1pSC5: Within-session stability of acoustic features of conversational and clear speech

Sarah Hargus Ferguson[1], Shae D. Morgan[1], Lydia R. Rogers[1], and Eric J. Hunter[2]

[1]Department of Communication Sciences and Disorders, University of Utah, Salt Lake City, UT
[2]Department of Communicative Sciences and Disorders, Michigan State University, East Lansing, MI

UNIVERSITY OF UTAH COLLEGE OF HEALTH

MICHIGAN STATE UNIVERSITY

Background

In several planned studies, talkers will perform a set of speech tasks several times in a given recording session. For example, one study will have talkers perform the tasks four times (in quiet, in two levels of noise, and in reverberation) in each of two sessions – one in which talkers will be instructed to speak conversationally, and another in which they will be told to speak clearly, as though they were talking to an individual with hearing loss.

It is unknown, however, whether or how much the acoustic details of speech are affected by simple repetition of a speech task. Over the course of a recording session, talkers' speech might become less careful due to fatigue, boredom, familiarity with the speech materials, or some combination of these factors. Such repetition effects could offset the effects of different speaking conditions or of speaking style instructions given at the beginning of a session.

The present study assessed speech acoustic changes over four repetitions of a speech production task set performed under conversational or clear speech instructions. Our hypotheses are that, over the course of four repetitions,

1) Vowel space dimensions will decrease.
2) Speaking rate will increase.
3) Median voice pitch will decrease, more so for clear speech than for conversational speech.
4) Voice pitch range will decrease.

We also predict that within-session changes will be smaller than between-session speaking style effects.

Participants

Talkers were recruited from the University of Utah Department of Psychology participant pool. A total of 19 talkers (11 males) were recorded.

A subset of 10 talkers aged 18 to 24 (5 males) were selected for the present analyses. These talkers completed both recording sessions and met the following criteria by self-report:

- They had normal hearing and no history of speech or language disorders.
- They had grown up in Utah and affirmed that "I talk like I'm from around here."

Materials and Procedures

Talkers participated in two recording sessions. Within a session a set of three speaking tasks were performed and repeated four times per test session:

1. The Rainbow Passage (Fairbanks, 1960)
2. A list of 110 sentences
   a) 50 "vowel sentences" (/bVd/ in neutral context; 5 tokens each of 10 vowels)
   b) 6 lists from the Hearing In Noise Test (HINT; Nilsson et al., 1994)
3. A picture description task

Although the speaking task order was fixed for the four repetitions, the sentences were in a different random order and a different picture was used each time.

Recordings were made in a quiet, sound-treated room using a headset microphone (Shure SM-10) and a Marantz PMD 670 digital recorder.

In both sessions, talkers were given speaking style instructions (Ferguson, 2004) and a list of 15 practice sentences. Conversational speech was always recorded in the first session, and clear speech in the second session. Talkers were given feedback about conversational speech but not about clear speech.

When talkers felt they were comfortable with the speaking style instructions, they performed the task set a total of four times. They were offered a short break and water to drink after each task set.

Acoustic analyses

For each talker, four measures were taken from the 1st & 4th repetitions of the speaking task set in each style:

1. Vowel space perimeter in Barks: The sum of four Euclidean distances between steady-state F1 & F2 values for the vowels /i/, /æ/, /ɑ/, & /u/. For each task set and vowel, F1 and F2 were extracted from the second, third, and fourth productions of individual vowel sentences using Praat and then averaged.
2. Rainbow Passage Speaking rate in syllables per second: duration of passage divided by # syllables.
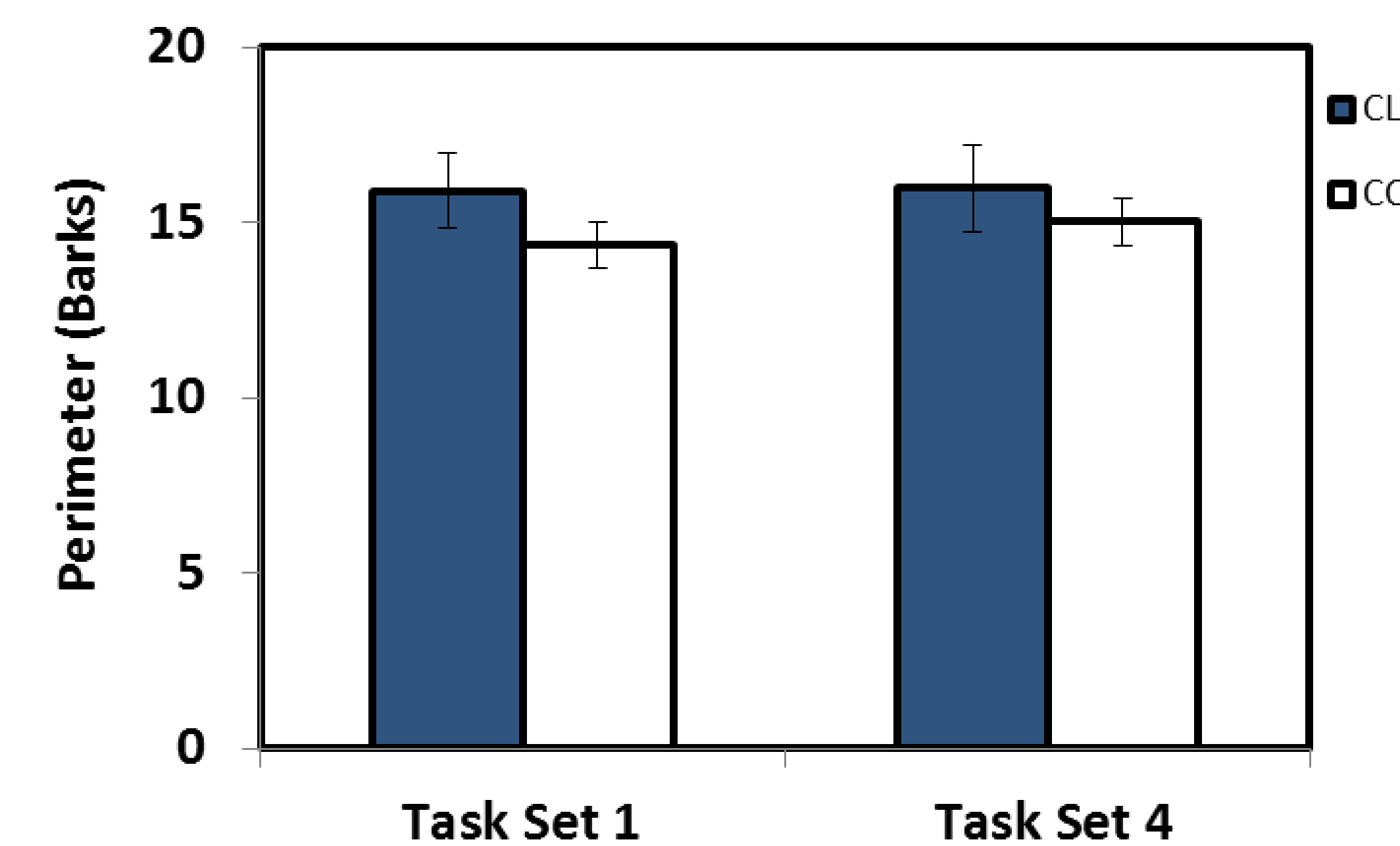
Two pitch measures were estimated from the Rainbow Passage using AudSwipePrime (Camacho, 2012):

2. Median voice pitch in Hz.
3. Pitch range in Hz: The difference between the 25th and 75th quartiles.
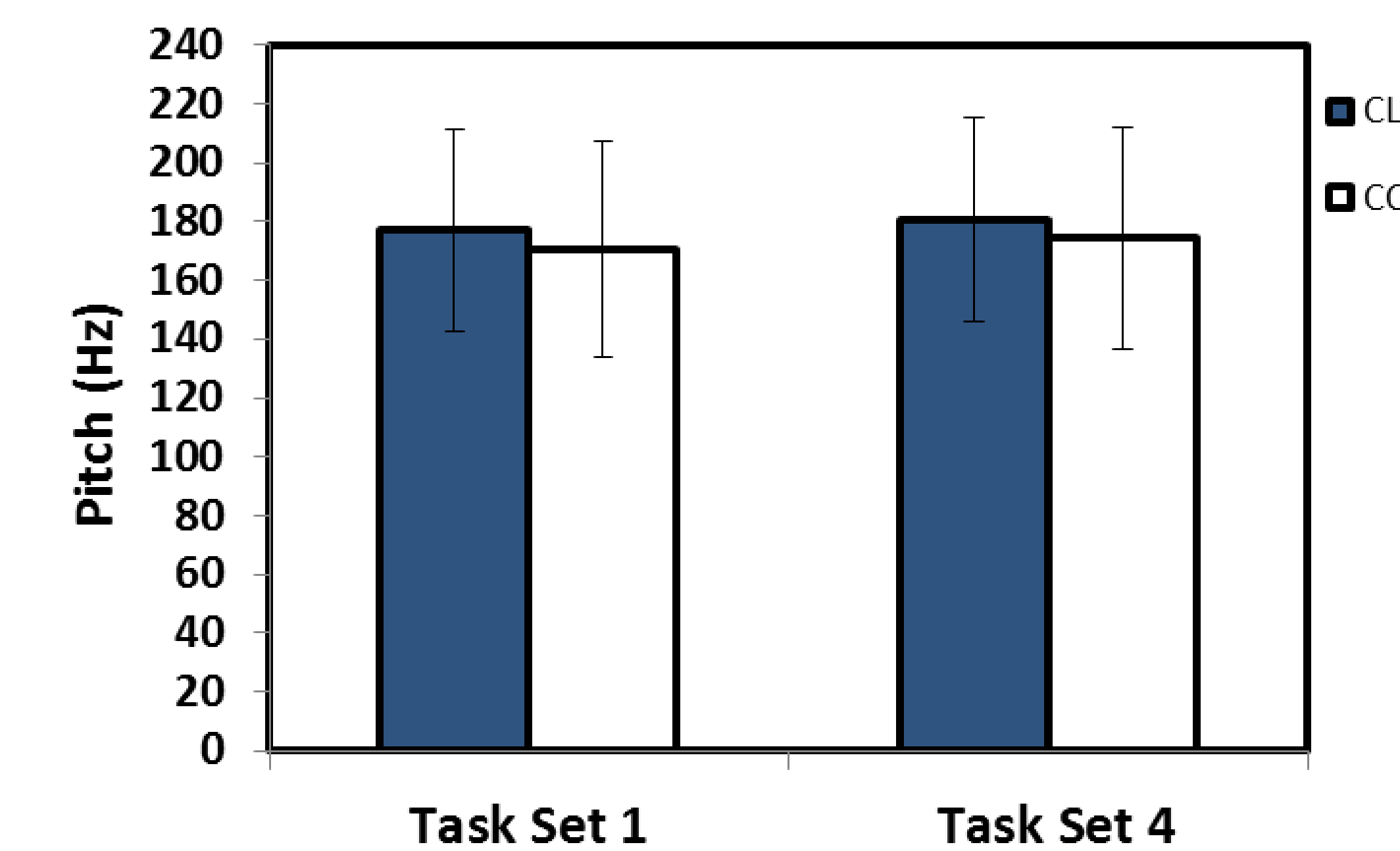
Results

For each acoustic metric, a two-way repeated measures ANOVA was carried out using SPSS to test the main effects of and interactions between speaking style (clear vs. conversational) and task set (1 vs. 4).
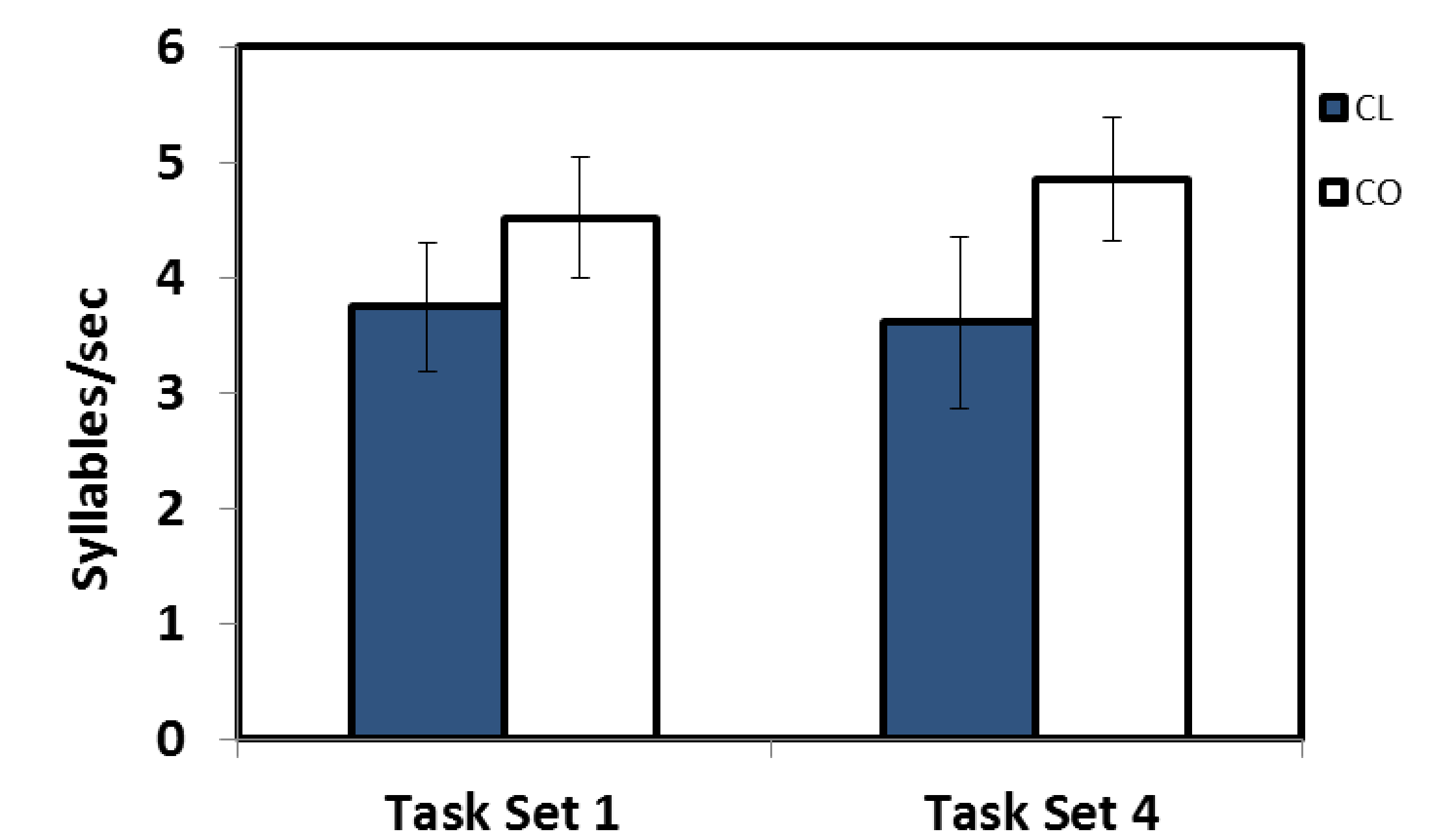
Vowel space: The main effects of speaking style and task set were significant [$F(1,9)$ = 9.172 and 7.581 respectively, $p<.04$]; the interaction between the two effects was not [$F(1,9)$ = 2.625, $p=.14$].

Speaking rate: The main effect of speaking style was significant [$F(1,9)$ = 14.66, $p<.005$]; the task set effect was not [$F(1,9)$ = 0.38, $p=.55$]. The interaction just missed significance [$F(1,9)$ = 4.88, $p=.054$].
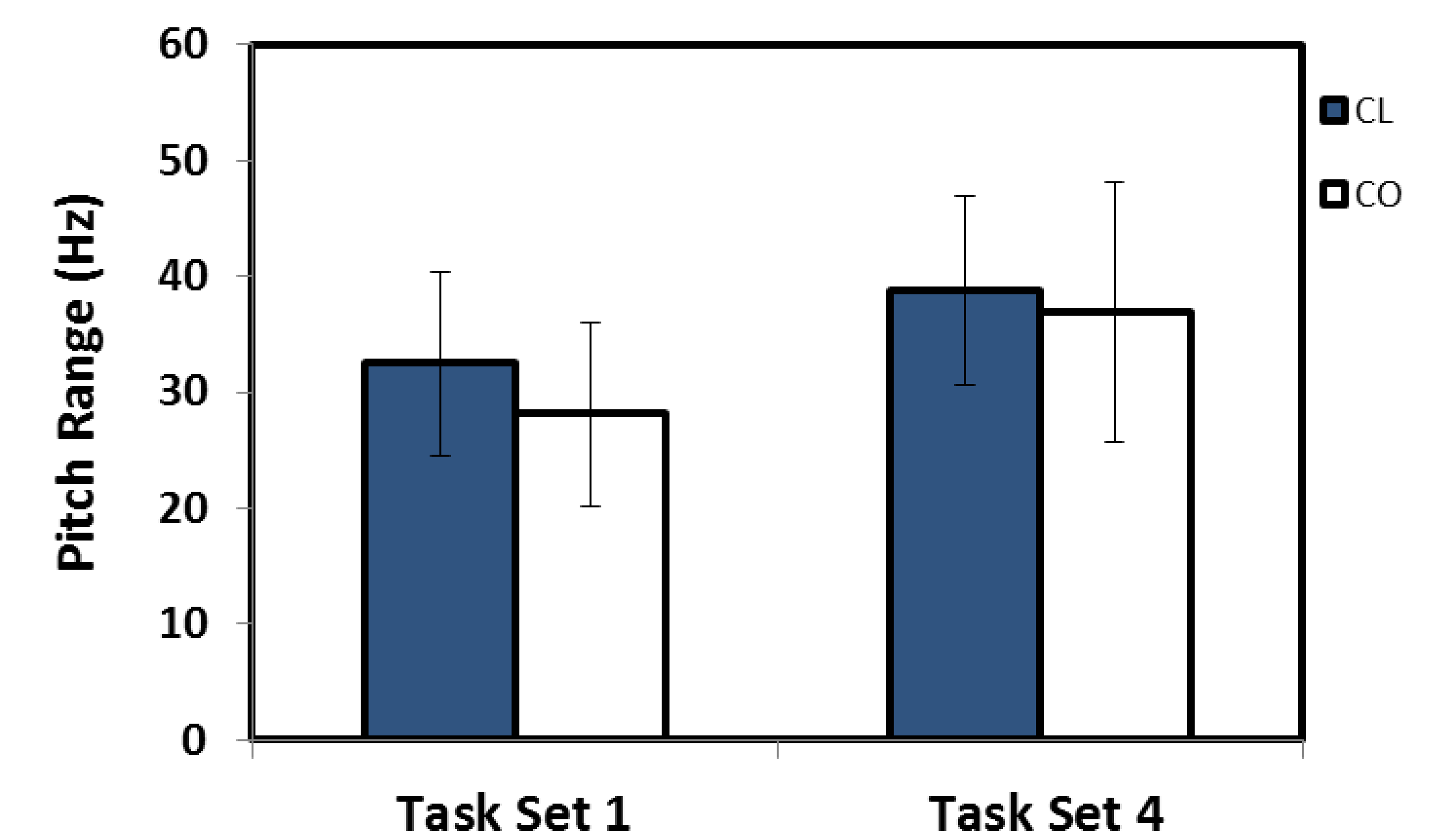

Median pitch: Neither the main effects of speaking style or task set nor the interaction between them were significant [$F(1,9)$ = 2.846, 2.040, and 0.0, respectively, $p>.10$].

Pitch range: The main effect of task set was significant [$F(1,9)$ = 13.15, $p<.01$]; the main effect of speaking style and the interaction were not [$F(1,9)$ = 2.192 and .506 respectively, $p>.17$].

Acknowledgments

References

Camacho, A. (2012). "On the use of auditory models' elements to enhance a sawtooth waveform inspired pitch estimator on telephone-quality signals," Information Science, Signal Processing and their Applications (ISSPA), 2012 11th International Conference on, IEEE, 1080–1085.

Fairbanks, G. (1960). *Voice and Articulation Drillbook* (2nd ed.). New York: Harper and Row.

Ferguson, S. H. (2004). "Talker differences in clear and conversational speech: Vowel intelligibility for normal-hearing listeners," J. Acoust. Soc. Am. 116, 2365-2373.

Nilsson, M., Soli, S. D., and Sullivan, J. A. (1994). "Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise," J. Acoust. Soc. Am. 95, 1085-1099.