



Ein lautes und professionell klingendes Vorlesen mithilfe einer Sprachausgabe ist vermutlich einer der strittigsten Punkte, die man sich im Umgang mit derselben vorstellen kann. Im Folgenden wird gezeigt, dass dies durchaus möglich ist, zumindest insofern einige Umstände Beachtung finden. Anzusprechen ist zunächst die konzeptionelle Bestimmung der Ausgaben des Screenreaders als gesprochene Schrift und die daraus resultierenden Umgangsweisen. Im Anschluss daran kann anhand eigener Daten aus Interviews mit routinierten Verwenderinnen und Verwendern sowie einem Alltagsbeispiel aufgezeigt werden, wie eine solche alternative Vorlesepraxis gelingen kann.

Schlagworte: Critical Blindness Studies; Hörweisen; Agency; Medienaneignung; auditives Lesen

Zitiervorschlag: Schulz, Miklas (2024). Vorlesen mit Sprachausgabe. Zum Gelingen einer unterrepräsentierten Praxis. *blind-sehbehindert*, 144(2), 65-76. Bielefeld: wbv Publikation. <https://doi.org/10.3278/BSB2402W002>

E-Journal Einzelbeitrag
von: Miklas Schulz

Vorlesen mit Sprachausgabe. Zum Gelingen einer unterrepräsentierten Praxis

aus: blind-sehbehindert 2/2024 (BSB2402W)
Erscheinungsjahr: 2024
Seiten: 65–76
DOI: 10.3278/BSB2402W002

Vorlesen mit Sprachausgabe. Zum Gelingen einer unterrepräsentierten Praxis.

Einleitung

Den Hörsinn als Lesesinn zu disziplinieren ist eine Herausforderung. Es ist weder einfach noch selbstverständlich. Wie man bei dem kanadischen Medientheoretiker McLuhan (1995) nachlesen kann, war es das allerdings im Falle des Auges auch noch nie. Nur liegt die Frühphase breiter Alphabetisierung und damit zusammenhängender Sinnesdisziplinierung lange zurück; was gleichfalls für die Brailleschrift gilt. Damit dies im Falle des Ohres und im Kontext technischer Stimmsynthesen nun auch gelingen kann, gibt es mindestens drei Dimensionen zu berücksichtigen (Schulz 2022; 2024a; 2024b), wobei diese in der blindenpädagogischen Forschung bislang keine hinreichende Aufmerksamkeit finden (Barclay 2012; Herlich 2012; Holbrook et al. 2017; Winter 2022), obgleich sie für den Umgang mit Screenreadern durchaus Relevanz besitzen: Gemeint sind das Phänomen einer individuellen Lesestimme, die eigendynamische Qualität der Maschinensprache sowie ihre vergleichsweise geringe Sinnstiftung an das technisch gesprochene Wort. Vor diesem Hintergrund soll vorliegend neben einer knappen konzeptionellen Erkundung des Untersuchungsgegenstands der Sprachausgabe die inklusionspädagogisch gerahmte Frage beantwortet werden, wie diese Aspekte in einem auditiven Leseprozess von den Anwenderinnen und Anwendern berücksichtigt und integriert werden. Wie lässt sich also eine Praxis des

Vorlesens mit Sprachausgabe genau beschreiben und schließlich verwirklichen?

Forschungen aus den vergangenen Jahren zur Verwendung einer Sprachausgabe verdeutlichen eine gewisse Skepsis, mit der ihr zuweilen im schulischen Kontext begegnet wird. In der ZuBra-Studie schneidet das über Screenreader erwirkte Hörverstehen beispielsweise im Kontrast zur Brailleschrift weniger überzeugend ab (Winter et al. 2019). In diesem Erhebungssetting war eine fortwährende und individuelle Kontrolle des Textflusses aus methodischen Gründen der Standardisierung jedoch nicht vorgesehen; vielmehr wurde „die Bedienung [...] durchwegs von der Testleitung übernommen“ und die Hörgeschwindigkeit „konnte während der Testung nicht mehr verändert werden“ (Winter 2022, 104). Damit scheint es naheliegend, einmal die tatsächlichen und alltäglichen Umgangsweisen mit einer Sprachausgabe genauer zu untersuchen und differenziert zu beschreiben. Auf diese Weise lässt sich klären, wie ein erfolgsversprechendes Lesen mit Sprachausgabe gelingen und dann entsprechend gelehrt werden kann.

Auch im Rahmen der interdisziplinären Leseforschung werden seit Längerem Modellbildungen präsentiert, die den Leseprozess erklären sollen (für einen Überblick Christmann 2015). In diese Modellierungen fließen vor allem quantitative Messdaten ein, die in Experimenten und weitgehend unter Laborbedingungen gewonnen werden. Ihnen liegt ein gemeinsames quantifizierendes Forschungsparadigma

zugrunde, über das bestimmte Annahmen geteilt, aber eben auch Leerstellen stabilisiert und Aporien produziert werden (Lösener 2015; Schulz 2024a). Neben anderen Aspekten ist dabei das hier interessierende Verhältnis einer inneren Stimme zur Subvokalisation und zu einem inneren Hören/Vernehmen relevant. Schließlich sprechen Menschen beim Lesen in fremdem Namen zu sich selbst, wie es der Literaturwissenschaftler Klaus Weimar (1999) treffend bezeichnet. Es bedarf demnach bei der zum Lesen gehörenden Wandlung von Schrift in Sprache einer eigenen – mitunter stumm und innerlich bleibenden – Lesestimme. Selbige decodiert die zunächst fremden, im Text formulierten Gedanken und macht sie einem subjektiven Verständnis zugänglich. Forschungspraktisch problematisch ist, dass das innere Sprechen beim Lesen „wohl grundsätzlich jeder nicht-reflexiven Beobachtung entzogen“ ist, was dazu führt, dass selbiges „lediglich als factum brutum in die Modellkonstruktion einbezogen“ (Weimar 1999, 49) wird. Zwar ist man sich dieser Umstände durchaus bewusst, „eine Lösung wird aber experimentell nicht zu finden sein und wird meines Wissens auch nicht eigentlich gesucht“ (ebd.).

Dieser kritische Befund hängt mit dem kognitionspsychologischen Fokus in der Leseforschung zusammen: In die entstehenden Modellierungen können nur solche Bestandteile integriert werden, die den Forschenden entweder kulturell selbstverständlich scheinen oder die in diesem Untersuchungsparadigma auch plausibel zu operationalisieren sind. Bezogen auf die kulturellen Selbstverständlichkeiten wird in einem Positionspapier jüngst konstatiert, dass die kognitionswissenschaftliche Leseforschung in einem selbst erfüllenden Kreislauf des okularzentristischen Desinteresses gefangen ist, den sie selbst maßgeblich mit produziert; nicht zuletzt, indem sie aus der Sehkultur stammende Annahmen

unreflektiert auf die Beforschung der Brailleschrift überträgt (Engelbredson et al. 2023, 402–403). Folgerichtig wurde der Messung der Augenbewegungen viel Aufmerksamkeit in der Leseforschung gewidmet. Sie lässt sich nicht nur forschungspraktisch erfolgreich umsetzen und passt in das vorherrschende Paradigma, darüber hinaus steht sie auch in Übereinstimmung mit einer weitreichenden Visualitätsorientierung (Hutmacher 2019). In diese Kerbe schlägt auch der Sprachwissenschaftler Markus Bader (2015), wenn er festhält, dass es vor allem methodische Gründe seien, die dazu führten, „dass es wesentlich schwieriger ist, zu gesicherten Erkenntnissen über den Verlauf des auditiven Sprachverstehens zu kommen“ (Bader 2015, 142). Das hat zur Konsequenz, dass erheblich weniger Forschung und Befunde zum Thema der auditiven Sprachverarbeitung vorliegen (Bader 2015, 142; Hutmacher 2019).

Verglichen mit anderen Bereichen handelt es sich beim Hören und Verstehen von Sprache somit um Desiderata. Obgleich es wesentliche Kompetenzen sind, die sowohl für die kindliche Entwicklung als auch für Schul- und Bildungserfolge zentral sind, sind diese Felder in der Forschung randständig. So heißt es auch in einem einschlägigen Buch zum Thema des Sprachhörens und -verstehens gleich zu Beginn, dass „[d]ie rezeptiven und kognitiv geprägten Anteile der menschlichen Sprache [...] in der Literatur, in der Forschung und im pädagogischen Alltag vernachlässigt [werden]“ (Günther 2008, 8). Dies stimmt mit Weimars Befund überein, dass „der Wahrnehmung der Schrift durch das Auge alle Aufmerksamkeit gilt und keine mehr für das ‚innere Hören‘ als Wahrnehmung des ‚inneren Sprechens‘ übrig bleibt“ (Weimar 1999, 50). Derartige Begrenzungen in der Konstruktion des Forschungsgegenstands zu überwinden, ist ein Ziel der hier vorgestellten Untersuchung.

1 Erhebungsmethode und Auswertungspraxis

Soll eine auf unterschiedliche Medialitäten zielende Lesepraxis aus einer soziologischen Perspektive detailliert beschrieben werden, geschieht dies scheinbar besser mithilfe anderer, nämlich qualitativer Forschungsmethoden. Denn im Rahmen eines interpretativen Paradigmas ist es möglich, die – wie von Weimar angemahnt – reflexiven Anteile im Erkenntnisprozess zu berücksichtigen. So können die Erlebnisqualitäten und Einsatzpunkte der eigenen Lesestimme genauer beschrieben und verstanden werden. In der vorliegenden Studie wurden daher 18 blinde, erblindete oder sehbeeinträchtigte Menschen interviewt, die allesamt geübt sind im Umgang mit einer Sprachausgabe. Neben dem Alltagswissen und den Routinen im Umgang mit verschiedenen Medienformaten wurden auch die darauf bezogenen subjektiven Konzepte erhoben. Dafür wurde ein Leitfaden entwickelt, der die teilnarrativen Interviews rahmen und besser vergleichbar machen sollte. Der Leitfaden gliederte die Interviews in drei zentrale Themenblöcke: Nach einem biografisch orientierten Einstieg, der abfragte, wann es zu alternativen Lesestrategien gekommen ist und wie diese sich im Laufe der Ausbildung (weiter-)entwickelt haben, wurde darum gebeten, verschiedene Medien und ihre Aneignungssituationen mit den jeweiligen Vor- und Nachteilen zu schildern. In einem letzten Block stand die Frage im Mittelpunkt, inwiefern unterschiedliche Formen von Medien im Alltag eingesetzt werden und welches subjektive Empfinden damit verbunden ist.

Die Gewinnung der Interviewpersonen lief über verschiedene Onlinekanäle und über einschlägige Mailinglisten u. a. von Selbsthilfevereinen und Selbsthilfeverbänden blinder und sehbeeinträchtigter Menschen. Mit Onlinerekrutierungen gehen notwendig Verzerrungen eines

Samples einher. Allerdings war dieser Selektionseffekt im vorliegenden Fall sogar erwünscht, denn da ein geübter Umgang mit der Sprachausgabe für die Interviewteilnahme relevant war, konnte eine gewisse Affinität zu Screenreader-Programmen so vorausgesetzt werden. Die Bildungsbiografien/-abschlüsse und Tätigkeitsfelder der Interviewpersonen wurden auf Basis des erheblichen Interesses an der Studie dann bewusst sehr unterschiedlich zusammengestellt. Gleichzeitig waren aber in allen Fällen die vorhandenen blindenspezifischen Arbeitstechniken erprobt und ausgereift. Die Interviewpersonen waren im Alter zwischen 28 und 67 Jahren. Alle bezeichneten sich selbst als blind oder sind blind im Sinne des Gesetzes. Wie bei der Geschlechterverteilung war auch die Zahl der Geburts- und der Späterblindeten nahezu gleich, was für die Frage nach der Selbstverständlichkeit von Brailleschriftkompetenzen wichtig ist; lernen doch viele späterblindete Menschen die Punktschrift entweder gar nicht oder sie erlangen keine attraktive Lesegeschwindigkeit, womit ihr Einsatz sehr begrenzt bleibt. Umso wichtiger wird in diesen Fällen dann der umfängliche Rückgriff auf die Sprachausgabe. Folglich handelt es sich bei dem Sampling um eine durchaus heterogene Gruppe, sodass es geboten war, bei der Datenauswertung nach Gemeinsamkeiten zu suchen (Kelle & Kluge 2010).

Für die Auswertung der Interviews wurde auf das an die Grounded Theory angelehnte integrative Basisverfahren als ein sprachanalytisches Instrument zurückgegriffen (Kruse 2014), um Aufschluss über Erfahrungen und Deutungen im Umgang mit verschiedenen Medien (Schwarzschrift, Punktschrift und Sprachausgabe) zu erhalten (ausführlich Schulz 2024a, 55–66). Es sollte die Frage nach dem Umgang mit den potenziell aufdringlichen Sprachzeichen des Screenreaders geklärt werden, wofür die

Daten mithilfe einer Agencyanalyse ausgewertet wurden (Leineweber 2021). Wie die Bezeichnung bereits andeutet, steht beim Agency-Konzept die Bestimmung von Handlungs- bzw. Wirkmächtigkeit im Mittelpunkt. Die von den Interviewpersonen jeweils gewählte Ausdrucksweise gibt Hinweise auf die Zuschreibung von Handlungsmacht, die neben Personen auch Ereignissen oder anonymen Mächten gelten kann. Ziel der Nutzung dieses heuristischen Auswertungsinstrumentes war es, die vielfältigen Bezüge zu den verschiedenen Ebenen von Zeitlichkeit zu erfassen, wie sie im Kontext auditiver Sprachwahrnehmung relevant werden.

Der vorliegende Fokus auf das Vorlesen mit Sprachausgabe als ein Teilaspekt der größeren Studie (Schulz 2024a) lässt sich durch das Zusammenwirken der einleitend angedeuteten drei Komponenten in dieser Praxis begründen. Deutlicher als beim Vorlesen mit Sprachausgabe kann der Rückgriff auf eine individuelle Lese-stimme kaum hervortreten. Darüber hinaus finden sich in dieser spezifischen Hör-/Sprechweise auch die anderen Dimensionen des kontrollierten Textflusses (Zeitregie) wieder, die es für eine erfolgreiche prosodische Übertönung der technischen und somit selbst nicht hinreichend bedeutungstragenden gesprochenen Schrift des Screenreaders bedarf. Bevor jedoch im Folgenden auszugsweise Ergebnisse aus dieser Studie dargelegt werden können, wird nun in einem weiteren Schritt der Gegenstand der Sprachausgabe genauer untersucht (2), um darauf aufbauend das Phänomen der Lese-stimme zu thematisieren (3) und dann die Praxis entlang verschiedener Beispiele detailliert zu schildern (4).

2 Die Bestimmung der Ausgabe von Screenreadern

Es gibt eine Reihe von Gründen, die einer Vorlesepraxis im Modus des Auditiven scheinbar zuwiderlaufen. Daher ist zunächst mit den drei folgenden Missverständnissen aufzuräumen, um sie empirisch begründet weiter unten übersteigen zu können:

- Die Sprachausgabe gebe gesprochene Sprache aus, der für ein auditives Lesen lediglich zuzuhören sei.
- Die Sprachausgabe ersetze voll umfänglich die für ein Lesen übliche Decodierungspraxis.
- Die Sprachausgabe sei eine externe Stimme, die eine – für die Decodierung erforderliche – individuelle Lesestimme zum Verstummen bringen müsse.

In der vorliegend eingenommenen Perspektive macht der Screenreader auf den Unterschied von Schrift und Sprache aufmerksam. Die Schrift ist lediglich ein Substrat; aus ihr wird von kompetenten Leserinnen und Lesern regelmäßig und so routiniert etwas herausgehört, dass dieser Vorgang im Alltag kaum mehr auffällt. Bei Rosebrock und Nix (2020) wird dieser Prozess als lokale Kohärenzbildung beschrieben, die nun mithilfe der Sprachausgabe durch die automatisierte Wort- und Satzerkennung unterfüttert wird. Die mit der Decodierung in Zusammenhang stehende Bedeutungskonstruktion übersetzt demnach Schrift zurück in lebendige Sprache und findet einen individuellen Sinn in derselben wieder. Genau genommen wandeln Screenreader die Schrift also in eine verlautlichte Rede. Man könnte auch von gesprochener Sprache sprechen, wobei das aus einer analytischen Perspektive als einigermaßen unpräzise abzulehnen ist. Denn mit der gesprochenen Sprache sind Qualitäten verbunden, die von

einer Maschinensprache nicht eingelöst werden können. Gemeint sind insbesondere Dimensionen einer bedeutungstragenden Prosodie: Der Sprachausgabe ist eine Varianz im Ausdruck ihres gesprochenen Wortes nicht möglich. Sie kann zwar – sofern gewünscht – gemäß ihrer Programmierung eine Betonung simulieren, es bleibt jedoch bei einer Simulation in nur einer möglichen Variante. Ohne eine Veränderung des Textes selbst bleibt die technische Ausgabe dauerhaft und 100 % identisch (was einem Menschen hingegen unmöglich ist). Hinter der technisch reproduzierten Sprache des Screenreaders steht schlicht kein Subjekt mit Bewusstsein, das mit den Schriftzeichen intentional, interpretatorisch und prosodisch spielen könnte. Eines solchen Bewusstseins bedürfte es allerdings, um von gesprochener Sprache sprechen zu können. Ähnlich verhält es sich im Falle von Lautstärke, Tempo oder Rhythmik des gesprochenen Wortes. Manches kann zwar als Einstellungen in der Software vorgenommen werden, was dann ausgegeben wird, verfügt jedoch immer noch nicht über einen subjektiv gemeinten Sinn.

Die Technisierung gesprochener Sprache hat folglich einen spezifischen Effekt: Die Sprachausgabe schiebt gewissermaßen Qualitäten, die aus der Schriftrezeption vertraut sind, in das maschinell gesprochene Wort. Es bedarf dann der Menschen, um die gesprochene Schrift wieder zum Leben zu erwecken. Sie müssen ihr wesentliche Elemente der Bedeutung wieder hinzufügen. Lesen heißt folglich auch im auditiven Modus, den Sinn eines Textes subjektiv und situativ zu erschaffen. Dass ein Screenreader nur gesprochene Schrift ausgibt und keine gesprochene Sprache, hat bedeutende Folgen für die Rezeption, denen weiter nachzugehen ist.

3 Die Lesestimme im auditiven Aneignungsmodus

Nach der bereits erwähnten Definition des Literaturwissenschaftlers Klaus Weimar (1999) heißt lesen, in fremdem Namen zu sich selbst zu sprechen. Der fremde Name, in dem zu sich selbst gesprochen wird, ist die Aussage des Autors oder der Autorin. Aus dieser Perspektive heraus trägt man sich das Geschriebene selbst vor. Das geschieht meist innerlich und leise; dann wird in fremdem Namen zu sich selbst gesprochen. Wird nun laut gelesen, verbalisiert die Lesestimme die Gedanken des Autors oder der Autorin für sich und für andere hörbar. In jedem Fall entsteht ein (innerer) Dialog über die Aussagen des Textes.

Der Einsatz einer individuellen Lesestimme im auditiven Modus folgt nun aus der Bestimmung der Ausgaben des Screenreaders als gesprochene Schrift. Wie im Falle der Schrift üblich fehlt der technischen Entstehung des akustischen Textes der Prozess der subjektiven Bedeutungskonstruktion. Einer solchen bedürfte es allerdings für ein tiefgreifendes Textverstehen im Sinne eines Lesens. Der Sinn in der gesprochenen Schrift muss – wie beim haptischen oder visuellen Lesen auch – aus der Ausgabe des Screenreaders herausgehört und individuell hinzugefügt oder wiederentdeckt werden. Andernfalls entsteht keine Aneignungspraxis, die sich als verstehendes Lesen bezeichnen ließe. Soll also auditiv gelesen werden, muss die eigene Lesestimme das von der Sprachausgabe Verlautlichte noch einmal (innerlich) wiederholen. Dies geschieht mit einer kurzen zeitlichen Verzögerung, ähnlich wie es beim Dolmetschen der Fall ist (ausführlich Schulz 2024a, u. a. 114–117, 137–158). Diesen Vorgang beschreibe ich als die Resemantisierung (Schulz 2018; 2022). Dieses (leise, innere oder auch laute) „Nachsprechen“ des zuvor Gehörten gewähr-

leistet, was hier nur andeutungsweise postuliert werden kann, nämlich dass mittels dieser Übersetzung in eine eigene Sprache die Anschlussfähigkeit an die subjektiven kognitiven Prozesse und Strukturen der Rezipierenden gewährleistet ist. Schließlich ist es kaum möglich, etwas länger korrekt nachzusprechen, das man selbst nicht versteht. Damit allerdings die Lesestimme in den Leseprozess eingefügt werden kann, bedarf es einer Steuerung des Textflusses der Sprachausgabe. Er muss gezielt und in zeitlicher Hinsicht kontrolliert werden. Diese Steuerung geschieht so, dass die individuelle Lesestimme ihren Raum beanspruchen und zu Gehör kommen kann. Selbstredend droht sie von der Maschinenstimme übertönt zu werden. Kein Mensch kann resemantisierend Schritt halten, wenn ein Screenreader begonnen hat, ein ganzes Buch vorzutragen. Diese Gefahr wird durch eine Praxis der Tastaturnavigation abgewendet, die von mir bereits verschiedentlich als Zeitregie beschrieben wurde.

Während beim visuellen oder beim haptischen Lesen die Schriftdecodierung sowie die Bedeutungskonstruktion unmittelbar eigenständig erfolgen, ist dies beim auditiven Lesen also ein zweigeteilter Prozess, an dem die Sprachausgabe beteiligt ist, indem sie die geschriebene Schrift zunächst in Akustik transferiert (Decodierung). Für die Bedeutungskonstruktion im engeren Sinne bedarf es anschließend der eigenständigen/individuellen Resemantisierung der akustischen Reize mithilfe der (inneren) Lesestimme. Die drei Elemente (gesprochene Schrift, Resemantisierung und Zeitregie) machen sodann den Unterschied zwischen einem Texthören der Sprachausgabe und einem auditiven Lesen (Schulz 2024b). Natürlich kann ein vom Screenreader ausgegebener akustischer Text weiterhin schlicht gehört statt gelesen werden. Eine solche – sich womöglich dem Überfliegen annähernde – Lesepraxis

wird als bloßes Texthören beschrieben. Dieses Texthören überspringt mitunter Teile oder nutzt weitere Navigationsmöglichkeiten wie die gezielte Suche relevanter Schlagworte oder das Springen zu (formatierten) Überschriftsebenen. Solch einfaches Texthören etabliert keine so rigide Zeitregie und die innere Lesestimme bleibt ausgeblendet, es erfolgt auch kein innerliches Nachklingen der Wörter. Die gesprochene Schrift wird dann nicht wie für einen Leseprozess erforderlich kleinteilig interpretiert.

4 Auditives Vorlesen mit Sprachausgabe

Anhand der erhobenen Daten lässt sich zeigen, dass ein akustischer Text durchaus einfach nur gehört werden kann. In diesem Fall bleibt jedoch das Charakteristikum der gesprochenen Schrift unberücksichtigt. Soll hingegen die Praxis auditiven Lesens verwirklicht werden, müssen zwei Bestandteile in den auditiven Modus integriert werden, die als Zeitregie und Resemantisierung bezeichnet wurden (Schulz 2018; 2022; 2024a; 2024b). Letztere benötigt eine individuelle Lesestimme, die entweder leise bleiben oder für ein Vorlesen verlautlicht werden kann. Anhand des nachfolgenden Interviewauszugs, bei dem ein späterblindeter Autor über die Lesungen aus seinen eigenen Büchern reflektiert, lässt sich dies illustrieren:

„Wenn ich also einen Text vorlese [...] mach ichs grundsätzlich so ich nehme also ein normales Dokument [mhm] wo eben halt mein Text drinsteht und geh [...] dann zeilenweise vor das heißt ich habe einen Kopfhörer auf [...] und ähm gehe dann also am Rechner mit meinen Pfeiltasten von Zeile zu Zeile runter und spreche den Inhalt der mir da also von der Sprachausgabe vorgelesen wird nach.“ (Ludwig Lust)

Wie hier verdeutlicht, kann er den akustischen Text in seinem Worddokument gezielt

über die Pfeiltasten steuern und über die entstehende zeitliche Lücke auch parallel seine eigene Lesestimme einsetzen. Das angedeutete Nachsprechen impliziert außerdem dabei die Etablierung einer eigenen Betonung.

Im Folgenden kommt eine zweite späterblinde Person zu Wort, die ihre ausgefeilte Vorlesetechnik schildert, die sie mit der Sprachausgabe entwickelt hat. Claudia Campe ist eine approbierte Psychotherapeutin, die seit Langem Erfahrungen damit macht, für sich als Lernstrategie oder für ihr Kind Hörbücher aufzusprechen. Das dafür erforderliche Setting beschreibt sie wie folgt:

„Naja also es_is EIGENTLICH so ich äh sitze dann mit KOPFHÖRERN am COMPUTER [mhm] äh STELL die SPRACHgeschwindigkeit bei JAWS LANGsamer weil wenn ich jetzt SO lese hab' ich die Geschwindigkeit RELATIV schnell [...] und lass dann- geh dann einfach mit dem CURSOR eine Zeile RUNTER [mhm] höre mir AN was der da VORliest [ja] (1,0) und WÄHREND der das LIEST spreche ich das eigentlich (1,0) NACH Joa und ähm (1,0) so LERNE ich auch manchmal also ich kann auch manchmal besser LERNEN so wie MANCHE Leute laut LESEN irgendn Text so [ja] keine Ahnung äh SPRECHE ich das dann manchmal auch LAUT mit während DER mir das VORLIEST das wirkt dann vielleichtn bisschen KOMISCH aber (1,0) dann verARBEITE ich das irgendwie nochmal anders_ich setz_es nochmal UM_ irgendwie in meine EIGENE SPRACHE oder_dann WÄHREND ich das spreche denke ich auch schon drüber NACH also [mhm] WENN ichs AUFLESE kann- fällt mir auch AUF wenn da irgendn BLÖDSINN steht oder sowas [mhm] also is jetzt NICHT dass ichs nur NACHplappere.“
(Claudia Campe)

Geradezu paradigmatisch wird hier das laute Vorlesen mit Sprachausgabe geschildert. Garant für einen Erfolg dieser Praxis ist die fortdauernde Kontrolle des dynamischen Textflusses.

Wird dieser nicht gezielt gesteuert oder werden bestimmte Vorkehrungen im vorzutragenden Text selbst nicht getroffen, wird entweder durch die Geschwindigkeit oder den ununterbrochenen Sprechfluss ein simultanes Mitsprechen tendenziell unmöglich. Beide konstitutiven Komponenten werden von ihr hervorgehoben: die innere Stimme, die auch bei ihrer Lektüre hörbar verlautlicht, und die Zeitregie, die den Zeilenwechsel bei angemessener Sprachgeschwindigkeit organisiert. Folgerichtig bezeichnet sie diese Praxis dann auch als ein „Rauslesen“ oder als ein Lesen mit dem Ohr. Das „Herauslesen“ verweist auf eine erfolgreiche Sinnentnahme, wie sie sie an anderer Stelle mit der Instanz eines inneren „Mitlesers“ beschrieben hat, die dann die eigene verlautlichende Interpretation und stimmliche Inszenierung anleiten kann. Hervorzuheben ist weiterhin ihre Schilderung, dass sie das Nachgesprochene in ihre „eigene Sprache“ umsetzt. Der Grund für sie ist ein anderes Gefühl des Durcharbeitens, das mit dem lauten Vorlesen verbunden ist. Sobald sie der Sprachausgabe resemantisierend folgt, wird das, was sie aufnimmt, von ihr noch einmal anders verarbeitet oder sie denkt bereits beim Mitsprechen kritisch darüber nach. Es handelt sich dabei allerdings nicht um ein bloßes Nachplappern, sondern um eine gesteigerte interpretative Aktivierung der Kognitionsleistungen, eben um den Vorgang des Lesens, der die eigene Lesestimme als Sinngenerator notwendig involviert und ein tiefgreifendes Verstehen ermöglicht.

Wie die Antwort auf eine Nachfrage verdeutlicht, kann die dabei entfaltete Aktivität kaum glaubhaft angezweifelt werden:

„GeRATE dann in ein kleines SELBSTgespräch mit mir oder so [ja] Ähm das_is halt EINFACH nochmal son anderer- ich kann dann meine AUFMERKSAMkeit besser bei der- (1,0) bei dem Text aus LASSEN und mich damit auch AKTIVER

beSCHÄFTIGEN irgendwie [mhm, mhm] Mh n
BISSCHEN auch wien DiaLOG vielleicht
manchmal wenn ich dann so-also MANCHMAL
stell ich mir auch vor DASS ichs jemanden
VORLESE [ja] und ähm ich HAB auch man- auch
schon für FREUNDE was aufgelesen oder vorge-
lesen und das gefällt mir eigentlich AUCH weil
ich denk mir dann ich MUSS es ja dann so LESEN
dass DIE'S verstehn [...] es soll ja nich so total
UNangenehm sein sich das ANzuhörn [ja] achte
ich auf die BeTONUNG.“ (Claudia Campe)

Sie verwickelt sich in einen Dialog mit dem Text und bezeichnet damit konkret das Wechselspiel unterschiedlicher Lesestimmen, die gemeinsam den Leseprozess konstituieren. Unter erblindeten Menschen gibt es offenkundig Erfahrungen mit einer Decodierung der gesprochenen Schrift, der über eine zu ergänzende Betonung noch eine individuelle Bedeutung hinzugefügt ist. Die weiter oben von Weimar für ein Lesen im engeren Sinne geforderte Rückübersetzung mithilfe einer individuellen Lesestimme von Schrift in die eigene Sprache findet sich nun im auditiven Lesemodus wieder. Auch dem oben aufgeführten Postulat (Jäger 2014), demzufolge die Stimmen, die einen Text präsentieren, innere Anteile zum Verstummen bringen müssten, wird deutlich widersprochen. Zuletzt soll dies unabhängig von den eigenen Interviews noch ein weiterer Fall belegen, der unmittelbar dem Alltag abgelauscht ist.

In einem Interview aus der Podcast-Reihe von Freedom Scientific (der Entwickler-Firma von JAWS, in Folge 103 und im Internet zum Nachhören verfügbar), berichtet Nas Campanella von ihren Erfahrungen und Herausforderungen, die diese spezielle Lesepraxis mit sich

bringt. Nas Campanella ist eine Nachrichtensprecherin, die bei Triple J arbeitet, einem australischen öffentlich-rechtlichen Hörfunksender, der von ABC betrieben wird. Sie ist eine JAWS-Nutzerin und liest mithilfe der Sprachausgabe die Nachrichten flüssig und authentisch vor (FS Cast 103: 2015).¹ Hier ein von mir transkribierter Auszug aus diesem Gespräch, in dem die für den vorliegenden Kontext wesentlichen Punkte ausgeführt werden:

Jonathan Mosen: So that brings me onto the technology side of this. You're a news reader with Triple J, so that requires a lot of fluency. How do you achieve that then without the use of braille?

Nas Campanella: I use JAWS. I have that downloaded onto my computer in front of me and that reads out everything onto the screen and so I just listen and hear things and, you know, that's how I know what I'm typing and whether I've written something incorrectly or that sort of thing.

JM: When it comes to the actual news bulletin itself. Do you have JAWS in your ear as you read or have you memorized the newscast ahead of time?

NC: Oh no, it's impossible to memorize any bulletin, because things change, even within a minute 30 seconds, two. So I go in and I have a bit of a pre-read and then, when I'm actually doing the bulletin, I listen to JAWS and read at the same time, so I'm listening and speak at the same time.

JM: Right, anybody of course does a pre-read before a news bulletin, so there is nothing unusual about that. But then, the ability to actually be hearing something and at the same

1 Der folgende Auszug wurde aus dem Audio-Podcast transkribiert, damit er hier wiedergegeben werden kann. Siehe FSCast Episode 103: <https://fscast.libsyn.com/2015/02>. Die kurze Demo, in der sowohl Frau Campanella zu hören ist als auch das, was sie von JAWS auf die Ohren bekommt, beginnt ziemlich genau nach 20 Minuten. Herzlich gedankt sei an dieser Stelle einer meiner Interviewpersonen für diesen Hinweis.

time repeating something you heard, maybe a second or two ago. Did that take a long time to develop that skill?

CN: It's still a skill that I think I'm still learning, really. I mean you don't ever perfect anything in a short amount of time. I really started to practice listening and reading at the same time about four years ago. And it's gotten a lot better but (you know) obviously there is still room for improvement.

JM: How fast are you running JAWS, are you reading like a line at a time? Are you like up and down arrowing through a bulletin as you read?

CN: I do up and down arrow through the bulletin. And it's so, that I can control the pace obviously. Cause you know, you have to take into account the time that you need to take a breath and allow for the ability to scroll back up, in case you stumble and make a little mistake. So yeah, I constantly got my hand on my computer.

JM: And is there ever a time, when it doesn't work out, when you feel like you have stumbled a little bit and have to go back and there's a little bit of a break in the fluency or do you find it's a technique, that works for you most of the time?

CN: Look, most of the time I think it's pretty okay, but there (you know) have been times and there probably will be in the future, many times where I sort of stumbled and had to read an entire sentence over again.

Mit diesem Interviewauszug ist ein weiterer Beleg erbracht, dass ein professionelles Vorlesen mit Sprachausgabe gut gelingen kann. Als größte Herausforderung benennt die Radionachrichtensprecherin, dass JAWS nicht wahrhaft sinnbezogen betont, somit Wörter häufig komisch ausspricht. Es bedarf also einer Übersetzung, in der diese falsch anmutende Verlautlichung der Schrift korrigiert wird. Diese Beobachtung und Erfahrung teilend, sprach ich

selbst von einem Dolmetschen (Schulz 2022) und meinte damit ebendiese simultan zum Hören zu erbringende Übersetzungsleistung. Interessanterweise griffen insgesamt vier weitere meiner Interviewpersonen selbst und unabhängig von der Kenntnis meiner Veröffentlichungen und dieser australischen Nachrichtensprecherin auf genau die gleiche Metapher zurück (Schulz 2024a, u. a. 114–117, 137–158). Sie verwenden in ihren Schilderungen der Praxis des lauten Vorlesens mit Sprachausgabe auch die Analogie zum Simultandolmetschen. Eine solche intersubjektive Übereinstimmung kann als Beleg für die Stimmigkeit dieser Befunde verstanden werden. Außerdem wird das entscheidende Erfordernis der Zeitregie deutlich.

Auch Nas Campanella legt offenkundig großen Wert darauf, den Textfluss zu kontrollieren. Nur insofern diese feingliedrige Navigation gewährleistet ist und keine unpassenden Pausen entstehen, ist es wirklich möglich, der Sprachausgabe nachzusprechen. Gleichzeitig erlaubt es dann die kleine erschaffene zeitliche Lücke, der gesprochenen Schrift eine eigenständig mit Sinn erfüllte Betonung hinzuzufügen bzw. sie subjektiv aus der Maschinensprache herauszuhören. Wie gezeigt, ist dafür der Textfluss der Sprachausgabe aufzubrechen, wofür die Interviewpersonen zuweilen individuelle Lösungen entwickelt haben. Entweder die Zeilenlänge wird über eine große Schriftgröße begrenzt oder es wird gezielt ein Zeilenwechsel als Absatz gesetzt, sodass nur eine Handvoll Wörter pro Zeile stehen bleibt. Während die Kürzung der Zeile über die Schriftgröße dann automatisch und ohne einen subjektiven Nachvollzug des Inhalts erfolgt (was Auswirkungen auf die Stimmigkeit der resemantisierenden Verlautlichung haben kann), ist im letzteren Fall im Vorfeld bereits eine tiefere Auseinandersetzung mit dem Sinn des Textes zu erbringen. Der

dann gezielt positionierte Zeilenumbruch erleichtert einen sinngemäßen Redefluss. In didaktischer Perspektive lässt sich dies an die Idee von Sinneinheiten in Form von Wortgruppen anschließen. Darüber hinaus bietet sich die Möglichkeit, den Einsatz einer Sprachausgabe weit über die Zielgruppe blinder und sehbeeinträchtigter Menschen hinaus als ein inklusionspädagogisches Leseförderverfahren weiterzuentwickeln (ausführlich Schulz 2024a, 202–220). Ähnliche Verzögerungen des eigentlich interessierenden Inhalts entstehen auch, wenn Sonderzeichen oder Scanfehler noch im Text enthalten sind. Sie werden gemäß der Texttreue von der Sprachausgabe vorgetragen, lenken aber eigentlich nur ab und sollten deshalb vermieden und im Vorfeld aus einem Dokument entfernt werden. Alle zusätzlichen Zeichen, die nicht auf den Inhalt des Textes bezogen sind, werden gelöscht, wodurch eine puristische Version eines Textdokuments entsteht. Dabei ist es wenig verwunderlich, dass auch der Hörsinn für seine Lesefähigkeit spezifische Bedingtheiten mit sich bringt; anders ist es beim visuellen Lesen schließlich auch nicht. Funktioniert letztere Praxis nicht im Dunkeln, operiert der Hörsinn notwendig im Medium der Zeitlichkeit.

Fazit

Das Konzept des auditiven Lesens sollte vorliegend in einem Detail spezifiziert und damit weiter profiliert und empirisch fundiert werden: das Vorlesen mithilfe der Sprachausgabe. Dafür wurden Interviews mit Menschen mit Blindheit/Sehbeeinträchtigung ausgewertet, die schon länger einen Umgang mit Screenreadern einüben. Die diskutierten Dimensionen einer verlautlichenden Lesestimme, die zur Resemantisierung eingesetzt wird, konnten in den Daten ebenso wiedergefunden werden wie auch ein

feingliedriges Vorgehen bei der Textnavigation. Letztere wurde als Zeitregie gekennzeichnet und ist unabdingbar für ein auditives (Vor-)Lesen. Vor diesem Hintergrund können das Konzept und seine zentralen Thesen nun als intersubjektiv und empirisch validiert gelten.

Die Schrift ist ein vergleichsweise stummes und statisches Medium, das aus sich heraus kaum eine abschließende Bedeutung kundzutun vermag. Aus einer konzeptionellen Perspektive betrachtet, existieren aufgrund der Technizität der Maschinensprache die substrathaften Qualitäten, die aus der Schrift vertraut sind, auch im akustischen Text der Sprachausgabe. Das Resultat ist die gesprochene Schrift des Screenreaders, die entsprechend neue und andere Reaktionen in den Umgangsweisen mit ihr nahelegt, aber eben auch ermöglicht. Wie aus anderen Schriftzeichen muss auch aus der gesprochenen Schrift ein subjektiver Sinn erst herausgehört werden bzw. ist ein solcher im individuellen Verstehen erst zu erschaffen und wiederzufinden. Das heißt aber auch, dass ein bloßes Hören der gesprochenen Schrift nicht immer hinreichend ist, sondern zuweilen eine spezifische Hörweise einzuüben ist, die die Qualitäten derselben ernst nimmt (ausführlich und sprachphilosophisch Schulz 2024a, 48–51). Folglich wird es für die gesprochene Schrift nötig, Lesarten zu entfalten, wie sie im Zusammenhang der Schwarz- oder Brailleschrift geläufig sind (aktive Bedeutungskonstruktion bzw. Resemantisierung). Die letztgenannten Zeichen sind zunächst stumm und statisch und entwickeln nicht wie die gesprochene Schrift durch die Sprachausgabe eine dann zu kontrollierende Zeichendynamik.

Die Möglichkeiten, die das Ohr mithilfe medientechnischer Unterstützung für ein Lesen zu bieten hat, sind bisher in weiten Teilen verkannt geblieben. Vor dem Hintergrund des geschilderten traditionell schriftkulturellen

Verweisungszusammenhangs von Decodierung, Lesestimme und Sinnesmodalität scheint dies wenig verwunderlich. Auch scheint diese Unterschätzung kein Zufall zu sein, hängt sie doch mit einem bislang nicht hinreichend präzisen Verständnis der Ausgaben von Screenreadern zusammen, die die eigentliche Decodierung der akustischen Schriftzeichen nicht vollumfänglich übernehmen können. Die Tatsache, dass verschiedene Menschen zu unterschiedlichen Zeiten und unabhängig voneinander identische Strategien im Umgang mit der gesprochenen Schrift entwickelt haben, zeugt von ihrem Bewusstsein gegenüber diesem Umstand. Es werden immer wieder eigene Lösungen für gleichbleibende Herausforderungen ersonnen. Demgegenüber wäre es wünschenswert, diese bislang unterrepräsentierte, in der Praxis entfaltete Kompetenz noch besser zu verstehen und weiter bekannt zu machen. Dabei gerät eine mangelnde Punktschriftkompetenz beispielsweise im Zusammenhang mit Späterblindung zur Triebfeder einer Experimentierfreudigkeit und produziert einen bemerkenswerten Erfindungsreichtum, der von der Pädagogik vermehrt in den Fokus genommen werden dürfte. Schließlich sind diese Menschen darauf verwiesen, ihren Hörsinn als Lesesinn zu disziplinieren, zu formen und zu fordern, um auf diesem Wege hauptsächlich oder sogar ausschließlich mithilfe der Sprachausgabe zu arbeiten; so eben auch im Falle eines Vorlesens.

Insgesamt ist jedoch der Einfluss von Screenreadern auf den Erfolg des Lese-/Hörverstehens noch wenig erforscht. Hier sind weitere Bemühungen zu unternehmen, die unterschiedliche Hörweisen und Umgangsformen berücksichtigen sowie die andauernde, individuelle Kontrolle des Textflusses (Zeitregie) auch ermöglichen sollten. Zugleich ist anzumerken, dass eine quantifizierende Vermessung der vorliegenden Thesen noch aussteht. Dafür wäre die

spezifische Hörweise des auditiven Lesens unter dem Einbezug der eigenen Lesestimme abseits der vorliegenden plausiblen Erfahrungsberichte in einer vergleichenden Untersuchungsanordnung zu überprüfen. Darüber hinaus lassen sich an die geschilderte Hör-/Sprechweise weitere Überlegungen knüpfen, die auch lesedidaktisch und für eine inklusionsorientierte Modellierung des Leseprozesses von Relevanz sind. Beispielsweise wurde am Material deutlich, wie eng verschränkt scheinbar ein Lesen/Hören und Verstehen von Sprache mit dem Sprechen ist.

Literatur

- Bader, Markus (2015). Leseverstehen und Sprachverarbeitung. In: Ursula Rautenberg/Ute Schneider (Hg.). Lesen. Ein interdisziplinäres Handbuch. Berlin/Boston, Walter de Gruyter, 141–168.
- Barclay, Lisbeth A. (2012). The Importance of Listening Instructions. In: Lisbeth A. Barclay (Hg.). Learning to listen, listening to learn: Teaching listening skills to students with visual impairments. New York, AFB Press, 3–23.
- Christmann, Ursula (2015). Kognitionspsychologische Ansätze. In: Ursula Rautenberg/Ute Schneider (Hg.). Lesen. Ein interdisziplinäres Handbuch. Berlin/Boston, Walter de Gruyter, 21–45.
- Englebreton, Robert/Holbrook, M. Cay/Fischer-Baum, Simon (2023). A position paper on researching braille in the cognitive sciences: decentering the sighted norm. *Applied Psycholinguistics* 44 (3), 400–415. <https://doi.org/10.1017/S0142716423000061>
- Günther, Herbert (2008). Sprache hören, Sprache verstehen. Sprachentwicklung und auditive Wahrnehmung. Weinheim/Basel, Beltz.
- Herlich, Stephanie (2012). Middle School and High School: Advanced Skill Development. In: Lisbeth A. Barclay (Hg.). Learning to listen, listening to learn: Teaching listening skills to students with visual impairments. New York, AFB Press, 153–195.
- Holbrook, M. Cay/D'Andrea, Frances Mary/Wormsley, Diane P. (2017). Literacy Skills. In: M. Cay Holbrook/Tessa McCarthy/Cheryl Kamei-Hannan (Hg.). Foundations of Education: Instructional Strategies for Teaching Children and Youths with Visual Impairments. New York, AFB Press, 374–426.

- Hutmacher, Fabian (2019). Why Is There So Much More Research on Vision Than on Any Other Sensory Modality? *Frontiers in Psychology* 10, 1–12. <https://doi.org/10.3389/fpsyg.2019.02246>
- Jäger, Ludwig (2014). Audioliteralität. Eine Skizze zur Transkriptivität des Hörbuchs. In: Natalie Binczek/Cornelia Epping-Jäger (Hg.). *Das Hörbuch. Praktiken audioliteralen Schreibens und Verstehens*. München, Wilhelm Fink Verlag, 231–253.
- Kelle, Udo/Kluge, Susann (2010). Vom Einzelfall zum Typus. Fallvergleich und Fallkontrastierung in der qualitativen Sozialforschung. Wiesbaden, VS Verlag für Sozialwissenschaften.
- Kruse, Jan (2014). *Qualitative Interviewforschung. Ein integrativer Ansatz*, Weinheim/Basel, Beltz.
- Leineweber, Heike (2021). Das integrative Basisverfahren – Netz und doppelter Boden für rekonstruktive Analysen? *Zeitschrift für sportpädagogische Forschung* 9 (1), 33–48.
- Lösener, Hans (2015). Was hören wir beim Lesen? Zu einer Aporie kognitionspsychologischer Lesemodelle. In: Britta Herrmann (Hg.). *Dichtung für die Ohren. Literatur als tonale Kunst*. Berlin, Vorwerk, 49–66.
- McLuhan, Marshall (1995). *Die Gutenberg-Galaxis. Das Ende des Buchzeitalters*. Düsseldorf, ECON Verlag.
- Rosebrock, Cornelia/Nix, Daniel (2020). *Grundlagen der Lesedidaktik und der systematischen schulischen Leseförderung*. Baltmannsweiler, Schneider Verlag Hohengehren.
- Schulz, Miklas (2018). *Hören als Praxis. Sinnliche Wahrnehmungsweisen technisch (re-)produzierter Sprache*. Wiesbaden, Springer VS.
- Schulz, Miklas (2022). Lesen und Hören von Texten. Das Verhältnis von Lektürepraktiken, Sprachausgabe und Sinneswahrnehmung aus Perspektive der Critical Blindness Studies. *blind-sehbehindert: Zeitschrift für das Blinden- und Sehbehindertenbildungswesen* 142 (1), 19–32.
- Schulz, Miklas (2024a). *Auditives Lesen. Hörweisen mit Sprachausgabe – konzeptionelle Überlegungen und empirische Befunde*. Weinheim/Basel, Beltz Juventa.
- Schulz, Miklas (2024b). Zum Unterschied vom Auditiven Lesen und Texthören. Eine Inspiration für die blindenpädagogische Diskussion. In: *Leben. Bildung. Partizipation. Bericht des 37. VBS Kongress*. Verlag: Edition Bentheim. (im Druck)
- Weimar, Klaus (1999). Lesen. Zu sich selbst sprechen in fremdem Namen. In: Heinrich Bosse/Ursula Renner (Hg.). *Literaturwissenschaft. Einführung in ein Sprachspiel*. Freiburg, Rombach, 49–62.

- Winter, Fabian/Hofer, Ursula/Lang, Markus (2019). Lese- und Schreibkompetenzen von Jugendlichen und jungen Erwachsenen mit dualer Schriftnutzung. Forschungsergebnisse aus dem Projekt „Zukunft der Brailleschrift“. *blind-sehbehindert: Zeitschrift für das Blinden- und Sehbehindertenbildungswesen* 139, 92–108.

- Winter, Fabian (2022). *Duale Schriftnutzung. Brailleschrift und Schwarzschrift. Eine sequenzielle, explanative Mixed-Methods Studie*. Wiesbaden, Springer VS.

Prof. Dr. phil. Miklas Schulz
Stiftung Universität Hildesheim
miklas.schulz@uni-hildesheim.de

