

The recognition of nasal-stop distinction in adverse listening conditions¹

Zahra Mahmoodzadeh²

Received: 2017/03/13

Accepted: 2017/05/07

Abstract

The automatic recognition of Persian numerals [sefr-se] “zero–three”, [do-noh] “two-nine” and [haft-haft] “seven-eight” is considered as a challenge for speech recognition systems. Mahmoodzadeh and Bahrani (2014) found that the acoustic reduction of telephone speech triggers lexical ambiguity for the automatic recognition of [sefr-se] pair. The numeral [sefr] is produced with the deletion of [r] at word final position and the weak labial friction of [f] is masked by the channel noise, which results in an increase of acoustic similarity between [sefr] and [se].

The automated recognition of phonological distinctions can be improved by using perceptual cues which listeners apply for the natural recognition of sounds in adverse listening conditions such as telephone speech or noisy environment. In this research for discovering efficient perceptual cues responsible for distinction of numerals [do]-[noh] “two-nine” in both natural and telephone speech, perception tests were used. The acoustic signal of [noh] shows a weak or practically deleted final [h], which is not audible and recognizable from background noise. Therefore, the acoustic differences of nasal-stop distinction and the co-articulatory effects of nasal on the following vowel play an important role in [do-no] recognition.

¹ . (DOI): [10.22051/JLR.2017.14294.1293](https://doi.org/10.22051/JLR.2017.14294.1293)

² . Assistant professor of General Linguistics, Iranian research institute for information science and technology (IRANDOC). Iran;
mahmoodzadeh@irandoc.ac.ir

Results show that the acoustic landmarks, nasal murmur and nasal-vowel transitions affected by various sources of signal disruption, undergo phonetic reduction, which leads to perceptual similarity of [do-noh] pair and ultimately listeners' lexical ambiguity. In optimal listening condition, natural recognition of [n] based on nasal murmur is about 40% which increases to 96% after the addition of 10ms of the beginning of vowel [o]. However in telephone speech, natural recognition of [n] based on murmur is about 29% and after the addition of 10ms of vowel [o] transitions rises to just about 48%. According to the outcomes, nasal murmur is not an influential perceptual cue for the recognition of [n] in both optimal and adverse listening conditions; however recognition rises significantly after adding only 10 ms of the beginning of following vowel. It is likely that listeners' lexical confusion and uncertainty due to lack and reduction of phonetic information reinforced by listeners' perceptual bias toward unmarked consonant, give rise to [d] responses and failure of [n] identification in adverse listening conditions. The automatic recognition of these words should be done by capturing further temporal and spectral information from neighboring vowel transitions.

Keywords: adverse listening conditions, Persian numerals, perception test, nasal-stop distinction, vowel transition.

Quarterly Journal of Research in Linguistics, Al-ZahraUniversity
VOL. 10,NO. 29, 2018-2019
<http://jlr.alzahra.ac.ir>