Chapter 5
Second Language Instruction:
Extrapolating From Auditory–Visual
Speech Perception Research

Doğu Erdener
Middle East Technical University, Northern Cyprus Campus, Cyprus

ABSTRACT

Speech perception has long been taken for granted as an auditory-only process. However, it is now firmly established that speech perception is an auditory-visual process in which visual speech information in the form of lip and mouth movements are taken into account in the speech perception process. Traditionally, foreign language (L2) instructional methods and materials are auditory-based. This chapter presents a general framework of evidence that visual speech information will facilitate L2 instruction. The author claims that this knowledge will form a bridge to cover the gap between psycholinguistics and L2 instruction as an applied field. The chapter also describes how orthography can be used in L2 instruction. While learners from a transparent L1 orthographic background can decipher phonology of orthographically transparent L2s –overriding the visual speech information – that is not the case for those from orthographically opaque L1s.

INTRODUCTION

Traditionally, foreign / second language (L2, hereafter) instructional materials are based predominantly, if not only, on auditory media such as audio recordings. While there are examples of the use of visual materials in the context of L2 teaching, they are mostly visual materials based on not the nature of speech perception being visual as well as auditory, but other facets of the visual aspects of speech, e.g., providing contextual information such as animated stories or songs. One such approach is Content and Language Integrated Learning (CLIL) in which the subject being taught is taught in a foreign language with visual aids (Lasagabaster & de Zarobe, 2010). However, the experimental findings that specifically pertain to the benefits of the use of visual speech information (lip and face movements during speech) have had little or no room so far in the curricula of L2 instruction. This is partly due to the paucity of

DOI: 10.4018/978-1-7998-2588-3.ch005
theoretical and applied bridges between experimental psychology and psycholinguistic studies and L2 instruction as an applied field thus the dominance of mostly auditory-based materials in L2 teaching settings which has been almost the tradition.

Heavy reliance on only auditory-based materials is problematic in terms of ecological validity because most natural language communication does not occur in an auditory-only modality but rather in a multi-modal fashion. Specifically, when we are involved in a face-to-face verbal communication, not only do we hear but we also see the incoming speech in the form of lip and face movements. Research of the past few decades has consistently shown that L2 instructional material based not just on auditory-only but auditory-visual materials would be a profitable and more effective exploitation of the outcomes of that body of research. In this chapter, a picture of this still-neglected link between psycholinguistic work on the auditory-visual speech perception and L2 instruction is presented. First, a quick glance over speech perception is presented, then followed by the nature and development of L2-relevant aspects of auditory-visual speech perception are presented.

SPEECH PERCEPTION: ITS BASICS, NATURE AND DEVELOPMENT

Since the 1950s one significant challenge in the enterprise of speech perception is to understand how seemingly continuous and uninterrupted speech signal – as one would observe in a spectrogram – is deciphered by humans. Research in the past half a century has shown that this question cannot be answered in a straightforward fashion. On the contrary, this question requires a multi-dimensional and a multi-disciplinary response. Given the focus of this chapter, only developmental and L2-relevant aspects will be referred to here.

From a developmental point of view, speech perception emerges as an acoustic/phonetic process early in life. During the first year of life infants have a language-general capacity to perceive phoneme contrasts across the world’s languages (for a review of earlier literature see Burnham, 1986). This language-general capacity is not only marked by perceptual sensitivity for the contrasts in the ambient language environment but also for the phonologically irrelevant ones. Within this first year, there is a re-organization of speech perception such that it gradually moves from a language-general mode to a language specific one; in other words, the language-general and acoustically and phonetically based speech perception mechanism gradually becomes language-specific and phonologically-based as a result of exposure to the ambient (i.e., L1) language environment and its phonological repertoire. This trend occurs earliest for vowels between 4 and 6 months of age (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992), followed by lexical tones between 6 and 9 months (Mattock & Burnham, 2006) and finally the re-organization of consonants sometime between 7 and 11 months (Best, McRoberts, & Goodell, 2001), and discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener’s native phonological system. Some of these models of early speech perception focused on how L2 phonemes are processed and revised in the L2 context (Best & Tyler, 2007). In a nutshell, these models and theories commonly suggest that our inborn perceptual abilities that are acoustically and phonetically-oriented become attuned to the L1 phonology as a result of exposure, forming the L1 phonological repertoire. This repertoire in the subsequent stages of speech and language development allows for the discrimination of native (and relevant) phoneme contrasts that are crucial for vocabulary acquisition – particularly helping the person attain those contrasts that can be referred to as “minimal pairs”. For instance, the /r/-/l/ discrimination by Japanese adults is probably one of the most studied contrasts in the literature since the 1970s (Goto,