Dialogue Act Analysis for Alternative and Augmentative Communication

E. Margaret Perkoff
University of Colorado Boulder
`margaret.perkoff@colorado.edu`

Abstract

Augmentative and Alternative Communication (AAC) devices and applications are intended to make it easier for individuals with complex communication needs to participate in conversations. However, these devices have low adoption and retention rates. We review prior work with text recommendation systems that have not been successful in mitigating these problems. To address these gaps, we propose applying Dialogue Act classification to AAC conversations. We evaluated the performance of a state of the art model on a limited AAC dataset that was trained on both AAC and non-AAC datasets. The one trained on AAC (accuracy = 38.6%) achieved better performance than that trained on a non-AAC corpus (accuracy = 34.1%). These results reflect the need to incorporate representative datasets in later experiments. We discuss the need to collect more labeled AAC datasets and propose areas of future work.

1 Introduction

Dialogue Act classification takes a conversation transcript as input and identifies the appropriate intent for each turn in a conversation. For example, the sentence "How are you?" might be classified as an Open Ended Question. The exact tags that are used to label sentences depend on the context. The Switchboard DAMSL tag set (Jurafsky et al., 1997) is frequently used as a standard initial classification model which has forty-two distinct classes. Once labeled conversational data is available, it can be used to create generative statistical systems that take a sentence and a prior Dialogue Act as input and provide the next most like Dialogue Act for the conversation. Prior research has used this information to analyze both human-human conversations and better facilitate human-machine conversations (Ahmadvand et al., 2019).

However, research in Dialogue Act classification has not included conversations with individuals who do not rely solely on verbal speech to communicate. As of the 2010 United States census, approximately 15.7 million adults were listed as having a communicative disability (Brault, 2012). The communicative disability domain includes individuals who identify as having either a visual, hearing, or speech impairment or some combination of the three. Many of these individuals communicate through non-verbal methods including Augmentative and Alternative Communication (AAC) technology; we will refer to this population as AAC communicators.

AAC communicators leverage a broad set of tools to supplement their verbal speech or to replace it entirely. Speech language pathologists may recommend AAC as part of a treatment plan for an individual in order to maximize their ability to effectively communicate in their environment. In addition to providing more communication methods, research has proven that AAC technology can actually improve language development skills in children (Light et al., 2019). These systems vary in technical sophistication from picture boards corresponding to concepts to tablets or application based speech-generating devices (Elsahar et al., 2019). Figure 1 depicts two such devices, both of which include touch-based text displays. There are standalone or dedicated AAC devices available that only provide a communication interface whereas application-based solutions may run on a personal tablet or mobile device. In addition to touch access, devices can also incorporate eye-gaze, switch, or brainwave input. Some devices will allow individuals to switch between different access modes to account for fatigue levels they may experience at different times. (Elsahar et al., 2019)

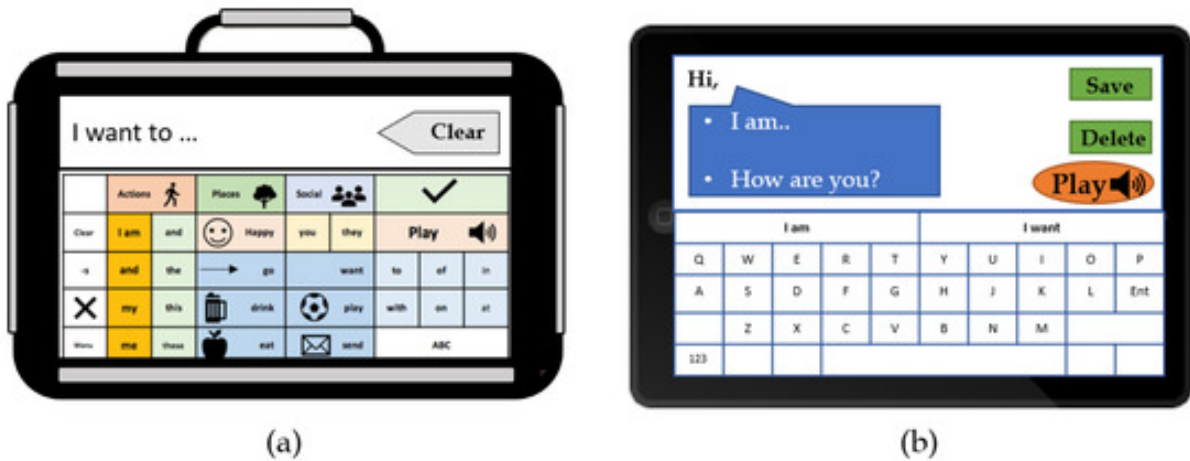The exact system used is tailored to the individual based on their cognitive, communicative, and

Figure 1: Two examples of AAC devices, a) is a dedicated AAC device using touch-input and b) is an AAC application running on a non-dedicated device. (Elsahar et al., 2019)

physical profile. An ideal system will maximize the individual's ability to express themselves while minimizing the cognitive and physical demands of using the system. Speech pathologists will perform an initial AAC evaluation to match a patient with the appropriate device to fit their immediate needs and long term communication goals and define an AAC intervention plan to track progress on these goals. Over the course of intervention, the system may be adjusted in order to better suit the needs of the individual whether it be physical changes to accommodate improved or worsening motor functionality or word selection adjustments to introduce more complicated vocabulary.

Over the last few decades, AAC devices have improved significantly, but satisfaction and retention rates for them remain low (Waller, 2019). There has been some effort to improve on-screen word prediction, but it has yet to provide sufficiently relevant suggestions during conversations or improve communication rates for AAC communicators. As we will discuss in the Related Work section, the application of NLP to AAC technology has been primarily limited to word prediction, despite the expansion of the field to a multitude of other tasks. We hypothesize that incorporating Dialogue Act information into AAC technology will improve the ease of use of these devices and in turn positively impact the ability of AAC communicators to participate in conversations.

In this paper, we address the potential benefits of applying Dialogue Act classification to conversations that include a participant communicating via an AAC device. We will start by presenting the previous NLP applications that have been used to enhance AAC software as well as currently available representative datasets. Then, we evaluate the performance of an existing state-of-the-art model on a small dataset of transcribed conversations between an AAC communicator and one of their daily communication partners. Finally, we present the challenges that inhibit work in this context. Ultimately, we hope that future researchers will recognize the value of applying language models to conversations with AAC communicators in order to improve their ability to independently participate in educational, social, and career settings.

2 Related Work

2.1 Natural Language Processing For AAC Users

There have been numerous efforts to incorporate different aspects of language processing into AAC. In 2011, (Higginbotham et al., 2011) conducted a review of the use of Natural Language Processing for Augmentative and Alternative Communication. These proposed enhancements are often aimed at improving the ease of use of devices or the rate of communication. The rate of communication for a device indicates how quickly an AAC communicator can respond using the technology. This is a critical element for being an active conversational participant. At the time, the relevant systems used optimized keyboards to improve input, word prediction, and speech recognition. There were different variations on improvements to word prediction including incorporating key noun phrases used by communication partners to enhance the

on-screen suggestions. (Wisenburn and Higgin-botham, 2009). However, none of the word prediction methods used were found to improve the rate of communication for AAC users.

Nearly a decade later, language processing research in AAC has not expanded much outside of the realm of word prediction. The research has focused on incorporating additional context to the word suggestions provided to users on device with the intent of improving communication rates as well as relevance of the suggestions themselves. Fried-Oken, M., Jakobs, T., & Jakobs, E. (2018) developed SmartPredict, an application-based AAC that leverages a statistical language model, the communicator's recent vocabulary, and content suggestions from their conversational partner via a partner application. Their hypothesis was that information provided by conversational partners would enhance the overall ease of use with the application. Initial findings from their experiments show a slight improvement in the number of selections that the AAC communicators required to indicate their desired intent, but these have not been expanded to a larger group yet. Garcia et al. (2015) investigated the use of location-aware language models for word and sentence prediction and found that they did not provide statistically significant improvements for participants' conversational rate. Location information was later used for pictogram prediction in a pictogram-based AAC device (Garcia et al., 2016) where the location based models also did not result in significant improvements in AAC usage. Outside of predictive models for word or pictogram-based devices, there has also been research into how NLP can be used to improve new AAC technology. Oken et al. (2014) were the first researchers to use NLP to enhance a Brain Computer Interface (BCI) system. Their system works by presenting the individual with a single letter for 2.5 seconds at a time and using non-invasive sensors to determine if this is the individual's target character. Instead of scanning through the entire alphabet, their statistical model presents the next letter based on what is most likely to occur following the previous letter. An enhanced BCI system has the potential to improve communication methods available to individuals with extremely limited or no voluntary motor control, including those with Locked-In Syndrome.

Research in the AAC space has remained limited to a small number of language processing tasks in the last several decades. Effort has been made to improve ease of use of AAC devices and communication rates by incorporating geographic, temporal, and contextual information into word prediction systems. Yet, as mentioned above, these additions have not significantly impacted the rate of communication or device retention rates of AAC communicators. Future work in this space needs to include experimental AAC designs that leverage a greater breadth of NLP applications to better meet the needs of this population.

2.2 Data including Individuals with Complex Communication Needs

In order to pursue further NLP applications for AAC, there is a need to collect or aggregate representative training data sets for these models. The most comprehensive dataset including conversational data for AAC communicators is the AAC and non-AAC Workplace Corpus (Friginal et al., 2013). This corpus includes transcripts of over two hundred hours of data captured with eight participants using AAC devices in their workplace environment. A single corpus of conversational transcripts is not sufficient to create statistical models that will provide significant benefit.

However, despite a lack of transcribed conversations, audio datasets have been greatly expanded upon to include speech samples that represent a variety of different language disorders. There are multiple corpora available that include samples of speech from adult Parkinson's patients (Tsanas et al., 2014) (Orozco et al., 2014) (Jaeger et al., 2019). Other audio datasets have been collected to study the dysarthric speech of individuals with Cerebral Palsy and Amyotrophic Lateral Sclerosis (Rudzicz et al., 2010). Little et al. (2007)'s corpus includes speech samples from individuals with a mixed set of language disorders. In addition to covering a range of language impairment types, the audio data that has been collected is also representative of individuals from different age groups. As an example, the Child's Pathological Speech Database (Ringeval et al., 2011) includes speech data from children with either autism spectrum disorder or a different language impairment. There is also the CSLU Autism Speech Corpus which contains data from speech pathology evaluations on forty-five children conducted from 2005-2012 (Gale et al., 2019) amounting to 1.5 hours of audio data with a total of 1,022 utterances.

Of the datasets mentioned above, the AAC and non-AAC Workplace Corpus is the only one that includes individuals communicating with an AAC device. It is also the only corpus that is coded for linguistic characteristics, including part-of-speech tagging. Even if speech-to-text applications were run to convert all of the audio corpora mentioned above to transcript formats, they would still need to be coded by linguistic features in order to be usable as training data for certain language processing tasks. For these reasons, additional effort to collect and label representative conversational data of AAC communicators is needed to make meaningful progress with NLP advancements.

3 Implications of Dialogue Act Analysis for AAC

Applying Dialogue Act classification to conversations including communicators reliant on AAC has the potential to improve their ability to communicate as well as enhance the AAC intervention and evaluation processes.

3.1 Benefits for AAC Communicators

The ultimate goal of AAC intervention is to increase the communicative competence of an individual. This covers not only the ability to communicate in the workplace or classroom setting, but also the ability to engage in personal conversations with friends and family. The current set of AAC devices and applications has yet to provide an adequate solution for individuals with complex communication needs. Many individuals who have been prescribed high-tech AAC devices end up abandoning them due to bad user interface, physical access limitations, the cognitive load required to learn them, or due to a lack of access to an expert (Waller, 2019). Those that continue to use their devices face limitations with conversational agency in terms of conversational, task, and device constraints (Valencia et al., 2020).

Incorporating Dialogue Act information into an AAC interface would improve the ability of AAC communicators to participate in conversations. A generative Dialogue Act model built into an AAC application would be able to predict the most likely next Dialog Act in a conversation. This information could then be used to provide the AAC communicator with partial or full phrases that correspond to the appropriate Dialog Act. Smart phrase recommendations may enhance the rate of communication,

making it easier for the AAC communicator to respond to the topic in a timely manner. For example, if their conversational partner asks, a *Wh-Question: What are you doing this weekend."* the system could provide partial phrase recommendations that conform to a *Statement-Non-Opinion* such as "I'm going to . . . ". Dialog Act suggestions also have the potential to impact ease of use with the device by reducing the amount of navigation required to find desired words or phrases. These improvements would reduce the cognitive and physical load imposed on the AAC communicator and potentially make them more motivated to continue to use their device.

3.2 AAC Evaluation and Intervention Improvements

Dialogue Act classification could also be used to quickly analyze speech pathology transcripts to improve both initial AAC evaluations and ongoing AAC intervention. As part of the initial AAC assessment, the conversations between the patient and members of the AAC team are coded for communicative functions such as requests, information sharing, and wh-questions. (Beukelman and Light, 2020) Speech language pathologists record these sessions and transcribe them on their own or send them to a transcription service. Once they have a written version, they review either the audio or written files and annotate them for the appropriate communicative function.

Communicative functions could be treated as Dialogue Act classes and annotated by speech pathologists on representative samples of atypical speech. A Dialogue Act classification model could then be trained on this gold standard data in order to automate this process in the future. This type of automation would make it easier for speech pathologists to evaluate patients for an initial AAC device as well as fitting them to a new device at a later stage in their treatment. As a result, their patients could gain access to an appropriate AAC device and improve their ability to communicate more quickly.

Following the initial assessment, Dialogue Act classification could then be used to track the progress of the patient with their initial evaluation goals. Current speech pathology research stresses the importance of evidence-based intervention for individuals with complex communication needs (Light et al., 2019). By using a Dialogue Act clas-

sification model, speech pathologists and conversational partners could quickly code interactions and identify how often the individual is able to express the communicative functions that correspond to their intervention goals. This provides more frequent feedback on goals and allows the AAC team to adjust appropriately. Additionally, models could be trained to identify the method by which the individual is communicating, either through vocalization or a device. Then multi-class models could associate particular communicative functions with communication methods. This would provide deeper insight into whether the individual can vocalize a particular communicative function or if they require a device to fulfill particular conversational needs. Automated transcription would also allow for conversations to be evaluated at home instead of in a speech pathologists office, reducing potential burden on the individual and their AAC team.

4 Experiments

To explore the potential of Dialogue Act classification for conversations including individuals using AAC, we will evaluate the accuracy of a state of the art model on a small representative dataset.

4.1 Data

The data used consists of written transcripts of unscripted conversation between an individual using a speech generating device and one of their regular communication partners. The data was collected at the University of Buffalo (Higginbotham, 2021). Each of the individuals involved in the original study had amyotrophic lateral sclerosis (ALS) which has impacted their ability to communicate vocally. Participants were prompted to discuss trips that they had taken in the past. There are ten unique transcripts which each correspond to a conversation between one of the communicative partner pairs In total, there are four hundred and thirty six utterances present in the dataset.

In addition to the AAC dataset, some of the models were trained on the Switchboard training corpus (Jurafsky et al., 1997). The Switchboard corpus contains labeled data from 1,155 5-minute conversations. The training set contains a total of 197,489 utterances. Both datasets were annotated with the Switchboard DAMSL tags which are described in detail below.

Dialogue Act Tag	Count
Statement-non-opinion	155
Statement-opinion	36
Yes-No-Question	26
Repeat-phrase	24
Open-Question	20
Other	19
Yes Answers	19
Agree/Accept	16
Response Acknowledgement	15
Backchannel in question form	14

Table 1: Counts of dialogue act tags in the ALS Dataset.

Dialogue Act Annotation   We annotated the sentences based on the Switchboard DAMSL Dialogue Act tags (Jurafsky et al., 1997) which are currently the standard benchmark for evaluating the accuracy Dialogue Act classification models. The DAMSL model consists of forty-two distinct classes of dialogue acts meant to represent the meaning of a particular utterance. The top ten most frequent set of dialogue act tags present in the dataset can be seen in Table 1. The standards followed are based on the examples provided in the Switchboard manual.

A sample conversation snippet with the associated dialogue act tags can be seen in the conversation below. AC refers to the AAC communicator and P is their conversational partner.

P: *20 years together you can't think of one thing?*
[Rhetorical-Question]
AC: *We don't take many trips*
[Statement-non-opinion]
AC: *Florida was cool when we went to Universal Studios* [Statement-opinion]
P: *Yeah* [Yes Answers]

4.2 Classification Model

To establish a baseline of model performance, we picked the top implementation currently available for Dialogue Act analysis based on existing leaderboards (Ruder, 2021). The classification model from Ravi and Kozareva (2018)'s is currently ranked as the highest performing solution with an accuracy of 83.1 on the Switchboard dataset. This approach avoids the need to use pretrained word embeddings and instead uses projection transformations to transform the input. This avoids the need for us to train word embeddings on

Training Set	Validation Set	Accuracy	Loss
AAC	AAC	0.386	2.655
SWBD	SWBD	0.341	2.787
SWBD	AAC	0.341	2.787

Table 2: We evaluated all three versions of our Dialogue Act Classification models with a subset of the AAC datset based on categorical accuracy and crossentropy loss. AAC indicates that the training or validation set was sampled from the AAC dataset whereas SWBD indicates that the sampling was from the Switchboard corpus.

our limited dataset. We used the publicly available implementation which closely follows the original algorithm, but achieves a maximum accuracy of 73.1 (Suarez, 2021). Our experiments used a neural network with 2 hidden layers with 256 units. The Dialogue Act labels were mapped to one-hot encoding vectors of size 42. All of the models are trained with stochastic gradient descent for 100 epochs.

We compared three methods of training a classification model based on different combinations of the AAC data and the Switchboard dataset. The goal of these experiments was to understand whether the AAC transcripts alone could be used to train a classification model as well as whether training a model on the Switchboard corpus would be sufficient for classifying AAC conversational data. The first model relies on only the ALS transcript data for training and validation. We randomly sampled 80% of the sentences from the transcripts to use for training data, 10% for validation, and 10% for testing. The second model was trained and validated with the Switchboard corpus and then tested with the same test set of AAC sentences. Lastly, we trained a model on the Switchboard training corpus and validated with samples from the AAC set. The same sample set of sentences was used for validation in the first and third models. For each model variation, we calculate the categorical crossentropy loss and the categorical accuracy on the test set of AAC sentences.

5   Results

The results in Table 2 show that the first model, which was trained and validated on the AAC corpus outperforms both of the others, which were trained on the Switchboard dataset. This indicates that the Switchboard dataset is an inadequate training set

for AAC conversations. It also suggests that the use of representative data in model training has a positive impact on classification accuracy. However, the low 38.6% accuracy of this model reflects the fact that our current AAC dataset is not adequate for training a statistical Dialogue Act classification model. A larger representative dataset would be needed to improve results for a model trained on AAC-data only. The use of AAC data only in the validation phase of model training, during which hyperparameters are set, seems to have minimal impact on the accuracy of the model. This could be due to the small size of the validation set or further evidence that the Switchboard training data does not generalize to the AAC sentence data.

The low accuracy scores could also be reflective of the need to use better annotation tags. The AAC conversational data used for future experiments could benefit from AAC context specific Dialogue Act tags. Although the DAMSL tags work well for comparison to results on existing datasets, it would be more beneficial to use a refined tag set that is specific to the AAC data that is being analyzed. Ideally, the tags should be modified depending on the context in which they are being used. A speech pathologist may choose to use a specific set of classes for coding sentences that are closely related to the goals of the individual's intervention. These tags are different from a set that would be used to provide real-time suggestions on a device during a typical conversation. Once a set of context appropriate tags has been established, the AAC data should be annotated by a group of annotators who are familiar with interpreting AAC output, such as speech pathologists in order to establish more accurate standard tags and inter-annotator agreement.

6   Conclusion

In this paper, we introduced the possibility of applying Dialogue Act classification to conversations that include one or more individuals communicating via AAC devices. We have discussed the potential benefits for applying this NLP technique to AAC applications as well as speech pathology transcriptions. In the experiments, we trained a Dialogue Act classification model on a small subset of AAC data and determined that a model trained on the Switchboard corpus does not perform as well on AAC conversational data as one that has been trained on a representative corpus. However, with the current AAC corpus available, the accu-

racy for the Dialogue Act classification model is far from the current benchmarks for these models on the Switchboard corpus. This indicates the need for more experiments to improve Dialogue Act classification accuracy in this context.

Future experiments will require labeled AAC conversational data for model training. The data collection task is a daunting proposition due to the fact that AAC devices are inherently personalized. Each AAC intervention session is tailored to the patient, the devices chosen, and goals set are meant to be the best choice for that individual. A representative dataset would need to be sufficiently large to incorporate individuals with different communication impairments, different degrees of impairment, and those from different age groups. The data must also be collected in a manner to preserve the privacy of the individuals' included. Furthermore, the current limitations of AAC devices may bias the data that could be collected. If a person is currently limited in their communicative ability due to poor user design with the device, the speech that they output will not be reflective of their full communicative desires. To mitigate this risk, further research needs to be done with this population to better incorporate their conversational goals into a representative dataset. Additionally, any model proposed should have a personalizable component. This component would allow for supplemental training data that is based on the individual's recent conversation history or a bootstrapping period. The inclusion of a personalization training period would provide enhanced conversational agency as well as incorporating an individual's communication style into their AAC device.

Once a sufficiently accurate Dialogue Act classification model is available, it can be embedded into a standalone AAC device or AAC application. This prototype should be designed alongside individuals who currently communicate via AAC as well as their AAC team. Through a collaborative research and design process, we hope to see improvements in usage rates for AAC devices and higher satisfaction rates by AAC communicators. More importantly, we anticipate that as AAC devices improve, AAC communicators will find themselves able to participate in more professional, educational, and social conversational opportunities.

References

Ali Ahmadvand, Jason Ingyu Choi, and Eugene Agichtein. 2019. Contextual dialogue act classification for Open-Domain conversational agents. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR'19, pages 1273–1276, New York, NY, USA. Association for Computing Machinery.

David R Beukelman and Janice C Light. 2020. *Augmentative et alternative communication : supporting children and adults with complex communication needs*. Paul H. Brookes Publishing Co.

Matthew W Brault. 2012. Americans with disabilities: 2010. Technical Report P70-131, U.S. Census Bureau, Washington, D.C.

Yasmin Elsahar, Sijung Hu, Kaddour Bouazza-Marouf, David Kerr, and Annysa Mansor. 2019. Augmentative and alternative communication (AAC) advances: A review of configurations for individuals with a speech disability. *Sensors*, 19(8).

Fried-Oken, M., Jakobs, T., & Jakobs, E. 2018. Smart-Predict: AAC app that integrates partner knowledge into word prediction. Annual Conference of the Assistive Technology Industry Association (ATIA).

Eric Friginal, Pamela Pearson, Laura Di Ferrante, Lucy Pickering, and Carrie Bruce. 2013. Linguistic characteristics of AAC discourse in the workplace. *Discourse Studies*, 15(3):279–298.

Robert Gale, Liu Chen, Jill Dolata, Jan van Santen, and Meysam Asgari. 2019. Improving ASR systems for children with autism and language impairment using Domain-Focused DNN transfer techniques. *Interspeech*, 2019:11–15.

Luís Filipe Garcia, Luís Caldas de Oliveira, and David Martins de Matos. 2016. Evaluating pictogram prediction in a location-aware augmentative and alternative communication system. *Assist. Technol.*, 28(2):83–92.

Luís Filipe Garcia, Luís Caldas De Oliveira, and David Martins De Matos. 2015. Measuring the performance of a Location-Aware text prediction system. *ACM Trans. Access. Comput.*, 7(1):1–29.

D Jeffery Higginbotham, Gregory W Lesher, Bryan J Moulton, and Brian Roark. 2011. The application of natural language processing to augmentative and alternative communication. *Assist. Technol.*, 24(1):14–24.

J Higginbotham. 2021. Communication and assistive device laboratory - AAC shared narrative corpora:ALS.

Hagen Jaeger, Dhaval Trivedi, and Michael Stadtschnitzer. 2019. Mobile device voice recordings at king's college london (MDVR-KCL) from

both early and advanced parkinson's disease patients and healthy controls.

Dan Jurafsky, Liz Shriberg, and Debra Biasca. 1997. *WS-97 Switchboard DAMSL Coders Manual*.

Janice Light, David McNaughton, David Beukelman, Susan Koch Fager, Melanie Fried-Oken, Thomas Jakobs, and Erik Jakobs. 2019. Challenges and opportunities in augmentative and alternative communication: Research and technology development to enhance communication and participation for individuals with complex communication needs. *Augment. Altern. Commun.*, 35(1):1–12.

Max A Little, Patrick E McSharry, Stephen J Roberts, Declan A E Costello, and Irene M Moroz. 2007. Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *Biomed. Eng. Online*, 6:23.

Barry S Oken, Umut Orhan, Brian Roark, Deniz Erdogmus, Andrew Fowler, Aimee Mooney, Betts Peters, Meghan Miller, and Melanie B Fried-Oken. 2014. Brain-computer interface with language model-electroencephalography fusion for locked-in syndrome. *Neurorehabil. Neural Repair*, 28(4):387–394.

Juan Rafael Orozco, Julian D Arias-Londoño, J Francisco Vargas-Bonilla, and Elmar Noeth. 2014. New spanish speech corpus database for the analysis of people suffering from parkinson's disease. In *International Conference on Language Resources and Evaluation*. unknown.

Sujith Ravi and Zornitsa Kozareva. 2018. Self-Governing neural networks for On-Device short text classification. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 887–893, Brussels, Belgium. Association for Computational Linguistics.

Fabien Ringeval, Julie Demouy, György Szaszak, Mohamed Chetouani, Laurence Robel, Jean Xavier, David Cohen, and Monique Plaza. 2011. Automatic intonation recognition for the prosodic assessment of Language-Impaired children. *IEEE Trans. Audio Speech Lang. Processing*, 19(5):1328–1342.

Sebastian Ruder. 2021. NLP-progress.

Frank Rudzicz, Aravind Kumar Namasivayam, and Talya Wolff. 2010. The TORGO database of acoustic and articulatory speech from speakers with dysarthria. *Language Resources and Evaluation*, 46(4):1–19.

Andres Suarez. 2021. SGNN.

Athanasios Tsanas, Max A Little, Cynthia Fox, and Lorraine O Ramig. 2014. Objective automatic assessment of rehabilitative speech treatment in parkinson's disease. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 22(1):181–190.

Stephanie Valencia, Amy Pavel, Jared Santa Maria, Seunga (gloria) Yu, Jeffrey P Bigham, and Henny Admoni. 2020. Conversational agency in augmentative and alternative communication. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pages 1–12, New York, NY, USA. Association for Computing Machinery.

Annalu Waller. 2019. Telling tales: unlocking the potential of AAC technologies. *Int. J. Lang. Commun. Disord.*, 54(2):159–169.

Bruce Wisenburn and D Jeffery Higginbotham. 2009. Participant evaluations of rate and communication efficacy of an AAC application using natural language processing. *Augment. Altern. Commun.*, 25(2):78–89.