

# Clear speech intelligibility: Listener and talker effects

*Rajka Smiljanic and Ann Bradlow*

Northwestern University

rajka@northwestern.edu, abradlow@northwestern.edu

## ABSTRACT

In this study, we investigated whether the intelligibility-enhancing mode of speech production, known as “clear speech” produced by native and non-native talkers influenced speech intelligibility equally for native and non-native listeners. In a series of three experiments, we explored the effect of clear speech for various native and non-native talker and listener pairs. Combined, the results showed that “native” speech is overall more intelligible than “foreign” accented speech for both native and non-native listeners. Importantly, the proportional intelligibility gain for clear speech produced by both native and non-native talkers was similar across listener groups suggesting common speech processing strategies across all talker-listener groups.

## 1. INTRODUCTION

Speech intelligibility and word recognition depend on a wide variety of talker-, listener- and signal-dependent factors. The goal of this paper was to examine how the native language backgrounds of listeners and talkers (native vs. non-native) influence communication for various listener-talker pairs. Furthermore, we wanted to investigate whether native and non-native hyperarticulation articulatory strategies provide similar intelligibility benefits for both native and non-native listener groups. To that end, we looked at intelligibility of plain and “clear” speaking styles in English as produced by American English (AE) talkers and by Croatian talkers for AE and Croatian listeners.

Clear speech is a distinct, intelligibility-enhancing mode of speech production that talkers naturally and spontaneously adopt under adverse listening conditions. It is characterized by a wide range of acoustic/articulatory adjustments, including a decrease in speaking rate, an expansion of pitch range and an enhancement of phonological category contrasts in language-specific ways [1, 2, 3, 4]. These plain-to-clear speech articulatory modifications enhance intelligibility for normal-hearing and hearing-impaired adults, children with

and without learning disabilities and non-native listeners, among others [1, 2, 3, 5].

In their cross-language study, Smiljanic and Bradlow [3] showed that clear speech produced by native speakers of English and of Croatian increased intelligibility by 17 and 15% for English and Croatian listeners, respectively. Moreover, the accompanying cross-language acoustic analyses have shown both similar and different clear speech production strategies across English and Croatian [3, 4]. In this paper, we extend these findings by exploring whether clear speech strategies by native talkers (in their L1) and non-native talkers (in their L2) are beneficial to native and non-native listener groups. We hypothesized that some of the clear speech enhancement strategies produced by native talkers are not fully beneficial to listeners who do not share the same background L1 sound structure [5]. Similarly, non-native talkers’ clear speech strategies (in their L2) may include some enhancement modifications that are specific to their L1 and may not provide intelligibility benefit to the native talkers of L2 but may benefit the listeners who share their background L1.

## 2. EXPERIMENT 1: NON-NATIVE LISTENERS AND NATIVE TALKERS

### 2.1. Method

In order to minimize the beneficial effect of sentence context on intelligibility, we constructed semantically anomalous sentences such as in (1):

(1) *Your tedious beacon lifted our cab.*

Four (3 female, 1 male) native AE talkers were recorded in a sound-attenuated booth reading the 20 sentences once in plain and once in clear speaking style. For the plain style, the talkers were instructed to read as if they were talking to someone familiar with their voice and speech patterns. For the clear speaking style, the talkers were instructed to read as if they were talking to a listener with a hearing loss or a non-native speaker.

In order to obtain equivalent overall amplitude levels, all speech files were equated for RMS amplitude and then mixed with speech-shaped noise at a +5 dB signal-to-noise ratio (SNR). We used the results for native listener-native talker (matched) pairs reported in [3] as a baseline in deciding the noise levels in the experiments reported here. We aimed to achieve the same average intelligibility range of 45-65% across native and non-native listeners but had to take into account factors such as using L1 vs. L2 and same or different background L1, all of which may have a detrimental effect on intelligibility. In this experiment, we, therefore, increased SNR for the miss-matched pairs of native talkers and non-native listeners from 0 dB SNR to +5 dB SNR.

Each participant in the perception experiment heard a total of 20 sentences produced by only one of the talkers. Half of the sentences heard were in plain style and half in clear style for each talker condition. The listeners never heard the same sentence twice. In each talker condition, clear speech sentences preceded plain sentences so that any effect of adjusting to the task during the experiment could not account for the intelligibility increase in clear speech.

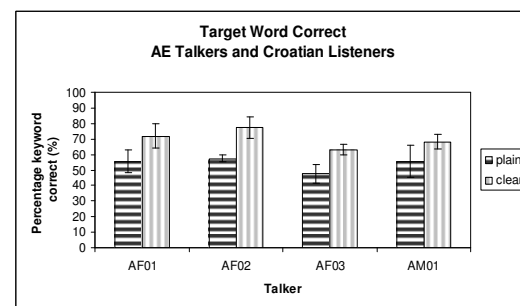
16 native Croatian listeners participated in the sentence-in-noise listening test. They were either undergraduate students of English at the University of Zagreb or had a significant amount of instruction in English in regular and specialized language schools. Their English proficiency was high as determined by a pre-test where they listened to 16 syntactically simple and meaningful sentences that included words highly familiar to non-native speakers mixed with noise at +5 dB SNR and wrote down what they heard. The average keyword intelligibility score for these sentences was 43/50 (range: 33-49). In the test condition, they were seated in front of a computer and heard one target sentence at a time over headphones. They could hear each sentence only once but could take as much time as needed between the sentences to record their answer. They were instructed to write down every word they heard. Each participant received a keyword correct score out of 40 for the 10 sentences they heard in each style (plain vs. clear). All content words were counted as keywords. All listeners identified the keywords as highly familiar in a post-test. Percentage correct scores were calculated and then

converted to rationalized arcsine transform units (RAU) for statistical analysis [6].

## 2.2. Results

The results showed a significant increase in intelligibility for clear speech when compared with plain speech for all 4 talkers (Figure 1). The average intelligibility score was 54% in plain and 70% in clear speech yielding the average clear speech intelligibility increase of 16%. The result of a paired-samples *t*-test showed a significant effect of style on intelligibility score:  $t(3) = -9.899$ ,  $p < .01$ .

**Figure 1:** Average intelligibility scores (percentage keyword correct) for non-native listeners in plain and clear speaking styles for each native AE talker.



The results showed that native talkers were successful in modifying their speech in a way that provided more salient acoustic cues for L2 processing by non-native listeners. Compared to the matched pairs' results [3], the amount of intelligibility gain by the non-native listeners in the current experiment was very similar (17% in [3] and 16% here). This suggests that native clear speech strategies are equally beneficial for native and for fairly fluent non-native listeners. Combined, the results show that in order to achieve a similar level of performance by native and non-native listeners, the level of noise has to be decreased by 5 dB. In other words, the added difficulty in speech processing of being a non-native listener is offset by an increase in SNR of 5 dB.

## 3. EXPERIMENT 2: NATIVE LISTENERS AND NON-NATIVE TALKERS

### 3.1. Method

In this experiment, materials, speech elicitation methods and the listening task were the same as in Experiment 1. Talker and listener groups differed

from those in 1. Here, native AE listeners listened to non-native speech produced by Croatian talkers.

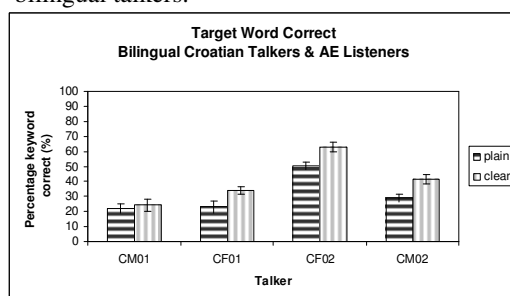
40 native AE listeners were recruited from the Northwestern University Linguistics Department subject pool. The talkers in this experiment were four (2 female, 2 male) non-native speakers of English whose first language was Croatian. They were all undergraduate students at Northwestern University and came to the US within five years prior to the recordings to pursue undergraduate degrees. They were fluent in English as confirmed by the General Record Examination (GRE) and Test of English as a Foreign Language (TOEFL) scores required for admission to a US university.

Since in this experiment native listeners were listening to their native, albeit “foreign”-accented, speech, we lowered SNR from +5 dB in Experiment 1 to 0 dB SNR. The 0 dB SNR allowed us to make a direct comparison with the results for matched pairs obtained in [3] and to estimate how detrimental “foreign”-accented speech is for speech intelligibility.

### 3.2. Results

The results showed an increase in intelligibility that accompanied plain-to-clear speech articulatory modifications by non-native talkers for native listeners (Figure 2). The average intelligibility for plain speech was 31% and for clear speech 41% resulting in the average intelligibility gain of 10%. The result of the paired-samples *t*-test showed a significant effect of style on the overall intelligibility score:  $t(3) = -3.749, p < .05$ .

**Figure 2:** Average intelligibility scores for native AE listeners in two speaking styles for all CRO/AE bilingual talkers.



Although there was some variability in how successful non-native talkers were in modifying their speech to accommodate native listeners (e.g., CM01 vs. CF02), overall their clear speech strategies benefited native listeners. When

compared with matched pairs, intelligibility for miss-matched pairs (non-native talkers and native listeners) is lower. Average intelligibility in plain and clear speech for AE matched-pairs, as reported in [3], was 46 and 63%, respectively. There was a decrease in intelligibility for non-native speech by 15 and 22% in plain and clear speech, respectively. The overall gain was lower by 7% for non-native clear speech. Combined, the results show that with the same noise levels, “foreign” accent is rather detrimental to speech perception. The effect of “foreign” accent may be offset by lowering noise levels, similar to the findings for non-native listeners listening to L2.

The effect of non-native clear speech was smaller compared to native clear speech for matched pairs (10 vs. 17%) and to the non-native listeners listening to the native speech in Experiment 1 (10 vs. 16%) although noise levels differed in Experiments 1 and 2.

## 4. EXPERIMENT 3: NON-NATIVE LISTENERS AND NON-NATIVE TALKERS

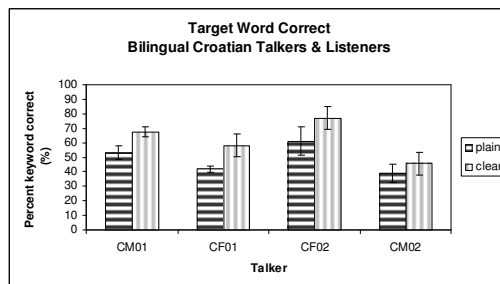
### 4.1. Methods

The materials, speech elicitation methods and listening test procedures were the same as in Experiments 1 and 2. The talkers were the same as in Experiment 2: four (2 female, 2 male) non-native speakers of English whose first language is Croatian. The listeners were 16 native Croatian listeners drawn from the same population as in 1 (different individuals). Their fluency in English was estimated in the same pre-test as in 1. The average keyword correct score for the pre-test sentences was 44/50 (range: 38-49). The SNR used here for mis-matched talker-listener pairs was +5 dB, the same as in Experiment 1.

### 4.2. Results

The results showed that there was a beneficial effect of clear speech on intelligibility, i.e., non-native speakers produced clear speech that increased intelligibility for listeners listening to their L2 (Figure 3). The average intelligibility scores were 49 and 62% for plain and clear speech, respectively. The intelligibility increase was 13%. The result of the paired-samples *t*-test showed a significant effect of style on the overall intelligibility score:  $t(3) = -5.649, p < .05$ .

**Figure 3:** Average intelligibility scores (percentage keyword correct) for non-native listeners in plain and clear speaking styles for each CRO/AE bilingual talker.



The results showed that there was a slight decrease in intelligibility for Croatian listeners listening to L2 by Croatian talkers (Experiment 3) compared with the results for Croatian listeners listening to L2 speech by native AE talkers (Experiment 1) with the same level of noise (49 vs. 54% in plain speech; 62 vs. 70% in clear speech, 13 vs. 16% gain). This suggests that sharing the same background L1 sound structure does not provide an additional level of benefit when listening to L2 (both plain and clear speech productions).

Similar levels of intelligibility were reported in [3] for native Croatian matched pairs (50 and 65% for plain and clear speech; 15% gain). This suggests that when listening to non-native speech (L2) the level of noise that allows the same performance as when listening to native (L1) speech has to be lower by about 5dB, i.e., unfamiliarity with L2 sound structure plus “foreign” accent in L2 can be offset by increasing SNR levels by 5 dB for these fluent non-native groups.

Finally, the performance of native AE listeners listening to “foreign-accented” speech (Experiment 2) was overall lower compared with non-native listeners listening to non-native speech with shared background L1 (Experiment 3). This difference could be in part due to a lower SNR in Experiment 2.

## 5. DISSCUSSION AND CONCLUSIONS

This study investigated how native language background interacts with clear speech strategies in determining levels of speech intelligibility. The results showed that “native” speech is preferred over “foreign” accented speech by both native and non-native listeners. Furthermore, listening to

“foreign” accented speech affects both native and non-native listeners regardless of whether they share the same background L1 or not. We also demonstrated that various talker-listener native language mismatches (which affect intelligibility negatively) can be offset by varying signal-to-noise ratio levels.

Finally, the results of this study revealed that clear speech is a beneficial articulatory modification regardless of the listener and talker L1 backgrounds. Moreover, if we examine proportional clear speech increase relative to the plain speech intelligibility (clear minus plain divided by plain intelligibility score), there is a remarkable similarity in intelligibility gain regardless of the native language background of either talkers or listeners. The average proportional intelligibility increase for native AE talkers and non-native Croatian listeners is 30% (Experiment 1), for non-native Croatian talkers and native AE listeners is 32% (Experiment 2) and for non-native Croatian listeners and non-native Croatian talkers is 27% (Experiment 3). These results are fairly close to the results for native-native AE and Croatian pairs: 39% and 31%, respectively [1]. These data provide strong evidence that clear speech as a listener-oriented and intelligibility-enhancing mode of speech production is helpful even when the overall intelligibility levels vary for various listener and talker groups.

Ultimately, we would like to develop a detailed understanding of how all of these factors interact in real-world listening situations and how we can aid listeners in unfavorable listening conditions.

## 6. REFERENCES

- [1] Uchanski, R. M. 2000. Clear speech. In D.B. Pisoni and R.E. Remez (Eds.), *The Handbook of Speech Perception*. Blackwell Publishing. 207-235.
- [2] Bradlow, A., Kraus, N., Hayes, E. 2003. Speaking clearly for learning-impaired children: Sentence perception in noise. *J. Speech Hear. Res.* 46, 80-97.
- [3] Smiljanic, R., Bradlow, A. 2005. Production and perception of clear speech in Croatian and English. *J. Acoustical Society of America*, 118(3), 1677-1688.
- [4] Smiljanic, R., Bradlow, A. In press. Stability of temporal contrasts across speaking styles in English and Croatian. *J. Phon.*
- [5] Ferguson, S., Kewley-Port, D. 2002. Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *JASA*, 112, 259-271.
- [6] Bradlow, A., Bent, T., 2002. The clear speech effect for non-native listeners. *J. Acoustical Society of America*, 112, 272-284.
- [7] Studebaker, G. 1985. A ‘rationalized’ arcsine transform. *J. Speech Hear. Res.* 28, 455-462.