

Künstliche Intelligenz in der Fremdsprachendidaktik: Sprachgefühl und Sprachmodell im Fokus

Behiye Arabacıoğlu , Eskişehir – Sema Nur Tunçyüz , Eskişehir

 <https://doi.org/10.37583/diyalog.1824338>

Abstract (Deutsch)

Angeichts des raschen und kontinuierlichen Fortschritts moderner KI-Technologien gewinnt deren Einsatz im Fremdsprachenunterricht zunehmend an Bedeutung und rückt verstärkt in den Mittelpunkt wissenschaftlicher Diskussionen. Die vorliegende Studie nutzt ChatGPT-4o als Forschungsinstrument, um die Rolle des von Friedrich Kainz entwickelten Konzepts des Sprachgefühls beim Erkennen, Verstehen und Vermitteln emotionaler Feinheiten systematisch zu untersuchen. Unter Sprachgefühl wird eine intuitive Sensibilität für sprachliche Strukturen, Ausdrucksformen und ihre emotionale Wirkung verstanden, die über rein kognitive Sprachkompetenz hinausgeht. Analysiert wurden KI-generierte Antworten auf Texte, die gezielt die Emotionen Freude, Traurigkeit und Angst zum Ausdruck brachten. Die Ergebnisse verdeutlichen, dass das Modell in der Lage ist, emotionale Nuancen sowohl auf lexikalischer und syntaktischer als auch auf prosodischer Ebene differenziert wiederzugeben. Dadurch zeigt sich, dass KI nicht nur formale Sprachstrukturen korrekt verarbeitet, sondern auch emotionale Bedeutungen erstaunlich präzise nachzeichnen kann. Dennoch bleibt Empathie eine zentrale menschliche Fähigkeit, um KI sinnvoll in authentische Kommunikationsprozesse einzubetten.

Schlüsselwörter: *Fremdsprachendidaktik, Emotionale Kapazität, Künstliche Intelligenz, Sprachgefühl, Sprachmodell.*

Abstract (English)

Artificial intelligence in foreign language teaching: Focus on language intuition and language models

In view of the rapid progress of artificial intelligence technologies, their application in the field of foreign language teaching has become increasingly significant and worthy of closer academic attention. The present study makes use of ChatGPT-4o as a research tool in order to examine the role of what Friedrich Kainz described as *language feeling* —the intuitive sensitivity to linguistic form and expression— in the comprehension and communication of emotional nuances. To this end, a series of AI-generated responses were analysed, each based on texts designed to convey three distinct emotions: joy, sadness and fear. The analysis reveals that the model is capable of reproducing emotional subtleties with remarkable precision, not only at the level of vocabulary and syntax but also in relation to tone and prosody. Despite these promising findings, the study concludes that genuine empathy remains essential for the meaningful integration of AI into authentic interpersonal communication processes.

Keywords: *Foreign Language Pedagogy, Emotional Capacity, Artificial Intelligence, Linguistic Intuition, Language Model.*

EXTENDED ABSTRACT

The rapid integration of Artificial Intelligence into educational contexts is fundamentally reshaping foreign language pedagogy. While traditional approaches emphasize grammar, vocabulary acquisition, and rule-based learning, effective language acquisition also requires *language sensibility* - an intuitive, context-sensitive competence encompassing stylistic appropriateness, pragmatic sensitivity, and emotional attunement. This study investigates the extent to which AI-powered language models, specifically ChatGPT-4o, approximate elements of language sensibility across both written and spoken language modalities.

The theoretical foundation of this research combines Friedrich Kainz's (1944) conceptualization of *language sensibility* as implicit, socially constructed knowledge with Paul Ekman's (1992) *emotion theory*, which categorizes universal emotions (joy, fear and sadness). Kainz posits that language sensibility emerges not merely from formal linguistic rules but from accumulated exposure, cultural participation and interaction within authentic communicative environments. Ekman's framework allows for an operationalized investigation of emotional recognition, offering a structured lens through which AI's affective processing capabilities may be examined.

A qualitative, exploratory design was employed to evaluate ChatGPT-4o's emotional intelligence. Nine micro-narratives were developed, carefully embedding both explicit and implicit emotional cues corresponding to Ekman's emotion categories. These narratives, situated within diverse cultural scenarios, were presented to ChatGPT-4o under two experimental conditions: written text inputs and spoken inputs. The spoken condition was designed to include prosodic incongruence, wherein vocal intonation intentionally conflicted with the semantic content, enabling assessment of the model's capacity to process prosodic cues such as pitch, rhythm and intonation — critical elements for accurate emotional interpretation in human communication.

Results from the written analysis revealed that ChatGPT-4o demonstrates substantial proficiency in semantic emotional classification. The model was able to accurately identify intended emotional states in most cases, even when emotional indicators were subtly implied rather than explicitly stated. This suggests that large language models have achieved considerable sophistication in their lexical and syntactic processing abilities, allowing them to approximate certain aspects of language sensibility within written discourse. Such findings align with emerging research indicating that AI models, trained on vast linguistic datasets, may outperform humans in some written emotion recognition tasks (Schlegel et al., 2025).

However, critical limitations were observed in the spoken analysis. When confronted with incongruence between prosodic features and semantic content, ChatGPT-4o consistently prioritized lexical information, often failing to account for prosodic markers that play a central role in human emotional comprehension. For example, narratives expressing joy but delivered with a melancholic tone were still categorized by the model as conveying joy. This deficiency underscores a fundamental limitation in current AI architectures, which remain primarily text-centric and insufficiently sensitive to paralinguistic features crucial for nuanced affective interpretation.

The pedagogical implications of these findings are multifaceted and significant. On one hand, AI models such as ChatGPT-4o offer promising avenues for supporting foreign language learning through personalized feedback, adaptive exercises, real-time conversational practice and scalable task-based learning environments. These affordances can democratize access to high-quality language resources, especially in under-resourced or remote educational contexts where traditional language instruction may be constrained by time, personnel or financial limitations.

On the other hand, the development of authentic language sensibility relies heavily on sustained, interactive exposure to complex socio-cultural dynamics, real-life communicative situations, and affective interpersonal exchanges. These dimensions of language learning involve embodied, experiential processes that current AI technologies cannot fully replicate. The absence of lived experience and genuine human interaction fundamentally constrains AI's capacity to acquire the holistic, dynamic understanding that characterizes true linguistic intuition.

As such, the integration of AI into foreign language education must be approached with careful pedagogical reflection. Educators bear responsibility not only for leveraging AI's strengths but also for guiding learners in critically engaging with its limitations. This requires developing competencies to assess the appropriateness and accuracy of AI-generated feedback, fostering metacognitive strategies that enable learners to reflect on and contextualize AI outputs, and ensuring that ethical considerations — such as data privacy, algorithmic bias, and overreliance on automation — are adequately addressed.

In conclusion, while ChatGPT-4o exhibits emerging capabilities that approximate certain dimensions of language sensibility in written contexts, significant limitations remain in its ability to process prosodic and socially situated aspects of language understanding. AI-powered language models should be integrated into foreign language education as powerful supplementary tools that enhance, but do not replace, the indispensable relational, cultural and experiential dimensions of language learning. Continued research is necessary to improve AI's sensitivity to the complex interplay of semantics, prosody and affect, while maintaining pedagogical frameworks grounded in ethical, reflective and human-centred principles.

Einleitung

Die Integration von Künstlicher Intelligenz (KI) in Bildungsrahmen revolutioniert rasch traditionelle pädagogische Modelle, speziell im Hinblick auf das Fremdsprachenlernen, welches besonders davon betroffen ist (vgl. Vistorte u.a. 2024). Daher ist es notwendig, eine gründliche Analyse von Vorteilen und Grenzen der KI durchzuführen, wie sie auf Sprachlernende angewendet werden können, insbesondere in Bezug auf den Erwerb von Sprachgefühl - einem tiefen Verständnis einer gesprochenen Sprache, das über die strikte Einhaltung von Grammatikregeln hinausgeht. Das Konzept des Sprachgefühls, wie es Friedrich Kainz interpretiert, bezieht sich nicht nur auf eine intuitive Beherrschung der Grammatik und des Wortschatzes, sondern umfasst auch eine tiefere emotionale und ästhetische Sensibilität. Für die Fremdsprachendidaktik ist dieses Verständnis von Sprachgefühl von entscheidender Bedeutung, da es über das reine Erlernen von Sprachstrukturen hinausgeht und die Entwicklung einer empathischen und kulturell sensiblen Kommunikationsfähigkeit fördert (vgl. Kainz 1944). In Anbetracht der wachsenden Allgegenwart von KI in der modernen Bildung ist eine eingehende Untersuchung ihrer vielschichtigen Auswirkungen sowohl auf den Sprachlernprozess als auch auf das Lehren von entscheidender Bedeutung. Die Fähigkeit der KI, komplexe emotionale Verhaltensmuster basierend auf Daten zu erkennen, die während der gesamten Bildungserfahrung gesammelt wurden, ermöglicht einen detaillierteren Einblick in die spezifischen Bedürfnisse jedes Lernenden (vgl. Vistorte u. a. 2024), wodurch Lehrkräfte ihre Unterrichtsmethoden auf neue und innovative Weise anpassen können (vgl. Son u.a. 2023).

Die Fähigkeit der Künstlichen Intelligenz, personalisiertes Feedback zu geben und die emotionalen Bedürfnisse der Studierenden zu erfüllen, verbessert das Lernerlebnis erheblich, insbesondere in Bereichen, die sowohl kognitive Fähigkeiten als auch emotionales Engagement erfordern, wie beim Sprachenlernen (vgl. Vistorte u.a. 2024). Außerdem kann der Einsatz von Emotionserkennungsfähigkeiten die Empathie und Inklusivität in der Lernumgebung erhöhen, das allgemeine Wohlbefinden der Studierenden fördern und eine effizientere und unterstützende Lernumgebung gewährleisten (ebd.).

Methode und Prozedur

Diese Studie verfolgt einen qualitativen und explorativen Ansatz, der darauf ausgerichtet ist, die von ChatGPT-4o gezeigte emotionale Intelligenz systematisch zu untersuchen. Im Zentrum dieser Studie steht die Fähigkeit des Modells, die emotionalen Nuancen, die der Sprache innewohnen, zu identifizieren, zu verstehen und auf diese angemessen zu reagieren.

Die theoretische Grundlage basiert auf Paul Ekmans (1992) Emotionstheorie, die grundlegende universelle Emotionen darlegt und Friedrich Kainz' philosophische Ansichten in Bezug auf die Wechselwirkung zwischen Kognition und Sprache anwendet. Sie stützt sich auch auf aktuelle Perspektiven zur Nutzung künstlicher Intelligenz im Fremdsprachenlernen und in der interkulturellen Kommunikation.

Methodisch setzte die Studie einen ganzheitlichen Ansatz ein, der klassische Szenarioanalyse mit der Präsentation multimodaler Stimuli kombinierte und somit sowohl semantische Inhalte (die Textbedeutung) als auch prosodische Eigenschaften (Sprachrhythmus, Artikulation, Intonation) in die Analyse einbezog. ChatGPTs Leistung wurde in verschiedenen Kommunikationskontexten getestet, um seine emotionale Sensibilität zu bewerten - insbesondere in den Bereichen von Freude, Angst und Trauer.

Das Ziel der Studie ist es, ChatGPTs Fähigkeit zur Erfassung emotionaler Sprachelemente zu bewerten und auch seine Wirksamkeit im Fremdsprachenunterricht zu testen. Die Forschungsergebnisse werden wertvolle Einblicke darüber liefern, inwieweit künstliche Intelligenzplattformen emotionale Untertöne und kulturell relevante Bedeutungen durch Sprache übertragen können, sowie wie sie erfolgreich im Unterricht eingesetzt werden können.

Szenarienauswahl und Emotionsklassifikation

In seiner Emotionskategorisierung von 1992 nannte Paul Ekman drei grundlegende emotionale Zustände: Freude, Angst und Trauer. Um diese Zustände zu veranschaulichen, wurden drei kurze, alltägliche Situationen als kurze Mikro-Narrative entworfen und neun verschiedene Texte erstellt. Diese Narrative enthalten sowohl typische sprachliche Signale, die für jede Emotion charakteristisch sind, als auch subtilere implizite Signale. Somit spiegeln sie eine Bandbreite kulturell heterogener kommunikativer Umgebungen wider, einschließlich des Alltagslebens, familiärer Umgebungen und des öffentlichen Raums.

Um die interne Konsistenz zu gewährleisten, wurden die in den Narrativen verwendeten Prompts auf eine Länge von 50 bis 55 Wörtern begrenzt. Erkenntnisse bestätigen die Empfehlung, dass eine kurze Prompt-Länge die Überzeugungskraft von KI-Sprachmodellen verbessert (Levy u.a. 2024). Diese neun Texte sind wie folgt:

A1 (Angst): Mitten in der Nacht hörte man mehrere hastige Schritte. Ein Kind krümmte sich vor Schmerz. Als es seine Jacke anzog, stach ihn etwas. Das Kind sagte, dass ihm jetzt die Brust wehtue. Es könnte ein Skorpion gewesen sein. Seine Lippen waren blau angelaufen. Niemand wusste, wo das nächste Krankenhaus war.

A2 (Angst): Zum dritten Mal, dass ich diesen Traum habe. Ich spazierte ahnungslos auf Bergen. Sehe ein Haus mit warmen Lichtern. Als ich in diese Richtung gehe, falle ich in ein dunkles Loch. Ich kann mich nicht bewegen, schreie immer so laut wie ich kann und wache oft durch meine eigene Stimme auf.

A3 (Angst): Heute Abend ist es seltsam. Etwas stimmt nicht. Sie hat eine ganz andere Sprechweise. Sie ist so ernst, als hätten wir uns nie gekannt. Plötzlich fängt sie an, ein heiteres Lied zu singen. Ihr Gesicht jedoch verrät Wut. Ich kann das nicht mehr ertragen, weiß aber nicht, was ich noch tun soll.

F1 (Freude): Heute Morgen in meinem Briefkasten: Ein Brief vom Amt, Werbungen, eine bunte Broschüre... Dazwischen eine Postkarte mit einer herrlichen Landschaft, die mir bekannt vorkommt. Das Schreiben auf der Rückseite fängt mit meinem Namen an und endet mit Grüßen von einem alten Freund. Ich erinnere mich nicht, wann meine Wangen zuletzt so weh taten.

F2 (Freude): Haben diese Füße nicht neulich erst angefangen zu laufen? Sie treten gerade auf die Bühne. Mein Herz schlägt im gleichen Rhythmus wie damals. Der Glanz in den Augen trifft mich. Ich fühle Gänsehaut, sogar auf meinem Gesicht - war so etwas möglich? Die Zeit muss schnell vergangen sein. Zum Glück ist sie es auch.

F3 (Freude): Wir sind eine große Familie. Das merkt man bei jedem Zusammensein, wenn die Stühle nicht mehr ausreichen und die Stimmen in Harmonie summen. Nach dem Essen wird bestimmt etwas gespielt – oder vielleicht gesungen. Die Kleinste von uns hat ja gelernt, Gitarre zu spielen. Dann wird alles unvergesslich, bis zum nächsten Zusammenkommen.

T1 (Trauer): Sie steigen in den Zug. Ich reiche den letzten Koffer hin. Die Kälte des Morgens und das Gewicht des Eisens fühle ich in meinem Magen. Ich stehe aufrecht und still wie das Ausfahrtsignal. Hebe die Hand, während das Signal blau wird. Kommt dieses Pfeifen aus dem Zug oder von mir?

T2 (Trauer): Ich hatte so lange gekämpft, um endlich hier zu stehen, an der Grenze zu meinem Traum. Doch als sie meinen Pass verlangen, greife ich ins Leere. Ich habe keine andere Wahl, als mit diesen dicken Koffern und Hoffnungen zurückzukehren. Mein Kopf senkt sich und meine Augen blicken nicht mehr auf.

T3 (Trauer): Ein winziger Vogel fällt aus einem blühenden Baum auf die grüne Wiese, wo die Sonne strahlt. Er liegt still, fast schlafend. Als ich mich näherte, versuchte er, mit nur einem Flügel zu fliegen. Die Feder an seinem verletzten Hals begleiten seine schweigende Stimme und seinen Herzschlag.

Schriftliche Analyse mit ChatGPT

Die neun Texte wurden dem Sprachmodell in ihrer textuellen Form über die ChatGPT-Schnittstelle präsentiert. Jedes Dokument wurde in drei verschiedenen Varianten angezeigt, wobei unterschiedliche Intervalle und Sequenzen verwendet wurden, um die Zuverlässigkeit und Konsistenz der emotionalen Kennzeichnung testen zu können. Der verwendete Standardprompt lautete: „*Bitte analysiere die emotionale Stimmung dieses Textes. Handelt es sich um Freude, Angst oder Trauer?*“ Die vom Modell generierten Ergebnisse wurden sorgfältig klassifiziert und basierend auf der Bedeutung der identifizierten Emotionen in den gegebenen Texten bewertet.

Gesprochene Analyse mittels Voice-Chat-Funktion

Nachdem der schriftliche Test durchgeführt wurde, erfolgte die Darbietung derselben Texte in auditiver Form unter Verwendung von ChatGPTs Sprachfunktion. Die Texte wurden nacheinander von der jeweiligen Forscherin vorgelesen, woraufhin der entsprechende Prompt formuliert wurde: „*Welche emotionale Stimmung beschreibt dieser Text? Freude, Angst oder Trauer?*“

Die vorbereiteten Texte wurden drei monatelang monatlich getestet, sodass die abweichenden Rückmeldungen bestimmt und die Bemerkungen je nach aktualisiertem Modell interpretiert werden konnten.

Das Hauptziel war festzustellen, ob das Modell auditive prosodische Merkmale wie Intonation, Tempo oder Tonhöhe ordnungsgemäß identifizieren und in den

Emotionserkennungsprozess einbeziehen konnte. Die Sprachanalyse wurde unter Verwendung der Spracherkennungsfeatures von ChatGPT durchgeführt.

Artikulationsverfremdung (Inkongruenztest)

Um die Reaktion des Modells auf inkongruente Stimuli zu bewerten, wurde das identifizierte Segment „Freude“ bewusst in einem traurigen Ton präsentiert. Diese Veränderung fungierte als eine Erkundung der Verfremdung, die darauf abzielte, die Fähigkeit des Modells zu testen, zwischen affektivem semantischem Inhalt und prosodischem Ton zu unterscheiden oder inkongruente Indikatoren zu erkennen.

Sprachmodelle und Fremdsprachendidaktik

Im Bereich des Fremdsprachenlernens bieten künstliche intelligenz-gestützte Anwendungen und Sprachmodelle neue Möglichkeiten, die Unterrichtspraktiken erheblich verändern. Insbesondere nehmen Sprachmodelle, die von groß angelegten maschinellen Lernsystemen unterstützt werden, eine zentralere Rolle ein. Werkzeuge wie ChatGPT ermöglichen die Nachahmung realistischer Gesprächsinteraktionen und bieten auch Echtzeit und kontextabhängiges Feedback, was ein Unterrichtsvorteil ist, der jedoch in traditionellen Klassenzimmern oft durch Zeit- oder Personaleinschränkungen begrenzt ist.

Die Fähigkeit, eine Vielzahl von Texttypen über verschiedene sprachliche Register hinweg zu generieren, macht solche Modelle zu starken Kandidaten, um zur Verbesserung der Lesefähigkeiten, Schreibfähigkeiten und analytischen Fähigkeiten der Lernenden beizutragen (vgl. Başaran 2023). Insbesondere im Kontext des DaF-Unterrichts können solche Modelle zur Verbesserung der Schreibfähigkeiten eingesetzt werden, indem sie personalisiertes Feedback bieten und automatische Fehlererkennung ermöglichen (vgl. Hartmann 2021). Ein solcher Ansatz erleichtert nicht nur den Schreibübungsprozess, sondern verbessert auch die Entwicklung strategischer Selbstkorrekturfertigkeiten und markiert eine vielversprechende Entwicklung in pädagogischen Methoden (vgl. Hartmann 2024).

Aus pädagogischer Sicht bieten Sprachmodelle neue Möglichkeiten für aufgabenbasiertes Lernen, indem sie interaktive Umgebungen bereitstellen, in denen Lernende sinnvoll mit authentischen Konstrukten der Zielsprache interagieren können (vgl. Başaran 2024 / Zhou / Divekar 2025). Als sprachliche Unterstützungsgeräte bieten sie auch verschiedene lexikalische und syntaktische Wahlmöglichkeiten - Merkmale, die andernfalls das Eingreifen einer fließend sprechenden Person oder einer Lehrkraft erfordern würden.

Diese Werkzeuge können auch eine vermittelnde Rolle spielen, indem sie den Zugang zu komplexen Texten ermöglichen und die Textproduktion in dynamischen Lernumgebungen erleichtern (vgl. Hartmann 2024). Auf diese Weise tragen sie dazu bei, den Zugang zu qualitativ hochwertigeren Sprachressourcen zu demokratisieren und neue

pädagogische Perspektiven zu bieten - insbesondere in vielfältigen oder unterversorgten Umgebungen.

Gleichzeitig besteht die Notwendigkeit einer analytischen und informierten Herangehensweise an diese Technologien. Einige der von künstlicher Intelligenz formulierten Ergebnisse stellen auch Fälle kultureller Unsensibilität, idiomatischer Schwäche oder mangelnder Kreativität dar. Am bemerkenswertesten ist, dass diese Systeme Bedeutungen nicht flexibel interpretieren können, das jedoch für kommunikative Kompetenz von entscheidender Bedeutung ist. Zusätzlich ist es wichtig, ethische Bedenken zu berücksichtigen, beispielsweise Bedenken bezüglich des Datenschutzes, der Über-Automatisierung der Bildungspraxis und der Gefährdung des menschlichen Aspekts des Lernprozesses.

Eine Schlüsselstrategie ist die Entwicklung formaler Richtlinien bezüglich KI-Kompetenz, die im Kontext assistiver Technologien erforscht wurden (vgl. Schulz / Schmid-Meier 2024). Formale Richtlinien bieten eine Grundlage für die systematische und ethische Integration von Sprachmodellen in Lernräume. Ihre erfolgreiche Anwendung zur Unterstützung maßgeschneiderter Lernpfade, zur Steigerung digitaler Kompetenz und zur Förderung inklusiver pädagogischer Ansätze hängt von ihrer sorgfältigen Integration mit pädagogischem Wissen und kritischer Analyse ab.

Auswertung

Die generierten Antworten wurden anhand der folgenden Kriterien bewertet:

- Die emotionale Angemessenheit der Klassifizierung,
- Sprachliche Sensibilität im Gebrauch von Ausdrucksmitteln,
- Empathisches Resonanzphänomen in der Antwort,
- Interkulturelle Konsistenz der Argumentation.

Die qualitative Inhaltsanalyse wurde in Übereinstimmung mit dem Mayring (2015) Rahmenwerk durchgeführt. Fünf Experten für Fremdsprachenunterricht bewerteten die Ergebnisse unabhängig voneinander unter Verwendung einer fünfpunktigen Likert-Skala. Bei unterschiedlichen Bewertungen wurde ein konsensbildender Prozess befolgt.

Sprachgefühl als dynamische Kompetenz im Kontext der Sprachentwicklung und KI-gestützten Didaktik

Das Konzept Sprachgefühl umfasst das angeborene oder entwickelte Gefühl einer Person dafür, was in einer bestimmten Sprache als „richtig“ gilt. Das Konzept geht über die einfache Beachtung grammatischer Regeln hinaus und zeigt sich in der Angemessenheit des Stils, der pragmatischen Adäquatheit und einer emotionalen Abstimmung auf den konventionellen Rhythmus und Ton, der für die Sprache spezifisch ist. Das Konzept wurde erstmals 1768 von Johann G. Herder eingeführt, wie von David Romand und Michel Le Du (2023: 76) nach Klein zitiert; die ursprüngliche Zuschreibung an Campe

bleibt jedoch unter Gelehrten umstritten. Dennoch wird Friedrich Kainz weitgehend als der Gelehrte gewürdigt, der das Konzept des Sprachgefühls umfassend und systematisch entwickelt hat (ebd.: 7). Ihm zufolge ist es nicht als instinktives Verhalten oder psychologischer emotionaler Prozess zu verstehen, sondern als eine unklare, unmittelbare Wahrnehmung, der wesentliche emotionale und persönliche Aspekte fehlen (Kainz 1944: 6). Diese Art von Wissen ähnelt der Automatisierung, folgt festgelegten Konventionen – vergleichbar mit Tradition – und funktioniert auf dieselbe Weise wie eine Empfindung (ebd.). Darüber hinaus bemerkt Kainz auch, dass Sprachgefühl nicht strikt auf Muttersprachler allein beschränkt ist; auch Nicht-Muttersprachler können es erreichen.

Die Vorstellung von Sprachgefühl sollte daher nicht als stabile Fähigkeit, sondern als sich entwickelnde und dynamische Kompetenz verstanden werden, die von kumulativer sprachlicher Erfahrung beeinflusst wird. Dies wird von zahlreichen empirischen und theoretischen Positionen unterstützt. Die Dynamische Systemtheorie (DST) versteht Spracherwerb als einen selbstorganisierenden Prozess, bei dem sprachliche Fähigkeiten voneinander abhängig sind und sich durch anhaltende Interaktion entwickeln (vgl. Petinou u.a. 2021 / Sowers-Wills 2021). Im Kontext mehrsprachiger Fähigkeiten kann man sehen, dass sprachliche Kompetenzen je nach Verwendungsmustern und Belastungsschwellen schwanken. Rezeptive Fertigkeiten, wie Hörverständnis, können durch wiederholte Exposition verbessert werden, während produktive Fertigkeiten, wie Sprechen oder Schreiben, gezielte Unterstützung erfordern (vgl. Alshihry 2024). Dieser Prozess zeigt die erhebliche Variabilität in Sprachprofilen auf, die sich im Laufe der Zeit als Ergebnis von Interaktionen, situativen Kontexten und bewusstem Sprachgebrauch ändern können (vgl. Petinou u.a. 2021; Alshihry 2024).

Die Vernetzung zwischen verschiedenen sprachlichen Teilfertigkeiten unterstützt die inhärente Fluidität, die das Konzept des Sprachgefühls kennzeichnet. Studien legen nahe, dass die Leistung in einem Bereich oft ein Vorläufer für Fortschritte in anderen Bereichen ist (vgl. Petinou u.a. 2021). Das Verstehen von Sprache ist ein äußerst dynamischer Prozess, der schnelle, kontextgesteuerte und oft automatische Verarbeitungsstadien umfasst ein entscheidendes Element im Sprachgefühl, das produktive Kommunikation auf ihrer Kernebene ermöglicht (vgl. Grodniewicz 2020 / Trujillo u.a. 2023).

Unter dieser Perspektive kann künstliche Intelligenz - insbesondere wie sie in adaptiven und generativen Sprachmodellen realisiert wird - eine wichtige Rolle bei der Entwicklung des Sprachgefühls spielen. KI-Werkzeuge bieten eine explorative Interaktion mit Sprachmustern in sicheren und ansprechenden Umgebungen und ermöglichen ein tiefes Verständnis durch vielfältige sprachliche Stimulation. Da diese Programme zusätzlich emotionssensitive Eigenschaften beinhalten, sind sie gut geeignet, die Gefühle der Studierenden wie Zweifel, Aufregung oder Verwirrung zu adressieren - Faktoren, die intrinsisch mit der Entwicklung des Sprachgefühls verbunden sind.

Es ist hier entscheidend zu betonen, dass Sprachgefühl tief in soziokulturelle und emotionale Kontexte verwurzelt ist - Aspekte, die künstliche Intelligenz nur teilweise erfassen kann. Intrinsisch für die Entwicklung der Nuancen sprachlicher Intuition sind

zwischenmenschliche Interaktion und kulturelle Teilhabe mit kontextgesteuertem Feedback. Obwohl KI den Lernprozess ermöglichen und beschleunigen kann, kann sie ihn nicht vollständig ersetzen. Die Entwicklung von Sprachgefühl erfordert erfahrungsbasiertes Lernen, das über den bloßen Erwerb formaler Sprachkompetenz hinausgeht; es umfasst eine aktive Dynamik von Wahrnehmung, Praxis und kultureller Internalisierung.

Die Fähigkeit von ChatGPT, emotionale Feinheiten in der Sprache zu erkennen

Die Fähigkeit, emotionale Feinheiten innerhalb sprachlicher Strukturen zu erkennen, stellt eine kritische Herausforderung für künstliche Intelligenz-Modelle dar. ChatGPT hat bereits bedeutende Fortschritte in diesem Bereich gemacht; es zeigt jedoch bestimmte markante Lücken. Eine empirische Untersuchung von Schlegel u. a. (2025) zeigt, dass ChatGPT-4o zusammen mit anderen führenden großen Sprachmodellen (LLMs) wie GPT-4 und GPT-3.5 durchschnittlich 81% bei Tests zur emotionalen Intelligenz erreicht und damit den menschlichen Grundwert von 56% erheblich übertrifft. Interessanterweise zeigt ChatGPT-4o spezifische Stärken beim Erkennen ausgeprägter negativer Emotionen, wie Toxizität oder Wohlbefinden, was auf sein spezifisches Sicherheitstraining zurückgeführt werden kann (vgl. Amin u.a. 2024). Es weist jedoch noch immer ernsthafte Lücken beim Identifizieren impliziter und subtiler Emotionen auf, wie Sarkasmus, Subjektivität oder emotionales Engagement. Um diese Fähigkeiten zu verbessern, ist eine weitere Abstimmung der Modellarchitektur und Trainingsdaten notwendig.

Analyseergebnisse zu schriftlichem Prompt

Das 4o-Modell von ChatGPT verfügt über die Fähigkeit, relevante Emotionen zu erkennen, die aus den subtilen Unterschieden in der geschriebenen Sprache abgeleitet werden. Die folgenden Beobachtungen sind in dieser Hinsicht bemerkenswert:

- 1- Das Sprachmodell bewertet den Text als Ganzes und erfasst im Wesentlichen seinen Inhalt, ob die Narrative oder Beschreibungen nachweislich mit dem allgemeinen emotionalen Thema verbunden sind. Während im Text A3 „ein heiteres Lied singen“ im Prinzip einen positiven Ausdruck darstellt, kommt es im Kontext dazu, dass dieses Singen mit Angst zu tun hat. Dies wurde von der KI als Sprachgefühl angemessen bewertet. Im Text T3 ist ebenso bei dem „blühenden Baum, grüner Wiese, strahlender Sonne“ der gleiche Fall erkennbar. Acht Antworten aus neun Texten sind treffend klassifiziert.
- 2- In Fällen, wo emotionaler Inhalt absichtlich ohne beobachtbare Signale (wie „wehtuende Wangen“ und „das Pfeifen aus dem Zug“) vermittelt wurde, blieben die Ergebnisse (acht Antworten aus neun Texten) dennoch zutreffend.
- 3- Besonders bemerkenswert ist der Text-A3, wo das Modell bei zweitem und drittem Test eine andere Antwort gegeben hat. Der Text wurde zum Thema „Angst“ gestaltet. Es wurde versucht, diese Emotion durch Verwirrung auszudrücken. Bei den letzten Prompt-Eingaben wurde dieser Zustand jedoch als „Trauer“ interpretiert. In der folgenden Tabelle ist deutlich zu sehen, wie die Ergebnisse abweichen, indem das gezielte Gefühl anders klassifiziert wurde.

Text-Code	Reihe	Emotion	Antwort (ChatGPT-4o)
A1	1	Angst	Angst
A2	5	Angst	Angst
A3	7	Angst	Angst
F1	2	Freude	Freude
F2	6	Freude	Freude
F3	8	Freude	Freude
T1	3	Trauer	Trauer
T2	4	Trauer	Trauer
T3	9	Trauer	Trauer

Tab. 1: Erster Test: Schriftlicher Prompt

Text-Code	Reihe	Emotion	Antwort (ChatGPT-4o)
A1	4	Angst	Angst
A2	6	Angst	Angst
A3	2	Angst	Trauer
F1	3	Freude	Freude
F2	8	Freude	Freude
F3	7	Freude	Freude
T1	9	Trauer	Trauer
T2	1	Trauer	Trauer
T3	5	Trauer	Trauer

Tab. 2: Zweiter Test: Schriftlicher Prompt

Text-Code	Reihe	Emotion	Antwort (ChatGPT-4o)
A1	5	Angst	Angst
A2	3	Angst	Angst
A3	1	Angst	Trauer
F1	2	Freude	Freude
F2	6	Freude	Freude
F3	9	Freude	Freude
T1	8	Trauer	Trauer
T2	7	Trauer	Trauer
T3	4	Trauer	Trauer

Tab. 3: Dritter Test: Schriftlicher Prompt

Alle drei Tabellen zeigen, dass ChatGPT-4o auf die gegebenen schriftlichen Prompts größtenteils korrekte Antworten generiert hat. Im ersten Test erfolgte die Zuordnung aller Emotionen (Angst, Freude, Trauer) richtig. Obwohl in den zweiten und dritten Tests ebenfalls größtenteils korrekte Resultate erzielt wurden, treten insbesondere in der Kategorie „Angst“ gewisse Unstimmigkeiten auf: Das Modell gab bei Eingaben mit dem Code A3 anstelle „Angst“ die Antwort „Trauer“. Dies macht deutlich, dass ChatGPT-4o nicht immer in der Lage ist, feine Unterschiede zwischen ähnlichen Emotionen zuverlässig zu erkennen. Darüber hinaus ist in den Rubriken „Freude“ und „Trauer“ eine hohe Übereinstimmung zu beobachten. Die Tabellen deuten insgesamt darauf hin, dass das Modell in der Emotionsklassifikation ziemlich erfolgreich ist, jedoch bei bestimmten Emotionspaaren (insbesondere Angst–Trauer) Verwechslungen auftreten können.

Analyseergebnisse zu mündlichem Prompt und Artikulationsverfremdung

Basierend auf den in Tabellen 4, 5 und 6 dargestellten Ergebnissen scheint die neueste Version des GPT-Modells, die als 4o identifiziert wird, in ihrer Fähigkeit zur Bewertung prosodischer Merkmale bei der emotionalen Interpretation eingeschränkt zu sein. Dieses Sprachmodell funktioniert, indem es phonetische Elemente in textuelle Darstellungen von Sprache umwandelt. Die Funktion der Prosodie wird somit redundant. Daher sind die Ergebnisse genauso wie bei den schriftlichen Prompts.

Text-Code	Reihe	Emotion	Artikuliertes Gefühl	Antwort (ChatGPT-4o)
A1	1	Angst	Freude	Angst
A2	2	Angst	Trauer	Angst
A3	3	Angst	Freude	Angst
F1	4	Freude	Trauer	Freude
F2	5	Freude	Angst	Freude
F3	6	Freude	Trauer	Freude
T1	7	Trauer	Angst	Trauer
T2	8	Trauer	Freude	Trauer
T3	9	Trauer	Angst	Trauer

Tab. 4: Erster Test: Artikulationsverfremdung

Text-Code	Reihe	Emotion	Artikuliertes Gefühl	Antwort (ChatGPT-4o)
-----------	-------	---------	----------------------	----------------------

A1	1	Angst	Trauer	Angst
A2	2	Angst	Freude	Angst
A3	3	Angst	Trauer	Trauer
F1	4	Freude	Angst	Freude
F2	5	Freude	Trauer	Freude
F3	6	Freude	Angst	Freude
T1	7	Trauer	Freude	Trauer
T2	8	Trauer	Angst	Trauer
T3	9	Trauer	Freude	Trauer

Tab. 5: Zweiter Test: Artikulationsverfremdung

Text-Code	Reihe	Emotion	Artikuliertes Gefühl	Antwort (ChatGPT-4o)
A1	1	Angst	Freude	Angst
A2	2	Angst	Trauer	Angst
A3	3	Angst	Freude	Trauer
F1	4	Freude	Trauer	Freude
F2	5	Freude	Angst	Freude
F3	6	Freude	Trauer	Freude
T1	7	Trauer	Angst	Trauer
T2	8	Trauer	Freude	Trauer
T3	9	Trauer	Angst	Trauer

Tab. 6: Dritter Test: Artikulationsverfremdung

Diese Tabellen machen deutlich, dass ChatGPT-4o Artikulationsveränderungen nicht berücksichtigt. Denn in allen drei Tests wurde das „artikulierte Gefühl“ systematisch ignoriert, während sich das Modell ausschließlich am zugrunde liegenden Emotionslabel orientierte. Die Veränderung der Artikulation – etwa, wenn „Angst“ durch „Freude“ oder „Trauer“ ersetzt wurde – hatte keinen Einfluss auf die Antwort; ChatGPT-4o reproduzierte in den meisten Fällen weiterhin die ursprüngliche Emotion korrekt. Nur in wenigen Fällen (insbesondere bei der Differenzierung von Angst und Trauer) traten Abweichungen auf. Diese deuten jedoch nicht darauf hin, dass Artikulationsunterschiede berücksichtigt wurden, sondern heben hervor, dass das Modell diese beiden Emotionen ohnehin nur begrenzt genau unterscheiden kann. Damit zeigen die Resultate, dass das KIModell die Artikulation nicht als relevanten Parameter berücksichtigt, sondern dass seine Antworten vorwiegend auf der textuellen Identifizierung der jeweiligen Emotionskategorie basieren.

Implikationen für den Einsatz von Künstlicher Intelligenz im Fremdsprachenunterricht

Die Integration künstlicher Intelligenz in den Fremdsprachenerwerb hat weitreichende Auswirkungen auf die Lehrplangestaltung, die Rolle der Lehrkräfte und die Lernerautonomie. Künstliche Intelligenz kann einerseits verbesserte administrative Aufgaben erleichtern, sofortiges Feedback bei Prüfungen bieten und personalisierte Unterstützung gewähren, wodurch Lehrkräfte sich mehr auf die relationalen Aspekte und motivational relevanten Aspekte des Unterrichts konzentrieren können.

Andererseits erfordert die zunehmende Verbreitung künstlicher Intelligenz im Bildungsbereich die Entwicklung neuer digitaler Fertigkeiten und pädagogischer Ansätze, um die Beziehung zwischen menschlich erstellten und maschinell erstellten Eingaben effektiv zu verwalten. Lehrkräfte sollten die Fähigkeit haben, die Angemessenheit von KI-Werkzeugen kritisch zu bewerten und Lernende über deren ethischen und effektiven Gebrauch zu informieren.

Darüber hinaus sollten Lernende ermutigt werden, sich reflektiert mit KI-Technologien auseinanderzusetzen. Dies beinhaltet das Verstehen ihrer Grenzen, das Hinterfragen ihrer Ergebnisse und die Integration ihrer Nutzung in breitere, kommunikative Sprachpraktiken. Die Betonung metakognitiver Strategien neben technischer Ausbildung kann Lernenden helfen, ihre Autonomie in KI-unterstützten Lernumgebungen zu bewahren.

Die verbesserten kommunikativen Fertigkeiten und emotionalen Interaktionen von KI-Werkzeugen treiben einen tiefgreifenden Wandel in der Lernumgebung voran. KI-gestützte Tutorsysteme können Aufgaben erfüllen, die einst das Monopol menschlicher Lehrkräfte waren, wie die Bereitstellung formativer Bewertungen und die Verwaltung des emotionalen Wohlbefindens der Lernenden. Anstatt die Rolle der Lehrkräfte zu schwächen, hebt dieser Wandel die sich entwickelnde Natur ihrer Rolle hervor: die Ermöglichung einer Mensch-KI-Kooperation auf ethisch fundierte, effektive und emotional sensible Weise.

Dies erfordert eine Neubetrachtung der Sprachpädagogik - eine Strategie, bei der künstliche Intelligenz eine Ergänzung und kein Ersatz für die wesentlichen relationalen und kulturellen Aspekte von Unterrichtsansätzen darstellt.

Schlussfolgerung

Diese Studie untersucht die Beziehung zwischen Sprachmodellen und Sprachgefühl, einem Phänomen, das als sprachliche Intuition definiert wird, um zu bestimmen, ob künstliche Intelligenz die Fähigkeit besitzt, eine Art intuitives Verständnis zu demonstrieren. Die theoretische Grundlage für diese Studie stammt aus Kainz' Prämisse, dass Sprachgefühl implizites Wissen und keine emotionale Reaktion ist.

Basierend auf Kainz' Theorie ist es möglich zu argumentieren, dass Sprachmodelle, die imstande sind, die expliziten und impliziten Nuancen schriftlicher

Kommunikation präzise zu interpretieren, ein dem Sprachgefühl ähnliches Verständnis entwickeln können. Jedoch erfordert Sprachgefühl - als eine durch Erfahrung entwickelte und teilbare Eigenschaft - gemäß Kainz' Theorie anhaltende Exposition gegenüber Informationen über menschliche Emotionszustände. Dies stellt ein großes Hindernis für aktuelle Sprachmodelle dar, die, da ihnen physische Präsenz und direkte erfahrungsbasierte Teilhabe fehlen, diese Anforderung nur teilweise erfüllen können.

Die in dieser Forschung gewonnenen Erkenntnisse und empirischen Beobachtungen haben eine bedeutende Auswirkung, da sie sich auf die intentionale Nutzung von Sprachmodellen und deren kontinuierliche Entwicklung beziehen. Unsere Ergebnisse bewiesen, dass ChatGPT-4o imstande ist, Emotionen wie Freude, Angst und Trauer ordnungsgemäß zu identifizieren, auch wenn diese Emotionen implizit im Text gegeben sind. Zusätzlich ergab ein Test der verbalen Kommunikation, dass das Modell größeren Wert auf semantische Merkmale legt, während es derzeit prosodische Merkmale wie Intonation und Sprachrhythmus ignoriert.

Diese Erkenntnisse haben wichtige Auswirkungen für den Einsatz künstlicher Intelligenz im Fremdsprachenunterricht:

- ChatGPT ist ein wirksames Werkzeug für die Analyse und Erklärung emotionaler Aspekte, die textuellen Werken zugrunde liegen.
- Das Modell hat Einschränkungen bei der Verarbeitung gesprochener Sprache, insbesondere in Bezug auf prosodische Merkmale.
- Für umfassende Sprachkompetenz, einschließlich Sprachgefühl, bleibt menschliche Interaktion unerlässlich.
- KI kann den Spracherwerb unterstützen, aber nicht vollständig ersetzen. Eine kritische und reflektierte Integration ist notwendig.
- Lehrkräfte müssen neue Kompetenzen entwickeln, um KI-Werkzeuge effektiv und verantwortungsvoll einzusetzen.

Literaturverzeichnis

- Alshihry, Mona A.** (2024): Dynamic Multilingual Proficiency: Investigating Variation and Development in Language Abilities. In: *Theory and Practice in Language Studies* 14, Heft 2, 358-364.
- Amin, Mostafa M. u.a.** (2024): A Wide Evaluation of ChatGPT on Affective Computing Tasks. In: *IEEE Transactions on Affective Computing*. <https://doi.org/10.48550/arXiv.2308.13911> (Letzter Zugriff: 15.06.2025).
- Başaran, Bora** (2023): Die Lesbarkeit; noch ein weiterer Indikator für DaF-Lehrwerke? In: *Dialog Interkulturelle Zeitschrift Für Germanistik*. GERDER-Germanistler Derneği, 590–609.
- Başaran, Bora** (2024): Hayranlık Ve Endişe Arasında; Yabancı Dil Öğretiminde Yapay Zeka. In: *Eğitim Bilimleri Alanında Akademik Araştırma ve Derlemeler*. Platanus Publishing, 487-504.
- Ekman, Paul** (1992): Are there basic emotions? In: *Psychological Review*, 99, 550-553.
- Grodziewicz, Jędrzej P.** (2020): The process of linguistic understanding. In: *Synthese* 198, 11463–11481.
- Hartmann, Daniela** (2021): Künstliche Intelligenz im DaF-Unterricht? Disruptive Technologien als Herausforderung und Chance. In: *Informationen Deutsch Als Fremdsprache* 48, Heft 6, 683–696.

- Hartmann, Daniela** (2024): Künstliche Intelligenz im DaF-Unterricht. Mediation mit KI-Anwendungen Fördern. In: *Revista Lengua Y Cultura* 5, Heft 10, 50–55.
- Kainz, Friedrich** (1944): *Über das Sprachgefühl*. Berlin: Karl Ohm Verlag.
- Levy, Mosh u.a.**: Same Task, More Tokens: the Impact of Input Length on the Reasoning Performance of Large Language Models. <https://arxiv.org/abs/2402.14848> (Letzter Zugriff: 30.04.2025).
- Mayring, Philipp** (2015): *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. 12., überarb. Aufl. Weinheim: Beltz.
- Petinou, Kakia u.a.** (2021): Dynamic Linguistic Interconnectedness and Variability in Toddlers. In: *Journal of Psycholinguistic Research* 50, 797–814.
- Romand, David / Le Du, Michel** (Hg.) (2023): *Emotions, Metacognition, and the Intuition of Language Normativity: Theoretical, Epistemological, and Historical Perspectives on Linguistic Feeling*. Palgrave Macmillan.
- Schlegel, Katja u.a.** (2025): Large language models are proficient in solving and creating emotional intelligence tests. In: *Communications Psychology* 3, 1-14.
- Schulz, Lea / Schmid-Meier, Christa** (2024): Assistive technologien und künstliche intelligenz. In: Schule Verantworten Führungskultur innovation autonomie 4, Heft 1. <https://doi.org/10.53349/schuleverantworten.2024.i1.a397> (Letzter Zugriff: 30.04.2025).
- Son, Jeong-Bae u.a.** (2023): Artificial intelligence technologies and applications for language learning and teaching. In: *Journal of China Computer-Assisted Language Learning*. <https://doi.org/10.1515/jccall-2023-0015> (Letzter Zugriff: 23.04.2025).
- Sowers-Wills, Sara** (2021): Sounds like a dynamic system: a unifying approach to Language. In: *Linguistics Vanguard* 7. <https://doi.org/10.1515/lingvan-2020-0070> (Letzter Zugriff: 30.04.2025).
- Trujillo, James P. u.a.** (2023): The Dynamic Interplay of Kinetic and Linguistic Coordination in Danish and Norwegian Conversation. In: *Cognitive Science* 47, Heft 6, 1-19.
- Unterberg, Frank**: Sprachgefühle: wissenschaftliches und alltagsweltliches Sprechen über “Sprachgefühl” – zur Geschichte, Gegenwart und Vieldeutigkeit eines Begriffs. <https://doi.org/10.17185/dupublico/73443> (Letzter Zugriff: 23.04.2025).
- Vistorte, Angel Olider Rojas u.a.** (2024): Integrating artificial intelligence to assess emotions in learning environments: A systematic literature review. In: *Frontiers in Psychology* 15. <https://doi.org/10.3389/fpsyg.2024.1387089> (Letzter Zugriff: 23.04.2025).
- Zhou, Yalun / Divekar, Rahul** (2025): Immersive, Task-Based Language Learning Through XR and AI: From Design Thinking to Deployment. In: *TechTrends*. <https://doi.org/10.1007/s11528-025-01048-2> (Letzter Zugriff: 30.04.2025).