International Journal of Computer-Assisted Language Learning and Teaching
Volume 10 • Issue 1 • January-March 2020

Fuzzy Logic Applied for Pronunciation Assessment

Halima Bahi, University Badji Mokhtar, Annaba, Algeria
https://orcid.org/0000-0002-1519-7075

Khaled Necibi, University of Constantine 2 (MISC Laboratory), La Nouvelle Ville Ali Mendjeli, Algeria

ABSTRACT

Pronunciation teaching is an important stage in language learning activities. This article tackles the pronunciation scoring problem where research has demonstrated relatively low human-human and low human-machine agreement rates, which makes teachers skeptical about their relevance. To overcome these limitations, a fuzzy combination of two machines scores is suggested. The experiments were carried in the context of Algerian pupils learning to read Arabic. Although the native language of Algerian pupils is a dialect of Arabic, Modern Standard Arabic remains difficult for them with difficult sounds to master and letters close in their pronunciation. The article presents a fuzzy evaluation system including both oral reading fluency, and intelligibility. The fuzzy system has shown that despite the disparities between human ratings, its scores correspond at least to one of their ratings and most of the time its ratings are in favor of learners. Therefore, fuzzy logic, more favorable than thresholding systems, encourages learners to pursue their training.

KEYWORDS

Arabic Speech Recognition, Automatic Speech Recognition (ASR), Computer-Assisted Language Learning (CALL), Fuzzy Logic, Pronunciation Assessment

INTRODUCTION

Pronunciation teaching is an important stage in language learning activities. Advances in automatic speech recognition (ASR) technology -in terms of models, algorithms, portability, etc.- have positively impacted the research in several fields (for more details, see (Baker et al., 2009)). Particularly, these advances have promoted the development of computer-assisted pronunciation learning systems (Lee et al., 2017; Sefara et al., 2017) and have enabled the automatic pronunciation assessment (Chen & Li, 2016). A few languages like English have benefited from this technological aspect but others such as Arabic remain poorly resourced. This work aims to alert teachers on pupils with specific difficulties to correct them early, it also aims to give pupils additional time and more comfortable environment for practicing their pronunciation.

Overall, research in Computer-Assisted Pronunciation Teaching (CAPT) systems follows two directions: the first one is error detection (Hu et al., 2015; Lee et al., 2016) while the second is pronunciation assessment and scoring (Cheng et al., 2014; Neumeyer et al., 2000). In both directions, pronunciation teaching finds applications in many fields such as language learning (Landini et al., 2017; Reeder et al., 2015) or speech therapy (Necibi et al., 2013).

DOI: 10.4018/IJCALLT.2020010105

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

A typical ASR-based CAPT system involves several stages depicted in figure 1. When the learner pronounces a given word (or another part of speech), acoustic observations are extracted from the incoming signal and are represented as a collection of acoustic vectors. After that, the system force-aligns this representation with the model of the correct pronunciation (e.g. native-like). Hidden Markov Models (HMMs) are the most used models to represent speech.

A Hidden Markov Model is a collection of states connected by transitions (transitions are labeled with probabilities). An HMM begins in a designated initial state. In each discrete time step, a transition is taken in a new state where an output symbol is generated (Rabiner, 1989). "When an HMM is applied to speech recognition, the states are interpreted as acoustic models, indicating what sounds are likely to be heard during their corresponding segment of speech" (Bahi & Benati, 2009).

In the force-alignment stage, the speech recognizer computes the probability $p(W|O)$, where O is the observation represented by the extracted features from the incoming signal and W the model of the word to pronounce. In CAPT context, the acoustical representation of the incoming speech is force-aligned to HMMs (one or more) that models the expected pronunciation. The system outputs a score representing how close is the incoming speech to the correct pronunciation.

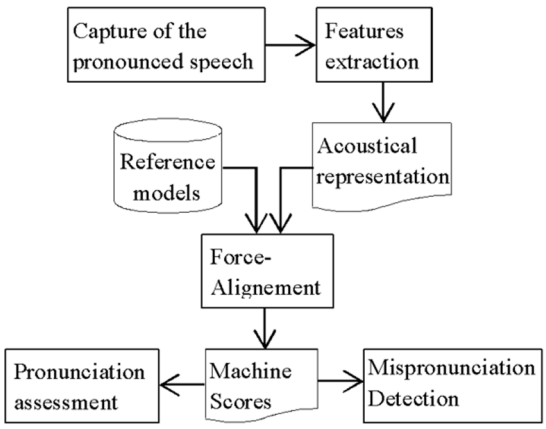
O'Brien et al. (2019) pointed that: "Early studies primarily dealt with pronunciation assessment and showed a relatively strong correlation between human pronunciation ratings and machine scores," and experiments in (Cucchiarini et al., 2000) have concluded that "raters who did not receive any instructions on the use of the rating scales may differ from each other in the absolute values of the scores assigned." Moreover, Luo et al. (2016) observed that "Many teachers remain skeptical about the fairness of automatic scores given by machines even with the most advanced scoring methods."

As fuzzy logic is known to promote soft frontiers between classes, its use would make more consensus between experts while still alerting teachers on possible difficulties of the learners. To overcome limitations related to the experts' rating disparities and to make teachers more comfortable with automatic scores, we developed a Fuzzy logic-based System for Pronunciation Assessment (FuSPA). The research question is:

Can automatic scores provided by FuSPA resemble those provided by human experts?

This is done by evaluating the performances of FuSPA on corpora learners' speech. In this study, FuSPA is evaluated in the context of Algerian pupils learning Arabic pronunciation, particularly, to assess their pronunciation. Although the native language of Algerian pupils is a dialect of Arabic, Modern Standard Arabic (MSA) remains a difficult language for them with difficult sounds to master, such as long vowels.

Figure 1. Several stages in ASR-based CAPT



11 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/fuzzy-logic-applied-for-pronunciation-assessment/243696

Related Content

Fuzzy Logic Applied for Pronunciation Assessment
Halima Bahiand Khaled Necibi (2020). *International Journal of Computer-Assisted Language Learning and Teaching (pp. 60-72).*
www.irma-international.org/article/fuzzy-logic-applied-for-pronunciation-assessment/243696

Communication Barriers and Conflicts in Cross-Cultural E-Learning
Rita Zaltsman (2009). *Learning Culture and Language through ICTs: Methods for Enhanced Instruction (pp. 318-330).*
www.irma-international.org/chapter/communication-barriers-conflicts-cross-cultural/25527

A Blended Chinese-as-a-Foreign-Language Short Course: Design and Perceptions
Shenglan Zhang (2016). *International Journal of Computer-Assisted Language Learning and Teaching (pp. 35-55).*
www.irma-international.org/article/a-blended-chinese-as-a-foreign-language-short-course/153894

Technology in Support of Self-Access Pedagogy
Hayo Reindersand Noemí Lázaro (2008). *Handbook of Research on Computer-Enhanced Language Acquisition and Learning (pp. 469-482).*
www.irma-international.org/chapter/technology-support-self-access-pedagogy/19827

Integrated Micro Learning During Access Delays: A New Approach to Second-Language Learning
Silvia Gstreinand Theo Hug (2006). *User-Centered Computer Aided Language Learning (pp. 152-176).*
www.irma-international.org/chapter/integrated-micro-learning-during-access/30647