordefinierten Mustern gemäß einem Algorithmus entsprechen, der dieses vordefinierte Element anhand vieler historischer Beispiele gelernt hat.

Spannend wird es wenn man mit Unterstützung der künstlichen Intelligenz über die Extraktion von Ontologien und die semantische Analyse hinaus denkt. So kann es vorkommen, dass ein Zeichen für Chinesisch nicht verstanden wird, jedoch aufgrund der Ähnlichkeit der Morpheme kein generell falsches Zeichen vom User übermittelt wurde, sondern ein in der Art und Güte sehr ähnliches Zeichen. Bisweilen konnte OCR dies nicht eigenständig mit anderen Definitionen und Werten in Verbindung bringen, da die Begriffe und Wörter keine Information mit Bedeutung haben und die Verknüpfung über Metadaten und Hintergrundinformation noch nicht implementiert ist

Um beim Beispiel der Liebe zu bleiben. Liebe hat keine Bedeutung für die Maschine. Es ist lediglich eine Zeichenfolge aus Nullen und Einsen. Um festzumachen ob das Zeichen z.B. Liebe und ein anderes Chinesisches Zeichen verwandt sind benötigt es einen Algorithmus der Wort Embedding (Wort Einbettung besitzt)

```

-6.1392349e-01, -1.6051197e+00, -6.3689017e-01, 3.8095194e-01,
2.1319206e+00, 9.6120477e-01, -6.2474030e-01, -5.7019120e-01,
-1.8183403e+00, 9.6917850e-01, -1.5792913e+00, 2.1291933e+00,
1.8838526e+00, -1.0935460e-01, -2.0293846e+00, 8.2380170e-01,
-7.2767907e-01, 1.7820674e+00, -1.3902879e+00, -8.4493607e-01,
-1.6358774e+00, 2.3685048e+00, -1.0215342e-01, 1.2315027e+00,
1.1138500e+00, 6.2029552e-01, 2.9188380e-01, -1.6836153e+00,
-8.3009660e-01, 4.4462427e-01, -9.7930026e-01, -1.0372717e+00,
3.7992764e-01, -2.6100000e-02, -1.8703303e+00, 8.8636756e-01,
1.9373804e+00, 1.3514622e+00, -1.7750920e+00, -1.2479116e+00,
9.4832259e-01, -1.0964856e+00, 1.3280897e+00, -1.2219310e+00,
-2.5767584e+00, 9.0413433e-01, 1.1110367e-01, -1.8491225e+00,
2.9016769e-01, 8.2326263e-02, -4.5731178e-01, 7.7771544e-01,
-1.5007248e+00, 3.7626073e-01, 2.7165787e+00, 2.1836729e+00,
5.1056012e-04, -9.2552716e-01, -6.9675356e-01, -1.8852206e+00,

```

Abbildung 36: Tesseract Parameter, Sicht aus dem Programm auf ein Zeichen

Mit Word Embedding wird ein Wort als Vektor abgebildet und dieser Vektor wird in einen n-dimensionalen Raum eingestellt. Anhand des obigen Beispiels sieht man die Definition und Matrix für das Wort. Die Maschine sucht jetzt für die bestimmte Sprache nach Distanzen zwischen diesem Vektor und anderen Vektoren für vorab in der Datenbank hinterlegte ähnliche Wörter. Dadurch ist essentiell das man jedes Wort in der gewählten Sprache bereits in Wort-Vektoren konvertiert hat. Hierzu kann man auf Datenbasen wie GloVe, BERT oder FastText zurückgreifen, die diese Abbildungen für ganze Sprachen bereits anwendungsbereit zur Verfügung stellen.

Dadurch kann dann das Programm und die Maschine prüfen, ob die Distanz des Wortes „Liebe“ zu anderen Wörtern größer oder geringer ist. Sie ist gering zu den Vektorabbildungen von Wörtern wie „Romantik“, „Herz“, „Küssen“, „Glück“ und ähnlichem. Man könnte eine neue Verständnisebene implementieren und auf dieser Basis Künstliche Intelligenz mit diesem reichhaltigen Informationskontext entwickeln und mittels User Aktion oder Cloud Referencing erlernbar machen. Dadurch können wir einer Maschine tatsächlich etwas beibringen (also Maschinelles Lernen), denn die Maschine entwickelt dann auf Datenbasis etwas das man als Verständnis bezeichnen könnte. Dennoch hängt die Wahrscheinlichkeit weiterhin grundsätzlich am Feedback der Nutzenenden und der Qualität, Ausprägung und den Anknüpfungspunkten der Datenbasis.

7.3 Herausforderungen mit Tesseract

Chinesisch als Zeichensprache und das Programm im Besonderen bieten für Tesseract eine große Herausforderung, die sich in der Konzeption Tesseracts als Programm begründet, da es für die Erkennung maschineller Zeichen und Texte konzipiert ist. Dies sind schriftartbedingt zwar generell mit Varianz versehen, aber dennoch durchgehend absolut regelmäßig und finden sich in stets geraden Zeilen mit konstantem Zeilenabstand. Deshalb führt eine Unregelmäßigkeit, welche bei handschriftlichen Schriftzeichen insbesondere in einem VR Umfeld unweigerlich auftritt zu Abweichungen in der Erkennung der Tesseract Engine und somit zu Herausforderungen in der Qualität.

8 Methodik

8.1 Auswahl der Methodik

Um die Forschungsfrage zu untersuchen und zu beantworten, wurde im Rahmen der Bachelorarbeit das Programm „Chinese Character Game“ entwickelt. Dieses Lernprogramm ist gerichtet an Anfänger oder Interessierte an der chinesischen Schriftsprache, um ihnen einen Einblick zu ermöglichen, wie Chinesisch als tonale und kontextbasierte Zeichensprache funktioniert. Die Schlüsselfunktionen hierbei sind das Eintauchen in eine VR Umgebung, wo Lernende in einem virtuellen Klassenraum chinesische Zeichen schreiben lernen und die Auswertung der Richtigkeit über OCR. Es folgt eine Begründung der Auswahl für die technischen Elemente, dies betrifft Unity als Entwicklungsumgebung, Oculus als VR HMD und Tesseract 4.0 als OCR Programm.

Als Forschungsmethodik wird das Experiment ausgewählt. Die Entscheidung wird begründet mit dem Vorteil aus beiden Methodiken. In dieser qualitative Forschungsuntersuchung wird das Nutzerverhalten von Experten im Umgang mit VR sowohl mit Laien.

8.2 Begründung der Auswahl technischer Elemente

8.2.1 Oculus/Meta Quest 2

Die Eigenschaften der Oculus/Meta Quest 2 wurden bereits in Abschnitt 5.6.3. erläutert.

Die Auswahl der Quest 2 zur Realisierung von „Chinesise Writing Game“ liegt primär an den Fakt, dass es im Standalone Betrieb genutzt werden kann, was dem markerlosen Inside-Out Tracking zugrunde liegt. Dadurch ergibt sich eine größere Bewegungsfreiheit der AnwenderInnen bei der Programmnutzung. Die Massentauglichkeit des HMDs durch den vergleichsweisen günstigen Preis der

Quest 2 stellt sich als Chance dar und soll die Einsteigerfreundlichkeit von VR abbilden. Dazu besteht die Möglichkeit der Nutzung von Fingertracking zum Zeichenschreiben, das letztendlich aus Zeiteffizienzgründen nicht in das Programm integriert wurde.

8.2.2 Unity

Die Unity Game Engine ist eine cross-plattform Game Engine, die bislang für Privatnutzende frei verfügbar ist. Seit der Veröffentlichung der Engine seit 2005 erscheinen regelmäßige Updates mit neuen Entwicklerfunktionen. Da die Plattform sowohl im Gamingbereich als auch im non-Gaming bzw. serious Gamingbereich Anwendungen findet wie z.B. für Visualisierungen und Simulationen weit verbreitet ist, ist eine große Anzahl von YouTube-Tutorials vorhanden, die für Einsteiger aber auch Experten geeignet ist.

Durch die komponentenbasierte Eigenschaft der GameObjects, die grundlegende, erweiterbare Elemente der Engine sind, kann der Zugriff und das Verhalten von einzelnen Bausteinen in Unity wie 3D Objekte angesteuert werden. Ebenso sind Licht, Audio oder selbst Physikelemente einfach im Editor oder durch Skripte zur Laufzeit verwaltbar und manipulierbar. Unity verwendet C# als objektorientierte Programmiersprache, die interoperabel mit anderen Programmiersprachen ist.

Ursprünglich für die Videospielkreation auf Desktopcomputern gedacht, unterstützt Unity mittlerweile plattformübergreifende Entwicklung, darunter Mobilplattformen (iOS, Android), Konsolenplattformen (z.B. PlayStation 4, 5) oder auch Virtual Reality Plattformen wie Oculus oder Steam VR.

Als Open Source Entwicklungsumgebung mit extensiven Bibliotheken ist Unity unter EntwicklerInnen gern gesehen.

In dieser Bachelorarbeit wurde mit der Unity Version 2021.3.15f1 gearbeitet.

Zusätzlich wird die Entwicklung mit Oculus VR HMDs in Unity wie in Abschnitt 8.2.2 genannt unterstützt. Dies wird durch das XR Interaction Toolkit realisiert, was ein komponentenbasiertes Interaktionssystem für VR Anwendungen ist. Es nimmt plattformübergreifende Controllerinputs an, womit Quest 2 Controller ebenfalls unterstützt werden. Einige andere Funktionen der Interaktion mit einer virtuellen Umgebung werden durch das XR Interaction Toolkit realisiert:

- grundlegende Interaktionen mit Objekten wie auswählen und greifen
- grundlegende Canvas UI Interaktion mit Controllern
- Haptisches Feedback durch VR Controller
- Visuelles Feedback, das Interaktionen der Controller indiziert
- Bewegungen der Controller und des HMDs in den sechs Freiheitsgraden mittels „XR Origin“ ein Kamera Rig für stationäres und Roomscale VR

Der Kern dieses zusätzlich integrierbaren Packages liegt allerdings bei der Verwaltung von Interaktoren, interagierbaren Komponenten und einem Interaktionsmanager, womit z.B. eine Priorisierung der interagierbaren Komponenten erfolgen kann.

Für das Programm wurde die XR Interaction Toolkit Version 2.4.3 genutzt.

Die in der Software beigebrachten Zeichen bestehen derzeit aus dem chinesischen Sprachlevel HSK 1. HSK steht für „Hanyu Shuiping Kaoshi“ und bedeutet „Prüfung des Chinesisch-Niveaus“, es bezeichnet das Sprachlevel der mündlichen und schriftlichen Kenntnisse in Mandarin-Chinesisch. HSK-Prüfungen sind standardisierte Tests für Menschen, die Mandarin-Chinesisch als Fremdsprache lernen und misst ihren Leistungsstand, vergleichbar wie TOEFL für die englische Sprache. Derzeit umfasst HSK drei Niveaus mit insgesamt 9 Stufen: Elementarstufe (HSK 1-3), Mittelstufe (HSK 4-6) und Oberstufe (HSK 7-9). Die Niveaus unterscheiden sich in den Anforderungen für 5 Sprachfertigkeiten: Hören – Sprechen – Lesen – Schreiben – Übersetzen. Inhaltliche geht es um die kommunikative Sprachfähigkeit, Inhalte zu chinesischer Kultur, chinesischem Alltag und den quantitativen Sprachindex, welcher den aktiven und passiven Schriftzeichenwortschatz in Anzahl der gelernten Zeichen bestimmt und voraussetzt.

Die HSK 1 Vokabelliste besteht aus 150 Wörtern und prüft die Anwendung von Chinesisch im Alltag. Das Sprachniveau ist vergleichbar wie A1 in anderen Sprachlevels. Prüfungsabsolventen der HSK 1 Prüfung können sehr einfache chinesische Sätze bilden und grundlegende Bedürfnisse zur Kommunikation äußern wie einfacher Smalltalk, der Wunsch nach Essen und Trinken oder das Benennen von Alltagsgegenständen und -personen.

Im Rahmen dessen soll das Programm „Chinese Writing Game“ als Einblick in Zeichen der Stufe HSK 1 dienen oder kann auch als Nachbereitung eines Sprachkurses dienen, um die Zeichen nochmal zu verinnerlichen. Der Fokus liegt auf das Lehren von einzelnen Zeichen, die für sich stehen und dessen Bedeutung in der Alltagssprache liegt, verglichen zu Wörtern, wo man sich mehrere Zeichenkombinationen merken muss.

8.2.3 Integration von Tesseract 4.0 als OCR

Im Folgenden werden die inhaltlich wichtigsten Quellcodeabschnitte gezeigt, welche maßgeblich für das grundlegende Tesseract API Setup und die Texterkennung erforderlich, erklärungsbedürftig und als wichtig betrachtet werden. Es wird darauf hingewiesen, dass nachfolgend dargestellte Quellcodes Codeauschnitte oder Zusammenfassungen zur grundlegenden Erklärung der Tesseract OCR Erkennung sind, daher kann sich der Inhalt leicht vom tatsächlichen Quellcode abweichen.

8.2.4 Grundlegendes Tesseract API Setup

Zuerst wird eine Klasse „TesseractWrapper“ erstellt, was als eine API Schnittstelle zwischen der Anwendung und der Tesseract DLL dient. Um auf die Funktionen von Tesseract zuzugreifen, werden sie durch die Methode „dllImport“ importiert.

Dann gibt es eine Klasse „TesseractDriver“ die als eine zusätzliche Schicht über dem Wrapper dient. Es soll die Tesseract-Umgebung einrichten und ist verantwortlich für den Eintrittspunkt des Users.

Mit der Methode „Init()“ wird eine Instanz von Tesseract initiiert, genauer gesagt werden zwei Funktionen gerufen „TessBaseApiCreate()“ and „TessBaseAPIInit3()“

```
[DllImport(TesseractDllName)]
private static extern IntPtr TessBaseAPICreate();

[DllImport(TesseractDllName)]
private static extern int TessBaseAPIInit3(IntPtr handle, string
dataPath, string language);

IntPtr tessHandle;

public TesseractWrapper()
{
    tessHandle = IntPtr.Zero;
}

public bool Init(string lang, string dataPath)
{
    try
    {
        tessHandle = TessBaseAPICreate();

        int init = TessBaseAPIInit3(tessHandle, dataPath, lang);
        if (init != 0)
        {
            Debug.LogError("Tess Init Failed");
            return false;
        }

        return true;
    }
    catch (Exception e)
    {
        Debug.LogError(e);
        return false;
    }
}
```

Tesseract benötigt trainierte Sprachdatendatei, die sogenannte „tessdata“.

Diese sind klassischerweise nach dem LSTM Model oder CNN trainiert. Im offiziellen Tesseract Github Repository werden diese zur Verfügung gestellt. Es wird zwischen den Typen „fast, best, normal“ unterschieden, welche jeweils Schnelligkeit gegen Genauigkeit abwägen.

Der Parameter „dataPath“ in der Methode „TessBaseAPIInit3()“ stellt den Datenpfad „tessdata“ her, während „lang“ den Namen der tessdata verlangt. Diese werden vom offiziellen Github Repository von Tesseract zur Verfügung gestellt. Im Fall von „Chinese Writing Game“ wurden die Trainingsdaten „chi_sim“ genommen, was für „simplified chinese“, spricht „vereinfachtes Chinesisch“ steht.

Da in TesseractDriver die Tesseract-Umgebung durch die Init() Methode initiiert wird, wird dort die Auswahl der Sprachtrainingsdaten gefordert.

```
public void Setup()
{
    _tesseract = new TesseractWrapper();
    string datapath = Application.streamingAssetsPath
        + "/tessdata/";

    if (_tesseract.Init("eng", datapath))
    {
        Debug.Log("Init Successful");
    }
}
```

Texterkennung

Für den Zweck der Texterkennung werden zuerst weitere API Funktionen importiert.

```
[DllImport(TesseractDllName)]
private static extern void TessBaseAPIsetImage(IntPtr handle, IntPtr
    imagedata, int width, int height,
    int bytes_per_pixel, int bytes_per_line);

[DllImport(TesseractDllName)]
private static extern void TessBaseAPIsetImage2(IntPtr handle,
    IntPtr pix);

[DllImport(TesseractDllName)]
private static extern int TessBaseAPIRecognize(IntPtr handle, IntPtr
    monitor);

[DllImport(TesseractDllName)]
private static extern IntPtr TessBaseAPIGetUTF8Text(IntPtr handle);

[DllImport(TesseractDllName)]
private static extern void TessDeleteText(IntPtr text);

[DllImport(TesseractDllName)]
private static extern void TessBaseAPIClear(IntPtr handle);
```

Es wird eine Methode „Recognize()“ geschrieben, die ein „Texture2D“ Bildobjekt annimmt sowie deren Eigenschaften wie Größe und Farbe.

```

public string Recognize(Texture2D texture)
{
    if (_tessHandle.Equals(IntPtr.Zero))
        return null;

    int width = texture.width;
    int height = texture.height;
    Color32[] colors = texture.GetPixels32();
    int count = width * height;
    int bytesPerPixel = 4;
    byte[] dataBytes = new byte[count * bytesPerPixel];

```

Nun wird dieses zu erkennende Bild aufgebaut durch ein Bytestream, das von der Textur produziert wird. Dann wird dieses Byte-Array in einen Speicher geschrieben und einen Zeiger auf den Speicher als Parameter an die TessBaseSetImage Methode von Tesseract übergeben.

```

int bytePtr = 0;
for (int y = height - 1; y >= 0; y--)
{
    for (int x = 0; x < width; x++)
    {
        int colorIdx = y * width + x;
        dataBytes[bytePtr++] = colors[colorIdx].r;
        dataBytes[bytePtr++] = colors[colorIdx].g;
        dataBytes[bytePtr++] = colors[colorIdx].b;
        dataBytes[bytePtr++] = colors[colorIdx].a;
    }
}

IntPtr imagePtr = Marshal.AllocHGlobal(count * bytesPerPixel);
Marshal.Copy(dataBytes, 0, imagePtr, count * bytesPerPixel);

TessBaseAPISetImage(_tessHandle, imagePtr, width, height,
    bytesPerPixel, width * bytesPerPixel);

```

Tesseract wird nach dem zu erkennenden Bild gefragt und anschließend durch (TessBaseAPIRecognize) und anschließend den zu erkennenden Text (TessBaseAPIGetUTF8Text). Dieser Text wird zunächst als Pointer wiedergegeben, der durch die Marshal.PtrToStringAnsi Methode in einen String umgewandelt werden kann. Anschließend kann der API Schnittstelle von Tesseract bereinigt werden (TessBaseAPIClear) und und Pointer „str_ptr“ gelöscht werden durch (TessDeleteText). Dass nur noch die Rückgabe des recognizedText-Objektes erfolgt.

```

    if (TessBaseAPIRecognize(_tessHandle, IntPtr.Zero) != 0)
    {
        Marshal.FreeHGlobal(imagePtr);
        return null;
    }

    IntPtr str_ptr = TessBaseAPIGetUTF8Text(_tessHandle);
    Marshal.FreeHGlobal(imagePtr);
    if (str_ptr.Equals(IntPtr.Zero))
        return null;
#if UNITY_EDITOR_WIN || UNITY_STANDALONE_WIN
    string recognizedText = Marshal.PtrToStringAnsi (str_ptr);
#else
    string recognizedText = Marshal.PtrToStringAuto(str_ptr);
#endif

    TessBaseAPIClear(_tessHandle);
    TessDeleteText(str_ptr);

    return recognizedText;
}

```

Diese im Wrapper geschriebene Funktionalität wird nun im TesseractDemoScript.cs aufgegriffen, indem eine Methode Recognize() ein Texture2D Bildobjekt „imageToRecognize“ annimmt.

```

[SerializeField] private Texture2D imageToRecognize;

private void Recognize()
{
    display.text += "\n" +
        _tesseractDriver.Recognize(imageToRecognize);
}

```

In TesseractDemoScript wird eine Objektinstanz von der Klasse TesseractDriver aufgerufen und auf die Recognize() Methode von TesseractDriver zugegriffen,

```

private void Start()
{
    _tesseractDriver = new TesseractDriver();
    display.text = _tesseractDriver.CheckTessVersion();
    _tesseractDriver.Setup();
    Recognize();
    display.text += "\n" + _tesseractDriver.GetErrorMessage();
}

```

während diese letztendendes auf TesseractWrapper.Recognize() kaskadiert.

```
public string Recognize(Texture2D imageToRecognize)
{
    return _tesseract.Recognize(imageToRecognize);
}
```

8.2.5 Darstellung des Textes

In der Methode „SaveTexture()“ ein Screenshot der „DrawCanvas“ erstellt, die RenderTexture dieses Screenshots wird dann auf eine Texture2D Textur gespeichert, welches durch die EncodeToPNG() zunächst in ein Byte Array gespeichert wird. Nachdem der Speicherort festgelegt wurde wird das Bild benannt als „SavedScreenshot.png“. „System.IO.File.WriteAllBytes()“ schreibt dann das Bytearray in die Datei und schließt sie. Falls der Dateiname vorhanden sein sollte wird es überschrieben.

```
void SaveTexture()
{
    characterExplainText.text = "";

    screenshotCam.enabled = true;

    RenderTexture rt = new RenderTexture(2048, 2048, 24,
RenderTextureFormat.ARGB32, RenderTextureReadWrite.sRGB);
    RenderTexture oldRT = screenshotCam.targetTexture;
    screenshotCam.targetTexture = rt;
    screenshotCam.Render();
    screenshotCam.targetTexture = oldRT;

    //RenderTexture rt = Selection.activeObject as RenderTexture;
    //RenderTexture rt = GameObject.Find("Draw
Camera").GetComponent<Camera>();
    RenderTexture.active = rt;

    Texture2D tex = new Texture2D(rt.width, rt.height,
TextureFormat.ARGB32, false);
    //Texture2D texture = new Texture2D(rt.width, rt.height,
TextureFormat.ARGB32, false);
    tex.ReadPixels(new Rect(0, 0, rt.width, rt.height), 0, 0);

    byte[] bytes = tex.EncodeToPNG();
    var dirPath = Application.dataPath + "/RenderOutput";
    if (!System.IO.Directory.Exists(dirPath))
```

```

        {
            System.IO.Directory.CreateDirectory(dirPath);
        }
        System.IO.File.WriteAllBytes(dirPath + "/SavedScreenshot.png",
bytes);
        Debug.Log(bytes.Length / 1024 + "Kb was saved as: " + dirPath);

        //statt @"C:\Users\FTZDR-Omen\TestVR - Kopie
2\Assets\RenderOutput\SavedScreenshot.png";
        //--> Application.dataPath + "/RenderOutput";
        string fileName = @"C:\Users\FTZDR-Omen\TestVR - Kopie
2\Assets\RenderOutput\SavedScreenshot.png";

```

Zuletzt wird auf die Datei aus dem Pfad zugegriffen und gelesen und an der Tafel dargestellt. Dabei folgt die Texterkennung durch die Recognize() Methode.

```
//Zuständig fürs lesen und darstellen an der großen UI Wand
```

```

var rawData = System.IO.File.ReadAllBytes(fileName);
tex.LoadImage(rawData);
imageToRecognize = tex;
tex.SetPixels32(imageToRecognize.GetPixels32());
tex.Apply();
Recognize(tex);
RenderTexture.active = null;
screenshotCam.enabled = false;

```

In Recognize() wird zudem die Methode „AddToTextDisplay()“ ausgeführt, welche für die Darstellung des erkannten Textes auf einer Textobjekt „displayText und das gezeichnete Bild in Form einer Texture2D „outputTexture“ zeigt, welches in der Unity Szene dargestellt wird. So wird das visuelle Feedback – welches Zeichen erkannt wurde und wie der gezeichnet Input aussieht - auf der Tafel sichtbar.

```

private void Recoginze(Texture2D outputTexture)
{
    ClearTextDisplay();
    AddToTextDisplay(_tesseractDriver.CheckTessVersion());
    _tesseractDriver.Setup();
    AddToTextDisplay(_tesseractDriver.Recognize(outputTexture));
    AddToTextDisplay(_tesseractDriver.GetErrorMessage(), true);
}

```

Anschließend folgen die Fälle für eine erfolgreichen Erkennung des Zeichens oder eine unerfolreichen Erkennung.

```

private void AddToTextDisplay(string text, bool isError = false)
{
    if (string.IsNullOrEmpty(text)) return;

    _text += (string.IsNullOrEmpty(displayText.text) ? "" : ") +
text;

    if (isError)
        Debug.LogError(text);
    else
        Debug.Log(text);

    //character recognition
    //□ KOU
    if (_characterVideoScript.videoPlayer.clip.name.Equals("kou_cut"))
    {
        if (text.Contains("A") || text.Contains("□"))
        {
            StopAllCoroutines();
            SetNextCharBool(true);
            Debug.Log("Erkanntes Zeichen " + "\n" + text);

            confirmRecognitionText.text = "Great! You've correctly drawn
the character!";
            characterExplainText.text = text + "means: \nmouth";

            StartCoroutine(PlayCoroutineCharacterKOURecognized());

            IncreaseCorrectCounter();

```

```

        buttonNextCharacter.gameObject.SetActive(true);
    }
    else
    {
        Debug.Log(text);
        StopAllCoroutines();

        StartCoroutine(PlayCoroutineCharacterNotRecognized());
        confirmRecognitionText.text = "Sorry, the character was drawn
or recognized incorrectly. Please try again.";
    }
}

```

Zwischendurch wurde eine Funktionalität eingebaut, wo erkannte Zeichen auf dem Output mit einer Kontur umrandet werden. Da dies nicht maßgeblich für den Zeichenerkennungsprozess ist, wird dies nicht erläutert.

8.2.6 CharacterVideoScript.cs

In dieser Klasse befinden sich alle Funktionalitäten, die mit der Darstellung des zu zeichnenden Schriftzeichens und den UI Elementen des Videos über dem linken Controllers, sowie der Ausführung der zu tun haben.

Das interagierbare Video der Zeichen über dem linken Controller ist eine Videodatei, die für die Nachvollziehbarkeit der einzelnen Zeichenstrichfolgen steuerbar gemacht wird, indem dieses Video nach Vollendung eines Zeichenstriches angehalten wird. Dafür werden für alle Zeichen Listen angelegt, in denen die jeweiligen passenden Frame-Momenten hinterlegt sind, an denen das Video stoppen soll. Zudem sind alle Zeichen in einem VideoClip Array gespeichert.

Während ShowCharacter() ist die Methode ist, die die passenden Videoclips lädt. Es wird zum Laden des jeweiligen Levels genutzt oder wenn ein nächstes Zeichen mit der Methode UpdateShowCharacter() angezeigt wird.

Am wichtigsten sind hier die Methoden ProcessStrokeOrder, die das Video anhalten, wenn ein Zeichenstrich fertig dargestellt wurde und pausiert wird UpdateVideoPlayerToFrame(), in der das angehaltene Video wieder abgespielt werden kann. Diese Methode wird bei den UI Buttons „previous stroke“ und „next Stroke“ ausgeführt.

8.2.7 SoundMananagerScript.cs

Alle Menu, UI und gesprochene Wörter und chinesisches und englisch sind hier als AudioClip verwiesen sowie deren Getter-Methoden, welche dann in TesseractDemoScript.cs und CharacterVideoScript.cs aufgerufen werden.

8.3 Vorstellung der Software aus Nutzersicht

Nach dem Start der Software erlangt der Nutzer zunächst das Startmenü der „TitleScene“, in dem man auch die Quellenverweise zu z.B. genutzten Github Repositories und YouTube Tutorials einsehen kann.



Abbildung 37: Startbild Chinese Writing Game

Quelle: Eigenes Bild

Das Bestätigen des Start Knopfes führt den Nutzer zu der neuen Szene „ChWritingTesseractMainScene“, einer Umgebung, die einen Klassenraum einer Schule darstellen soll. Dem Nutzer wird ein Level Menü vorgezeigt und eine Auswahl zwischen den Levels gegeben:

- Tutorial
- Level 1, 2 und 3
- und das letzte Level „Test your skills!“



Abbildung 38: Virtueller Classroom aus dem Chinese Writing Game

Quelle: Eigenes Bild

Alle im Level vorgestellte Zeichen sind Teil des Sprachlevels HSK 1. Die Auswahl dieser Zeichen wurden nochmal in verschiedene Levels unterteilt basierend auf eigene Einschätzung der Schwierigkeit der Anzahl der Zeichenstriche und Komplexität, da auch simple Zeichen in Alltagssprache schon viele Zeichenstriche haben können, was einen Einsteiger überfordern könnte. Das Tutorial soll dem Nutzer nicht nur einsteigerfreundliche Zeichen lehren, die nur zwei bis drei Striche beinhalten, sondern auch einen Einstieg in das Programm geben und ihm vertraut mit der Bedienung der Software machen. Daher sind im Tutorial-Level nur 3 Zeichen gefordert, um das Level zu bestehen, statt wie in allen anderen Levels mit 5 Zeichen. In den Level 1-3 werden die Zeichen hinsichtlich der Anzahl Striche und Komplexität der Struktur herausfordernder, allerdings wird bei jedem Level erst einfach gestartet und von Zeichen zu Zeichen arbeitet der Nutzer sich in der Schwierigkeit auf, sodass das letzte Zeichen eines Levels das am schwierigsten zu schreibendes ist. In jedem Level hat der Nutzer die Möglichkeit es abubrechen und ins Level Menü zurückzukehren. Das letzte Level „Test your skills!“ hat ebenfalls eine Auswahl von 5 Zeichen soll eine Prüfung der in den 3 Levels davor vorgekommene Zeichen sein. Ziel dieses Levels ist es den Nutzer die Verinnerlichung der beigebrachten Zeichen abzufragen. Zusätzlich bilden diese 5 Zeichen einen Satz, der einen persönlichen Bezugspunkt zum Nutzer herstellen soll und die Anwendung der Zeichen und Wörter in einem Satz beibringen soll. Daher wird dieses Level erst freigeschaltet, nachdem die Level 1 bis 3 mindestens angeschaut wurden.

Nachdem der Nutzer ein Level ausgewählt hat erscheint direkt vor dem Nutzer eine quadratische, weiße Fläche, die in alle Himmelsrichtungen und Diagonale durch gestrichelte schwarze Linien unterteilt ist. Diese Fläche stellt die Malwand dar, auf der der Nutzer die Schriftzeichen schreiben kann. Die fünf Zeichen, die pro Level zu zeichnen sind, werden nacheinander vorne auf der auf der schwarzen Tafel und über dem linken VR-Controller in Form eines Videos dargestellt. Es stellt beides das gleiche Video bzw. das gleiche Zeichen dar. Damit soll der Nutzer nicht immer wieder auf dem Controller schauen müssen, sondern kann sich das Zeichen auch größer auf der Tafel erkennen. Das Video ist aber erstmal angehalten, da alle chinesische Zeichen einer bestimmten Schriftzeichenstrichreihenfolge unterliegen.

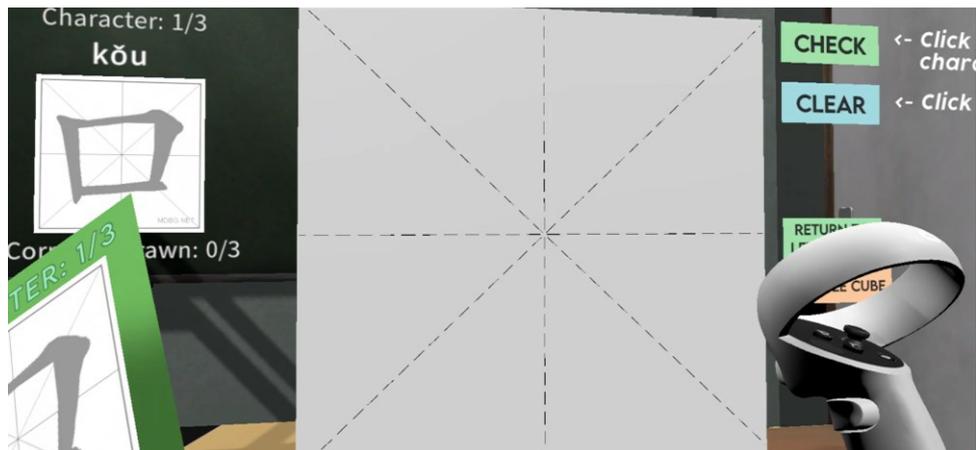


Abbildung 39: Blick auf die Malwand

Quelle: Eigenes Bild

Der Nutzer soll den rechten Controller die Trigger-Knopf (Zeigefingertaste) nutzen, um beim linken Controller die Buttons „show next stroke“ und „show previous stroke“ den nächsten oder vorherigen Zeichenstrich anzuzeigen und das Video fortsetzen bzw. von einem vorherigen Zeichenstrich wieder zu beginnen. So kann er jederzeit die Reihenfolge der Zeichenstriche in seinem eigenen Tempo verfolgen und Strich für Strich das Schriftzeichen zeichnen. Nachdem alle Striche eines Zeichens angezeigt wurden, wechselt der Button „show next stroke“ auf „show all strokes“. Das Klicken dieses Buttons stellt das Video der Schriftzeichenstrichreihenfolge fließend ohne Unterbrechung dar für eine eventuelle bessere Nachvollziehbarkeit. Gleichzeitig wechselt der Button „show previous stroke“ auf „show single strokes“, womit das Video der Schriftzeichenstrichreihenfolge wieder unterbrochen wird an der Stelle, wo der letzte Strich angezeigt wurde. Jetzt können alle Zeichenstriche wieder einzeln angezeigt werden.

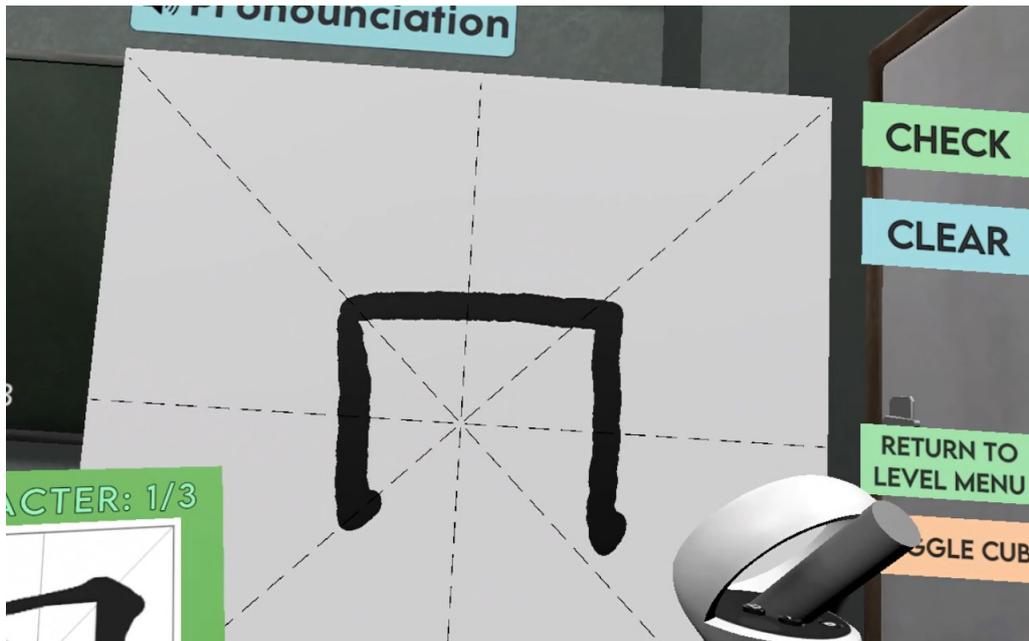


Abbildung 40: Zeichnungen auf der Malwand

Quelle: Eigenes Bild

Der Nutzer kann zu seiner rechten Seite einen Stift mit dem Grab bzw. Grip-Knopf greifen, der bei Oculus Controllern beim Mittelfinger liegt und diesen zur Malwand führen. Sobald die Stiftspitze die Malwand berührt, hinterlässt es einen Strich. So kann der Nutzer zeichnen. Das Loslassen des Stiftes lässt diesen in der Luft verharren, sodass sich der Nutzer Zeit nehmen kann um den nächsten Zeichenstrich anzeigen lassen kann. Verharrt der Stift nicht in seiner Ursprungsposition, wird dieser nach 7.5 Sekunden dorthin teleportiert. Der Nutzer kann durch das Radiergummi nicht erwünschte Striche wegradieren oder durch Klicken des „Clear“-Buttons rechts neben der Malwand die ganze Malwand zurücksetzen

Ist das Zeichen fertig gezeichnet oder sobald erwünscht, kann der Nutzer durch Klicken des „Check“-Buttons neben der Malwand das Zeichen auf Korrektheit prüfen lassen. Erkennt das System das gezeichnete Zeichen, gibt es eine positive auditive Rückmeldung und auf der Tafel wird das korrekte Zeichen ausgegeben. Danach wird zusätzlich die Bedeutung des Zeichens erklärt, die chinesische Silbe mündlich ausgesprochen. Auch wird das Zeichen auf Englisch mündlich und schriftlich übersetzt. Der Nutzer erhält einen von fünf möglichen Richtigkeitspunkten (drei beim Tutorial-Level) für die erfolgreiche Erkennung und kann nun ins nächste Zeichen übergehen, indem der Button „next character“ über dem linken Controller geklickt wird.



Abbildung 41: Auswertung eines korrekt gezeichneten Schriftzeichens

Quelle: Eigenes Bild

Sollte das System das Zeichen nicht erkennen, gibt es eine negative auditive Rückmeldung und der Nutzer darf sich erneut versuchen. Falls der Nutzer das Zeichen überspringen möchte, kann das über dem linken Controller der Button „skip character“ erfolgen. Im Gegensatz zu einer erfolgreichen Erkennung des gezeichneten Zeichens wird bekommt der Nutzer beim Überspringen allerdings keinen Richtigkeitspunkt. Jedoch geht sowohl auf der Tafel und beim linken Controller der Fortschritt weiter, sodass der Nutzer sehen kann wie viele Zeichen noch zu zeichnen sind.

Nach dem fünften Zeichen kann der Nutzer der Level beenden, indem entweder rechts neben der Malwand der Button „Finish“ geklickt wird oder das zu zeichnende Zeichen übersprungen wird. Es wird angezeigt wie viele Zeichen korrekt erkannt wurden und abhängig davon gibt es eine andere auditive Ausgabe. Nun kann der Nutzer zum Level Menu zurückgehen um ein weiteres Level beginnen oder zum Main Menu zurückkehren.

8.3.1 Experiment mit Versuchspersonen

Um die praktische Wirksamkeit und Lerneffektivität des Programms „Chinese Writing Game“ und ebenfalls die in Abschnitt 1 (Einleitung) genannte Forschungsfrage zu untersuchen, wurde ein Versuch mit Versuchspersonen durchgeführt.

Hierbei gab es zwei Versuchsgruppen: Versuchspersonen mit Erfahrungen in Virtual Reality, die schon mal ein HMD benutzt haben, regelmäßig eins nutzen oder sogar VR-EntwicklerInnen sind. Diese Gruppe wird die Expertengruppe genannt. Die zweite Versuchspersonengruppe hat keine Erfahrungen im Umgang mit VR, diese Gruppe wird die Laiengruppe genannt. Sie soll die Menschen repräsentieren, an die das Computersprachlernprogramm gerichtet ist. Die Expertengruppe wurde nochmal in zwei weitere Gruppen aufgeteilt. Eine, bei der eine ausführliche mündliche Einweisung in die Nutzung des Programms mit Vorzeigen der Interaktionsmöglichkeiten stattgefunden hat. Und die andere, bei der keine Einweisung stattfand, sodass sich diese Gruppe nur auf die Benutzeranleitung verlassen musste, der beim Start in die Klassenraum-Szene angezeigt wird.

8.3.1.1 Versuchsablauf:

Beide Versuchspersonengruppen wurde die Möglichkeit gegeben „Chinese Writing Game“ mit einer Versuchszeit von 10min auszuprobieren. Es wurde deren Erfolg gemessen in Anzahl geschaffter Zeichen und Anzahl durchschnittlicher Versuche, bis ein Zeichen geschafft wurde. Alle Versuchspersonen begannen beim Tutorial-Level. Bei Fragen zur Nutzung wurde geholfen. Anschließend gab es eine Feedbackrunde, wo den Versuchspersonen Fragen zum Nutzungsverhalten und zum beigebrachten Inhalt gestellt wurden.

8.3.1.2 Versuchsbeobachtungen:

Bei der Expertengruppe gab es keine Gruppe oder Versuchsperson, die mit der Handhabung des Programms Schwierigkeiten hatten, die Nutzung der Controller war für sie einleuchtend und nachvollziehbar. Die UI Interaktionen zur Charakterzeichenstrichfolge und der Malwand war für sie intuitiv. Allerdings erreichten viele es nicht mehr als über das zweite Zeichen im Tutorial zu kommen. Häufiger Grund war eine falsche Erkennung vom System, obwohl das Zeichen sehr ähnlich zum angezeigten Video gezeichnet wurde. Es wurde durchschnittlich drei bis vier Male versucht das Zeichen zu wiederholen, bis es übersprungen wurde.

Bei der Laiengruppe gab es wenige, aber nicht keine Versuchspersonen, die bei der Handhabung des Programms Schwierigkeiten hatten. Die UI Interaktionen und Controllersteuerung wurden ebenfalls

relativ schnell gemeistert. Ausschließlich eine Person hat die Knöpfe für „Trigger“ (Zeigefinger) und „Grab“ (Mittelfinger) verwechselt. Auch hier kamen die Versuchspersonen nicht über das Tutorial, diesmal wurde sogar kaum beim ersten Zeichen des Tutorials gescheitert. Ein weiterer Frustrationsfaktor war, dass die Striche nicht so schnell gezeichnet werden konnten, da sonst statt einer Linie viele kleine Punkte erscheinen würden. Trotz vieler scheiternden Versuche blieben sie bemüht, bis sie aufgegeben hatten. Bemerkenswert war, dass viele viel Zeit investierten die Zeichen nicht nur nachzuzeichnen, sondern auch im Detail auch nachzumalen. Die Einweisungshilfe empfanden sie als unterstützend.

8.3.1.3 Versuchsergebnisse:

Als Positives und Chancen wurde von der Expertengruppe angemerkt, dass das Programm eine spielerische motivierende Möglichkeit sei, um Chinesisch beizubringen. Die Controllersteuerung sei einfach zu nutzen. Allerdings sei dieser Lernweg aufwändiger und fehleranfälliger. Es fehlte eine Art Fortschrittsanmerkung zwischendurch und Feedback zur Richtigkeit des Zeichens, da die fehlerhafte Erkennung einem nachvollziehbar erschien. Dazu sei die Tafel schlecht erkennbar, sodass das erlernte Zeichen nicht erkennbar war und der auditive Hinweis der Bedeutung unzureichend war. Eine Person merkte, die an die Stifte am Anfang kurz übersehen zu haben.

Die Laiengruppe teilte ebenfalls, dass man nach einer kurzen Zeit ein gutes Gefühl für die Controller bekäme und die Klassenraumumgebung sich sehr immersiv anfühle. Trotz der Frustration des Nichterfolges wurde oft ein Ansporn durch Ehrgeiz verspürt das Zeichen weiter zu malen. Hier wurde ebenfalls der auditive Hinweis bei einem erfolgreichen Zeichen nicht wahrgenommen und der Tafelhinweis nicht gesehen, da die Tafel größtenteils von der Malwand verdeckt wird.

8.3.1.4 Interpretation der Ergebnisse

Obwohl die Bedienung des Programms für beide Versuchsgruppen zugänglich war, konnten sie keinen großen Lernfortschritt bei der Nutzung erreichen. Dies wurde an verschiedenen Parametern bemerkt, die folgendermaßen erklärt werden können.

Die fehlerhafte Erkennung der gemalten Zeichen kann sehr wahrscheinlich daran liegen, dass Tesseract als OCR eher nicht geeignet ist für die Analyse von Handschrift oder bestimmten Schriftarten wie . Dies liegt an qualitativ unzureichenden Image Processing Algorithmen, die selbst bei Tesseract 4.0 noch auftreten. Ein anderer Grund ist die besonders kalligrafische Natur von chinesischer Handschrift. Selbst in der Regelschrift finden sich noch große Unterschiede durch Serifen der Schreibrift, welche ebenfalls im Vorzeigevideo in „Chinese Writing Game“ gefunden werden.

Die Verzögerung beim Zeichnen auf der Malwand, dass besonders beim schnellen Zeichnen Punkte statt Linien entstehen könnte auf eine Latenz hinweisen. Allerdings ist bisher unklar, woher die Latenz

entsteht. Naheliegender wäre die Berechnung vieler graphischer Elemente wie die custom shader und render Texturen der Malwand als auch die Recognize() Methoden in den Tesseract-Klassen, welche eher berechnungsintensiv sind, was an einer kurzen Ladeverzögerung des Programms beim Drücken von „Check“ für die Auswertung des Bildes merkt.

Sowohl die Frustration als auch der Ehrgeiz das Zeichen erfolgreich zu schaffen können mit den Reizen des behaviouristischen CALLs zusammenhängen sich nach Lob und Anerkennung sehnen. Um diesen Reiz nicht zu stark in Versuchspersonen hervorzurufen, wurde im Sinne des integrativen CALLs eine Person mit Aufgabe des Mentors neben den Versuchspersonen gesetzt, um sie zu leiten und begleitend zum computer-assistiertem Sprachlernen sie mit weiterem Wissen und Fakten über den Lehrinhalt zu informieren.

8.3.1.5 Einschätzung eines Zeitrahmens bei Softwarenutzung des Chinese Writing Game

Das Erlernen von 2500 chinesischen Schriftzeichen erfordert viel Zeit und Engagement. Je nach Lerngeschwindigkeit kann es mehrere Jahre dauern. Es ist wichtig, konsistent zu bleiben und täglich zu üben. Mit VR-gestütztem Lernen könnte es jedoch schneller Fortschritte geben, da die Motivation und die interaktiven Elemente die Lernenden ermutigen, sich intensiver mit der Sprache auseinanderzusetzen.

Dennoch ist es schwer vorherzusagen, wie lange es dauern wird, bis jemand 99,5% einer chinesischen Zeitung verstehen kann. Dies hängt von vielen Faktoren ab, einschließlich des Ausgangsniveaus, der Lernmethode und des persönlichen Engagements. Es ist ein anspruchsvolles Ziel, das Geduld und Zeit erfordert, aber mit der Unterstützung von VR und einem gut geführten Lerntagebuch ist es definitiv erreichbar

8.4 Vorstellung der Software

Die Lernsoftware „Chinese Writing Game“ soll eine Lernsoftware sein, die gerichtet ist an Anfänger oder Interessierte an der chinesischen Schriftsprache, um ihnen einen Einblick zu ermöglichen wie Chinesisch als kontextbasierte Zeichensprache funktioniert.

Lernziel: Zeichen nachzeichnen können, Striche erkennen und Anwenden, Bedeutung der einzelnen Zeichen einzuordnen

Die in der Software beigebrachten Zeichen bestehen derzeit aus dem chinesischen Sprachlevel HSK 1. HSK steht für „Hanyu Shuiping Kaoshi“ und bedeutet „Prüfung des Chinesisch-Niveaus“, es bezeichnet das Sprachlevel der mündlichen und schriftlichen Kenntnisse in Mandarin-Chinesisch. HSK-Prüfungen sind standardisierte Tests für Menschen, die Mandarin-Chinesisch als Fremdsprache lernen und misst ihren Leistungsstand, vergleichbar wie TOEFL für die englische Sprache. Derzeit umfasst HSK drei Niveaus mit insgesamt 9 Stufen: Elementarstufe (HSK 1-3), Mittelstufe (HSK 4-6) und Oberstufe (HSK 7-9). Die Niveaus unterscheiden sich in den Anforderungen für 5 Sprachfertigkeiten: Hören – Sprechen – Lesen – Schreiben – Übersetzen. Inhaltlich geht es um die kommunikative Sprachfähigkeit, Inhalte zu

chinesischer Kultur, chinesischem Alltag und den quantitativen Sprachindex, welcher den aktiven und passiven Schriftzeichenwortschatz in Anzahl der gelernten Zeichen bestimmt und voraussetzt.

Die HSK 1 Vokabelliste besteht aus 150 Wörtern und prüft die Anwendung von Chinesisch im Alltag. Das Sprachniveau ist vergleichbar wie A1 in anderen Sprachlevels. Prüfungsabsolventen der HSK 1 Prüfung können sehr einfache chinesische Sätze bilden und grundlegende Bedürfnisse zur Kommunikation äußern wie einfacher Smalltalk, der Wunsch nach Essen und Trinken oder das Benennen von Alltagsgegenständen und -personen.

Im Rahmen dessen soll das Programm „**Chinese Writing Game**“ als Einblick in Zeichen der Stufe HSK 1 dienen oder kann auch als Nachbereitung eines Sprachkurses dienen, um die Zeichen nochmal zu verinnerlichen. Der Fokus liegt auf das Lehren von einzelnen Zeichen, die für sich stehen und dessen Bedeutung in der Alltagssprache liegt, verglichen zu Wörtern, wo man sich mehrere Zeichenkombinationen merken muss.

Nach dem Start der Software erlangt der Nutzer zunächst das Startmenü der „TitleScene“, in dem man auch die Quellenverweise zu z.B. genutzten Github Repositories und YouTube Tutorials einsehen kann.

8.5 Ablauf des Programms

Das Bestätigen des Start Knopfes führt den Nutzer zu der neuen Szene „ClassroomScene“, einer Umgebung, die einen Klassenraum einer Schule darstellen soll. Dem Nutzer wird ein Level Menü vorgezeigt und eine Auswahl zwischen den Levels gegeben:

- Tutorial
- Level 1, 2 und 3
- und das letzte Level „Test your skills!“

Alle im Level vorgestellte Zeichen sind Teil des Sprachlevels HSK 1. Die Auswahl dieser Zeichen wurde nochmal in verschiedene Levels unterteilt, basierend auf der eigenen Einschätzung der Schwierigkeit der Anzahl der Zeichenstriche und Komplexität. Auch simple Zeichen in der Alltagssprache können schon viele Zeichenstriche haben können, was einen Einsteiger überfordern könnte. Das Tutorial soll dem Nutzer nicht nur einsteigerfreundliche Zeichen lehren, die nur zwei bis drei Striche beinhalten, sondern auch einen Einstieg in das Programm geben und ihn vertraut mit der Bedienung der Software machen. Daher sind im Tutorial-Level nur 3 Zeichen gefordert, um das Level zu bestehen, statt wie in allen anderen Levels 5 Zeichen. In den Levels 1-3 werden die Zeichen hinsichtlich der Anzahl Striche und Komplexität der Struktur herausfordernder. Allerdings wird bei jedem Level erst einfach gestartet und von Zeichen zu Zeichen arbeitet der Nutzer sich in der Schwierigkeit auf, sodass das letzte Zeichen eines Levels das am schwierigsten zu schreibendes ist. In jedem Level hat der Nutzer die Möglichkeit diese abubrechen und ins Level Menü zurückzukehren.

Das letzte Level „Test your skills!“ hat ebenfalls eine Auswahl von 5 Zeichen und stellt eine Prüfung der in den 3 vorangegangenen Levels erlernten Zeichen dar. Ziel dieses Levels ist es den Nutzer die Verinnerlichung der beigebrachten Zeichen abzufragen. Zusätzlich bilden diese 5 Zeichen einen Satz, der einen persönlichen Bezugspunkt zum Nutzer herstellen soll und die Anwendung der Zeichen und Wörter in einem Satz beibringen soll. Daher wird dieses Level nur freigeschaltet, nachdem die Level 1 bis 3 mindestens bearbeitet wurden.

Nachdem der Nutzer ein Level ausgewählt hat erscheint direkt vor dem Nutzer eine quadratische, weiße Fläche, die in alle Himmelsrichtungen durch gestrichelte schwarze Linien unterteilt ist. Diese Fläche stellt die Malwand dar, auf der der Nutzer die Schriftzeichen schreiben kann. Die fünf Zeichen, die pro Level zu zeichnen sind, werden nacheinander vorne auf der auf der schwarzen Tafel und über dem linken VR-Controller in Form eines Videos dargestellt. Es stellt beides das gleiche Video bzw. das gleiche Zeichen dar. Damit soll der Nutzer nicht immer wieder auf dem Controller schauen müssen, sondern kann sich das Zeichen auch größer auf der Tafel darstellen lassen. Das Video ist aber erstmal angehalten, da alle chinesische Zeichen einer bestimmten Schriftzeichenstrichreihenfolge unterliegen.

Der Nutzer soll am rechten Controller die Trigger-Knopf (Zeigefingertaste) nutzen und beim linken Controller die Buttons „show next stroke“ und „show previous stroke“ um den nächsten oder vorherigen Zeichenstrich anzuzeigen und das Video fortsetzen bzw. von einem vorherigen Zeichenstrich wieder zu beginnen. So kann er jederzeit die Reihenfolge der Zeichenstriche in seinem eigenen Tempo verfolgen und Strich für Strich das Schriftzeichen zeichnen. Nachdem alle Striche eines Zeichens angezeigt wurden, wechselt der Button „show next stroke“ auf „show all strokes“. Das Klicken dieses Buttons stellt das Video der Schriftzeichenstrichreihenfolge fließend ohne Unterbrechung dar für eine eventuelle bessere Nachvollziehbarkeit. Gleichzeitig wechselt der Button „show previous stroke“ auf „show single strokes“, womit das Video der Schriftzeichenstrichreihenfolge wieder unterbrochen wird an der Stelle, wo der letzte Strich angezeigt wurde. Jetzt können alle Zeichenstriche wieder einzeln angezeigt werden.

Der Nutzer kann zu seiner rechten Seite einen Stift mit dem Grab bzw. Grip-Knopf greifen, der bei Oculus Controllern beim Mittelfinger liegt und diesen zur Malwand führen. Sobald die Stiftspitze die Malwand berührt, hinterlässt es einen Strich. So kann der Nutzer zeichnen. Das Loslassen des Stiftes lässt diesen in der Luft verharren, sodass sich der Nutzer Zeit nehmen kann um den nächsten Zeichenstrich anzeigen lassen kann. Verharrt der Stift nicht in seiner Ursprungsposition, wird dieser nach 7.5 Sekunden dorthin teleportiert. Der Nutzer kann durch das Radiergummi nicht erwünschte Striche wegradieren oder durch Klicken des „Clear“-Buttons rechts neben der Malwand die ganze Malwand zurücksetzen.

Ist das Zeichen fertig gezeichnet oder vom User ein Check erwünscht, kann der Nutzer durch Klicken des „Check“-Buttons neben der Malwand das Zeichen auf Korrektheit prüfen lassen. Erkennt das System das gezeichnete Zeichen, gibt es eine positive auditive Rückmeldung und auf der Tafel wird das

korrekte Zeichen ausgegeben. Danach wird zusätzlich die Bedeutung des Zeichens erklärt und in Englisch übersetzt angezeigt. Der Nutzer erhält einen von fünf möglichen Punkt für Richtigkeit je nach Erfolg der Erkennung und kann nun ins nächste Zeichen übergehen, indem der Button „next character“ über dem linken Controller geklickt wird.

Sollte das System das Zeichen nicht erkennen, gibt es eine negative auditive Rückmeldung und der Nutzer darf sich erneut versuchen. Falls der Nutzer das Zeichen überspringen möchte, kann das über dem linken Controller der Button „skip character“ erfolgen. Im Gegensatz zu einer erfolgreichen Erkennung des gezeichneten Zeichens wird bekommt der Nutzer beim Überspringen allerdings keinen Richtigkeitspunkt. Jedoch geht sowohl auf der Tafel und beim linken Controller der Fortschritt weiter, sodass der Nutzer sehen kann wie viele Zeichen noch zu zeichnen sind.

Nach dem fünften Zeichen kann der Nutzer der Level beenden, indem entweder rechts neben der Malwand der Button „Finish“ geklickt wird oder das zu zeichnende Zeichen übersprungen wird. Es wird angezeigt wie viele Zeichen korrekt erkannt wurden und abhängig davon gibt es eine andere auditive Ausgabe. Nun kann der Nutzer zum Level Menu zurückgehen um ein weiteres Level beginnen oder zum Main Menu zurückkehren.

9 Fazit

Der Umfang der Bachelorarbeit beschränkte sich ursprünglich ausschließlich auf die Entwicklung von „Chinese Writing Game“, Grundlagen zu VR und auf die Untersuchung an Testpersonen. Durch eigenen Ehrgeiz entstand Anreiz eine wissenschaftlich fundierte und umfassende Arbeit zu verfassen.

Die Ergänzung der Themen CALL und Chinesisch haben die Bachelorarbeit zwar fachlich angereichert, aber den Bearbeitungszeitraum strapaziert.

Allerdings konnten alle Aspekte, die mit der Forschungsfrage zusammenhängen dadurch gut beleuchtet werden und besonders der Versuch konnte einen großen Teil zur Beantwortung der Forschungsfrage beitragen.

Der Einsatz von modernen Medien in CALL ist heutzutage unabdingbar. Durch den technischen Wandel und den schnellen Fortschritt erscheinen multimediale Medien in der Lehre als attraktiv gegenüber traditionellen Lehrmethoden und helfen besonders Lernende motiviert zu bleiben, da der soziale Aspekt von Kommunikation miteinander ebenso die Sozialkompetenz und den kulturellen Austausch stärkt. Dies ist insbesondere bei der Sprachlehre vom Mandarin-Chinesisch vom großen Vorteil, da durch den Unterschied gegenüber anderen nicht logographischen Sprachen das Erlernen sich als herausfordernd stellt. Dies wird weiter erschwert durch den Nachteil, dass das alleinige Lernen von einzelnen Komponenten wie z.B. das Lernen von Silben bereits kompliziert, aber immer noch nicht zielführend zum Zweitspracherwerb ist. Sondern alle Komponenten der Schrift (Schreibrichtung, Schriftfolge),

Sprache (Tonalität, Silben) und Bedeutung (Homophonie/-graphie) müssen parallel erlernt werden. Dafür kann ein kultureller Austausch mit Experten oder Einheimischen helfen, um soziokulturellen Bezug und zur Sprache zu erhalten und motiviert zu bleiben. Schließlich kann OCR eine gute Möglichkeit bieten um bei logographischen Sprachen den Selbstspracherwerb zu unterstützen, hierzu sind aber ausreichend Erkennungsalgorithmen nötig, die Handschrift verlässlich erkennen können als Feedbackmöglichkeit im Rahmen vom modernen, computerunterstützten Lernen, um durch Frustration beim Lernen nicht die Motivation zu nehmen. Hierzu wäre eine detailliertere und häufigere Feedbackmöglichkeit im Programm hilfreich, um Lernenden nicht erst bei der Auswertung, sondern bereits beim falschen Strichzeichnen zu warnen oder die Stelle zu markieren, welche falsch gezeichnet wurde. Dies benötigt allerdings eine Echtzeitauswertung, was die Rechenleistung weiter in Anspruch nehmen würde und einen Abgleich zu Trainingsdaten, die nicht nur aus Druckschrift, sondern aus Bildschrift bestehen. Darüber hinaus wird Computertechnologie selbst mit dem starken Wandel in Richtung Automatisierung durch neuronale Netzwerke heutzutage die Rolle einer Lehrenden Person im Bildungswesen nie ersetzen, da gerade in einem interaktiven Bereich wie Sprachlehre das Reagieren auf die einzelnen Bedürfnisse von Lernenden individuell zugeschnittene pädagogische Hintergründe hat. Die Kombinatorik von künstlicher Intelligenz und virtueller Realität ist zwar unterstützend zum Zweitspracherwerb von Mandarin-Chinesisch, jedoch reicht es nicht aus, da bislang der soziale Aspekt und insbesondere die Kommunikation in der chinesischen Sprache essenziell für den Lernerfolg sind.

Literaturverzeichnis

1. Abrams, J. (2005). *The Art of Virtual Reality*. Publisher.
2. Berkenbusch, E. (1997). *Praktisches Chinesisch 1. Übungsbuch der chinesischen Schriftzeichen*. China Book Trading GmbH.
3. Berkenbusch, E. (1997). *Praktisches Chinesisch 1. Übungsbuch der chinesischen Schriftzeichen*. China Book Trading GmbH.
4. Bowman, D. A., Kruijff, E., LaViola, J. J., & Poupyrev, I. (2004). *3d-user interfaces: theory and practice*. Addison Wesley Longman Publishing Co., Inc.
5. Brooks, F. P. (1999). What's real about virtual reality? In: *IEEE Comput Graph Appl* 19 (6): 16–27.
6. Cruz-Neira, C. (1993). *Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE*.
https://www.researchgate.net/publication/220721398_Surround-Screen_Projection-Based_Virtual_Reality_The_Design_and_Implementation_of_the_CAVE
7. Ellis, S. R., Breant, F., Manges, B., Jacoby, R., & Adelstein, B. D. (1997). Factors influencing operator interaction with virtual objects viewed via head-mounted see-through displays: viewing conditions and rendering latency. In: *Proc IEEE Virtual Reality*, S 138–145.
8. Hummel, J., Wolff, R., Dodiya, J., Gerndt, A., & Kuhlen, T. (2012). Towards interacting with force-sensitive thin deformable virtual objects. *Joint Virtual Reality Conference of ICAT – EGVE – EuroVR, 2012*. Eurographics Association.
9. Karlgren, B. (2001). *Schrift und Sprache der Chinesen*. 2. Aufl., Berlin u. a.: Springer.
10. Krashen, S. D., & Terrell, T. D. (1983). *The natural approach: Language acquisition in the classroom*. New York: Pergamon Press.
11. Kumar, E. S., & Sreehari, P. (2009). *Computer Assisted Language Learning (CALL)*. In *A Handbook for English Language Laboratories* (pp. 3–14). Foundation Books.
12. Levy, M. (1997). *Computer-assisted language learning: Context and conceptualization*. Oxford University Press.
13. Lindner, P. (2021). Better, Virtually: the Past, Present, and Future of Virtual Reality Cognitive Behavior Therapy. *J Cogn Ther*, 14, 23–46. <https://doi.org/10.1007/s41811-020-00090-7>
14. Milgram, P., Kishino, F. (1994). A Taxonomy Of Mixed Reality Visual Displays. *Jeice Trans. Inf. Syst.*, E77-D, 1321–1329.

- https://scholar.google.com/scholar_lookup?title=A+Taxonomy+Of+Mixed+Reality+Visual+Displays&author=Milgram,+P.&author=Kishino,+F.&publication_year=1994&journal=Ieice+Trans.+Inf.+Syst.&volume=E77-D&pages=1321%E2%80%931329
15. Mohring, M., & Frohlich, B. (2011). Effective manipulation of virtual objects within arm's reach. *Proc. Virtual Reality Conference (VR)*, 131,138, IEEE. doi: 10.1109/VR.2011.5759451
 16. Polfuß, J. (20.03.2023). Wie viele chinesische Zeichen gibt es? <https://china-wiki.de/wie-viele-chinesische-zeichen-gibt-es/>
 17. Postert, P., Wolf, A. E. M., & Schiewe, J. (2022). Integrating Visualization and Interaction Tools for Enhancing Collaboration in Different Public Participation Settings. *ISPRS International Journal of Geo-Information*, 11(3), 156. <https://doi.org/10.3390/ijgi11030156>
 18. Qiu, X. (2000). *Chinese writing*. Berkeley, California.
 19. Riva, G., Baños, R. M., Botella, C., Mantovani, F., & Gaggioli, A. (2016). Transforming Experience: The Potential of Augmented Reality and Virtual Reality for Enhancing Personal and Clinical Change. *Front. Psychiatry*, 7:164. doi: 10.3389/fpsyt.2016.00164
 20. Sauer, O. (2023). Digitaler Zwilling – das Schlüsselkonzept für Industrie 4.0. <https://www.iosb.fraunhofer.de/de/geschaeftsfelder/automatisierung-digitalisierung/anwendungsfelder/digitaler-zwilling.html>
 21. Schieffelin, B. B., & Ochs, E. (1986). Language socialization. *Annual Review of Anthropology*, 15, 163-191.
 22. Schott, W. (1837). Mandarin und Bonze. In: *Magazin für die Literatur des Auslandes* 12:2, S. 492. (eingeschränkte Vorschau in der Google-Buchsuche).
 23. Schmidt, W. G. A. (1990). *Einführung in die chinesische Schrift- und Zeichenkunde*. Bushke.
 24. Simonson, M. R., & Thompson, A. (1997). *Educational computing foundations*. Merrill/Prince Hall. (k CALL, quelle von davor)
 25. Slater, M., Usoh, M., & Steed, A. (1994). Depth of presence in virtual environments. *Presence: Teleoperators and virtual Environments*, 3, 130–144.
 26. Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing Our Lives with Immersive Virtual Reality. *Frontiers in Robotics and AI*, 3, 74. doi: 10.3389/frobt.2016.00074
 27. Warschauer, M. (1996). Computer-assisted language learning: An introduction. *Multimedia language teaching*, 3-20.

28. Warschauer, M. (1997). Computer-Mediated Collaborative Learning: Theory and Practice. *Modern Language Journal*, 81, 470-481. <https://doi.org/10.1111/j.1540-4781.1997.tb05514.x>
29. Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language Teaching*, 31, 57-71.
<http://dx.doi.org/10.1017/S0261444800012970>
30. Warschauer, M., & Healey, D. (1998). Computers and language learning: an overview. *Language Teaching*, 31(2), 57–71.
<http://doi.org/10.1017/S026144480001297>
31. Wilhelm Schott (1837). Mandarin und Bonze. In: *Magazin für die Literatur des Auslandes* 12:2, S. 492. (eingeschränkte Vorschau in der Google-Buchsuche).
32. Xigui, Qiu (2000). *Chinese writing*. Berkeley, California.
33. Yamaguchi, T. (2005). *Vocabulary Learning with a Mobile Phone*.
<http://lt.msu.edu/vol10num1/emerging/default.html>, accessed Jul. 2010.
34. Yuan, Z. (2007). Problems in researching e-learning: The case of computer-assisted language learning (pp. 416-36).