Automated speech analysis enables MCI diagnosis

Charalambos Themistocleous[1], Marie Eckerström[2], Dimitrios Kokkinakis[3]

[1]Department of Neurology, Johns Hopkins University, USA
[2]Department of Psychiatry & Neurochemistry, University of Gothenburg, Sweden
[3]Department of Swedish, University of Gothenburg, Sweden

https://doi.org/10.36505/ExLing-2020/11/0050/000465

Abstract

Mild Cognitive Impairment (MCI) is a condition characterized by cognitive decline greater than expected for an individual's age and education level. In this study, we are investigating whether acoustic properties of speech production can improve the classification of individuals with MCI from healthy controls augmenting the Mini Mental State Examination, a traditional screening tool, with automatically extracted acoustic information. We found that just one acoustic feature, can improve the AUC score (measuring a trade-off between sensitivity and specificity) from 0.77 to 0.89 in a boosting classification task. These preliminary results suggest that computerized language analysis can improve the accuracy of traditional screening tools.

Keywords: acoustic analysis, machine learning, cognitive impairment, MMSE.

Introduction

Mild Cognitive Impairment (MCI) is a syndrome characterized by cognitive decline greater than expected for an individual's age and education level. The Mini-Mental State Exam (MMSE) is a screening tool for cognitive impairment often employed for evaluating individuals for MCI. Although, it has been considerably validated and is extensively used, it has received criticism as not being effective in detecting MCI. In an earlier work, Fraser, Lundholm Fors, Eckerström, Themistocleous, and Kokkinakis (2018), examined the utility of augmenting MMSE scores with automatically extracted linguistic information from a narrative speech task to better differentiate between individuals with MCI and healthy controls in Swedish. The study found that with the addition of just four linguistic features, the AUC score (measuring a trade-off between sensitivity and specificity) is improved from 0.68 to 0.87 in logistic regression classification. These results suggested that the accuracy of traditional screening tools may be improved through the addition of computerized language analysis. In this study, our goal is to determine whether cognitive decline can be estimated using acoustic information, as we have shown that speech production is impaired in patients with MCI (Themistocleous, Eckerström, & Kokkinakis, 2018, 2020; Themistocleous, Kokkinakis, Eckerström, Fraser, & Fors, 2018). Thus, contributing to the work of identifying automatic diagnostic markers that have the potential to facilitate clinical evaluation and therapy.

Methodology

Participants for this study were recruited from the Gothenburg MCI study (Wallin, Nordlund, Jonsson, & others, 2016) a large-scale longitudinal study that aims to advance the nosological understanding in AD and other types of dementia (Wallin et al., 2016). Additional assessment tests were conducted for the purposes of the Riksbankens Jubileumsfond – The Swedish Foundation for Humanities & Social Sciences "Linguistic andextra-linguistic parameters for early detection of cognitive impairment" research grant (NHS 14-1761:1) where speech recordings from the cookie theft picture description task were elicited. Thirty healthy controls and 25 MCI—between 55 and 79 years old (M=69, SD=6.4) participated in the study (see Table 1). The two groups did not differ with respect to age [t(52.72)= −1.8178,p=n.s.] and gender (W=1567.5, p=n.s.) significantly. The recordings were conducted in an isolated environment at the University of Gothenburg. The recordings were analyzed acoustically, using advanced acoustic analysis and signal processing algorithms (see for the methodological analysis (Themistocleous et al., 2020)).

For the speech part, participants were asked to describe what they could see in the "Cookie Theft" picture from the Boston Diagnostic Aphasia Examination (Goodglass, Kaplan, & Barresi, 2001), widely employed picture that aims to elicit narrative speech. The picture description task was audio recorded using a Zoom H4N audio recorder, and the audio recordings were analyzed acoustically using the open source software for acoustic analysis Praat (Boersma & Weenink, 2018). Specifically, we analyzed speech sounds and measured acoustic properties related to voice quality and speech fluency. Measurements of the number of syllables, number of pauses, average syllable duration, phonation time, speech rate, articulation rate, and speaking time. We employed a machine learning approach to classifying patients with MCI from HC. We trained two Boosting Classification models (i) a Base Boosting Classification model with the MMSE, speaker age, gender, and education and (ii) an Augmented Boosting Classification Model that includes the added acoustic measures. Both Boosting Models add predictors to a decision tree ensemble, each one correcting its predecessor. Boosting fits the new predictor to the residual errors made by the previous prediction, instead of modifying the weights for every incorrectly classified observation.

Results

The estimated accuracy of each run is shown in Figure 1. The Base Boosting Model provided 75% classification accuracy and the Augmented Boosting

Model had an increased accuracy, namely 79%. These measures are reflected in the precision and recall of the two models (shown in Table 1). Figure 1 shows the AUC of the Based Boosting Model and the Augmented Boosting Model, the AUC for the MCI was increased in the Augmented Boosting Model from 76% to 89%. The Based Boosting Model employed primarily MMSE and education whereas the Augmented Boosting Model employed Articulation Rate in addition to MMSE, Education, and Gender to perform the classification task. The relative influence of the predictors of the Base Bosting Model and the Augmented Boosting Model are shown in Figure 1. In addition to the MMSE and education, the articulation rate and Gender affect the Augmented Boosting Model.

Table 1. Evaluation Metrics of the Base Boosting Classification model and the Boosting / Acoustics Classification model.

		Precision	Recall	F1 Score	AUC
Base Boosting	HC	0.647	0.917	0.759	0.802
	MCI	**0.909**	**0.625**	**0.741**	**0.766**
	Average / Total	0.797	0.750	0.748	0.784
Aug. Boosting	HC	0.800	0.800	0.800	0.995
	MCI	**0.769**	**0.769**	**0.769**	**0.887**
	Average / Total	0.786	0.786	0.786	0.941

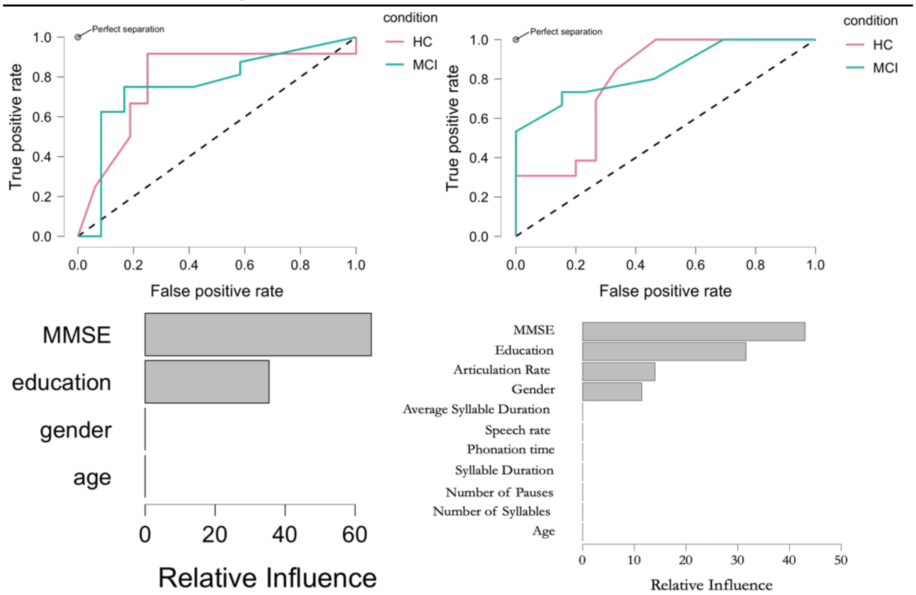

Figure 1. Shows the ROC curve plot for the Base Bosting Model (Panel A) and Boosting Model / Acoustics (Panel B) and relative influence of predictors of the Based Boosting Model (Panel C) and the Augmented Boosting Model (Panel D).

Discussion

In this study, we aimed to determine the use of acoustic measures to improve MCI classification, relative to using MMSE, education, and age scores alone. The results were positive, showing that the Augmented Boosting Model had an improved AUC over the Base Boosting Model by including just one acoustic measure. In the Augmented Boosting Model, the MCI AUC was improved from 0.76 in the Base Model using MMSE, age, gender and education to 0.89 by allowing the classifier to include the Articulation Rate in addition to the MMSE score. Speech contains information about the cognitive state of an individual and can reveal impairment in individuals with MCI as was shown in several earlier studies. This study provides promising results for automated measures of cognitive decline. Automated diagnostic measurements can be conducted at primary care centers and memory clinics quickly and easily assisting the work of clinicians, speech and language therapists and caregivers. The main limitation of the study is that we considered mainly temporal acoustic measures, as these are known to be affected in MCI, yet in our future research we aim to include measures of voice quality and prosody as well. Moreover, we plan to combine these measures with measures of grammar and discourse, as the Cookie Theft picture description task reduces the possibilities of eliciting a variability of linguistic expressions manifested using phonetic and grammatical means, such as various types of speech acts, questions, etc. We expect that tasks eliciting discourse and conversation will enable the assessment of linguistic and cognitive skills of individuals. Nonetheless, this study provides promising results towards augmenting and automating traditional evaluation tests using computational techniques from computational linguistics and machine learning.

References

Boersma, P., Weenink, D. 2018. Praat. Retrieved from http://www.praat.org

Fraser, K., Lundholm Fors, K., Eckerström, M., Themistocleous, C., Kokkinakis, D. 2018. Improving the sensitivity and specificity of MCI screening with linguistic information. LREC, Miyazaki, Japan.

Goodglass, H., Kaplan, E., Barresi, B. 2001. BDAE-3: Boston Diagnostic Aphasia Examination–Third Edition: Lippincott Williams & Wilkins Philadelphia, PA.

Themistocleous, C., Eckerström, M., Kokkinakis, D. 2018. Identification of Mild Cognitive Impairment From Speech in Swedish Using Deep Sequential Neural Networks. Frontiers in Neurology, 9, 975. doi:10.3389/fneur.2018.00975

Themistocleous, C., Eckerström, M., Kokkinakis, D. 2020. Voice quality and speech fluency distinguish individuals with Mild Cognitive Impairment from Healthy Controls. PLoS One, 15(7), e0236009. doi:10.1371/journal.pone.0236009

Themistocleous, C., Kokkinakis, D., Eckerström, M., Fraser, K., Fors, K. L. 2018. Effects of Cognitive Impairment on vowel duration. ExLing 2018, 113.

Wallin, A. et al. 2016. The Gothenburg MCI study. Journal of Cerebral Blood Flow & Metabolism, 36(1), 114-131. Retrieved from https://doi.org/10.1038/jcbfm.2015.147.