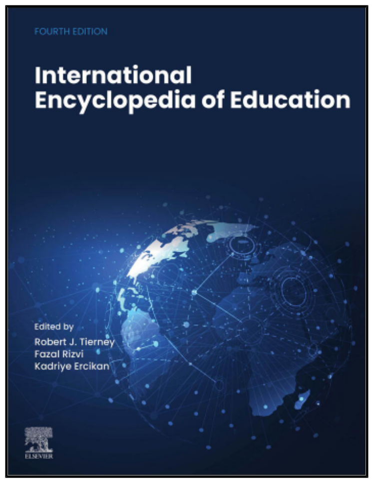

From Massaro, D., Bernstein, J., 2023. In: Tierney, R.J., Rizvi, F., Erkican, K. (Eds.), International Encyclopedia of Education, vol. 10. Elsevier.
https://dx.doi.org/10.1016/B978-0-12-818630-5.07005-6.
ISBN: 9780128186305

Artificial intelligence in literacy

Dom Massaro[a] **and Jared Bernstein**[b], [a] Department of Psychology, University of California, Santa Cruz, CA, United States; and [b] Analytic Measures Inc., Palo Alto, CA, United States

Introduction to Artificial Intelligence and Literacy

Literacy involves knowledge and skills, which develop and deepen over time. This includes knowledge of print conventions on paper and on screen, of how characters relate to words, and of how words and other language structures relate to meaning. Full adult literacy entails some facility in applying this knowledge at a useful rate that yields accurate understanding of the intent and content of a variety of text forms from tables and graphs, to recipes, and to extended narratives and advanced textbooks.

This article addresses some current impacts on literacy instruction that arti cial intelligence is having now and some foreseeable impacts that AI can have in the future as literacy evolves. The focus here is on early literacy development in young children with no perceptual or cognitive impairments, and on instruction as practiced in school and at home in a North American context. The impacts are illustrated within a three-period timeframe: traditional practice, current AI applications in practice, and future applications and practice. Each time frame is illustrated with reference to a triad of Teacher, Student, and Student Performance. Fig 1 shows a simple instruction cycle within the basic triad structure.

The article does not address any recent controversies in reading pedagogy or current approaches to situate reading instruction and literacy development within social or cultural contexts. It also does not focus on the impact or amelioration of physical or cognitive disabilities that affect reading development. The assumption will be that most reading instruction has practical goals and teachers typically work with an eclectic set of resources that may be commonly at hand.

Components of reading skill

A simple theory of reading [historically associated with Charles Perfetti (Perfetti and Hogaboam, 1975) and Philip Gough (Gough and Tunmer, 1986)] posits that reading skill has two main components, decoding (D) and listening comprehension (LC), and that some level of skill in both is required to achieve useful reading comprehension (RC). Decoding is the process of transforming

International Encyclopedia of Education Fourth Edition 2023 529–542

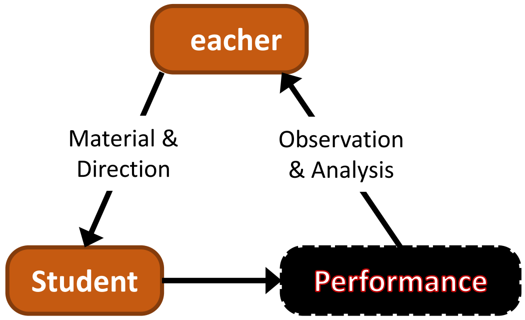

Fig. 1 A basic reading instruction cycle schematized within the teacher-student-performance triad.

a sequence of one or more characters into a likely phonological sequence and/or into a lexical item, that is, into a word. Listening comprehension is the process of transforming a spoken utterance, such as a conversational turn, into a likely sequence of linguistic forms (words, phrases, sentences), and with these linguistic forms, into the meanings that the speaker intended to convey. The theory is summarized in one simple equation: $RC \quad D \times LC$.

Because the operation between D and LC is multiplication, if either D or LC is equal to zero, then RC will also be zero. A more general term that subsumes decoding is Word Recognition. Word recognition includes decoding, as such, along with direct apprehension of word forms, known as sight-word recognition. Both components are needed for reading, and each of the two main components of reading skill in turn has several components, listed below for convenience.

- Word Recognition: knowledge of (alphabet, letter-to-sound relations, alphabetic and non-alphabetic conventions, af xes and their conventional spelling rules, etc.).
- Listening Comprehension: knowledge of (sound-to-lexical map, phrase structure, semantics, rhetorical conventions, etc.)

Spoken and written language vs. formal/informal

Learning to understand speech is never an issue as it is for written language. Conceivably our evolutionary history of spoken language relative to written language is responsible for this difference (Merlin, 1991; Pinker, 2007). Writing was originally met with some skepticism from early philosophers who thought that our humanity was compromised because memory was less essential with writing than with speech.

Table 1 illustrates examples of communication media in which the formality of the language and its modality can be categorized independently of one another. Although we usually think of informal spoken language and formal written language, a TED talk is formal and spoken whereas texting is informal and written. For example, the formal nature of picture books allows the writer to make more deliberate word and grammatical choices that are not always feasible in child-directed speech (CDS). CDS and adult-directed speech (ADS) requires a spontaneity (Grice, 1975) that necessarily limits word choice and allows deixis to substitute for various words and grammatical constructions.

Artificial Intelligence (AI) delineated

Note that machine learning (ML) is different from Al, and we can consider how the two relate to reading. Note that both ML and Al are relatively recent terms, so their meanings are still in flux. Briefly put:

> Al refers to automated actions that seem appropriate to biological intelligence;

Table 1 Gives a taxonomy of the potential independence of language modality and formal non-conversational) versus informal conversational) dialog.

	poken	*Written*
Formal	TED talks, lectures, MOOCs	Books, articles Newspapers
Informal	Conversations, interviews TV media dialogs	Texting, messaging Email

ML is an engineering eld that develops technology to build AI systems.

ML refers to algorithmic processes that operate on data sets to produce algorithms that cluster, classify, or identify patterns in new (unseen) data sets. Typically, an algorithm infers a function (or a model) that assigns labels to new data from an analysis of labeled training data. For example, thousands of transcribed voice recordings are analyzed by an ML procedure, which produces an algorithm that transcribes new voice recordings.

Arti cial Intelligence (AI) is an ability of automated systems to perform tasks that until recently required human or other biological information processing, often including signal analysis, decision, and sometimes, verbal and/or mechanical response. People think of AI as the use of automated systems that perform tasks which traditionally have required human, or at least vertebrate, intelligence. For example, cars were designed to be driven by people, but now can drive themselves as well as an average human driver. As we delve into AI and reading, note that self-driving cars are possible in part because many millions of miles of car travel (on the road with multiple sensors) have been logged and analyzed with machine learning algorithms. Obtaining and analyzing large databases for machine learning has been essential in autonomous vehicle development, as in other domains. Thus, we expect that large databases will be necessary for the development of AI applications in reading too.

Traditional, current, future approaches to reading instruction

The remainder of this article is divided into three sections:

- Traditional reading instruction,
- Current reading instruction as augmented with AI tools/applications, and
- Future reading instruction, including future literacy practices more generally.

The rst section is a brief overview that may help a reader locate the needs that current AI tools and applications address and the bene ts that accrue from these. In the second section, we learn that most of the current AI tools do not disrupt the flow of reading instruction. Instead, they facilitate current practice and render it more fun, more ef cient, more accurate, or more effective. AI may radically change the future of reading and reading instruction altogether. The last section describes a few current research initiatives that may further improve current instructional practice and a foreseeable future in which AI capabilities that are likely to change the nature of information access generally and change the nature of reading development and instruction along the way.

Traditional reading instruction

Pre-digital reading instruction

In the United States, between 2000 and 2020, reading instruction in schools largely followed patterns that were informed by a model of reading that was set forth in a report from the NICHHD (2000), which Tim Shanahan (2006) usefully summarized in 2006, setting out a rough set of ve main layers of skill/instruction that support successful reading. These ve layers are shown in **Fig 2**.

We will assume that students have suf cient phonemic awareness to start learning to read. First comes the establishment of print awareness skills the student can recognize and name all the lower- and uppercase letters and say what their most common sound value is. Then, students are introduced to the most common sight words (e.g., the, in, can, no, how), and to the basics of phonics, or letter-sound correspondences, going beyond the most common sound value for each letter. These foundational skills are usually established before the "Instruction Loop" shown in **Fig 3** gets started.

When the instruction loop starts, teachers direct students to read from leveled text material starting with very simple materials that have a picture and only a few words or a short sentence per page. The teacher then informally, monitors the literacy

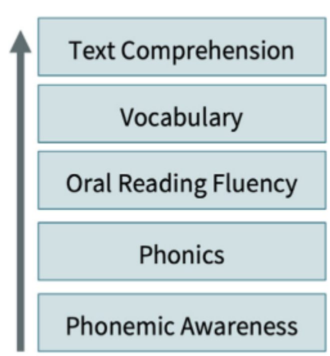

Fig. 2 Five elements of reading.

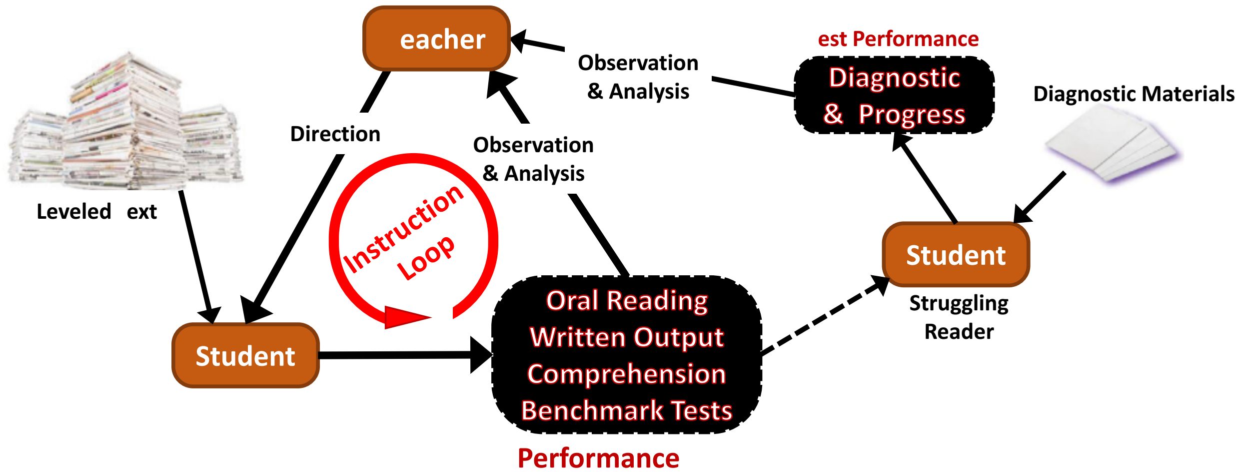

Fig. 3 A fuller version of a typical reading instruction cycle in grades 1 3, the dashed line shows an added path for students who are identi ed as struggling readers. Two static resources are shown: leveled text and diagnostic materials.

development of each student by observing as the student performs in a variety of tasks that may include reading aloud, writing, understanding texts, and sometimes in periodic formal benchmark tests that may be required of all students. As time goes by, if students perform at about the level expected for their age and grade, teachers gradually increase the dif culty of the leveled text material that students read, and the instruction loop continues, perhaps over 3–5 years, until students can read ef ciently enough to understand and enjoy age-appropriate stories and books, and can independently read grade-appropriate academic text to learn new material.

However, for struggling readers who perform below expectation on benchmark tests, who are not performing at grade level, or who simply lag behind the rest of the class, a teacher or a reading specialist may administer a set of formal tests and score an individual student s reading performance. Diagnostic analysis of the individual test performance helps determine which, if any, of the component reading skills may be hindering the student s reading development. A teacher or reading specialist will then select suitable materials and give them to the struggling student with directions and encouragement, as needed.

Given reasonable curriculum materials and suf cient time with a quali ed teacher over a 3- or 4-year period, most young readers will emerge with basic literacy skills and a solid foundation for gaining broader and deeper literacy with the practice they encounter in late elementary and middle school (grades 5–9, or ages 9–13 years). By the end of 12th grade (age 17–18), after many hours of reading in the instruction loop and then across the curriculum and in life outside school, many skilled readers can comprehend text at a rate of 200 words per minute or more (Average Reading Speed, 2022). This is faster than the rate of most spoken language they encounter. The next section reviews (2021) AI applications that extend the range or the nature of literacy instruction, improve the ef ciency of reading development, or improve accuracy in measuring reading and literacy-related skills.

Current AI tools applied in current contexts

The extended mind (& AI-in-loop, human-in-loop)

Arti cial intelligence (AI) is performed by machines, but it usually involves a human participant. Some applications of arti cial intelligence can be viewed as contributing to the extended mind (Clark and Chalmers, 1998), which involves the situational milieu we inhabit. We now take for granted that many AI tools and apps facilitate a variety of human cognitive processes, as they become more accurate and commonplace. In the period 2000–2020, several technologies have off-loaded complex cognitive demands, including, complex information retrieval, text-to-speech synthesis, automatic speech recognition (ASR), and automatic language translation.

When early text-to-speech (TtS) synthesizers were rst commercialized in the late 1970s, they were crude but suf ciently intelligible for successful commercial application in reading machines for the blind (Klatt, 1987; History of TtS, 2022; van Santen et al., 1997). By 40 years later, TtS systems had achieved wide acceptance in many applications as the systems evolved to emit highly intelligible and much more natural sounding speech. Speech synthesis is still viewed as an AI technology, and two core components of TtS are products of machine learning: rst, the text-analysis methods for word disambiguation and for construction of prosody; and then the augmented ASR technology that nds optimal sequences of signal snippets to concatenate for clear and natural sounding words.

Going back further, the invention of visual symbols and print changed what was possible and might be considered an early instance of AI for language and communication of information. The magic of making physical marks that could be seen at a future time by others enabled communication in a manner previously impossible it allowed one person to communicate to another person across time. Writing preceded the recording of spoken language and it leaves out some of the collateral information in spoken language, such as personal identity and voice-borne expression of emotion. On the other hand, written medium also enabled important representations of information such as tables and graphs that are not easily understood in their spoken form.

Viewing AI as an elaboration of the extended mind might put AI in a more acceptable framework in that it is then a product of the human mind that has the human mind as an essential partner in various applications. As Neil Stephenson, the legendary science ction writer has stated, "Ideally we ask AI to function as an aid or assistant in our journey toward knowledge" (Michael Shermer interview, 2021). This framework doesn t eliminate the recent fearful premonitions of runaway AI but perhaps puts them in a better light.

Technology aided assessment

Given that live teacher time is limited, robots (and we believe virtual tutors) can substitute for the teacher just as effectively if not more effectively. Our ideal virtual tutor would have assessment built into the tutoring regimen, and the goal is to have the student as engaged with the testing as she is with the reading and mentoring itself. Engagement of a virtual agent or a physical robot has been studied in a variety of learning contexts, but not systematically in reading and learning to read. Although lacking in systematic reading contexts, there are helpful ndings on re ning measures of students engagement, how various verbal and nonverbals cues from the agent are picked up, and how engagement can fluctuate during the interaction. If researchers and teachers explore virtual or robot mentors for reading instruction, there is a wealth of guidelines in the literature (Belpaeme et al., 2018).

Students don t necessarily enjoy being tested particularly if they are somewhat unskilled, as most learners by de nition are. Furthermore, it s been found that in simple multiple-choice tasks in which a student selects one response from multiple options, many students respond so quickly that they could not have completely understood the question and the gallery of answers. One answer to this challenge has been proposed: the Technology Enhanced Item or TEI (Bryant, 2017; TEIs, 2021), with creative testing formats. An example of a TEI might ask the student to highlight the relevant text in answering a question. In a very large systematic study, TEIs have shown the lowest rapid guessing rate, which traditionally reflects lack of student engagement: 10 times more often on multiple choice items than on TEIs.

Electronic tutoring systems endorse intelligent programming to create innovative formats in the testing environment. A student might be asked to drag and drop content from a text to answer the question. Another possibility would be to highlight the text or select multiple items from the text. In this sense TEIs offer the promise that students can be kept engaged in a reading environment even though assessment is part of that reading environment. Given that there is strong evidence that TEIs improve engagement, we can push the envelope even further by requiring more sophisticated interactions from the students such as open-ended answers or having the student actually perform a simulation or instructing the virtual mentor to perform a simulation. Ideally, an engaging tutoring session in reading would blur any existing boundaries between learning and assessment.

One challenge in applying social robots to develop reading skills and to include intermittent testing is to maintain the student s engagement with the tutor. There is some indication that students interest diminishes when the novelty effect evaporates. One key to maintaining interest is to establish a meaningful relationship between the student and the tutor, and to reinforce the student s memory of this relationship (Kasap and Magnenat-Thalmann, 2012). Using RoboTutor, Boote et al. (2021) tracked the tutee s face to monitor engagement, and the partially successful goal of predicting whether the student would complete one of the exercises. This offers one source of information among many others such as properties of the voice, response latencies, and accuracy in the task that could increase the effectiveness of monitoring engagement. Interestingly, they were able to de ne an early bailout when the child didn t like the activity and a later bailout which occurred during the interactions with the tutor. This distinction is a valuable one because different types of intervention would be necessary for success.

Automating Oral Reading Fluency ORF) assessment

Periodic assessment of reading performance is important for guiding instruction, but a recurring problem is that students spend too much time in testing, and teachers spend too much time on activities that surround testing in-service training, administering and scoring tests, and then interpreting results (Bernstein et al., 2017). Thus, automatic systems to administer and score ORF performances have gained acceptance and proliferated in the period from about 2015 to 2022 and onward. This has been accomplished with student-friendly online presentation of material in a browser, which is combined with AI capabilities on servers in the cloud.

The server-side AI comprises child-appropriate ASR and automatic scoring and scaling of the recorded read-aloud responses. Such systems can now (2021) return accurate ORF scores and running records of a student s performance within one second after test completion. The most salient advantages of current automated ORF tests are their convenience and reliability. They are convenient for the student, because the tests can be self-administered at the student s convenience, and an automated ORF test may occupy less student time. Such tests are more reliable in that they apply acceptance and scoring rules quite consistently, and they score readers without favoritism toward individual readers or groups. Automated ORF measurement has also been shown to be indifferent to irrelevant dialect differences and reading pronunciations of readings by speakers of English as a second language (Balogh et al., 2012; Bernstein et al., 2017).

A key disadvantage in current automated testing and scoring is the absence of real-time system-initiated interaction with the student during the test administration. Such interactions from the system might include interruptions to encourage or give guidance, or simply to present the test-texts at a more appropriate level so they are within the student s zone of proximal development.

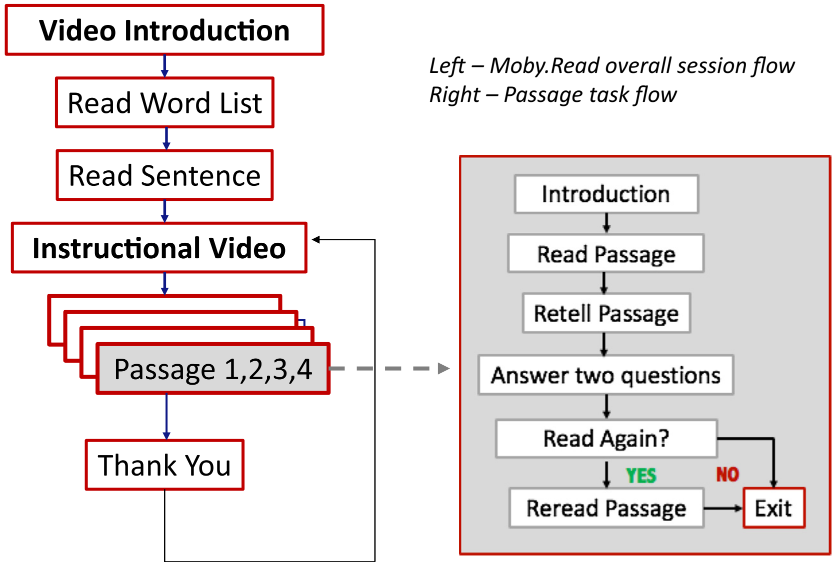

Fig. 4 Session and passage presentation flow for each passage 1 practice, 3 scored) in Moby.Read.

Moby.Read: an automated ORF test

The Moby.ReadTM (2022) system is a fully automated and self-administered oral reading fluency (ORF) benchmark assessment designed and built in 2016–18 and introduced commercially in January 2019. Moby.Read is designed for benchmarking early reading performance three times a year. Inside the Moby.Read system, an ASR was optimized for children s readings and their spontaneous speech in retelling text passages, and an augmented natural language processing (NLP) module enables immediate score reporting of ve key oral reading fluency skills:

1. Reading Level,
2. Accuracy,
3. Accurate Reading Rate (in Words Correct Per Minute),
4. Comprehension, and
5. Expression.

During the assessment, as shown in Fig 4, students read four passages aloud, summarize the passage content, and respond to short-answer questions using their own voice. Moby.Read embeds model readings and opportunities for students to "go back" and re-read a passage to hone their oral reading skills (see Fig 4). An assessment takes about 12 min and runs in Chrome or as a native app on iPads.

Scores are available immediately so teachers can apply the most current possible information to individualize reading instruction. Additionally, audio recordings allow teachers to diagnose performance issues and to share them with reading specialists, parents, and with the students themselves (see Fig 3). Fig 5 gives a sample presentation flow for each passage (1 practice, 3 scored) in Moby.Read.

Automated tutoring

Once interactive screen time became feasible, various commercial and school curricula of speech and reading tutoring were implemented in electronic exercises in speech, phonics, and vocabulary (Phonics Exercises, 2022). These involved very little AI other than utilizing speech and graphics within a given instructional-game design. Typically, children learn the alphabet and then how to sound out various spelling patterns. This method isn t foolproof, however, as experienced by a grandmother with her grandchild sounding out a word written under a picture of an insect. The child expertly sounded out bug for the word but the word was actually *ant*. Another method has children spell various sounds. One commercial game reinforcing the sound of speech is iSPY in which the child is given a sound and matches words that begin with that sound (iSPYPhonics Fun, 2021).

Although most early reading instruction is grounded in phonics, a case can be made for increasing the beginning reader s awareness of the language s orthography. It seems reasonable that drill and practice in phonics also heightens a child s awareness of these spelling patterns to facilitate letter and word recognition during reading. There is evidence that expert readers and children learning to read use spelling constraints to more easily identify letter patterns and words (Venezky and Massaro, 1979, 1987; Massaro and Hestand, 1983).

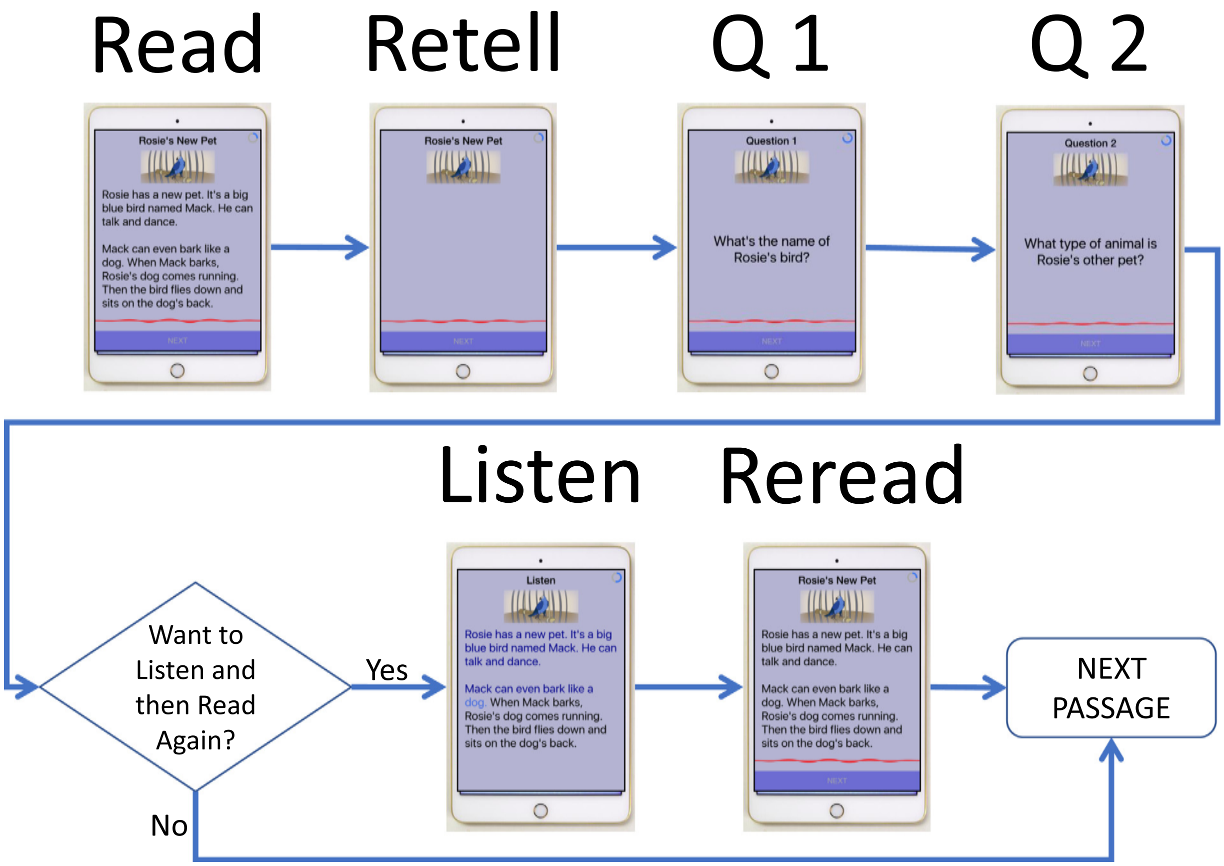

Fig. 5 A sample presentation flow for each passage 1 practice, 3 scored) in Moby.Read.

Phonics with Baldi

Given dramatic progress in creating a rich sensory world in educational domains, it is surprising that most current decoding exercises do not exploit multiple modalities. Given that seeing the speaker s face can contribute signi cantly to understanding and learning language (Massaro, 1998), and developments in computer animation of talking heads, it is now feasible to carry out phonics instruction and assessment bimodally. Various characteristics of the vocal tract can also be viewed from different perspectives, which might help the child (or second language learner) master the phonics and spelling of the language being learned.

The 3D computer-animated talking person, Baldi, was developed to produce accurate visible speech (Massaro et al., 2005; Baldi Tutoring, 2022a). Using this technology, a game was developed called Phonics with Baldi. This involves a matching game in which the child is required to match the animated pronunciation of a sound with a possible spelling. In the matching game, a successful match is rewarded by eliminating the matching pair from the display. The goal of the child is to complete all of the matches and move on to higher levels.

Read With Me

Picture books are a foundation for a child s inauguration into a literate world (Gurdon, 2019). Although seldom put to rigorous experimental tests, they appear to be a prevailing influence on children s cognitive and linguistic development. Content analysis of picture books and child directed speech (CDS) documents that the linguistic and cognitive sophistication of picture books far exceeds that found in CDS (Massaro, 2015, 2017; Montag et al., 2015; Williams et al., 2022). Although electronic TV media consume more of the child s time, picture books best demonstrate to the child that a static medium can elicit language and ideas that can engage their attention. The caregiver-child dyad appears to be a prominent characteristic of the impact of reading these books to engage children.

What might be surprising is that this shared picture book reading contributes very little to a child s acquisition of the mechanics of reading (Massaro, 2012, 2017). By this we mean phonics, letter recognition, and written word recognition. Children do learn that books contain words and stories (or at least stories) that can transport them to another place and time. However, picture books aren t designed to highlight written language but rather primarily illustrate in pictures the story being told. Because of the emphasis on pictures, the print is often either considered unimportant or embellished to look less like print and more like picture design. The design of picture books encourages children to focus on the pictures rather than the words in shared reading-aloud situations (Evans and Saint-Aubin, 2013). Perceiving words requires focused viewing of the words. The print in most children s picture books is not usually appropriate for a child learning the mechanics of reading (There is also some evidence, however, that children can learn some reading mechanics with picture books that emphasize the legibility of print, Massaro, 2017; Piasta et al., 2012).

We might ask how AI can be utilized to complement the shared reading experience of picture books in order to reinforce reading mechanics. To promote learning the mechanics of reading during shared picture book reading, Dom Massaro created an Apple iPad app called Read With Me (2022). At various times during shared reading, the adult reader can choose to display some text from the physical book in clear font on the iPad (Read With Me Demo, 2022).

To use the app, the caregiver and the child choose one of their favorite books from the app s library. The caregiver reads the book to the child, both of them enjoying the rich sharing of emotion and adventure. Read With Me adds to this experience by allowing the adult reader to dictate a sentence from the book to initiate the presentation of this sentence on the iPad screen. The app uses a robust method of ASR because it knows which book is being read.

The application allows the caregiver and child to share attention between the picture book and the text presented on the iPad. The child should have a good view of the screen when the dictated sentence is present, The iPad can be held by the caregiver, placed on a table, or placed inside a transparent holder sewn into a shirt that the caregiver would wear Read With Me Demo (2022). This experience allows the child to experience written language in the manner that they experience spoken language. The Read With Me app adds print without requiring the caregiver to provide explicit instruction in prereading skills. Many caregivers tend to have limited knowledge of or motivation to teach appropriate literacy skills. It remains to be tested whether this app or similar ones can help bootstrap a child s reading in the same manner that the speech toddlers hear aids spoken language acquisition.

Understand My World

Ideally, we would embrace an AI tutor reinforcing the mechanics of reading while simultaneously the child is primarily engaged in new and active learning. Using advances in AI, one example app, Understand My World (2022), immerses learners in spoken and written language as they are learning about their world. Equipped with a camera, microphone, and an internet connection, the student has direct access to unlimited information. Images captured by the camera are labeled and can be de ned in both speech and writing. Participants can also speak and record their speech or ask questions. The app then presents the dictated words or answers to questions in either written or both spoken and written form. Some example questions are "Tell me about Sesame Street." "Who is Ben Stiller?" "How many miles in 7 km?"

Because evaluation is central to language interventions, the app records a history of the student s interactions with all three functionalities. This history can be utilized in several ways, such as recording time on task, the vocabulary encountered, the student s speech, and the content of the student s interactions. One possible evaluation, for example, would be to compare the student s mastery of the content experienced within the app to comparable content the student experienced outside the app.

Reading practice

Jack Mostow and colleagues, situated within a campus with a premier speech science department, developed a computer program that uses speech recognition to listen to children read. Beginning roughly 3 decades ago, their core technology was ASR, well before it evolved to have the quality it has today (Mostow, 2013; Project LISTEN s Reading Tutor, 2022). The goal was to use ASR to listen to children read and guide each student individually in oral reading. Their reading tutor proved successful in a number of independent studies when compared to independent reading practice and even tutoring by certi ed teachers (Mostow, 2013). Dependent measures included word identi cation, oral reading fluency, and comprehension.

More recently, a number of reading tutors have been deployed. Reading with companion robots will de nitely become more commonplace, with their prices within the range of other appliances such as TVs. In one representative study (Sauppé and Mutlu, 2014, 2015), a reading companion/mentor robot, Minnie, interacted with children during the child s reading. Minnie was not programmed with the latest speech technology as in applications assessing oral reading fluency but she did make book suggestions, guided and monitored the child s reading goals, and made comments to support their understanding and interest in reading and science. In the near future, we expect assessment and instruction to be embedded in tutoring interactions. Given Minnie s portability, she could provide time on task for reading at home as well as in school. Based on their pilot studies (Sauppé and Mutlu, 2014, 2015), children appeared to make a strong social connection with Minnie, which agrees with ndings in other domains. We found this to be true of Baldi, an animated tutor, for both hard of hearing and children on the spectrum (Massaro, 1995).

Research foundations for future AI

Electronic databases

Some of the major achievements of AI are critically dependent on acquiring and analyzing electronic information. We all know how the internet has consolidated many unfathomable sources spanning the spectrum of knowledge. Students and professionals alike can easily access Wikipedia and other websites to nd the latest state of knowledge in almost any domain. The AI programming to make this information ef ciently retrievable and useable has led to impressive intelligent AI systems such as IBM s Watson, which was able to beat the champion Ken Jennings in a Jeopardy competition (Jeopardy, 2022).

AI s intelligence has increased exponentially in effectiveness hand-in-hand with computer s speed and memory. Data mining recording and analyzing linguistic content has also grown exponentially. We ve learned that the amount and quality of language children experience is central to the path to literacy (Hart and Risley, 1995). This observation has motivated researchers to determine important academic words to use in schooling (at perhaps the expense of non-essential words, acknowledging that time-on task is limited).

Children are talked to, read to, and in many cases watch TV Media. Different estimates of the amount of input 2–5-year-old children experience from these three different contexts are roughly 150 min for both CDS and TV Media, and only about 15 min of parents reading (picture books) to their children (Media, 2022). Massaro (2015; unpublished) analyzed potential vocabulary and grammar differences in these three linguistic environments. Content from each of these three media were analyzed to record their vocabulary and the reading level. To determine vocabulary differences among the three media, the three samples were assessed against two independent measures of frequently used spoken and written words (Brysbaert and New, 2009; COCA, 1990–2012; SUBTLEX, 2021). Children experienced over twice as many rare word types with picture books than with the other two media (Massaro, 2015; Williams et al., 2022).

The formal nature of picture books allows the writer to make more deliberate word and grammatical choices that are not always feasible in CDS and written in TV scripts. As mentioned previously, the spontaneous nature of conversational speech constrains word and grammar selections and encourages use of gesture and deixis. Using Hiebert s (2005) list of academic words, these rare words actually corresponded to important words for schooling. Picture books had about 2 or 3 times the number of these academic words than did Parent Speech and TV Media. Examples of academic words found in the Picture Book database but not in the other two media are *centers*, *heated*, *lowering*, *and winding*.

Words are also more valuable for schooling if they have high dispersion, that is, found in different content areas. Picture Books provided about twice as many high dispersion words relative to the other two media. Examples of words with a high dispersion value in picture books that did not occur in CDS and TV Media were *behave circulation exchange intricate strain* and *switch*. Using data mining, we learned that picture book reading provides valuable vocabulary for future schooling.

Mining databases to optimize reading complexity

Paving the road to literacy would ideally have linguistic and cognitive complexity of spoken language that is representative of material that the child will be reading in school. We expect that more complexity would be valuable. Grade level readability measures (Readability Formulas, 2021) were used to determine the reading grade level of the three media just discussed, as well as ADS. Children s picture books had a higher reading grade level than CDS and TV Media, as well as ADS. The data mining revealed that reading picture books to children is bene cial, as it appears to familiarize them with more challenging language and a more extensive vocabulary than they would otherwise encounter (e.g., van Kleeck et al., 2003).

Mentoring by social robots

A robot track man that calls balls and strikes in baseball more accurately than human experts is a prototypical instance of arti cial intelligence (AI Umpire, 2022). It mimics intelligence of humans but in this situation also does it better. It picks up information from the trajectory and location of the ball and computes if each pitch is in the strike zone. Given high-quality sensors and a simple AI algorithm, an accurate call is assured, signi cantly better than a live umpire.

Imagine a child learning to read with a personalized robot companion with knowledge of the child s reading level, including vocabulary and grammar, and general knowledge. In addition, the robot will have goals including one to advance the child s literacy, and will continuously learn about the child and the appropriate tutoring required to achieve the goals. With this knowledge, the robot can then anticipate where the child could be tripped up by the text and could modify and even embellish it appropriately to make it compatible with the understanding of the child. In this case, the robot would provide the affordances necessary for acquired literacy. This would also exploit the zone of proximal development described by Vygotsky (2022).

Two exploratory studies provide some incentive to determine whether social robots are facile enough to encourage and aid a child at various levels of reading skill, such as phonics, vocabulary learning, and fluent reading aloud. Michaelis and Mutlu (2018) found that a social robot engaged early adolescent children in a variety of guided reading activities. The children claimed that the robot helped with comprehension and provided them with motivation to read. Yueh et al. (2020) used a social Robot Julia to facilitate library literacy activities with elementary school children.

There has been a plethora of studies evaluating the effectiveness of robots as tutors in a variety of domains other than reading instruction and assessment (Belpaeme et al., 2018). According to a Belpaeme et al. meta-analysis of hundreds of studies, tutoring robots have been cognitively and affectively effective in many different domains. Initially, there seemed to be an advantage to having physically-present robot tutors in place of virtual tutors on a screen. A meta-analysis of 33 experimental studies found that "a co-present robot to be more persuasive, receive more attention, and be perceived more positively than a virtual simulation of the agent being tele-present" (e.g., presented on a screen, Li, 2015). However, there is also more recent evidence that virtual tutors can be just as persuasive as co-present tutors (Deng et al., 2019).

Although it is not certain that any difference between virtual and physically present tutors would occur with a reading tutor, it would be more practical and less costly to program tutors on screens than to provide every child with a physically present robot (Deng et al., 2019; Li, 2015). Students might also demand that their robot be unique in both physical and psychological characteristics. Advances in Augmented and Virtual Reality might diminish this gap between physically-present and on-screen social robots so that it would be more feasible for each student to have a personalized robot. On the other hand, we might learn that something like physical presence is important, perhaps a carryover from the child sharing picture books with loved ones. The nal word about virtual tutors and co-presence has yet to be uttered or written but their destiny will have a role in reinforcing literacy.

Although this review covered normative reading instruction and practice, it is worth mentioning that robotic tutors should have even greater potential for children with challenges such as hearing loss or existing on the autism spectrum. Baldi, introduced earlier, proved a successful tutor for both these groups of children. One major advantage of Baldi, like any robot, is his availability 24/7 and programmed extreme patience and encouragement. In one-on-one tutorials, both populations of children learned and retained new vocabulary and grammar. Children on the spectrum, who tended not to look at faces, also learned to perceive visible speech more accurately. Using a within-subject design, the research also convincingly showed that Baldi s intervention was responsible and not some other explanation such as time on task. Demonstrations of children in the tutoring situation are available (Baldi Tutoring, 2022b).

Predicting the future

While developing Baldi, a computer animated talking head, to provide accurate visible as well as audible spoken language, showing it was effective in communicating spoken language, and applying it to situations in language learning (Massaro et al., 2005), Massaro (1995) wrote a science ction scenario on a future without written language. English had become the world s dominant language and societies could no longer afford to continue English literacy. Literacy learners had dif culty conquering the irregularity of English s spelling, with many inconsistencies between spoken words and their spelling. A popular example is that it would be acceptable to spell sh as ghoti (gh as /f/ in tough, o as /I/ in women, and ti as / / in nation. Critiques and revisions of English writing have been advocated by illustrious celebrities including George Bernard Shaw and Mark Twain (English Spelling, 2022).

As with most antecedents of change, there were other factors leading to literacy s demise. Having overpopulated the earth with diminished raw materials, it was convenient to use existing books as insulation for high density housing. Most importantly, however, was the technological development of computer-animated characters with realistic speech and language. Although Face Time, audible books, and podcasts were not yet on the horizon, the prediction was that communication was primarily spoken whether between live participants or with animated characters.

Traveling from this 30-year-old prediction of the future to today, written language might still be in jeopardy. English no longer has to be the dominant language, so English will no longer be necessary for successfully navigating the international world. Huge increases in computational power and wireless communication with small portable devices, ef cient algorithms in speech synthesis, spoken language recognition, and language translation are currently available for many languages and many more around the corner (Google Read Along is currently available in 180 countries and supports 9 languages). Given this technology, real-time translation will allow access to any language, so one only requires their family s language learned from birth.

Spoken language might again gain prominence after just a few centuries of text dominance since the invention of the printing press. Being lost in a book or going to that special place with text may have a shorter lifespan than currently anticipated. As observed by Farhad Manjoo, "Yet there are just as many books that achieve a resonance via the spoken word that their text alone cannot fully deliver" (Manjoo, 2021). The narrator can add life to the writing and capture what Jaynes (1976) described as the voice capturing, or even kidnapping the listener s attention. As technology advances, the audio book will continue to be embellished creating sound effects, surround sound, and features yet to be imagined. For memoirs, your favorite celebrities, and stand-up comedians, the likely choice would be to listen (and perhaps watch if possible). In 2019, almost one in four Americans listened to an audiobook in the previous year, only slightly less than the number who read an electronic book (Reading or Listening to Books, 2022). In his dystopian novel of the future, not anticipating advances in speech technology. Neal Stephenson s Diamond Age (1995) found an essential role for live narrators of texts (called ractors) in contrast to the many technological advances in his multiverse.

For many of us, an illiterate society is a sobering thought. There must exist convincing reasons that written language has a role to play in the future. As one possible reason would be ef ciency: we read at least twice the rate as we utter and hear spoken language. But many of us have already learned that various audio and video applications have options to speed up (or slow down) the playback rate. Text-to-speech can be easily understood at much faster than normal, rates comparable to skilled silent reading. Podcasts (now in the millions) might be predicting the future. Yes, texting is pervasive but many texters use speech recognition to send these messages, and speech could be used to understand them. We can expect to have a technology translating adumbrated texts to coherent speech, ensuring spoken language its dominance. We think of text demanding much less memory storage but given advances in spoken language technology, the electronic footprint of text to speech can be reduced to be more similar to text.

Understanding tables, graphs and equations will demand some written form. As Manjoo (2021) observed, "I have also had trouble listening to dense, especially technical books." Written language won t disappear but for many, spoken communication will be preferred. We might fall back on print when the material is dense and complex, and simply to take a break from a voice that may be commanding more obedience than we are willing to fork over. But we predict the voice (even if it is synthetic) will carry the day. We know also that the real voice coming from the body will have a game-winning advantage, conveying emotion from multiple sensors including heartrate, sweat glands, and facial expressions that are lost in print. The voice will have a multimodal advantage, when it is in YouTube video guides of information to solve the most esoteric challenge. With one giant leap beyond virtual and augmented displays, an AI implant will connect us with what we need now because it is in constant communication with our own hardware. We are, however, humble in knowing how many predictions go awry.

Our private reading might be preempted by listening, but we won t want to sacri ce reading aloud to our children. This became particularly poignant when we observed the richer language our preschoolers experience in picture books relative to CDS and TV Media (Massaro, 2015; Williams et al., 2022). To preserve this tradition, the pictures in picture books can exist as they do now or they can be presented on screens. Since it is unlikely that we will memorize the many stories, the caregiver could listen along with the child or would have to somehow shadow the audio book while reading it to the child. But the caregivers could listen along with the child, embellishing the story in a variety of ways as they see appropriate. Thus, although listened to rather than read, picture books will continue to enrich our toddlers brain with the richness of formal language.

A potential danger is that dominance of spoken language might lower the quality of our stories. That is, we know writing is mostly rewriting following Nabokov s prescription that reading is re-reading and this dictum might fail to impress the spoken language storyteller that scripting is re-scripting. Books will not have the formal language we value if the author doesn t have the opportunity to embellish and revise their drafts. With the expected advancement of controls over spoken language, the author could choose to stick with the spoken medium (especially if she is illiterate). The rst generation experiencing the impending demise of written language will stick with writing even though the majority of "readers" will be listeners.

Sudowriters

After cautioning storytellers about AI s impending demise of written language, it might seem ironic to question whether AI could aid in our storytelling. To sample what AI might contribute, we input our initial rst draft sentence in a just discovered AI writing application Sudowrite (2022).

> The intersection of Arti cial Intelligence (AI) and Literacy should certainly be a show stopper for the casual reader.

Sudowrite is powered by (and we mean powered by) a supercomputer with 285,000 processors to learn from the existing multiverse of content using, of course, a multilayered neural network. Given a single topic sentence, it spits out some additional text corresponding to a variety of genres that include our senses and various metaphoric domains. Here s what we were able to select from Sudowrite s output delivered more or less instantaneously to our seedling rst sentence.

> One of the most advanced pieces of technology in existence, arti cial intelligence has evolved to reach comprehension for both spoken and written language. It is now able to understand simple questions, like "what is the de nition of the word 'cat ?", and can tell you what it is, or anything else you d ask of it.

> This technology is still in its infancy, but arti cial intelligence is quickly improving. More and more, it s becoming more engrained into our daily lives.

> The AI s purpose is to read things, but it is not a physical piece of machinery. Despite being made of machinery, it has no physical form.

> The concept of Arti cial Intelligence is that, if you have enough information, you can teach it to learn. Eventually, you will have an intelligence that is more sophisticated than humans. At the moment, arti cial intelligence is still in its infancy, but it s quickly becoming more common.

Although the program expects multiple iterations of feeding each output back to Sudowrite, we stopped here. Our gracefully extended deadline became hard and it was somewhat aerie to include Sudowrite as a co-author.

Challenges to AI s role in literacy

Limitations in embodying us

Should we expect a social robot to embody all of the unique characteristics of live tutors anytime soon? Some are optimistic that this goal is within reach (Kurzweil, 2021), whereas others resound a strong negation (Moore, 2016). Linguistic exchanges create a milieu of anxiety shared by both parties, and we are uncertain how to embody this characteristic into robots (Rowe, 2020). Reducing this anxiety during exchanges in conversations and even narratives is married to rewards. A non-linguistic example involves the little hit of dopamine that the infant receives when she wiggles her foot to jiggle the overhead happy birds mobile. Similarly, an interlocutor thrives on resolving uncertainty in the conversational exchange (Ramscar and Port, 2015, 2016).

However, we can expect that an arti cial tutor who only approximates embodiment of a human tutor to be effective in guiding the child or student to learn both language and new knowledge. Can an AI companion ever achieve the linguistic and cognitive capacities that humans nd so natural? Christiansen and Chater (2022) make a strong case that humans are unique in pursuing an iterative language game in which utterances take on meaning whereas AI systems are nothing more that mindless fact-collecting automata that don t have an inkling of what s up. Moore (2016) argues that improvements in speech (and animation/robotics) technologies can outpace a social robot s cognitive and linguistic capacities. This "uncanny valley" would dismantle productive interactions because participants would overestimate the social robot s capability. Humans are fairly robust interlocutors, however, and we might be well-equipped to converse productively with unequal partners. Real life successful examples include interlocutors with language and cultural differences, as well as children, chimpanzees, and dogs. Children are also successful in conversations with single word utterances (Herr-Israel and McCune, 2011; Massaro, 2020). The infamous chimpanzee Nim convinced his caregivers that they were productively conversing even while they were implicitly cuing him for appropriate replies (Massaro, 2020).

Although we are both predicting the future, one of us might actually be correct. As evidence for his pessimism, Moore in 2016 claimed that intelligent assistants, such as Siri, were not very popular but the demographics have changed signi cantly another instance in which it is dif cult to predict the future. As we write this, Siri receives about a billion questions a week; in 2016 Alexa had only 130 interactive skills but now the number is well above 100,000; Google s Home assistant is also popular and of course many people google search by voice and often receive voice replies or links to spoken and video answers. It s expected that over half of the US households have or will soon have smart devices that allow human-machine dialogs via voice. We might also expect that these devices will become more embodied, with virtual agents or even a physical robotic presence such as the intriguing applications of Furhat (2021).

We can expect that virtual tutors rather than physically-present robots would be prevalent because sensory motor interactions are much easier to program virtually than in actual physical substance. Successful interactions between humans and machines have certainly increased dramatically in the last few years, with no sign of diminishing. We nd it more than natural to address Siri or Alexa with a plethora of questions, for example. Also, there is active research and applications extending human-robot interactions to be embodied (Furhat, 2021). As social robots incorporate multisensory perception and action, we can expect much more rewarding dialogs and outcomes.

Retrospective

Arti cial Intelligence and Literacy is expected to continue to provide a synergistic partnership in both reading instruction, reading, and access to written sources spanning the range of human knowledge. Improved machine learning algorithms, technical advances, and computational speed and ef ciency are responsible for these impressive developments. Surprisingly for one of the authors was the realization that speech science was the primary AI contributing to this progress. Cognitive Science tended to partition research in separate speech and reading bins, and most research adhered to this separation. Much earlier, St. Augustine might have noticed the tight relationship between spoken and written language when he was surprised observing his mentor, St. Ambrose, reading silently without moving his lips (Manguel, 1996). Bochner and Bochner (2009) make a strong case against print alone accessing, acquiring, and utilizing linguistic input, based on deaf individuals learning to read (see also Lichtenstein, 1998; Yang, 2008). Some phonological source is needed which is most easily obtained via hearing. Other sources such as visible speech and touch from the face, and cued speech with the hands can help but appear to fall short of audible speech.

One reason that reading may have an advantage over listening is that since writing elicits that little voice in your head, it has two sources of information relative to just one. Utilizing multiple sources of sensory and contextual information in language processing has been a hallmark of recent theories (e.g., Massaro, 1998). Having speech, even if it s internal, as well as the written word on the page, anchors the language processor in more elaborate constraints than either modality alone (There are also rare instances of synesthesia in which speech can elicit written text, as in the mnemonist studied by Luria (1968, 1987). Writing also reduces language change which might be somewhat problematical in a speech limited metaverse.

For reading instruction, AI can automate reading experiences that can combine instruction and assessment in a seamless manner. AI can also optimize what the child is reading in terms of reading dif culty and important content to be learned. Assessment can also be made more ecologically valid and engaging by creative forms of integrated testing during reading. AI also can easily embody a reading tutor to help maintain the student s interest in reading and learning. With individually guided tutoring, reading to learn curricula can exploit the knowledge that has been organized and consolidated in various domains.

References

AI Umpire, 2022. https://sportsnaut.com/robot-umps-to-call-balls-strikes-in-triple-a-in-2022/.

Average Reading Speed, 2022. https://scholarwithin.com/average-reading-speed.

Baldi Tutoring, 2022a. https://cpb-us-e1.wpmucdn.com/sites.ucsc.edu/dist/0/158/ les/2022/02/raylay.mov.

Baldi Tutoring, 2022b. https://www.youtube.com/watch?v BeWxjMhAm-s; https://www.youtube.com/watch?v mDfKdRh5h_g.

Balogh, J., Bernstein, J., Cheng, J., Van Moere, A., Townshend, B., Suzuki, M., 2012. Validation of automated scoring of oral reading. Educ. Psychol. Meas. 72 (3), 435–452.

Belpaeme, T., et al., 2018. Guidelines for designing social robots as second language tutors. Int. J. Soc. Robot. 10 (3), 325–341.

Bernstein, J., Cheng, J., Balogh, J., Rosenfeld, E., 2017. Studies of a self-administered oral reading assessment. In: Proceedings of the 7th ISCA Workshop on Speech and Language Technology in Education, Stockholm, Sweden, pp. 180–184.

Bochner, J.H., Bochner, A.M., 2009. A limitation on reading as a source of linguistic input: evidence from deaf learners. Read. Foreign Lang. 21 (2), 143–158. ISSN 1539-0578.

Boote, B., Agarwal, M., Mostow, J., 2021. Early Prediction of Children s Task Completion in a Tablet Tutor Using Visual Features. The Thirty-Fifth AAAI Conference on Arti cial Intelligence (AAAI-21). Pp. 15761-15762.

Bryant, W., 2017. Developing a strategy for using technology-enhanced items in large-scale standardized tests. Pract. Assess. Res. Eval. 22, 1. https://doi.org/10.7275/70yb-dj34. Available at. https://scholarworks.umass.edu/pare/vol22/iss1/1.

Brysbaert, M., New, B., 2009. Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. Behav. Res. Methods 41, 977–990.

Christiansen, M.H., Chater, N., 2022. The Language Game: How Improvisation Created Language and Changed the World. Basic Books, New York.

Clark, A., Chalmers, D., 1998. The extended mind. Analysis 58 (1), 7–19.

COCA, 1990–2012. Corpus of Contemporary American English (COCA). (1990–2012). Words Added Between 1990 and 2012. Retrieved from. http://corpus.byu.edu/coca/.

Deng, E., Mutlu, B., Mataric, M.J., 2019. Embodiment in socially interactive robots. FNT Robot. 7 (4), 251–356. https://doi.org/10.1561/2300000056.

English Spelling, 2022. https://deepblue.lib.umich.edu/bitstream/handle/2027.42/62008/words.html.

Evans, M.A., Saint-Aubin, J., 2013. Vocabulary acquisition without adult explanations in repeated shared book reading: an eye movement study. J. Educ. Psychol. 105 (3), 596–608. https://doi.org/10.1037/a0032465.

Furhat, 2021. https://furhatrobotics.com/blog/5-ways-social-robots-are-innovating-education/.

Gough, P.B., Tunmer, W.E., 1986. Decoding, reading, and reading disability. Remedial Spec. Educ. 7, 6–10. https://doi.org/10.1177/074193258600700104.

Grice, H.P., 1975. Logic and conversation. In: Martinich, A.P. (Ed.), Philosophy of Language. Oxford University Press, New York, NY, pp. 165–175.

Gurdon, M.C., 2019. The Enchanted Hour: The Miraculous Power of Reading Aloud in the Age of Distraction. Harper, New York.

Hart, B., Risley, T., 1995. Meaningful Differences in the Everyday Experience of Young American Children. Paul H. Brookes, Baltimore, CA.

Herr-Israel, E., McCune, L., 2011. Successive single-word utterances and use of conversational input: a pre-syntactic route to multiword utterances. J. Child Lang. 38 (1), 166–180. https://doi.org/10.1017/S0305000909990237.

Hiebert, E.H., 2005. In pursuit of an effective, ef cient vocabulary curriculum for the elementary grades. In: Hiebert, E.H., Kamil, M. (Eds.), The Teaching and Learning of Vocabulary: Bringing Scienti c Research to Practice. LEA, Mahwah, NJ, USA, pp. 243–263.

History of TtS, 2022. https://www.kurzweiledu.com/about-kurzweil/history-of-text-to-speech.html.

iSPYPhonics Fun, 2021. https://www.thriftbooks.com/series/i-spy-a-book-of-picture-riddles/37646/.

Jaynes, J., 1976. The Origin of Consciousness in the Breakdown of the Bicameral Mind (PDF). Houghton Mifflin, New York. ISBN 0-618-05707-2.

Jeopardy, 2022. https://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html.

Kasap, Z., Magnenat-Thalmann, N., 2012. Building long-term relationships with virtual and robotic characters: the role of remembering. Vis. Comput. 28 (1), 87–97. https://doi.org/10.1007/s00371-011-0630-7.

Klatt, D., 1987. Review of text-to-speech conversion for English. J. Acoust. Soc. Am. 82 (3), 737–933. https://doi.org/10.1121/1.395275. Bibcode:1987ASAJ...82..737K. PMID 2958525.

Kurzweil, R., 2021. https://www.inverse.com/article/30913-8-staggering-predictions-ray-kurzweil.

Li, J., 2015. The bene t of being physically present: a survey of experimental works comparing copresent robots, telepresent robots and virtual agents. Int. J. Hum. Comput. Stud. 77, 23–37. https://doi.org/10.1016/j.ijhcs.2015.01.001.

Lichtenstein, E.H., 1998. The relationships between reading processes and English skills of deaf college students. J. Deaf Stud. Deaf Educ. 3, 80–134.

Luria, A.R., 1968. The Mind of a Mnemonist. Washington Square Press, New York.

Luria, A.R., 1987. The Mind of a Mnemonist. A Little Book About a Vast Memory, With a New Foreword by Jerome S. Bruner. Harvard University Press, Cambridge, MA.

Manguel, A., 1996. A History of Reading. https://www.historyo nformation.com/detail.php?entryid 4341.

Manjoo, F., 2021. https://www.nytimes.com/2021/10/06/opinion/audiobooks-better-than-reading.html.

Massaro, D.W., 1995. From speech-is-special to talking heads: the past to the present. In: Solso, R.L., Massaro, D.W. (Eds.), The Science of the Mind: 2001 and Beyond. Oxford University Press, New York, pp. 203–220.

Massaro, D.W., 1998. Perceiving Talking Faces: From Speech Perception to a Behavioral Principle. MIT Press, Cambridge, Massachusetts.

Massaro, D.W., 2012. Acquiring literacy naturally: behavioral science and technology could empower preschool children to learn to read naturally without instruction. Am. Sci. 100, 324–333.

Massaro, D.W., 2015. Two different communication genres and implications for vocabulary development and learning to read. J. Lit. Res. 47 (4), 505–527. https://doi.org/10.1177/1086296X15627528.

Massaro, D.W., 2017. Reading aloud to children: bene ts and implications for acquiring literacy before schooling begins. Am. J. Psychol. 130 (1), 63–72. https://doi.org/10.5406/amerjpsyc.130.1.0063.

Massaro, D.W., 2020. My words fly up, my thoughts remain below. Words without thoughts never to heaven go. Review of why chimpanzees can t learn language and only humans can (Leonard hastings schoff lectures) by Herbert S. Terrace. Am. J. Psychol. 133, 378–392.

Massaro, D.W., Hestand, J., 1983. Developmental relations between reading ability and knowledge of orthographic structure. Contemp. Educ. Psychol. 8, 174–180.

Massaro, D.W., Ouni, S., Cohen, M.M., Clark, R., 2005. A multilingual embodied conversational agent. In: Sprague, R.H. (Ed.), Proceedings of 38th Annual Hawaii International Conference on System Sciences (HICCS 05) (CD-ROM, 10 Pages). IEEE Computer Society Press, Los Alimitos, CA.

Media, 2022. https://www.commonsensemedia.org/.

Merlin, D., 1991. Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition. Harvard University Press, Cambridge, MA.

Michaelis, J.E., Mutlu, B., 2018. Reading socially: transforming the in-home reading experience with a learning-companion robot. Sci. Robot. 3 (21). Springer robotics.sciencemag.org).

Moby.ReadTM, 2022. www.analyticmeasures.com/moby-read.

Montag, J.L., Jones, M.N., Smith, L.B., 2015. The words children hear: picture books and the statistics for language learning. Psychol. Sci. 26 (9), 1489–1496. https://doi.org/10.1177/095679761559436.

Moore, R.K., 2016. Is Spoken Language All-or-Nothing? Implications for Future Speech-Based Human-Machine Interaction. International Workshop on Spoken Dialogue Systems (IWSDS) 13–16 January 2016, Riekonlinna, Finland.

Mostow, J., 2013. Lessons from project LISTEN: what have we learned from a reading tutor that listens?. In: Arti cial Intelligence in Education, pp. 557–561.

Perfetti, C., Hogaboam, T., 1975. The relationship between single word decoding and reading comprehension skill. J. Educ. Psychol. 67, 461–469.

Phonics Exercises, 2022. https://www.spellingcity.com/phonics.html.

Piasta, S.B., Justice, L.M., McGinty, A.S., Kaderavek, J.N., 2012. Increasing young children s contract with print during shared reading: longitudinal effects on literacy achievement. Child Dev. 83, 810–820. https://doi.org/10.1111/j.1467-8624.2012.01754.x.

Pinker, S., 2007. The Language Instinct: How the Mind Creates Language. Harper, New York.

Project LISTEN s Reading Tutor, 2022. https://www.ri.cmu.edu/project/project-listens-reading-tutor/.

Ramscar, M., Port, R., 2015. Categorization (without categories). In: Dawbroska, E., Divjak, D. (Eds.), Handbook of Cognitive Linguistics. DeGruyter Mouton, Berlin, Germany, pp. 75–99.

Ramscar, M., Port, R., 2016. How spoken languages work in the absence of an inventory of discrete units. Lang. Sci. 53, 58–74.

Read With Me Demo, 2022. http://psyenti cmind.com/wp-content/uploads/2012/07/Keagen2.mov.

Read With Me 2022. https://apps.apple.com/us/app/read-with-me/id494909666.

Readability, 2021. Automatic Readability Checker, a Free Readability Formula Consensus Calculator. Readability Formulas: Automatic Readability Checker. https://readabilityformulas.com/free-readability-formula-tests.php.

Reading or Listening to Books, 2022. https://www.bbc.com/culture/article/20200104-audiobooks-the-rise-and-rise-of-the-books-you-dont-read.

Rowe, B., 2020. The marvelous gadget. Cognitive gadgets: the cultural evolution of thinking by Cecilia Heyes. Am. J. Psychol. 133 (3), 341–378 (Fall 2020).

Sauppé, A., Mutlu, B., 2014. Effective task training strategies for instructional robots. In: Proceedings of Robotics: Science and Systems (RSS 2014).

Sauppé, A., Mutlu, B., 2015. Effective task training strategies for human and robot instructors. Aut. Robots 1–17. https://doi.org/10.1007/s10514-015-9461-0.

Shanahan, T., 2006. The National Reading Panel Report: Practical Advice for Teachers. Learning Point, Naperville, Illinois.

Stephenson, N., 1995. Diamond Age.

Stephenson, N., 2021. Michael Shermer Interview.

SUBTLEX, 2021. https://www.ugent.be/pp/experimentele-psychologie/en/research/documents/subtlexus.

Sudowrite, 2022. https://gpt3demo.com/apps/sudowrite.

TEIs, 2021. https://edulastic.com/blog/technology-enhanced-items-showcase/.

Understand My World, 2022. https://apps.apple.com/us/app/understand-my-world/id1298250806.

van Kleeck, A., Stahl, S.A., Bauer, E.B., 2003. On reading books to children: parents and teachers. Lawrence Erlbaum Associates, Mahwah, NJ.

van Santen, J.P.H., Sproat, R.W., Olive, J.P., Hirschberg, J., 1997. Progress in Speech Synthesis. Springer. ISBN 978-0-387-94701-3.

Venezky, R.L., Massaro, D.W., 1979. The role of orthographic regularity in word recognition. In: Resnick, L.B. (Ed.), Theory and Practice of Early Reading, vol. 1. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 85–107.

Venezky, R.L., Massaro, D.W., 1987. Orthographic structure and spelling-sound regularity in reading English words. In: Allport, A., MacKay, D., Prinz, W., Scheerer, E. (Eds.), Language Perception and Production: Relationships Between Speaking, Reading and Writing. Academic Press, pp. 159–179.

Vygotsky, L., 2022. https://en.wikipedia.org/wiki/Zone_of_proximal_development.

Williams, A., Tudjinovic, E., Baylon, T., Alarcon, J., Bose, R., Massaro, D.W., 2022. Mining Big Data to Analyze Influences on Children s Language Acquisition.

Yang, J.H., 2008. Sign language and oral/written language in deaf education in China. In: Plaza-Pust, C., Morales-Lopes, E. (Eds.), Sign Bilingualism: Language Development, Interaction and Maintenance in Sign Language Contact Situations. John Benjamins, Amsterdam, pp. 297–331.

Yueh, H.P., Lin, W., Wang, S.C., 2020. Reading with robot and human companions in library literacy activities: a comparison study. Br. J. Educ. Technol. 51, 1884–1900.