SPEECH RECOGNITION IN MODULATED AND UNMODULATED MASKERS FOR DIFFERENT TALKERS

Jingjing Xu   Mentor: Robyn Cox

Hearing Aid Research Laboratory, University of Memphis, Memphis, TN

Presented at the American Auditory Society Convention, Scottsdale, AZ, March 2007

ABSTRACT

This study explored whether the differences in speech recognition with modulated and unmodulated maskers are consistent across talkers who produce speech of varying intelligibility. The Speech Pattern Contrast (SPAC) test (Boothroyd, 1985) was used to quantify speech recognition overall and for speech features. Sentences were naturally produced by three female talkers, who had high, average, and low intelligibility (Cox et al, 1987). The maskers were modulated and unmodulated speech-shaped noises. Binaural sound field tests were carried out with the speech level of 55 dB SPL and a masker level of 62 dB SPL. Twenty-four young normal-hearing adults served as subjects. Results indicated that masker modulations are more valuable for understanding some talkers than others.

INTRODUCTION

Recent studies (Gatehouse et al., 2003; Lunner & Sundewall-Thoren, in press) showed that hearing-impaired listeners with higher cognitive ability get more speech recognition benefit than their lower cognitive counterparts from modulation in background noises. A similar study carried out in the Hearing Aid Research Lab (HARL) did not replicate this result. It was noticed that the talker used in the earlier studies produced very clear speech while the study in HARL used naturally produced sentences. These findings suggested that talker characteristics might make a significant contribution to the different results among studies. The present study was undertaken in attempt to explore whether the differences in speech recognition with modulated and unmodulated maskers are consistent across talkers who produce speech of varying intelligibility.

METHODS

Normal-hearing listeners were tested in a sound treated booth. Speech scores were measured:
- initial/final consonant voicing (ICV/FCV)
- initial/final consonant continuance (ICC/FCC)
- initial/final consonant place (ICP/FCP)
- composite score (COM)

Figure 1. Frequency responses of speech from the three talkers (55 dB SPL) and both maskers (62 dB SPL), measured in 1/3 octave bands.
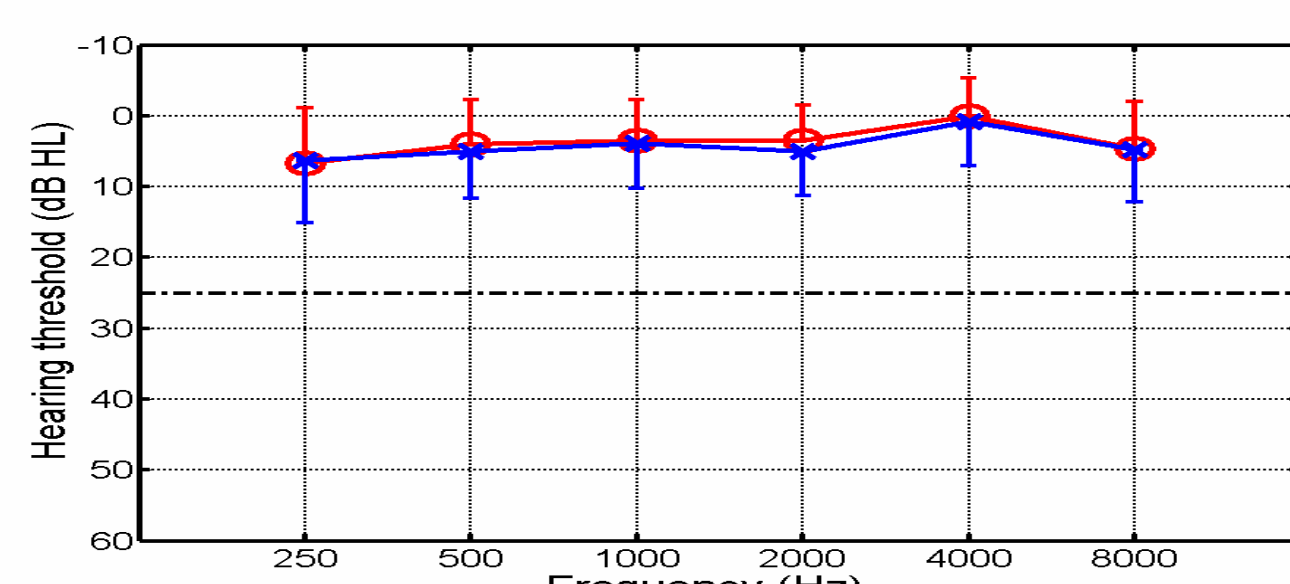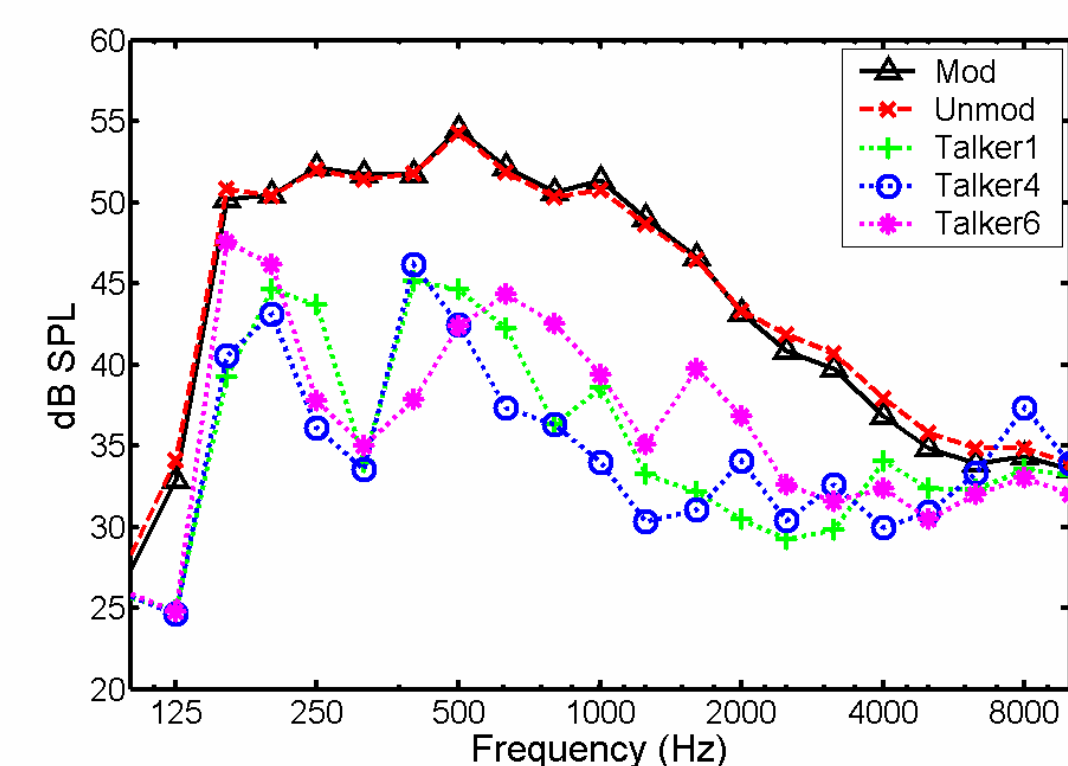
Figure 2. Mean audiometric data for the 24 normal hearing subjects. The error bars for mean data indicate 1 standard deviation.

Figure 3. Test setup.

RESULTS

In each speech feature, subjects with extreme scores were excluded due to the ceiling/floor effect. Hence, speech features of FCC and ICP were taken away because of the high concentration of full scores. Therefore, composite scores were averaged from the scores of the rest of the four features. Statistical analyses were performed on the speech features of ICV, FCV, ICC, FCP, and COM, using SPSS GLM within-subject design, in which recognition scores were used as dependent variable, while talkers and maskers were used as independent variables. Four specific questions aiming at our major research purpose were answered.

A. Are there differences in speech recognition among the talkers for each masker?

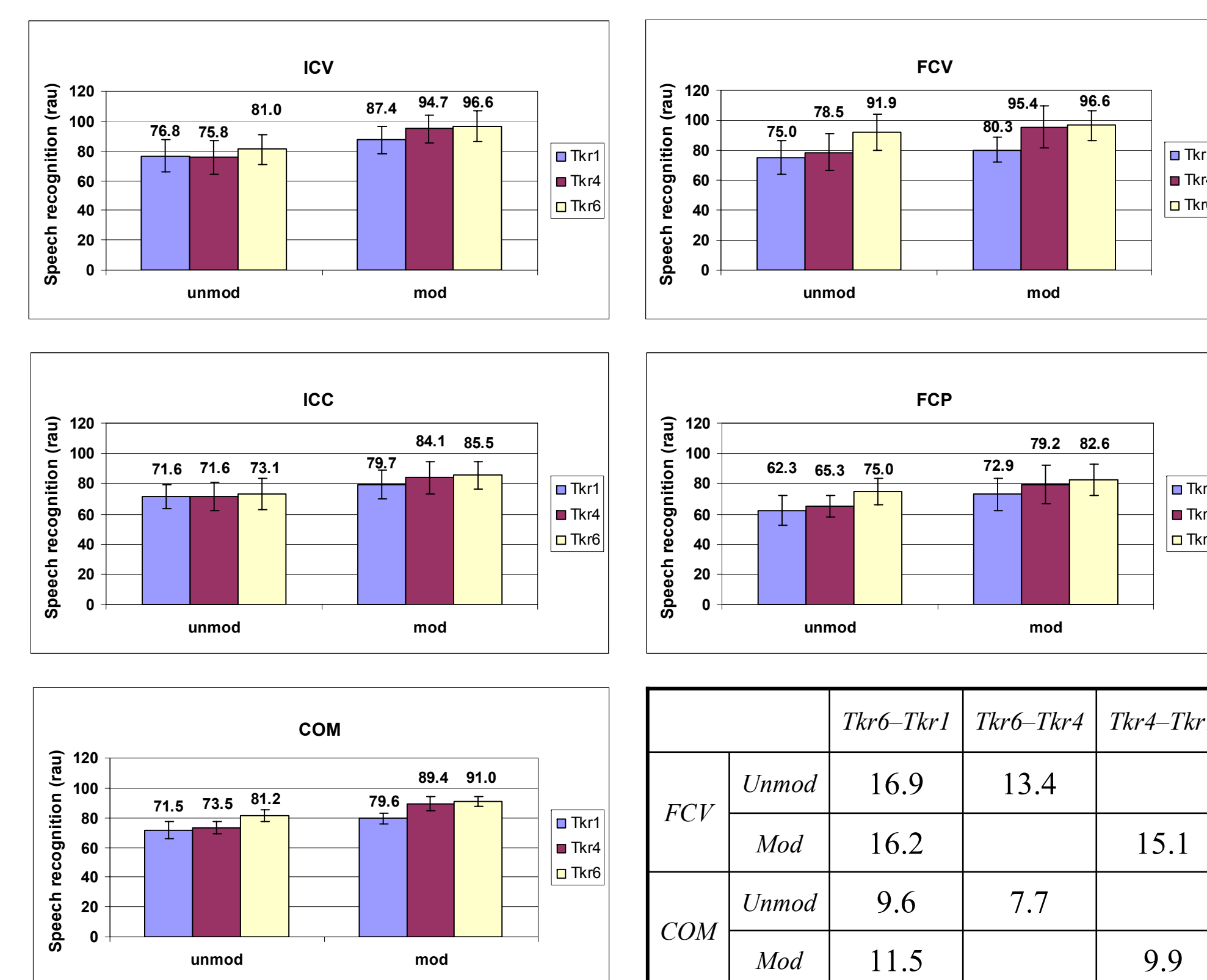|	Tkr6–Tkr1	Tkr6–Tkr4	Tkr4–Tkr1
FCV Unmod	16.9	13.4	
FCV Mod	16.2		15.1
COM Unmod	9.6	7.7	
COM Mod	11.5		9.9

Figure 4. Comparisons of speech recognition scores for the three talkers in each masker. Error bars indicate standard deviation. The table shows the significant differences (rau) between talkers in speech recognition score in each masker (p<.05).

Yes. Statistical analyses indicated that the interaction between talker and masker were significant in FCV and COM. For each masker, talker 1 yielded the lowest recognition scores among the three. Talker 6 yielded the highest scores in both masker conditions, but it was not significantly different from talker 4 in modulated masker condition (Fig.4). A similar trend was also found in other insignificant features. Therefore, the ranking of the three talkers, based on the intelligibility level they produced, is talker 6, talker 4, and talker 1.
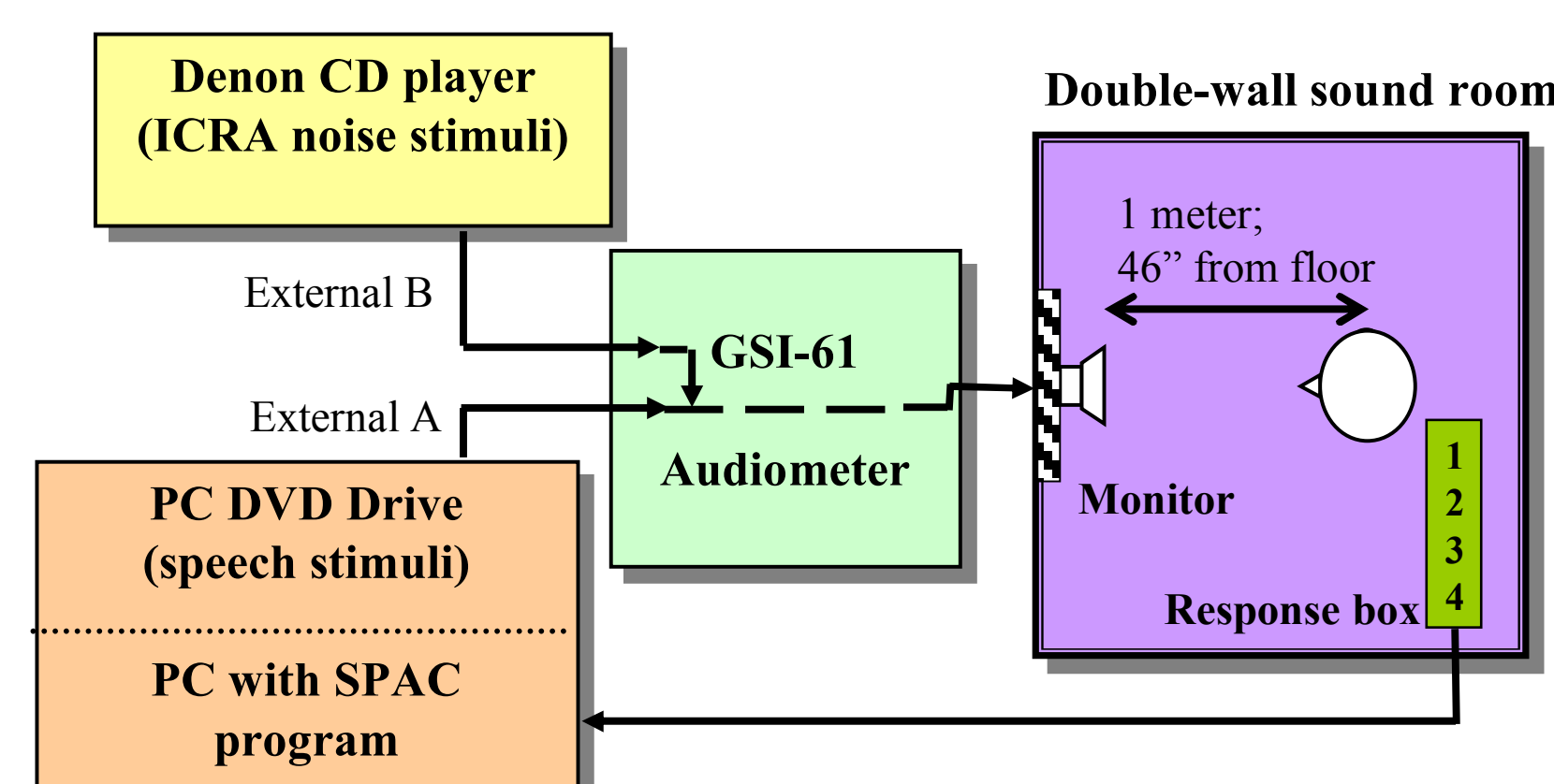
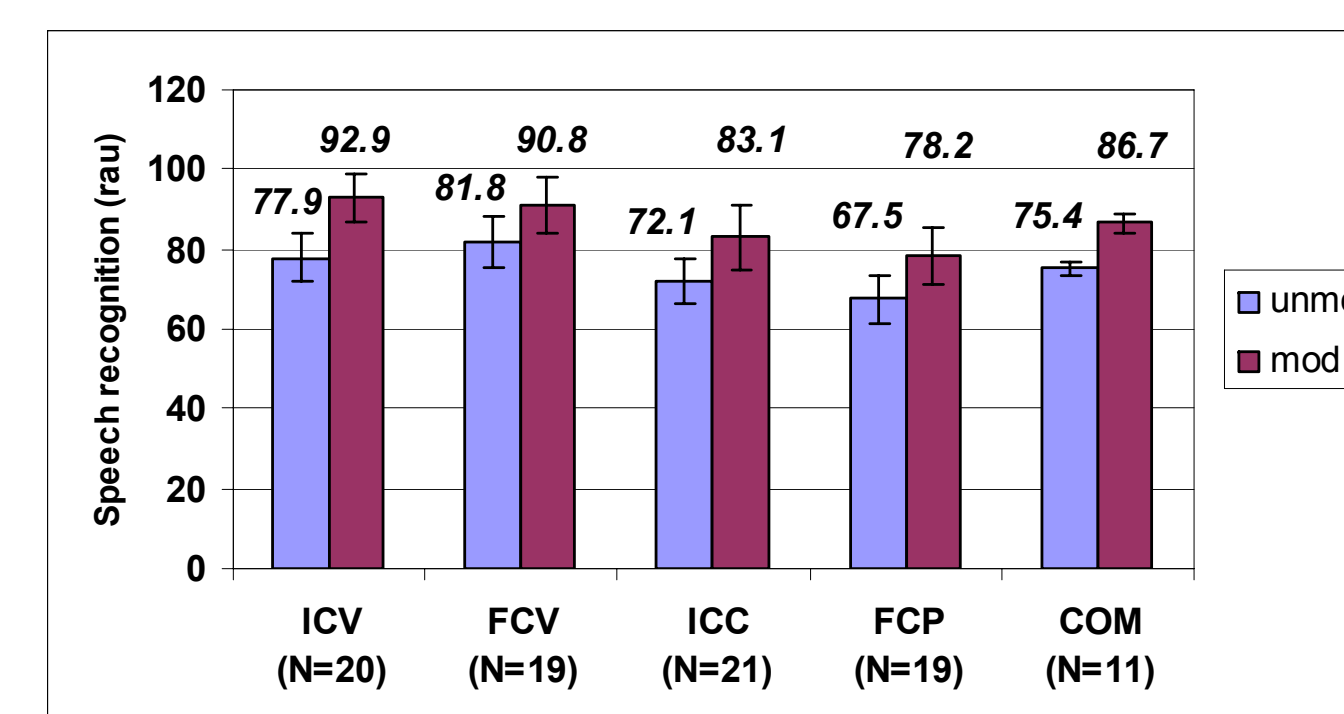### B. Do listeners obtain better speech recognition in the presence of the modulated masker?

Figure 5. Recognition scores and the corresponding standard deviations for unmodulated and modulated masker conditions (p<.05). N represents the number of valid subjects used.

Yes. By averaging the recognition scores across the three talkers, significantly better speech recognition performance was obtained with the modulated masker. The mean recognition improvement from modulation across all features and composite score is 11.4 rau.

C. Does masker modulation provide identical recognition improvement among the speech features?

No. Analyzing the averaged scores for the three talkers revealed that ICV was the most sensitive feature in modulation among the four chosen features (improvement: 15 rau) and FCV was the least sensitive (improvement: 9 rau). (Fig.5)

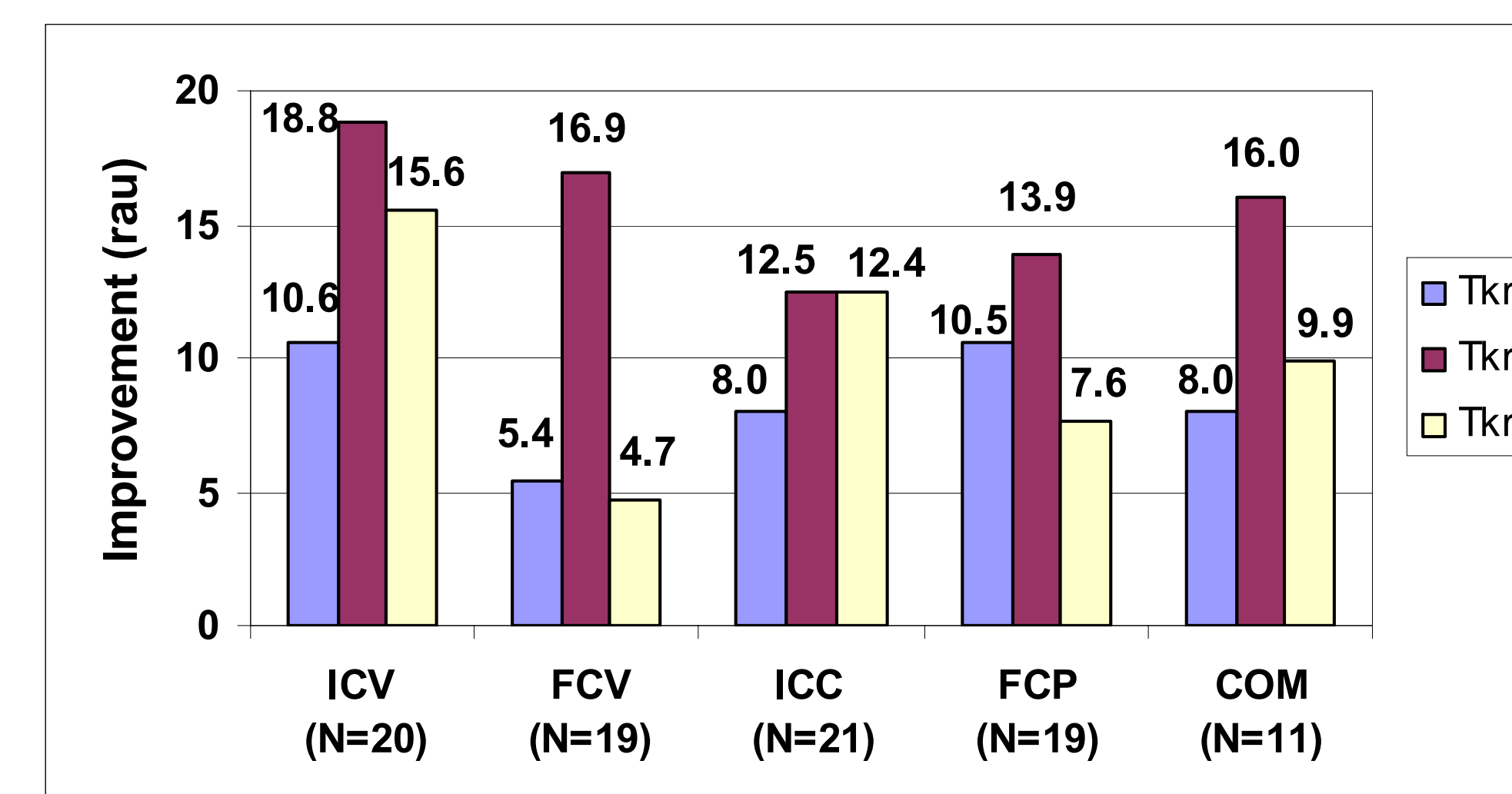### D. Do talkers with different intelligibility provide equal speech recognition improvement in modulation?

Figure 6. Speech recognition improvement between modulated and unmodulated maskers for each speech feature and the composite score.

No. The speech recognition improvement from modulation were significant in COM and FCV (p <.05). In COM, talker 4 yielded 8 rau higher than talker 1 and 6 rau higher than talker 6. In FCV, talker 4 yielded 11.5 rau higher than talker 1 and 12.2 rau higher than talker 6 (Fig.6). For other features, a similar trend of larger recognition improvement between modulated and unmodulated maskers with talker 4 was observed. Therefore, talker 4 provided a larger speech recognition improvement for normal hearing listeners from masker modulation than the other two talkers.

DISCUSSION

This study suggests that talkers varying in intelligibility could greatly affect speech recognition in the presence of a masker. Talker characteristics and talker-masker interaction could improve or deteriorate speech recognition performance. In Gatehouse et al. (2003) and Lunner and Sundewall-Thoren (in press), the speech materials were Four-Alternative Auditory Feature (FAAF) speech and Dentale II, respectively, both of which speech was produced in a very articulated manner. Thus, talker is one of the possible explanations to the unmatched speech recognition test results from different studies. Therefore, the impact of talker should not be neglected when testing speech recognition in modulated and unmodulated maskers.

Additionally, it is interesting that the maximum improvement in speech recognition from masker modulation was not yielded from the most intelligible talker (talker 6). However, the reason is still unclear.

CONCLUSION

When testing speech recognition in modulated versus unmodulated maskers for normal hearing listeners, talkers with different intelligibility levels do not result in consistent improvement. That is to say, masker modulations are more valuable for understanding some talkers than others. In order to generalize this finding, a similar test is needed for hearing impaired listeners.

REFERENCE

Boothroyd A. (1985). Evaluation of speech production in the hearing impaired: some benefits of forced-choice testing. *J. Speech Hear Res 28*: 185-196.

Cox, R.M., Alexander, G.C. and Gilmore, C.A. (1987). Intelligibility of average talkers in typical listening environments. *J. Acoust. Soc. Amer, 81*: 1598-1608.

Gatehouse, S., Naylor. G., & Elberling, C. (2003). Benefits from hearing aids in relation to the interaction between the user and the environment. *Int. J Audiol 42*:1S77-1S86.

Lunner, T. & Sundewall-Thoren, E. (in press). Interactions between cognition, compression, and listening conditions: effects on speech-in-noise performance in a two-channel hearing aid. *J. Am Acad Audiol*.

Please contact jxu3@memphis.edu for further information.
PDF-version of this poster can be obtained at
http://www.ausp.memphis.edu/harl/publications.html

HARL
Hearing Aid Research Lab