

Speech-to-Text für Schüler:innen mit Lese-Rechtschreibstörung

Chancen und Herausforderungen der Technologie für die Textproduktion in der fünften Klasse

Silvana Flütsch Keravec

Zusammenfassung

Speech-to-Text (STT) kann das Verfassen von Texten erleichtern. Doch eignet es sich auch für Schüler:innen mit Lese-Rechtschreibstörung (LRS) und funktioniert das Diktieren im Schulalltag? In einer Studie wurden handschriftliche und diktierete Texte von Fünftklässler:innen mit LRS aus der Deutschschweiz hinsichtlich ihrer Länge und Korrektheit verglichen. Dabei wurde untersucht, wie sich diese Variablen mit der Zeit verändern. Zudem berichteten Lehr- und Fachpersonen sowie Schüler:innen über ihre Erfahrungen mit STT im Unterricht. Die Ergebnisse zeigen einen Vorteil zugunsten des Diktierens, doch auch Herausforderungen, die beim Gebrauch in der Schule beachtet werden müssen.

Résumé

La synthèse vocale permet de faciliter la rédaction de textes. Mais est-elle également adaptée aux élèves ayant une dyslexie et peut-elle s'insérer dans le quotidien scolaire ? Une étude a comparé la longueur et l'exactitude de textes manuscrits et dictés par des élèves dyslexiques de cinquième année primaire de Suisse alémanique. Elle a examiné de quelle manière ces variables évoluent au fil du temps. De plus, des enseignantes et enseignants, des spécialistes et des élèves ont fait part de leur expérience avec la synthèse vocale en classe. Les résultats indiquent un attrait favorable à la synthèse vocale, mais également des défis à prendre en compte lors de son utilisation à l'école.

Keywords: Dyslexie, assistive Technologie, Speech-to-Text, Schreibfluss / dyslexie, technologie d'aide, synthèse vocale, écriture fluide

DOI: <https://doi.org/10.57161/z2026-01-05>

Schweizerische Zeitschrift für Heilpädagogik, Jg. 32, 01/2026



Lese-Rechtschreibstörung und Speech-to-Text als assistive Technologie

Speech-to-Text (STT) bezeichnet Technologien, die gesprochene Sprache automatisiert in geschriebenen Text umwandeln. Wenn man einen Text mit STT verfasst, das heisst, wenn man ihn diktieret, entfallen die sogenannten hierarchieniedrigen Schreibprozesse. Dazu gehören Rechtschreibung und Handschrift. Die hierarchiehohe Schreibprozesse bewältigt die diktierende Person weiterhin selbst, indem sie beispielsweise Ideen für den Textinhalt sammelt und den Text strukturiert (Almgren Bäck et al., 2024).

Die Kernsymptome einer Lese-Rechtschreibstörung (LRS) auf der Ebene der hierarchieniedrigen Schreibprozesse sind Schwierigkeiten mit der korrekten Rechtschreibung, Grammatik und Zeichensetzung (WHO, 2024). STT war zwar ursprünglich nicht oder nicht ausschliesslich für Personen mit besonderen Bedürfnissen bestimmt, passt für Personen mit LRS aber zum Konzept einer assistiven Technologie (Matre & Cameron, 2024). Eine solche zeichnet sich gemäss Edyburn (2015, S. 9, eigene Übersetzung)

durch das Potential aus, es Personen mit besonderen Bedürfnissen zu ermöglichen, selbständiger zu leben, zu lernen und zu arbeiten, da die Technologie die Auswirkung einer Behinderung reduziert, beseitigt oder mindert.

Kraft et al. (2019) begründen das Potenzial von STT aus theoretischer Sicht damit, dass durch die mündliche Spracheingabe kognitive Ressourcen gespart werden, die sonst für die Verschriftung eingesetzt werden müssten. Diese freiwerdenden Ressourcen könnten demnach für hierarchiehohe Schreibprozesse verwendet werden. Bedenkt man die Wichtigkeit des Schreibens für die schulische und berufliche Laufbahn sowie den Alltag, könnten insbesondere Schüler:innen mit LRS durch STT entlastet werden. Im Gegensatz dazu stellt das Diktieren Schüler:innen möglicherweise vor andere Herausforderungen als das Schreiben von Hand oder mit Tastatur. STT erfordert beispielsweise ein höheres Tempo für die Sprachplanung oder eine deutliche Aussprache. Herausforderungen gibt es bei der Integration von STT in den Unterricht auch für die Lehr- und Fachpersonen. Knezek und Christensen (2016) betonen, dass dies viel Wissen und Willen braucht.

Das Potenzial von STT im Vergleich zur Handschrift oder zum Tastaturschreiben wurde bereits mehrfach untersucht, allerdings nur vereinzelt bei Schüler:innen mit LRS. Im deutschen Sprachraum hat Schüler (2021) den Nutzen des Diktierens bei Achtklässler:innen untersucht und sowohl eine höhere Korrektheit als auch längere Texte festgestellt. Zudem wurden Texte von besseren und weniger guten Schreiber:innen verglichen. Dabei unterschieden sich die diktier-ten Texte weniger stark als die getippten und handgeschriebenen. Die Studie von Schüler (2021) weist deshalb darauf hin, dass STT Potenzial für schwächere Schreiber:innen hat. Spezifische Forschung zur Wirkung von STT für deutschsprachige Schüler:innen mit LRS gibt es jedoch nicht. In Skandinavien hingegen haben zum Beispiel Almgren Bäck et al. (2024) und Kraft et al. (2019) den Nutzen von STT spezifisch für Primarschüler:innen mit LRS untersucht. Ihre Ergebnisse sind allerdings teilweise uneinheitlich, beispielsweise in Bezug auf den Effekt auf die Textlänge oder die Vielfalt des Wortschatzes. Hinsichtlich der Rechtschreibung herrscht in beiden Studien Einigkeit: Diktierte Texte sind im Vergleich zu handgeschriebenen oder getippten Texten korrekter.

Die nun vorgestellte Teilstudie schliesst diese Lücke und untersucht den Nutzen von STT als assistive Technologie für Fünftklässler:innen mit LRS in der Deutschschweiz. Der linguistische Kontext in der Deutschschweiz ist geprägt durch die beiden Varietäten Schweizerdeutsch und Standarddeutsch. Dies könnte eine zusätzliche Herausforderung sein und sich auf das Diktieren mit STT, respektive die Transkriptionsgenauigkeit, auswirken.

Studiendesign und Forschungsfragen

Die Daten für diesen Beitrag stammen aus einer kontrollierten, quasi-experimentellen Interventionsstudie mit Mixed-Methods-Design und wurden zwischen Herbst 2022 und Frühling 2023 erhoben. Der SNF finanzierte die Studie über ein Projekt, welches an der PH Zürich und der Universität Freiburg angesiedelt war. Es handelt sich um eine Teilstudie, denn in der Interventionsstudie wurden neben dem hier präsentierten Medienvergleich zwischen STT und Handschrift auch Transfereffekte von STT auf den Erwerb der Schriftsprache und die Schreibmotivation untersucht (Fluetsch Keravec, 2026). Die Frage, wie STT als assistive Technologie bei Fünftklässler:innen mit LRS wirkt, wurde auf der Grundlage der Daten einer einzigen Studiengruppe untersucht. Die 36 Schüler:innen dieser Gruppe diktierten während einer 18-wöchigen Intervention Texte im Schreibunterricht. Vor und nach der Intervention wurden diktier-te Texte mit den handgeschriebenen Texten derselben Schüler:innen verglichen. Es handelte sich bei den Datenanalysen daher um In-tragruppenvergleiche.

Zentral in dieser Studie ist, dass die Forschungsfrage im natürlichen Schulumfeld untersucht wurde. Einzig die Einführung in die Nutzung der Software, *Dragon Home 15*, fand im Einzel- oder Kleingruppensetting in der Logopädie oder im Rahmen der Schulischen Heilpädagogik statt. Die Intervention selbst erfolgte im Schreibunterricht. Neben Erkenntnissen hinsichtlich der Wirksamkeit sollten so auch Herausforderungen bei der Nutzung des Tools im Unterricht aufgezeigt werden.

Forschungsfragen

Die Studie untersuchte erstens, ob die Schreibflüssigkeit in diktierten Texten von Schüler:innen mit LRS signifikant besser ist als in handschriftlichen Texten. Zweitens wurde gemessen, ob die Diktierflüssigkeit in den Texten dieser Schüler:innen zwischen Prä- und Posttest signifikant zunahm. Drittens wurden die Erfahrungen der Lehr- und Fachpersonen und die Einschätzungen der Schüler:innen erfragt.

Forschungsmethodisches Vorgehen

Die Schüler:innen verfassten Schreibproben in den beiden Modalitäten Diktieren und Handschreiben. In diesen Texten wurden die Schreibflüssigkeitsvariablen Produktivität und Korrektheit verglichen. Die Anzahl handgeschriebener oder diktiert Wörter (<total words written> [TWW] respektive <total words dictated> [TWD]) bildet die Produktivität ab. Die Rechtschreibung wurde anhand der Korrektheit von jeweils zwei aufeinanderfolgenden Wörtern beurteilt. Dies ergab die Anzahl korrekter Wortsequenzen (<correct word sequences> [CWS]), die weitere Elemente der Rechtschreibung, wie Satzzeichen, einbezieht. Um Tempounterschieden zwischen dem Handschreiben und dem Diktieren entgegenzuwirken, wurde der Prozentwert der korrekten Wortsequenzen (% CWS) benutzt.

Um die erste Frage nach dem Vergleich der Modalitäten zu beantworten, wurden die handgeschriebenen mit den diktierten Texten verglichen. Für die zweite Frage nach der Entwicklung der Diktierflüssigkeit wurden die Texte des Prätests und des Posttests miteinander verglichen. In beiden Fällen wurden gepaarte T-Tests berechnet. Zur Beantwortung der dritten Frage wurden elf Leitfaden-Interviews mit Klassenlehrpersonen und Fachpersonen (Logopädie und Schulische Heilpädagogik) geführt. Den Schüler:innen wurden vier Fragen gestellt. Die Interviews und Antworten wurden mittels qualitativer Inhaltsanalyse kodiert und quantifiziert.

Die Stichprobe umfasste zu Beginn der Intervention 42 Fünftklässler:innen mit einer LRS-Diagnose. Sechs Schüler:innen beendeten die Intervention aus unterschiedlichen Gründen nicht: Nutzung der Tastatur anstelle der Diktierfunktion (4)¹, nicht nach Vorgabe durchgeführte Intervention (1), nachträglicher Ausschluss aufgrund unklarer Diagnose (1). Für die folgende Berechnung standen jeweils 32 bis 34 Datensätze zur statistischen Analyse zur Verfügung.

Ergebnisse

Forschungsfrage 1: Vergleich zwischen den Modalitäten – Diktieren versus Handschreiben

Die deskriptiven Ergebnisse in Tabelle 1 und die inferenzstatistischen Werte² zeigen, dass Fünftklässler:innen mit LRS zu beiden Messzeitpunkten in der Diktiermodalität längere Texte verfassen als von Hand. Zudem ist in den diktierten Texten ein höherer Prozentanteil an korrekten Wortsequenzen zu verzeichnen.

¹ Die Lehrpersonen oder Logopädinnen gaben an, dass diese Schüler:innen Texte kaum mit der Diktiersoftware verfassten, sondern die Texte tippten.

² Dieser Vorteil zugunsten des Diktierens ist für beide Variablen und zu beiden Zeitpunkten statistisch bedeutsam (Produktivität Prätest: $M = 18,88$, $SD = 27,74$; $t(31) = 3,85$, $p < .001$, $d = .68$; Produktivität Posttest: $M = 40,58$, $SD = 51,97$; $t(32) = 4,49$, $p < .001$, $d = .78$; Korrektheit Prätest: $M = 38,57$, $SD = 3,45$; $t(31) = 11,18$, $p < .001$, $d = 1.98$; Korrektheit Posttest $M = 34,00$, $SD = 18,23$; $t(32) = 10,71$, $p < .001$, $d = 1.87$).

Tabelle 1: Deskriptive Statistik anhand der Mittelwerte (M) und Standardabweichungen (SD) zum Vergleich der Produktivität und Korrektheit in diktierten und handgeschriebenen Texten

Variable	Modalität	Messzeitpunkt	N	M	SD
TWW	handschriftlich	Prätest	32	32,72	15,66
TWD	diktirt	Prätest	32	51,59	25,74
TWW	handschriftlich	Posttest	33	33,55	15,25
TWD	diktirt	Posttest	33	74,12	50,69
% CWS	handschriftlich	Prätest	32	44,17	19,27
% CWS	diktirt	Prätest	32	82,73	14,71
% CWS	handschriftlich	Posttest	33	46,57	19,73
% CWS	diktirt	Posttest	33	80,56	17,64

TWW = total words written, TWD = total words dictated, CWS = correct word sequences

Während die Effektstärken der höheren Produktivität einen mittleren Effekt aufweisen, zeichnen sich die Unterschiede zwischen den Modalitäten hinsichtlich Korrektheit durch sehr starke Effekte gemäss Cohen (1988) aus. Insgesamt verfassen die Schüler:innen mit STT signifikant mehr und zugleich korrektere Texte als handschriftlich.

Forschungsfrage 2: Entwicklung von Diktierflüssigkeit (Produktivität und Korrektheit) über die Zeit

Der Vergleich der Texte (N = 34) zwischen Prä- und Posttest zeigt unterschiedliche Ergebnisse für Produktivität und Korrektheit. Die Schüler:innen diktierten am Ende der Intervention während fünf Minuten signifikant mehr Wörter, knapp 23 mehr als vor der Intervention.³ Im Gegensatz dazu unterscheiden sich die prozentualen Anteile korrekter Wortsequenzen (% CWS) zwischen Prä- und Posttest statistisch nicht signifikant (M = -3,05%, SD = 17,17).⁴ Zusammengefasst bedeutet dies, dass die Schüler:innen im Diktiermodus über die Zeit hinweg längere Texte produzieren, während die Korrektheit stabil bleibt.

Betrachtet man die beiden Variablen der Schreibflüssigkeit, zeigen sich verallgemeinert Vorteile des Diktierens gegenüber dem Handschreiben, erstens im Vergleich der Modalitäten (Forschungsfrage 1) und zweitens in Hinblick auf die Entwicklung (Forschungsfrage 2). In diesem Sinne ist STT für Fünftklässler:innen mit LRS eine hilfreiche assistive Technologie.

Forschungsfrage 3: Chancen und Herausforderungen der Spracherkennung im Schreibunterricht

Die Schüler:innen beantworteten am Ende der Intervention Fragen zu ihrer Tool-Vorliebe, zu ihrer Erwartung, wie häufig sie in Zukunft diktieren werden, und zu ihren Gefühlen beim Diktieren. Die Antworten zeigen eine gewisse Ambivalenz. Obwohl über dreimal so viele Schüler:innen von positiven Gefühlen beim Diktieren berichteten als von negativen (Flütsch Keravec, 2026), gab beispielsweise knapp die Hälfte aller Schüler:innen (d. h. inkl. Dropouts) an, am Ende der Intervention das Handschreiben gegenüber dem Diktieren zu bevorzugen (vgl. Tab. 2).

Tabelle 2: Tool-Vorliebe

Aussage	Anzahl	Prozentanteil
Ich schreibe lieber von Hand.	20	48 %
Von Hand schreiben und diktieren mache ich gleich gern.	13	31 %
Ich diktiere lieber.	9	21 %

³ Prätest im Mittel 49,97 Wörter (SD = 25,83) in fünf Minuten, Posttests im Mittel 72,71 Wörter (SD = 50,60). Dieser Zuwachs (M = 22,74; SD = 48,86) ist signifikant mit einem knapp mittleren Effekt: $t(33) = 2.71, p = .011, d = .47$.

⁴ Leichte Reduktion von 3,05 % ist statistisch nicht signifikant: $t(33) = 1.04, p = .307$.

Die interviewten Lehr- und Fachpersonen befürworteten den Einsatz von STT für Schüler:innen mit einer LRS. Eine Mehrheit nannte subjektiv wahrgenommene positive Einflüsse des Diktierens auf die mündliche und schriftliche Sprache der Schüler:innen. Allerdings beschrieben die Lehr- und Fachpersonen auch verschiedene Herausforderungen. Am häufigsten betrafen diese Herausforderungen Faktoren auf Seite der Schüler:innen, insbesondere die mündliche Sprachfähigkeit und die Strukturiertheit der Arbeitsprozesse. Die Lehr- und Fachpersonen schätzten deshalb die Automatisierung von Abläufen in der Sprachplanung und die einzelnen Schritte des Diktierens (Satz überlegen, Mikrofon einstellen, Diktieren, Mikrofon ausschalten, Satzzeichen setzen, Kontrollieren) für diese Gruppe als zentral ein.

Eine weitere häufig genannte Herausforderung war das Problem der schlechten Transkriptionsgenauigkeit. Dies erforderte Anpassungen von den Schüler:innen und Begleitpersonen. Sie mussten beispielsweise ganze Satzteile statt Einzelwörter sprechen, ihr räumliches Umfeld anpassen oder mit Frustration umgehen lernen.

Die Lehr- und Fachpersonen gehen davon aus, dass die Schüler:innen genügend Zeit und Begleitung brauchen, um das Diktieren zu erlernen.

Diskussion

Die Ergebnisse der ersten beiden Fragen zeigen Vorteile des Diktierens für hierarchieniedrige Schreibleistungen. Für hierarchiehöhere Fähigkeiten – wie die Textstruktur oder die Vielfalt der Ideen – lassen sich jedoch keine Aussagen ableiten. Wie oben erwähnt, informieren diese Ergebnisse auch nicht darüber, wie STT den Erwerb der Schreibfähigkeit beeinflusst (vgl. dafür Flutsch Keravec, 2026).

Die technologischen Entwicklungen waren und sind bedeutend. Aus diesem Grund sind die vorliegenden Ergebnisse eine Momentaufnahme (Peng et al., 2025).

Hinsichtlich Schreibflüssigkeit zeigen sich signifikante Vorteile von Speech-to-Text gegenüber dem Handschreiben.

Betrachtet man den Nutzen von STT als assistive Technologie für Fünftklässler:innen mit LRS hinsichtlich der Schreibflüssigkeit, zeigen sich signifikante Vorteile mit knapp mittelgrossen bis grossen Effekten gegenüber dem Handschreiben. Somit bestätigen die Ergebnisse für die spezifische Gruppe von Fünftklässler:innen mit LRS in der Deutschschweiz die früheren Erkenntnisse für Achtklässler:innen der Regelschule aus Deutschland (Schüler, 2021). Diese Vorteile treffen also trotz der allenfalls zusätzlichen Herausforderung durch die zwei Varietäten in der Deutschschweiz zu. Obwohl die statistisch festgestellten Vorteile für die untersuchte Gruppe verallgemeinerbar sind, bevorzugen nicht alle Schüler:innen das Diktieren. Dies kann verschiedene Gründe haben. Vorstellbar ist, dass die Schüler:innen nicht auffallen möchten und sich die Nutzung des Tools, wie von Almgren Bäck et al. (2023) beschrieben, stigmatisierend anfühlt. Die Interviews mit Lehr- und Fachpersonen weisen weiter darauf hin, dass das Diktieren nebst Vorteilen auch mehrere Herausforderungen mit sich bringt. STT wird deshalb als Tool eingeschätzt, das für manche Schüler:innen bereits in der 5. Klasse einen grossen Nutzen hat, für andere aber – zumindest zu diesem Zeitpunkt – kaum gewinnbringend ist. Da derzeit zu wenig Wissen über Subgruppen oder beispielsweise notwendige Mindestwerte der Transkriptionsgenauigkeit vorhanden ist, kommen Praktiker:innen nicht darum herum, die Eignung des Tools mit Schüler:innen zu testen. Zudem muss das Textverfassen mit STT geübt werden, was Zeit und allenfalls gewisse Anpassungen erfordert. Zusammenfassend sprechen die Resultate dafür, dass STT als assistive Technologie für Schüler:innen mit LRS einen bedeutenden Nutzen haben kann. Die Eignung sollte aber individuell eingeschätzt werden.



Dr. Silvana Flütsch Keravec
Lektorin
Departement für Sonderpädagogik
Universität Fribourg
silvana.fluetschkeravec@unifr.ch

Literatur

- Almgren Bäck, G., Lindeblad, E., Elmqvist, C. & Svensson, I. (2023). Dyslexic students' experiences in using assistive technology to support written language skills: A five-year follow-up. Disability and rehabilitation. *Assistive technology*, 19 (4), 1–11. <https://doi.org/10.1080/17483107.2022.2161647>
- Almgren Bäck, G., Mossige, M., Bundgaard Svendsen, H., Rønneberg, V., Selenius, H., Berg Gøtttsche, N., Dolmer, G., Fälth, L., Nilsson, S. & Svensson, I. (2024). Speech-to-text intervention to support text production among students with writing difficulties: a single-case study in nordic countries. Disability and rehabilitation. *Assistive technology*, 19 (8), 1–20. <https://doi.org/10.1080/17483107.2024.2351488>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed). Erlbaum Associates.
- Edyburn, D. L. (2015). Expanding the Use of Assistive Technology while Mindful of the Need to Understand Efficacy. In D. L. Edyburn (Ed.), *Advances in Special Education Technology. Efficacy of Assistive Technology Interventions* (pp. 9–20). Emerald Group Publishing Limited.
- Flütsch Keravec, S. (2026). *Speech-to-Text Technology for Students with Dyslexia: Benefits and Challenges of the Tool in the Writing Classroom. Sprachliche Bildung – Studien*. Waxmann.
- Knezek, G. & Christensen, R. (2016). Extending the will, skill, tool model of technology integration: adding pedagogy as a new model construct. *Journal of Computing in Higher Education*, 28 (3), 307–325. <https://doi.org/10.1007/s12528-016-9120-2>
- Kraft, S., Thurfjell, F., Rack, J. & Wengelin, Å. (2019). Lexikala analyser av muntlig, tangentbordsskriven och dikterad text producerad av barn med stavningssvårigheter. *Nordic Journal of Literacy Research (NJLR)*, 5 (3), 102–122. <https://doi.org/10.23865/njlr.v5.1511>
- Matre, M. E. & Cameron, D. L. (2024). A scoping review on the use of speech-to-text technology for adolescents with learning difficulties in secondary education. Disability and rehabilitation. *Assistive technology*, 19 (3), 1103–1116. <https://doi.org/10.1080/17483107.2022.2149865>
- Peng, J., Wang, Y., Li, B., Guo, Y., Wang, H., Fang, Y., Xi, Y., Li, H., Li, X., Zhang, K., Wang, S. & Yu, K. (2025). A Survey on Speech Large Language Models for Understanding. *IEEE Journal of Selected Topics in Signal Processing*, 1–32. <https://doi.org/10.1109/JSTSP.2025.3640535>
- Schüler, L. (2021). Schreibflüssigkeit im Medienvergleich: Handschrift – Tastaturschreiben – Diktieren mit Spracherkennung. *Zeitschrift für Angewandte Linguistik*, 75, 330–363. <https://doi.org/10.1515/zfal-2021-2077>
- WHO (2024). *ICD-11: for Mortality and Morbidity Statistics*. <https://icd.who.int/browse/2024-01/mms/en#1516623224> [Zugriff: 17.11.2025].