Chapter 8

Second Language Learners’ Spoken Discourse: Practice and Corrective Feedback through Automatic Speech Recognition

Catia Cucchiarini
Radboud University, The Netherlands

Helmer Strik
Radboud University, The Netherlands

ABSTRACT

This chapter examines the use of Automatic Speech Recognition (ASR) technology in the context of Computer Assisted Language Learning (CALL) and language learning and teaching research. A brief introduction to ASR is first provided, to make it clear why and how this technology can be used to the benefit of learning and development in second language (L2) spoken discourse. This is followed by an overview of the state of the art in research on ASR-based CALL. Subsequently, a number of relevant projects on ASR-based CALL conducted at the Centre for Language and Speech Technology of the Radboud University in Nijmegen (the Netherlands) are presented. Possible solutions and recommendations are discussed given the current state of the technology with an explanation of how such systems can be used to the benefit of Discourse Analysis research. The chapter concludes with a discussion of possible perspectives for future research and development.

INTRODUCTION

Research on L2 learning has indicated that although exposure to the target language and usage-based learning are essential elements in the learning process, these are not always sufficient to guarantee target-like proficiency (Ellis, 2008; Ellis & Bogart, 2007). Focus on linguistic form provided through corrective feedback may help improve form accuracy in L2 spoken discourse. Unfortunately, in traditional teacher-fronted lessons there is generally not enough time for sufficient practice and feedback on speaking performance.

DOI: 10.4018/978-1-4666-4426-7.ch008
Second Language Learners’ Spoken Discourse

In this setting, the interest in applying ASR technology to L2 learning has been growing considerably in recent years (Eskenazi, 2009). ASR-based CALL systems would make it possible to offer sufficient amounts of practice in L2 speaking and to provide automatic feedback on different aspects of L2 spoken discourse. In this sense, ASR-based CALL systems would constitute an interesting supplement to traditional L2 classes. In addition, such systems can provide speaking practice in a private environment, which is a considerable advantage as speaking tasks are known to cause anxiety in L2 learners (Young, 1990). Moreover, L2 learners can practice at their own pace whenever they want.

Against this background a number of projects were started at our lab which was aimed at conducting research and developing technology that would be conducive to the realization of ASR-based CALL systems that support practice and automatic feedback on L2 spoken discourse and that can be used in the context of meaningful communicative tasks. The projects are intended to provide integrated formative evaluation of ASR-based CALL systems in the context of meaningful communicative tasks and to conduct research on improvement strategies to enhance the effectiveness of ASR-based CALL systems. The projects are designed to provide different forms of feedback on the accuracy of L2 spoken discourse and to address L2 learners’ requirements for real-time feedback on L2 pronunciation.

BACKGROUND: AUTOMATIC SPEECH RECOGNITION (ASR)

Standard ASR systems are generally employed for future research and development perspectives for future research and development. The ASR system consists of a decoder (the search algorithm) and three ‘knowledge sources’: the language model, the lexicon, and the acoustic models. The language model (LM) contains probabilities of words and sequences of words and describes the probability of words and sentences in a particular language. The language model is used to estimate the likelihood of a sentence given the input speech. The ASR system consists of a decoder (the search algorithm) and a knowledge source. The decoder identifies the sequence of words that best matches the input speech and generates a transcript. The knowledge source is used to generate a hypothesis about the input speech and to generate a hypothesis about the output language.

In order to support spoken interaction and natural-like communication, ASR systems need to be able to recognize connected speech. In addition, ASR systems need to be able to recognize connected speech in the context of meaningful communicative tasks. The ASR system needs to be able to recognize the context of meaningful communicative tasks and to generate a transcript that is consistent with the context of meaningful communicative tasks.

The aim of this chapter is to inform the reader about recent developments in the field of ASR-based CALL research and to indicate how these can lead to new methods and paradigms for the acquisition of spoken discourse in a second language. We then conclude with a discussion of possible implications for future research and development. We then conclude with a discussion of possible implications for future research and development.