

Schweizerische Koordinationsstelle
für Bildungsforschung

Centre suisse de coordination pour
la recherche en éducation

Centro svizzero di coordinamento
della ricerca educativa

Swiss Coordination Centre for
Research in Education

Information Bildungsforschung

Permanente Erhebung über Projekte der schweizerischen Bildungsforschung

Information sur la recherche éducationnelle

Enquête permanente sur la recherche éducationnelle en Suisse

Informazione sulla ricerca educativa

Inchiesta permanente sulla ricerca educativa in Svizzera

Information about research in education

Permanent inquiry into educational research in Switzerland

ISSN 1013-6258

25:100

Laufzeit des Projekts: 2023–2024

Thema des Projekts:

Entwicklung eines sprachbasierten Chatbots zur Förderung mündlicher
Sprachkompetenz im Englischunterricht (Projekt ChaLL)

Durée de la recherche: 2023–2024

Thématique de la recherche:

Conception d'un chatbot basé sur la langue pour promouvoir
la compétence orale dans l'enseignement de l'anglais (projet ChaLL)

Institution: Pädagogische Hochschule Zürich (PHZH), Bildung und Digitaler Wandel, Zürich (1); Zürcher Hochschule für Angewandte Wissenschaften ZHAW, Centre for Artificial Intelligence (CAI), Winterthur (2); Universität Zürich, Institut für Computerlinguistik, Zürich (3)

Bearbeitung | Mise en œuvre: Luzia Sauer, Dr. (1); Michael Geiss, Prof. Dr. (1); Manuela Hürlimann, MSc Cognitive Science, Language and Multimodal Interaction (2); Mark Cieliebak, Prof. Dr. (2); Gerold Schneider, Prof. Dr. (3); Johannes Graen, Dr. (3); Jan Milan Deriu, Dr. (2); Janick Michot, MSc Data Science (2); Katsiaryna Mlynchik (2, Kompetenzzentrum für Generative KI)

Kontaktperson | Personne à contacter: Manuela Hürlimann (manuela.huerlimann@zhaw.ch)

Kurzbeschreibung: Im Fremdsprachenunterricht ist häufiges Sprechen der Zielsprache zentral und kommt oft zu kurz. Neuere Sprachtechnologien könnten Möglichkeiten schaffen, dialogisches Sprechen öfter und unabhängiger zu trainieren. Im Projekt «[Chall](#)» («Towards a Voice-Based Chatbot for Language Learners») wurde der Prototyp eines interaktiven *Chatbots* entwickelt, um Primarschülerinnen und -schülern mehr Gelegenheiten zu bieten, Englisch zu sprechen. Ein wichtiges Modul eines solchen *Chatbots* ist das automatische Spracherkennungssystem (ASR – *Automatic Speech Recognition*). ASR-Modelle werden (1) v.a. auf Sprachdaten erwachsener Personen (Erstsprache) trainiert und korrigieren (2) Fehler direkt bei der Transkription. Für didaktische Zwecke sollte ein ASR-Modell Fehler bewahren (um korrigierendes Feedback zu ermöglichen) und mit spontaner Sprache (gelernter Fremdsprache) einer jungen Schülerschaft funktionieren. Als Daten für die Modellentwicklung wie auch für die Testung sammelte das Forschungsteam 85 Stunden Sprachaufnahmen (Lernaktivitäten; 45000 Einzeläusserungen) von 4.–6.-Klässlerinnen und -Klässlern (n = 327) aus 8 Schulen. 7 bestehende ASR-Modelle wurden mit diesen Daten getestet. Ein weiteres Modell wurde mit 80% der Daten trainiert (*Finetuning*) und mit 20% ihm noch unbekannter Daten getestet. Die Ergebnisse zeigen, dass aktuelle ASR-Systeme die Sprache von Kindern/Jugendlichen bzw. Sprachlernenden nur unzureichend verarbeiten und systematisch Fehler korrigieren. Das feinabgestimmte Modell «Chall-300M» erreichte im Vergleich zu einem ähnlich grossen Standardmodell eine um 12 Prozentpunkte höhere Fehlererhaltungsquote. In einer detaillierten Analyse der 20 häufigsten Fehlertypen bewahrte Chall-300M durchschnittlich 82,7% der Fehler, während die beiden besten Vergleichsmodelle lediglich 67,3% bzw. 72,3% erreichten. Insgesamt belegen die Ergebnisse, dass eine gezielte Feinabstimmung auf Sprachdaten von Kindern/Jugendlichen resp. Sprachlernenden entscheidend ist, um ASR-Modelle so anzupassen, dass diese für den sprachdidaktischen Einsatz geeignet sind und Lernprozesse wirksam unterstützen.

Brève description de la recherche: Dans le cadre de l'enseignement des langues étrangères, l'usage fréquent et actif de la langue cible à l'oral représente un objectif fondamental, mais il demeure trop rarement atteint. Les technologies linguistiques plus récentes pourraient offrir la possibilité de s'exercer à la parole dialogique plus souvent et de manière plus autonome. Dans le projet «[Chall](#)» («Towards a Voice-Based Chatbot for Language Learners»), un prototype de *chatbot* interactif a été développé afin d'offrir aux élèves de l'école primaire davantage d'occasions de parler anglais. Un module important d'un tel *chatbot* est le système de reconnaissance automatique de la parole (ASR – *Automatic Speech Recognition*). Les modèles ASR sont (1) principalement entraînés sur des données vocales d'adulte-s (langue première/maternelle) et (2) corrigent les erreurs directement lors de la transcription. Un modèle ASR utilisé à des fins didactiques devrait conserver les erreurs (afin de permettre une rétroaction corrective) et fonctionner avec une parole spontanée (langue étrangère apprise) d'un jeune public scolaire. Pour le développement comme pour l'évaluation du modèle, l'équipe de chercheurs et chercheuses a recueilli 85 heures d'enregistrements vocaux (activités d'apprentissage; 45000 énoncés individuels) d'élèves de 4^e à 6^e année (n = 327) provenant de 8 écoles. 7 modèles ASR existants ont été testés avec ces données. Un autre modèle a été entraîné avec 80% des données (ajustement fin) et testé avec 20% de données lui étant encore inconnues. Les résultats montrent que les systèmes ASR actuels traitent de manière insuffisante la parole d'enfants/d'adolescent-e-s resp. d'apprenant-e-s de langues et qu'ils corrigent systématiquement les erreurs. Le modèle affiné «Chall-300M» a atteint un taux de conservation des erreurs supérieur de 12 points de pourcentage à celui d'un modèle standard de taille comparable. Dans une analyse détaillée des 20 types d'erreurs les plus fréquents, Chall-300M a conservé en moyenne 82,7% des erreurs, tandis que les deux meilleurs modèles de comparaison n'en conservaient que 67,3% et 72,3%. Dans l'ensemble, les résultats montrent qu'un ajustement ciblé sur des données vocales d'enfants/d'adolescent-e-s resp. d'apprenant-e-s de langues est déterminant pour adapter les modèles ASR de manière à les rendre aptes à un usage didactique et à soutenir efficacement les processus d'apprentissage.

Veröffentlichungen | Publications: Michot, J., Hürlimann, M., Deriu, J., Sauer, L., Mlynchyk, K. & Cieliebak, M. (2024). *Error-preserving automatic speech recognition of young English learners' language*. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 6444–6454). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.acl-long.348>, PDF siehe <https://arxiv.org/pdf/2406.03235>

Hier erwähnte Publikationen sind über den Buchhandel oder die durchführende Institution bzw. die Kontaktperson zu beziehen, nicht bei der SKBF.

Les publications mentionnées dans l'Information sur la recherche éducationnelle ne sont pas disponibles au CSRE; veuillez vous adresser à votre libraire ou à l'institution de recherche ou à la personne de contact mentionnée.

Methoden | Méthodes de recherche: quantitativ und qualitativ, deskriptive und vergleichende Evaluationsmetriken, Mittelwert, Standardabweichung, Audioaufnahmen, Transkriptionen, Cross-Validation, deskriptive Fehleranalyse, Modelltraining, Modellvalidierung, Modellvergleich

Geografischer Raum | Délimitation géographique: Deutschschweiz

Art des Projekts | Type de recherche: Kooperationsprojekt

Auftrag | Mandat de recherche: kein Auftrag

Finanzierung | Financement: Innosuisse, <https://www.innosuisse.admin.ch>

Schlüsselbegriffe: Primarschule, obligatorische Schule, Sprachlernen, Englischunterricht, Fremdsprachenerwerb, Sprechkompetenz, Sprachdidaktik, *Chatbot*, Sprachverarbeitung, Künstliche Intelligenz, KI, automatisierte Spracherkennung, ASR, *Automatic Speech Recognition*, Lernunterstützung, mündliche Kommunikation, Sprachfeedback, Fehlerkorrektur, Sprachtechnologie, Lerntechnologie, sprachbasierte Lernunterstützung

Mots-clés: école primaire, école obligatoire, apprentissage des langues, enseignement de l'anglais, acquisition des langues étrangères, compétence orale, didactique des langues, technologie linguistique, technologie d'apprentissage, *chatbot*, traitement du langage, intelligence artificielle, IA, reconnaissance automatique de la parole, ASR, *Automatic Speech Recognition*, soutien à l'apprentissage, communication orale, rétroaction linguistique, correction des erreurs, soutien à l'apprentissage fondé sur la langue