

Jonathan Dalby - Second-language speech intelligibility training

Friday 14 October 2011

At Communication Disorders Technology, Inc. (CDT), in Bloomington, IN several English speech-intelligibility training systems have been developed. The earliest were language-specific systems based on phonological error analyses of the accented English speech of native speakers of several different L1's, Mandarin Chinese, Spanish, and Japanese (Rogers and Dalby, 2005). It was believed that what is usually called 'foreign accent' might better be referred to as incomplete mastery of the phonological system (first) and the phonetic details (second) of the target language. The first of these is probably the most important, and perhaps also the most easily remediated. For example, most Mandarin learners of English are difficult to understand because they tend to speak English with an 'impoverished' vowel system—one that might be more like their native language than it is like American English. An underlying assumption of the CDT training systems was that learning new phonological contrasts requires intensive training in both production and perception practice through minimal-pairs drill (Dalby and Kewley-Port, 2007). The speech production component of these systems employs feedback to the client that is provided by automatic speech recognition technology. Evidence that such computer-based training can be effective in learning new phonological contrasts will be presented (Burlison and Dalby, 2007). But why bother? It has been shown that native listeners can adapt to 'foreign accents' quite rapidly and effectively, at least in laboratory experiments (Clarke and Garrett, 2004, Bradlow and Bent, 2008). Since that is so, can it not be assumed that accented speech is just as effective in real communicative situations as native speech? The first evidence that may be relevant to answering this question comes from work done in David Pisoni's lab at IU with early synthetic speech, where it was found that synthetic speech that was nearly equal in intelligibility to natural speech under ideal listening conditions, was less intelligible when listeners faced simultaneous cognitive tasks of different sorts. (See Winters and Pisoni, 2003 for a summary.) More direct evidence comes from testing with accented speech that measured intelligibility in sub-optimal listening conditions. Rogers, Dalby and Nishi (2004) showed that short sentences spoken by even highly proficient Mandarin learners of English were degraded in intelligibility more than was native speech when mixed with background noise. More recently, attempts to replicate this result for isolated

words have been conducted and will be discussed (Dalby and Rogers, 2009).