EFFECTS OF NOISE AND LANGUAGE EXPOSURE ON SPEECH INTELLIGIBILITY PERFORMANCE

Karen Yu and Donald G. Jamieson

Hearing Health Care Research Unit, School of Communication Sciences and Disorders
The University of Western Ontario, London, Ontario CANADA N6G 1H1

BACKGROUND
Native speakers of Korean often have difficulty producing and perceiving the distinctions between English sounds [1]. Particular problems may arise when communicating against a background of noise. Non-native listeners may display greater difficulty than native listeners under difficult listening conditions [6]. In view of the foregoing, the present study was designed to: (1) characterize the speech perception abilities of adult native speakers of Korean who have had very extensive exposure to the English language; (2) examine the degradation in word identification under difficult listening conditions. Performance was assessed in relation to the performance of native speakers of English.

METHOD AND PROCEDURE
We tested 30 native speakers of Korean, aged 35 to 58 (M = 49 years). All were native to Korea but had resided in Canada for between 10 and 28 years (M = 20 years) immigrating between the ages of 20 to 40 (M = 29 years). All subjects had studied English (mean time = 7 years) before coming to Canada. Twenty subjects had also studied in ESL programs in Canada (mean time = 6 months). Nine native Canadian English speakers with normal hearing also participated.

Our test battery consisted of three tests of speech perception: (1) the adaptive SRT (ASRT; [2]) provides an efficient, accurate, and reliable estimate of a listener's speech reception threshold for spondees. The SRT was administered in quiet and in a background of white noise; (2) the UWODFD [3] is a test of speech intelligibility, standardized for central Canadian English. It consists of 21 nonsense syllable stimuli; the target is the middle consonant of a VCVC word. All consonants are presented in the same context (A_IL). The UWODFD was administered in quiet and in noise at 4 Signal to Noise Ratios (SNRs; +10, +5, 0, and -5 dB); and (3) a two alternative [r-l] forced-choice identification task was used to assess the ability to distinguish English /r/ versus /l/ contrasts. This test used a subset of the stimuli used by Logan et al. [5]. This [r-l] test consisted of five minimal pairs (e.g., rock-lock) within each of five phonetic environments: initial singleton (IS), initial consonant cluster (IC), medial (M), final consonant cluster (FC), and final singleton (FS). Five native English speakers (3 male and 2 female) produced each of the ten words, contrasting /r/ and /l/ in the phonetic environments mentioned. Each of the five experimental tests was administered in 5 listening conditions: (1) in quiet and (2) in noise at 4 SNRs (10, 5, 0, and -5 dB). For both the UWODFD and the [r-l] tests, listening conditions were presented in sequence from most favourable (i.e., in Quiet) to the most difficult (i.e., -5 SNR) with the block order and stimuli randomized within each level. All aspects of stimulus sequencing and presentation (in both noise and quiet conditions), response recording, and experimental control were carried out using the experiment generator and controller utility contained in the CSRE 4.5 software package [4]. During all aspects of testing and training, signals were presented to listeners through Etymotic Research ER-2 insert phones.

RESULTS AND DISCUSSION
Accuracy Under Optimal Listening Conditions. The results from the three tests, in Quiet, show that the overall level of accuracy under optimal listening conditions was high for both native Korean listeners and native English listeners, with native English speakers having an advantage of ~5dB in SRT performance and ~18% and ~27% correct overall, in the UWODFD and [r-l] tests, respectively.

Figure 1. Identification of English consonants in the UWODFD test as a function of the signal to noise ratio.
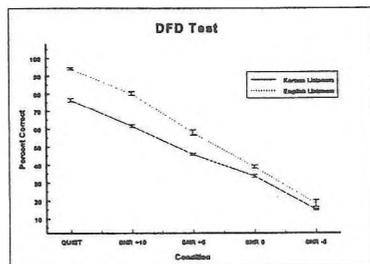
Figure 2. Identification of English /r/ and /l/ for target sounds in various syllable positions, by native Korean listeners (lower fan of curves), and by native English listeners (upper fan of curves), as a function of the signal to noise ratio
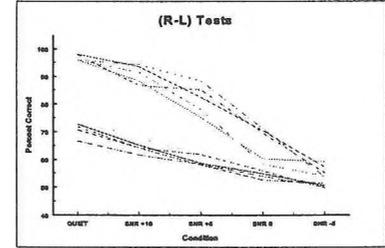
Accuracy Under More Difficult Listening Conditions. Under less optimal listening conditions, the performance of the Korean listeners declined rapidly for the UWODFD and [r-l] tests. Performance also declined with both tests for the native speakers of Canadian English, but this decline was more precipitous, reflecting the higher initial levels of performance in quiet. There was no indication that performance declined more rapidly in a noise background for our native Korean subjects on the ASRT.

Distribution of Identification Errors. When the format of the task does not constrain subjects' responses to just "L" and "R", the observed perceptual confusions are not limited to confusing /r/ with /l/ and vice versa. In fact, when listening in quiet during the UWODFD test, when /r/ targets were presented, 59% of the confusion errors Korean listeners made were with "W" responses, 34% were "L" responses. Under the same conditions, when /l/ targets were presented, 59% of confusion errors were presented were "R" responses, while 37% were "N" responses.

GENERAL DISCUSSION
The present study provides clear evidence that non-native listeners operate at a moderate to large disadvantage to native listeners even after many years of intensive experience with the English language. However, the disadvantage is seen most clearly under favourable listening conditions, not under degraded conditions. Furthermore, examination of the details of the confusion responses of Korean listeners shows that the two-alternative forced choice paradigm fails to characterize /r/-/l/ perception for non-native listeners. Future work may profit from the use of a wider range of response alternatives when assessing speech perception as well as for training listeners to perceive non-native contrasts.

REFERENCES
[1] **Borden, G, Gerber, A and Milsark, G** (1983). Production and perception of the /r/-/l/ contrast in Korean adults learning English. *Language Learning, 33*(4), 499-526.
[2] **Cheesman, MF** (1992). An automated technique for estimated speech-reception thresholds in multi-talker babble. *J Speech-Lang Path and Aud, 16,* 223-7.
[3] **Cheesman, MF and Jamieson, DG** (1996). Development, evaluation and scoring of a nonsense word test suitable for use with speakers of Canadian English. *Canadian Acoustics, 24,* 3-11.
[4] *Computerized Speech Research Environment (CSRE) Version 4.5* (1996), Avaaz Innovations Inc., London, Canada.
[5] **Logan, JS, Lively, SE and Pisoni, DB** (1991). Training Japanese listeners to identify English /r/ and /l/: A first report, *J Acoust Soc of Amer, 89*(2), 884-86.
[6] **Nábelek, AK and Donahue, AM** (1984). Perception of consonants in reverberation by native and non-native listeners, *J Acoust Soc of Amer, 75*(2), 632-4.