AAC and Artificial Intelligence (AI)

Samuel C. Sennott, Ph.D. [Assistant Professor of Special Education],
Universal Design Lab Director, College of Education, Portland State University, Post Office Box 751, Portland, Oregon, 97207

Linda Akagi, B.S. [Research Analyst],
Universal Design Lab, Portland State University

Mary Lee [NIH BUILD EXITO Scholar],
Portland State University

Anthony Rhodes, M.S. [Machine Learning Engineer, University Lecturer]
Maseeh Department of Mathematics and Statistics, Portland State University

Abstract

Artificially intelligent tools have given us the capability to use technology to address ever more complex challenges. What are the capabilities, challenges, and hazards of incorporating and developing this technology for augmentative and alternative communication (AAC)? *Artificial Intelligence* can be defined as the capability of a machine to imitate human intelligence. The goal of artificial intelligence is to create machines that can use characteristics of human intelligence to solve problems and adapt to a changing environment. Harnessing the capabilities of AI tools has the potential to accelerate progress in serving individuals with complex communication needs. In this article, we discuss components of AI, including: (a) knowledge representation, (b) reasoning, (c) natural language processing, (d) machine learning, (e) computer vision, and (f) robotics. For each AI component, we delve into the implications, promise, and precautions of that component for AAC.

The purposes of this paper are to (a) describe artificial intelligence (AI), (b) discuss its five main parts, and (c) explore the capabilities, challenges, and hazards associated with incorporating and developing AI for alternative and augmentative communication (AAC) systems. We believe that there is urgency for the AAC field to consider the incredibly powerful computing tools afforded by AI. New (and old) AI tools have the capability of transforming AAC systems as low-tech as non-digital communication boards with words and symbols and as high-tech as computers that employ a human voice for output (Beukelman & Mirenda, 2013). In Table 1 we present examples of different aspects of AI, how these can be used to enhance AAC systems and devices, as well as cautions when considering integrating AI with AAC systems and devices. This information is described in more depth below.

sennott@pdx.edu.

No Disclosures

By gaining knowledge about AI, we hope the reader can claim a seat at this metaphorical table where AAC and AI intersect. Individuals with complex communication needs, their families, their clinicians, and their teachers can and must be at the table to co-design with and inform AI specialists about the participation needs and values of people with complex communication needs.

Introduction to Artificial Intelligence

The ethics of AI are critical to consider from the beginning. We turn to Kai-Fu Lee's words, who in his recent exploration of the fundamental concepts of AI, expressed thoughts about harnessing this powerful technology towards social good and the importance of using love and empathy as a core design principle (Lee, 2018). This parallels the focus on dignity, inclusion, and empowerment that the AAC field has espoused in harnessing assistive technology tools (Brady et al., 2016; Blackstone, Williams, & Wilkins, 2006) for individuals who have communication disorders and needs.

Augmented or artificial intelligence.

Some time ago, Pattie Maes' MIT Media Lab web page included the term *augmented intelligence*, which has since been replaced by newer terminology, *assistive augmentation*. These terms, augmented intelligence and assistive augmentation, speak to how people harness these tools (Huber, Shilkrot, Maes, & Nanayakkara, 2017). The AI tools that are currently available and the rapid innovation that is occurring in the field of AI can help people with complex communication needs overcome barriers. Yet, these terms also serve as a reminder that it is the individual at the heart of efforts to use AI productively to enhance the lives of persons with complex communication needs. In summary, the principle is mindful service towards the communication needs and requirements of people with complex needs.

Intelligence as a concept.

To understand artificial intelligence, it might seem useful to first discuss human intelligence, but human intelligence itself has long been a topic of debate in the field of psychology (Conway & Kovacs, 2015). The concept of intelligence is complicated both by the fact that it is broad and multifaceted, and that it has been used to systematically discriminate against people with disabilities, sometimes resulting in institutionalization, eugenics, forced sterilization, and barriers to education and employment (Wehmeyer, 2013).

Despite the problem of defining and quantifying human intelligence, most important in this context is that artificial intelligence seeks to imitate some characteristics of human intelligence like creativity, language, emotion, self-awareness, learning, reasoning, planning, problem-solving, adaptation, and/or logic (Russell & Norvig, 2010). Those components of intelligence can be important to leverage towards serving people with complex communication needs, as discussed below.

Artificial intelligence defined.

Artificial Intelligence (AI) can be defined as the capability of a machine to imitate aspects of human intelligence. The goal of artificial intelligence is to create machines that can use characteristics of human intelligence to solve problems and adapt to a changing environment (Boden, 2018). Artificial intelligence has a long history with roots in many disciplines including mathematics, philosophy, psychology, neuroscience, linguistics, and economics, as well as computer engineering (Domingos, 2018; Russell & Norvig, 2010). Alan Turing's (1950) landmark paper introduced many concepts that would become the basis for fields associated with AI, including machine learning and natural language processing (Russell & Norvig, 2010). In Table 2, we provide a short list of varied resources for the interested reader to learn more about AI.

Components of Artificial Intelligence

For the purposes of this paper, AI can be segmented into several components: (a) knowledge representation, (b) reasoning, (c) natural language processing, (d) machine learning, (e) computer vision, and (f) robotics (Russell & Norvig, 2010; Turing, 1950). Although there is much overlap between these categories, and while it's not a perfect classification scheme, these were the main concepts that Alan Turing (1950) predicted would be necessary for a computer to pass his famous artificial intelligence test, the Imitation Game (Russell & Norvig, 2010). Although it's no longer considered the gold standard of testing for artificial intelligence, the framework he used remains a useful way to look at AI. For each component, we discuss the implications, promise, and cautions for AAC.

Knowledge representation, reasoning, and AAC

One part of AI is concerned with translating information into knowledge and storing that knowledge in a way that programs can understand and extract meaning. We call these components of AI *knowledge representation* and *reasoning*. Once a program has information translated into knowledge, it can organize objects into categories that make sense to it, and we can ask programs to reason about and make decisions and predictions based on that knowledge (Russell & Norvig, 2010). For example, a logic application used to translate information might appear as: If it is raining, then it is not mostly sunny. If it is raining, then the grass is wet. If the grass is wet, one must wear boots. Given these simplified facts and an input (raining or mostly sunny), the program can reason through them and decide whether we should wear the boots or not. In this way, bit by bit, knowledge bases are formed and the computer can make ever more complex and useful decisions.

Knowledge representation is an important part of creating a program that can make decisions and predict outcomes. Examples are virtual assistants like Apple's SIRI, Amazon's Alexa, and Google Assistant and their capacity for storage and retrieval of knowledge representations like current movies, businesses, and restaurants, and ability to keep track of contextual information, as the system must use aspects of reason to apply the contextual information it has stored to optimize for things such as if you are searching for a restaurant to go to in the morning, you probably are looking for breakfast or brunch, and not lunch or dinner.

AAC opportunities.—In considering AAC opportunities, knowledge representation and reasoning is a vital and promising component for assisting in the areas of language representation, AAC system vocabulary, cognitive supports, skill building, improving communication partner knowledge, virtual coaching, and treatment adherence (Light et al., 2019; Topol, 2019). Consider again that on a fundamental level, AAC systems are at their essence a form of symbolic knowledge representation (Beukelman & Mirenda, 2013; Baker & Nyberg, 1990).

How can the AAC community consider knowledge representation AI tools? One example speaks to how there are now many popular ways to organize computer- and paper-based AAC systems and the storage and retrieval of words, phrases, and longer utterances. From core vocabulary systems to pragmatically driven systems, vocabulary developers have carefully mapped the knowledge representations in the form of language structures designed for communicative expression (Beukelman & Mirenda, 2013). Another more specific example is Waller's (2018) innovations describing multiple ways that individuals with complex communication needs can store and use narratives or personal stories. Waller described a system in which there is capacity for tagging people, objects, and locations, and the system generates simple phrases that can be used to help create a personal narrative. The advantages of a system like this are that it provides scaffolding of the conversational exchange by offering suggestions for utterance turns that reflect the natural back and forth of conversation, and that it reduces the cognitive and motoric loads of generating text.

Another example is the promising idea of automating the knowledge representation and reasoning process of AAC system selection, which has been explored in an expert system model of AAC assessment (e.g., Napper, Robey, & McAfee, 1989). Expert systems models emulate what an AAC specialist may perform during an AAC assessment. Clinically, specialists or experts make decisions based on their knowledge of the person they are serving and a set of goals or rules that they believe are reasonable. AI expert systems attempt to model the work environment in order to provide suggestions during the evaluation process.

However, a precaution with using such an expert system is that choosing AAC systems involves a complex web of factors inherent in the participation model (Beukelman & Mirenda, 2013). The participation model includes considerations based on the identified participation patterns and requirements of the individual with complex communication needs. Nevertheless, in identifying barriers arising from team member knowledge, skills, and attitudes, an AI approach could include suggestions for questions to ask about to address such barriers or suggest resources to use to mitigate the barriers. Additionally, this same approach could be used to consider access barriers that include motor, linguistic, cognitive, literacy, and sensory/perceptual capabilities.

From a system perspective, given the large number of variables and considerations an AAC specialist must consider, this classic or symbolic approach to developing an artificial intelligence tool is daunting. Just like when a human makes a decision, AI programs do better if they have some background knowledge about the subject, goals, and can decide to complete actions based on past results. Finding an appropriate way to represent knowledge

to the program is vital in avoiding what is called the "frame problem" (Boden, 2018). The frame problem is what happens when the computer doesn't have a frame of reference for whatever it is facing. This happens because computers aren't very good at determining relevance in situations. This problem could be avoided by explicitly programming the computer with instructions for every possible contingency, but of course, we don't always know everything that might happen. This is why the sort of logic we talked about in the rainy and sunny discussion becomes useful: we don't need to explicitly tell the program what to do if it's sunny. The program is able to deduce a solution based on what it already knows. To reiterate, AAC is complex and the frame problem is very much one to pay attention to ethically when considering expert systems models.

Moving from assessment to intervention, another thought-provoking example area is that of communication partner instructional intervention, which is emerging as an evidence-based practice (Beukelman & Mirenda, 2013; Sennott, Light, & McNaughton, 2016). AI tools can be used to aid in the knowledge and skill building process for individuals who interact with an individual with complex communication needs who also uses AAC. Why is this important to consider in AAC? As Topol (2019) describes, the rising costs and diminishing returns of healthcare is a motivator for considering the use of AI technology to optimize and improve patient outcomes. As we face similar challenges in AAC applications, we must question how we are planning to meet the needs of people who use AAC. For instance, in schools, paraprofessionals who spend the most amount of time with children with complex communication needs who use AAC are often lacking in training and support for AAC implementation. In seeking synergy between communication partners by employing AI tools, we could help bridge the gap in their training, similar to the approach Schlosser et al. (2015) describe as just-in-time supports. In that approach, a practitioner watches a video about a specific intervention technique just prior to working with a child. This mirrors the approach used by Sennott, Ferrari, Crest, Fogarty, and Hix-Small (2017) with their check-in, check-out for AAC system that provides intervention support to communication partners just prior to or during their engagement with a client, including video modeling of communication strategies to achieve successful interactions.

These first fruits of AI-based automated knowledge and reasoning in AAC are even more interesting when considering how modern machine learning techniques are applied (Russel & Norvig, 2010; Boden, 2018). With a machine learning system, the results of the system can further enhance the capabilities of AAC. For instance, if a certain video model is manually tagged during check-in, check-out for AAC as useful or watched by a large number of users with subsequent user success, this gives an opportunity to train the computer to learn that this video is useful. In summary, this gives us the opportunity to create a learning system that improves with usage.

Natural Language Processing and AAC

As a field of study, *natural language processing* is defined by computer processes that focus on recognizing and generating natural human language. Historically, there has been much difficulty in determining how to represent the complexities of language in a way that computers could understand. In order to truly understand language, a program would need to

have at least a general understanding of the world in which it exists, as most language is concerned with objects and their relation to each other and the environment (Russell & Norvig, 2010). Older iterations of natural language processing involved closely analyzing grammar and syntax in order to solve the problems posed by context and relevance. These solutions did not work very well, as it led to too much complexity for the computer to handle. Modern natural language processing has partially abandoned this technique, and now relies on machine learning and statistical analysis to find and predict linguistic patterns (Boden, 2018).

AAC opportunities.—AAC system developers have been harnessing natural language processing tools in the form of keyboard arrangements, word and message prediction, word completion, icon-based prediction, and various methods for automated part of speech tagging for decades and new machine learning-based capabilities with voice recognition are revolutionizing speech recognition for people with dysarthria and other speech sound production disorders (Dudy & Bedrick, 2018; Fager et al., 2019; Higginbotham, Lesher, Moulton, & Roark, 2012; Langer & Hickey, 1999). One aspect of natural language processing with renewed interest involves contextually driven AAC system adaptation and prediction. For instance, AAC systems can be adapted or triggered by specified contextual elements such as location, time, prior language use, communication partner factors, conversation content, and internet based data (Higginbotham, Lesher, Moulton, & Roark, 2012; Judge, Hawley, Cunningham, & Kirton, 2015).

An appealing aspect of contextually driven prediction models in AAC is leveraging the speaking partner to aid in the generation of contextually relevant words and utterances, such as the approach used by Wisenburn and Higginbotham (2008). They describe research about their Converser program, which cleverly captured and parsed noun phrases from partners' communication turns, which could then be used by the person with complex communication needs. While the study documented a confirmation of some conversational turn rate improvement, the authors noted limitations with the automatic speech recognition, which negatively impacted the prediction available for the individual with impairment to use. Speech recognition has improved exponentially in the last decade since their study was published and the field eagerly anticipates new possibilities, such as that offered by Fager, Fried-Oken, Jakobs, and Beukelman (2019), who describe a recent example in the Smart Predict app developed by Invotek where a communication partner can send words, phrases, and sentences to the user's AAC prediction interface.

Whole-utterance approaches in AAC have been explored (e.g., Higginbotham & Wilkins, 2006; Todman, Alm, Higginbotham, & File, 2008), but recently there has been an emphasis on high frequency or core word approaches to vocabulary organization based on syntactic and categorical organization. Todman, Alm, Higginbotham, and File (2008) describe the balance between social interaction goals, conversational turn rate enhancement benefits, and precision and suggested that new natural language processing tools and internet connectivity could power utterance-based systems in novel ways. For instance, recently, the research institute OpenAI unveiled a general language model that it trained on over 40GB of text scraped from the web that is able to predict the next word in a given phrase (Radford et al., 2019). The OpenAI model, called GPT-2, can additionally generate much longer

continuations of a prompt, and match the style and content of the original prompt. It is not always successful, offering reasonable samples up to 50% of the time (Radford et al.). These results are promising, and could help improve conversational agents and computer speech recognition, but if it were to fall into the wrong hands, it could be used for malicious purposes, such as generating false news reports.

Voice input remains difficult for natural language processing programs, as there are so many subtleties and differences in human speech, though ever increasingly some personal assistant systems are able to learn an individual's unique voice, speech pattern, and accent (Boden, 2018). Voice recognition for people with dysarthria is another exiting domain that leverages innovations in machine learning and natural language processing. For example, Google's new Project Euphonia attempts to use modern machine learning techniques to create unprecedented levels of accuracy for this population (Cattiau, 2019). In a demonstration at their recent conference this past spring, Google presented a video featuring a man with dysarthria whose voice was recognized with a high degree of accuracy when using Project Euphonia. Google posted a website where community volunteers can sign up to be a part of the project (www.blog.google/outreach-initiatives/accessibility/impaired-speech-recognition).

Machine Learning and AAC

Although there are many kinds of *machine learning* algorithms and strategies, a program can be said to be learning if the system is observing its performance on tasks and using that knowledge to perform better in the future (Russell & Norvig, 2010). Alan Turing (1950) described some of the difficulties with creating programs that know everything, such as the difficulty in storing all the information required and the length of time it would take to program it into a machine. He suggested that, rather than trying to replicate an adult human brain, we instead create a "child machine" which could be taught and guided in its studies using rewards and punishments. This is the basic idea behind machine learning and has strong parallels to the fundamental principles of special education.

Approaches to machine learning.—There are several approaches to the process of creating a machine learning program that can learn and decide outcomes based on its input. A supervised machine learning model is given examples of already labeled training data (an example would be machine scoring of writing samples using a rubric that has been previously used to score a training set of papers and these data are used by the computer to score more papers using the same rubric), and then uses what it knows to categorize other data that is unlabeled. Unsupervised machine learning occurs when the program is given unlabeled data, and by recognizing similarities and differences, it finds patterns itself. There is also a reinforcement learning model, in which the program learns by judging if a decision is favorable based on its results.

Artificial neural networks include a method of machine learning that is loosely based on the organization of the human brain and uses mathematical models to predict outcomes and categorize items. Artificial neural networks or just, neural networks, are composed of nodes linked together, and information is passed through the nodes. Each node has a different

weight, or level of importance, and they are constantly being updated as new information is processed (Boden, 2018). When a node is updated, the network makes a guess about what to do next. As it makes guesses, some of which produce better output than others, the network learns that certain decisions lead to more errors, and adjusts the weights of the nodes accordingly. This process continues throughout the network until the output is as error-free as possible (Boden, 2018). Neural networks are used for categorization, predictions, and decision making, among many other things. In particular, speech-language pathologists may be familiar with connectionist or parallel distributed processing models of language learning and language acquisition: These models employ neural networks for modeling sentence processing, semantic learning, and phonological processing, among other areas (Joanisse & McClelland, 2015).

Deep learning is a method of machine learning that uses the artificial neural network technique, but has more than one, and often many, hidden "layers" of abstraction. Each layer goes through its own learning process, and the outputs from the learning process of one layer become the input for the next layer (Boden, 2018). Deep learning is useful in situations in which there is a great deal of unstructured and unlabeled data, as it is able to create its own categories from the raw data. Deep learning has many applications in image and speech recognition, as well as categorization.

AAC opportunities.—For AAC, machine learning and, specifically, deep learning approaches have impacted multiple domains including speech synthesis and alternative access to AAC. The mind set for this component of AI and AAC is to simply conceptualize areas where learning is needed. Speech synthesis work in AAC is providing increased options for multilingual speech synthesis voices, custom voice options, and the capacity to accommodate language growth over time (Mills, Bunnell, & Patel, 2014; Pullin, Treviranus, Patel, & Higginbotham, 2017). The VocalID project targets using machine learning approaches to creating personalized voices for augmented communicators and combines vocal qualities from the person with complex communication needs and voice donors to create a novel voice (Mills, Bunnell & Patel, 2014). The concept of personalized voices was made popular when Cereproc created a custom voice for the famous film critic Roger Ebert (2012), who notably abandoned the custom voice in favor of a generic voice. Recently, the Acapela Group has promoted their my-own-voice service, where people can leverage a deep learning toolset to (a) record themselves reading between 350 to 1500 sentences, (b) listen to the synthetic voice that is created, and then (c) purchase access to this voice to use across select software operating systems and supported applications (Malfrere et al., 2016).

This work related to speech synthesis is not confined to the AAC space; for instance, Adobe's Project VoCo focuses on rapidly creating novel voices and audio that can be used in a myriad of ways, including editing a video to change what the speaker says (Jin, Mysore, Diverdi, Lu, & Finkelstein, 2017). From a precautionary and ethical standpoint in both AAC and beyond, the socially positive contributions of the ability to easily create a speech synthesis voice are predicated on trust that this transformative technology will be used for social good. Obviously, speech synthesis technology could be harnessed to spread disinformation, fake news, and overall distrust in media. Additionally, for users of AAC, the ability to copy a voice quickly and affordably could lead to bullying, where individuals

could use the voice selected by a person who uses AAC for malicious intent. Other precautions in adopting machine learning are the equity issues around the current cost of these specialized voices, similar to overall equity issues around acquiring expensive AAC systems and devices. Certainly, the need and interest is there for customized voices (Mills, Bunnell, & Patel, 2014; Pullin, Treviranus, Patel, & Higginbotham, 2017). The potential for voices that change over time and can be refined by the user to meet their unique preferences is exciting given considerations of individuality, growth, and culture. Yet, currently it is expensive for people to employ, creating barriers to access for many.

Machine learning is a toolset that could improve alternative access strategies in AAC, including traditional alternative access techniques and new brain-computer interfaces. Alternative access to AAC is defined as modalities or methods that are used when touch or pointing is not possible (Boster & McCarthy, 2017; Higgenbothem, Shane, Russell, & Caves, 2007). The Higginbotham et al. (2007) review provides definitions of access, an overview of physical demands of access, and a description of various AAC access options set in a historical context. The paper sets parameters for considering motor skills and development as a means of accessing AAC and highlights important challenges that machine learning tools are well suited to help overcome. While we work clinically with the AAC tools we have today, rapid change has the potential to bring brain-computer interfaces to the forefront of AAC (Fager et al., 2019).

Specifically, Higgenbothem et al. (2007) and Fager et al. (2019) both describe the role of machine learning in movement sensing technologies and the tuning/calibration elements that can be critical to alternative access modalities. While this topic of sensors, alternative access, and AAC is certainly deserving of at least a book or two, one exciting area in AAC is motion and gesture recognition that can be used as a trigger for message selection. Higgenbothem et al. (2007) described a key benefit of this technology is that machine learning or manual calibration can be used to interpret unique movements of a person with complex communication needs instead of having the individual struggle to adopt to a standard interface tool. Fager et al. (2019) described more specific additional benefits in this area, such as detecting unintentional versus intentional movements and 3D sensors that allow for a wearable device to eliminate the need for such precise placement of a sensor that is so common for users of alterative access. In our experience, these little tweaks of the technology accrue large benefits, such as with young children we serve who use switch access and need precise positioning. While the ideal of these types of adaptive alternative access tools has been discussed for decades, very recently, Google, at their annual developer conference, announced a major gesture recognition project targeting the communication and access needs of individuals with motor challenges (Cattiau, 2019).

As an alternative access method, scanning has always been considered a useful, but slow method for many people with complex communication needs. One quantitative metric for average speeds is shared by Koester and Arthanat (2018), who report mean text entry rates of onscreen keyboard scanning at 1.7 words per minute in their systematic review. Speed in communication is an important challenge in AAC using alternative access. Huffman scanning (Roark, Fried-Oken, & Gibbons, 2015), which is an approach that potentially provides rate improvement through the use of natural language processing innovations made

possible by machine learning, instead of using the typical linear or row and column scanning patterns, uses a set of binary choices for pairing down the selection field systematically powered by a prediction model. Recently, the Huffman scanning approach appeared as an option in Google's Android operating system. An additional example of machine learning enhancements for alternative access is the Google (2018) collaboration with Tania Finlayson, who is a software developer with complex communication needs. The Morse code interface that runs on the Google Android keyboard allows for machine learning-powered word prediction to be used with Morse code.

Precautions for machine learning approaches to AAC alternative access include considering the automaticity that individuals with complex communication needs may acquire and how the use of machine learning-powered models may interact with that fluency. For instance, the team that redesigned Professor Stephen Hawking's AAC system from Intel Labs described how not every AI powered innovation was preferred by Hawking (Medeiros, 2015). Basically, users develop motor memory for their preferred access method. Using a new access approach was recently described to one of the authors as moving the brakes in a car to a new position and then expecting someone to drive that car effortlessly. Other important precautions include considering the cultural and developmental appropriateness of machine learning models, which also applies to much of our consideration of AI in this paper. One example of this risk is if word or message prediction models were consistently generating words or phrases that were culturally or geographically inappropriate. With this same example, developmental appropriateness could be compromised if the person is given access to words that just would not be suitable for the individual's age, such as a child using technical jargon. This is a very sensitive concept that must be considered on a case-by-case basis.

Computer Vision and AAC

Computer vision is the ability of a computer program to understand and process visual information. This is done by using neural networks to process information pixel by pixel. Processing visual information like this has all sorts of applications, from being able to recognize and identify faces, to extrapolating information from flat images taken at different angles into a 2D or 3D model (Russell & Norvig, 2010). Computer vision can be used to classify images, based on characteristics that they share. You might see this happening automatically if you use photo management services like Google Photos, which employs object identification and facial recognition to group albums, suggest who to share the photos with based on the people in them, or other content specific actions. As another novel example, in 2018, Matt Reed, a creative technologist, created a robot that uses computer vision to find Waldo in the popular children's book, *Where's Waldo* (Lee, 2018). The device uses a camera to scan images and look for matches in Google's AutoML Vision service, a machine learning model that can recognize the character. If a match is found, a hand attached to a robot arm which is connected to the camera will physically point out where Waldo occurs in the image.

AAC opportunities.—In AAC, the ability for computers to process and learn from visual information unlocks unprecedented language learning and access tools in the domains of

visual scene displays, learning materials, symbol systems, and eye tracking. Although the Where's Waldo example mentioned above is a simplistic, tools like this could have great potential to be employed to enhance AAC system development. Visual scene displays (VSDs) are prime candidates for making use of innovations in computer vision, specifically using this type of image recognition technology to automate parts of generating and organizing scenes (Light et al., 2019). Attainment Company's *GoVisual* app for Apple iOS demonstrates an early attempt at harnessing this type of computer vision-powered VSD creation as it gives suggestions for items over which to create hotspots.

Tintarev, Reiter, Black, Waller, and Reddington (2016) describe an innovative approach to leveraging computer vision by gathering photos and videos combined with other sensor data and natural language processing tools to aid story creation by a child with complex communication needs in their system called, *How was school today?* One of the exciting aspects of this toolset is the potential for independence and self-determination that it unlocks. To take this idea further, imagine being a child with complex communication needs entering a classroom with the capability of snapping a photo and the image is used to automatically suggest words that could be used by the child to communicate. Computer vision and image recognition makes this possible.

An important precaution with computer vision tools and the storytelling implementation described above is that the automation, if left unchecked, could actually be sharing information that the individual would rather keep private. Privacy regarding photos is an important overall precaution when it comes to AI. Additionally, this delicate balance between the benefits of automation and the potential drawbacks of lowering self-determination must be considered.

Robotics and AAC

We would not have necessarily thought to include a section on robotics, but very recently, one of the authors visited a speech therapy clinic and observed a social robot that is a core part of the therapeutic approach. Fear among clinicians that they will be replaced by robots is not warranted, at least not for the foreseeable future, as the use of robots in the provision of services to individuals who use AAC is quite limited at this point.

Robot defined.—A robot can be defined as a physical agent that can interact with and affect its environment (Russell & Norvig, 2010). Robots don't necessarily possess artificial intelligence, but robots that lack AI are very limited in their actions, as they can only perform a constrained set of actions that are specifically programmed. For instance, one could program a robotic toy car to drive in a circle, but if it hits a wall, the wheels would simply keep moving forward until it runs out of battery power. The toy car could be pre-programed to try to back up when it hits a wall, but what happens next time? It will just run into the wall again. This is where AI in robotics comes in handy. Intelligent robots can complete tasks, such as locomotion, by perceiving the environment with various input sensors and using that information to plan and control their actions. Robots have applications in many sectors, such as healthcare, transportation, and manufacturing, and are

predicted to become more advanced and commonplace in the coming years (Torresen, 2018).

AAC opportunities.—One of the areas that seemingly impacts AAC the most is social robotics and robot-assisted language learning (Breazeal, 2004; van den Berghe, Verhagen, Oudgenoeg-Paz, van der Ven, & Leseman, 2019; Dawe, Sutherland, Barco, & Broadbent, 2019. These areas harness the capabilities of AI, specifically natural language processing, to create robots that people can interact with and learn with in a naturalistic way. Research in creating social robots is growing overall, and for people with autism and other disabilities, social robots can be a unique and engaging tool for learning and communication (Breazeal, 2004). Robot-assisted language learning has been used to target vocabulary, reading skills, expressive spoken language, and sign language (van den Berghe et al., 2019). In a related fashion, for individuals using AAC, recent research using popular technologies such as the Amazon Echo and Show demonstrate some degree of proof of concept that conversations with natural language processing tools can be useful for individuals with complex communication needs by creating engaging contexts for communication such as turning on music and other entertainment, discussing current events, and something as simple as discussing a riddle or joke (Allen, Shane, & Schlosser, 2018; O'Brien et al., 2017). In summary, social robots hold promise in AAC, overall language learning, and in healthcare (Breazeal, 2004; van den Berghe et al., 2019; Dawe, Sutherland, Barco, & Broadbent, 2019).

However, an important precaution to consider is that professionals don't always necessarily see robots in the role of aiding with communication and we may hold many questions, reservations, and concerns about the limits of what robots can do to facilitate language acquisition (Diep, Cabibihan, & Wolbring, 2015). Instead, many currently see robots capable of helping with mechanical tasks. In one such example of that use, Galloway, Ryu, and Agrawal (2008) published a brilliantly titled paper, *Babies driving robots: Self-generated mobility in very young infants,* which described a feasibility study of infants using a joystick to drive themselves around connected to a small "friendly" robot. Galloway and colleagues later developed the Go-Baby-Go program where they adapt toddler electric ride-on cars to be switch controlled. Logan et al. (2017) describe the Go-Baby-Go cars and a related technology, *Throw-Baby-Throw,* which adapts an automatic ball throwing mechanism. This example serves a positive proof-of-concept for using robots to aid in participation, which both creates contexts for language learning and use and gives the myriad of benefits from self-powered mobility. However, caution should be taken as using these types of tools could create potentially problematic or even dangerous situations if the individuals or the technology led to reckless behavior. For instance, while switch-adapted ball throwing could be an incredibly enriching activity to build communication and social skills, it could also lead to someone or something fragile being hit by a ball.

Conclusion

Harnessing the capabilities of AI tools has the potential to accelerate the progress in serving individuals with complex communication needs who require AAC. Through our study of AI-enhanced tools, we have become very excited with the almost magical capabilities of the technologies available. The concept of assistive augmentation that Huber, Shilkrot, Maes,

and Nanayakkara (2017) describe is one where tools that allow "machine learning [to] seamlessly integrate with a user's mind, body and behavior in this very way–providing enhanced physical, sensorial, and cognitive capabilities (p.1)."

Yet, we also have been humbled into realizing that these powerful tools are tools nonetheless, and it is the purpose, intent, and cultural sensitivity with which they are applied that should drive their consideration, application, and innovation. We believe that people with complex communication needs and those families and practitioners who support them, including speech-language pathologists and AAC specialists, benefit from learning about artificial intelligence because awareness opens opportunities to leverage these impactful tools. One thing to realize as the field takes next steps into AI-enhanced AAC systems and devices is that innovators have been harnessing AI tools since the inception of AAC (Baker & Nyberg, 1990; Higginbotham, Lesher, Moulton, & Roark, 2012; Langer & Hickey, 1999; Napper, Robey, & McAfee, 1989; Vanderheiden, 2002). It is just that now, the tools have grown exponentially in power and widespread availability, giving us urgency as a field to help give voice in shaping how these tools will be applied in AAC (Domingos, 2018; Russell & Norvig, 2010).

Let us close out our discussion by looking at a final example. In this example we consider AAC that involves using the telephone, which can often pose a significant barrier to individuals with complex communication needs. Let us empathize with Tyler, an AAC user, whose experience is shared by Howery (2017):

> Yet the voice from the device may also mask my presence. One time I tried to use my device to call Handi-Bus. I called them and somebody picked up the phone at Handi-Bus. I said I want to be picked up this Friday at 1:30. My address is 3–4–5–3 Apple Way. The Handi-Bus person said, "What do you want?" I repeated my message: I want to be picked up this Friday at 1:30. My address is 3–4–5–3 Apple Way. The line went dead. Maybe they thought I was a crank pot? I don't know. Anyway, I thought okay, that didn't work… next… I guess I wait till Mom comes home and she can call them. I think they will know she is a real person (Howery, 2017, p. 140).

Speaking from the second author's perspective as an adult woman with cerebral palsy who has complex communication needs and who has always been challenged by phone conversations, I can relate to Tyler's experience. The frustrations that boil up in moments like that described can drastically impact someone's confidence.

We secretly wish for a society where someone answering a phone would nearly instantly realize that they were speaking with someone using AAC and have the requisite patience. Fortunately, we can look to a recent AI announcement from Google that tackles this challenge. Google Duplex is an AI system with cutting edge AI for accomplishing real-world tasks over the phone. AI technologies like this have the potential to positively impact pragmatic communication. This new AI tool allows for, when prompted, initiating a phone call and helping set up appointments on an individuals' behalf using deep learning, natural language processing, and speech synthesis. This "assistive agent" could be expanded to address a range of tasks for individuals with complex communication needs. While there

may be other ways to meet this challenge that do not involve AI (and we would agree), the critical point is that advances in computing are bringing these capabilities to mass markets. Now marks a very important time for the AAC field to claim a seat at the table, helping shape how these tools will be created and used to meet the diverse needs of consumers.

In summary, AAC can change the course of development, socialization, education, vocation, and community inclusion (Beukelman & Mirenda, 2013). So, from an ethical perspective, the social importance of AAC is very clear. AAC systems and devices powered by various AI tools hold potential to help give people with complex communication needs enhanced pathways to solve the participation challenges they face when their speech and/or language capabilities do not allow them to fulfil their communication needs. Hopefully, readers will join us at the metaphorical table, as we get set to create the future with tools and examples of AAC powered by AI.

References

Baker B, & Nyberg E (1990, 11). Semantic compaction: A basic technology for artificial intelligence in AAC. Presentation at the 5th Annual Minspeak Conference.

Boden MA (2018). Artificial intelligence: A very short introduction Oxford, United Kingdom: Oxford University Press.

Boster JB, & McCarthy JW (2017). When you can't touch a touch screen. Seminars in Speech and Language, 38(4), 286–296. 10.1055/s-0037-1604276 [PubMed: 28892835]

Blackstone SW, Williams MB, & Wilkins DP (2007). Key principles underlying research and practice in AAC. Augmentative and Alternative Communication, 23(3), 191–203. 10.1080/07434610701553684 [PubMed: 17701739]

Brady NC, Bruce S, Goldman A, Erickson K, Mineo B, Ogletree BT, … Wilkinson K (2016). Communication services and supports for individuals with severe disabilities: Guidance for assessment and intervention. American Journal on Intellectual and Developmental Disabilities, 121(2), 121–138. 10.1352/1944-7558-121.2.121 [PubMed: 26914467]

Breazeal C (2004). Designing sociable robots Cambridge, MA: MIT Press.

Cattiau J (2019, May 7). How AI can improve products for people with impaired speech Retrieved May 15, 2019, from Google website: https://www.blog.google/outreach-initiatives/accessibility/impaired-speech-recognition/

Conway RA, & Kovacs K (2015). New and emerging models of human intelligence. WIREs Cognitive Science, 6(5), 419–426.10.1002/wcs.1356 [PubMed: 26267702]

Dawe J, Sutherland C, Barco A, & Broadbent E (2019). Can social robots help children in healthcare contexts? A scoping review. BMJ Paediatrics Open, 3(1) 1–16. 10.1136/bmjpo-2018-000371

Diep L, Cabibihan J-J, & Wolbring G (2015). Social robots: Views of special education Teachers. Proceedings of the 3rd Workshop on ICTs for Improving Patients Rehabilitation Research Techniques, 160–163. 10.1145/2838944.2838983

Dudy S, & Bedrick S (2018, 7). Compositional language modeling for icon-based augmentative and alternative communication. Proceedings of the Workshop on Deep Learning Approaches for Low-Resource NLP, 25–32. Melbourne: Association for Computational Linguistics. Retrieved from aclweb.org/anthology/papers/W/W18/W18-3404/

Domingos P (2018). The master algorithm: How the quest for the ultimate learning machine will remake our world New York: Basic Books.

Ebert R (2012). Life itself: A memoir New York: Grand Central.

Fager S, Fried-Oken M, Jakobs T, & Beukelman DR (2019). New and emerging access technologies for adults with complex communication needs and severe motor impairments: State of the science. Augmentative and Alternative Communication, 35(1), 13–25. [PubMed: 30663899]

Galloway JC, Ryu J-C, & Agrawal SK (2008). Babies driving robots: Self-generated mobility in very young infants. Intelligent Service Robotics, 1(2), 123–134. 10.1007/s11370-007-0011-2

Google (2018). Hello Morse Retrieved from experiments.withgoogle.com/collection/morse

Higginbotham DJ, Lesher GW, Moulton BJ, & Roark B (2012). The application of natural language processing to augmentative and alternative communication. Assistive Technology, 24(1), 14–24. 10.1080/10400435.2011.648714

Higginbotham DJ, Shane H, Russell S, & Caves K (2007). Access to AAC: Present, past, and future. Augmentative and Alternative Communication, 23(3), 243–257. 10.1080/07434610701571058 [PubMed: 17701743]

Higginbotham DJ, & Wilkins DP (2006). The short story of Frametalker: An interactive AAC Device. Perspectives on Augmentative and Alternative Communication, 15(1), 18–22. 10.1044/aac15.1.18

Howery KL (2017). The lived experience of using a speech-generating device. Unpublished doctoral dissertation 10.7939/R32N4ZW4Q

Jin Z, Mysore GJ, Diverdi S, Lu J, & Finkelstein A (2017). VoCo: Text-based insertion and replacement in audio narration. ACM Transactions on Graphics, 36(4), 1–13. 10.1145/3072959.3073702

Joanisse MF, & McClelland JL (2015). Connectionist perspectives on language learning, representation and processing. Wiley Interdisciplinary Reviews. Cognitive Science, 6(3), 235–247. 10.1002/wcs.1340 [PubMed: 26263227]

Judge S, Hawley MS, Cunningham S, & Kirton A (2015). What is the potential for context aware communication aids? Journal of Medical Engineering & Technology, 39(7), 448–453. 10.3109/03091902.2015.1088091

Katz DS, & Some RR (2003). NASA advances robotic space exploration. Computer, 36(1), 52–61. 10.1109/MC.2003.1160056

Koester HH, & Arthanat S (2018). Text entry rate of access interfaces used by people with physical disabilities: A systematic review. Assistive Technology, 30(3), 151–163. [PubMed: 28368689]

Knight W (2019). An AI that writes convincing prose risks mass-producing fake news. MIT Technology Review Retrieved from https://www.technologyreview.com/s/612960/an-ai-tool-auto-generates-fake-news-bogus-tweets-and-plenty-of-gibberish/

Latson J (2015, 2). Did Deep Blue beat Kasparov because of a system glitch? Time Retrieved from http://time.com/3705316/deep-blue-kasparov/

Lee D (2018). This robot uses AI to find Waldo, thereby ruining Where's Waldo. The Verge Retrieved from www.theverge.com/circuitbreaker/2018/8/8/17665268/wheres-waldo-finding-robot-google-cloud-automachinelearning-ai

Lee K (2018). How AI can save our humanity [Video file] Retrieved from https://www.ted.com/talks/kai_fu_lee_how_ai_can_save_our_humanity

Light J, McNaughton D, Beukelman D, Fager SK, Fried-Oken M, Jakobs T, & Jakobs E (2019). Challenges and opportunities in augmentative and alternative communication: Research and technology development to enhance communication and participation for individuals with complex communication needs. Augmentative and Alternative Communication Advance online publication. 10.1080/07434618.2018.1556732

Lu H, Li Y, Chen M, Kim H, & Serikawa S (2018). Brain intelligence: Go beyond artificial intelligence. Mobile Networks and Applications, 23(2), 368–375. 10.1007/s11036-017-0932-8

Malfrere F, Deroo O, Franques E, Hourez J, Mazars N, Pagel V, & Wilfart G (2016). My-Own-Voice: A web service that allows you to create a text-to-speech voice from your own voice. Proc. Interspeech 2016, 1968–1969.

McCorduck P (2004). Machines who think: A personal inquiry into the history and prospects of artificial intelligence (25th ed.). Natick, Mass: A.K. Peters.

Medeiros J (2015). How Intel gave Stephen Hawking a voice. Wired Retrieved from https://www.wired.com/2015/01/intel-gave-stephen-hawking-voice/

Mills T, Bunnell HT, & Patel R (2014). Towards personalized speech synthesis for augmentative and alternative communication. Augmentative and Alternative Communication, 30(3), 226–236. 10.3109/07434618.2014.924026 [PubMed: 25025818]

Napper S, Robey B, & McAfee P (1989). An expert system for use in the prescription of electronic augmentative and alternative communication devices. Augmentative and Alternative Communication, 5(2), 128–136. 10.1080/07434618912331275116

Pullin G, Treviranus J, Patel R, & Higginbotham J (2017). Designing interaction, voice, and inclusion in AAC research. Augmentative and Alternative Communication, 33(3), 139–148. 10.1080/07434618.2017.1342690 [PubMed: 28675073]

Radford A, Wu J, Child R, Luan D, Amodei D, & Sutskever I (2019). Better language models and their implications [Blog post] Retrieved from https://blog.openai.com/better-language-models-/#content

Roark B, Fried-Oken M, & Gibbons C (2015). Huffman and linear scanning methods with statistical language models. Augmentative and Alternative Communication, 31(1), 37–50. [PubMed: 25672825]

Russell SJ & Norvig P (2010). Artificial intelligence: A modern approach (3rd ed.). Upper Saddle River: Prentice Hall.

Schlosser RW, Shane HC, Allen AA, Abramson J, Laubscher E, & Dimery K (2015). Just-in-time supports in augmentative and alternative communication. Journal of Developmental and Physical Disabilities, 28(1), 1–17. /10.1007/s10882-015-9452-2

Sennott SC, Light JC, & McNaughton D (2016). AAC modeling intervention research review. Research and Practice for Persons with Severe Disabilities, 41(2), 101–115.

Sennott SC, Ferarri R, Crest C, Fogarty JL, & Hix-Small H (2017). MODELER AAC intervention during shared reading and play in early childhood. Journal on Technology & Persons with Disabilities, 5, 270–285.

Todman J, Alm N, Higginbotham J, & File P (2008). Whole utterance approaches in AAC. Augmentative and Alternative Communication, 24(3), 235–254. 10.1080/08990220802388271 [PubMed: 18830912]

Topol EJ (2019). High-performance medicine: The convergence of human and artificial intelligence. Nature Medicine, 25(1), 44–56. 10.1038/s41591-018-0300-7

Torresen J (2018). A review of future and ethical perspectives of robotics and AI. Frontiers in Robotics and AI, 4, 1–10. 10.3389/frobt.2017.00075

Turing AM (1950). Computing machinery and intelligence. Mind, 59, 433–460. 10.1093/mind/LIX.236.433

Wehmeyer ML (Ed.). (2013). The story of intellectual disability: An evolution of meaning, understanding, and public perception Baltimore, MD: Paul H. Brookes.

Wisenburn B, & Higginbotham DJ (2008). An AAC application using speaking partner speech recognition to automatically produce contextually relevant utterances: Objective results. Augmentative and Alternative Communication, 24(2), 100–109. 10.1080/07434610701740448 [PubMed: 18465364]

van den Berghe R, Verhagen J, Oudgenoeg-Paz O, van der Ven S, & Leseman P (2019). Social robots for language learning: A review. Review of Educational Research, 89(2), 259–295. 10.3102/0034654318821286

Table 1.

Components of AI with common examples and some opportunities and precautions associated with integrating AI with AAC.

Component	Definition	Mainstream Example	Some AAC Opportunities	Some AAC Precautions
Knowledge Representation and Reasoning	The process of organizing information in a way that a computer can understand and use reasoning to make automated decisions	Classification systems, diagnostic engines, and prediction systems	Augmented intelligence such as personalized AAC assessment and language knowledge and skill building	Bias in assessment and prediction could create discrimination (e.g., AAC misconceptions could become part of the model, such that AAC could artificially hold a person with complex communication needs back by limiting their opportunities)
Machine Learning	Capability of a program to learn patterns and themes from data and apply that knowledge to other data	Many recommender and prediction algorithms used in entertainment apps and search engines	Customized word-prediction, custom speech synthesis (voices), tuning/ calibration of alternative access systems	Ability to manage/curate training and learning systems
Computer Vision	Combines components of AI to process images and make decisions about them	Identification of objects, navigation	Automated visual-scene displays, vocabulary capture tools	Privacy regarding photos and facial recognition
Natural Language Processing	Ability to process and generate human-like text or speech based on knowledge of previous uses of language	Spam filters, smart assistants, word prediction on mobile devices	Word- and message-prediction, automated storytelling, voice recognition of dysarthric speech	Autonomy and individuality compromises if AI systems cannot be checked manually by users (e.g., inappropriate vocabulary choices yielded via predication during visit to preschool versus out with friends)
Robotics	A physical agent that can move about and interact with its environment	Industrial robots for manufacturing, social robots for interaction	Social robots and robot-assisted learning; intelligent powered mobility	Awareness of the role of human interaction and conversation in social development

Table 2

Ways to learn more about AI.

Resource Type	Resource Exemplars
Books	• Artificial Intelligence: A Very Short Introduction by Margaret Boden
	• The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World by Pedro Domingos
	• Artificial Intelligence: A Modern Approach (3rd Edition) by Stuart Russell and Peter Norvig
Internet	• AI newsletter MIT Technology Review: https://go.technologyreview.com/newsletters/the-algorithm/
	• Open AI blog: https://blog.openai.com/
	• Google AI: https://ai.google
	• Technical tips: https://towardsdatascience.com/
Courses	• Elements of AI: https://www.elementsofai.com
	• Google AI for Social Good: https://ai.google/education/social-good-guide
	• Machine Learning Crash Course: https://developers.google.com/machine-learning/crash-course/ml-intro