RELATIONSHIP BETWEEN MEASURES OF INTELLIGIBILITY AND PHONETIC ACCURACY IN CHILDREN WITH AND WITHOUT CLEFT PALATE

Carrie L. Gotzke and Megan M. Hodge
Dept. of Speech Pathology and Audiology, University of Alberta, Edmonton, AB megan.hodge@ualberta.ca

1. INTRODUCTION

The Speech Intelligibility Probe for Children with Cleft Palate (SIP-CCLP) Ver. 3[1] is a computer-administered, 124-word imitation measure of a speaker's ability to make phonetic contrasts of English understandable to unfamiliar listeners. The contrasts sample the expected speech error patterns for children with cleft palate[1].These include manner and place preference, sibilant, voicing and glottal errors. For each error type, minimal word pairs that varied in their constituent consonants were selected as *SIP-CCLP* stimulus words. A child's recorded *SIP-CCLP* word productions can be judged using software-administered open- and closed-set tasks. In the open-set task, listeners identify the child's words, without knowledge of the target utterance, to yield an overall intelligibility score (percent words identified correctly) that reflects severity of the speech disorder. In the closed-set task, listeners first identify what sound is heard in a target position for each item and then rate the sound as "clear" or "distorted". The identification component captures errors of perceived sound substitutions or omissions, while the rating component captures perceived subphonemic differences that affect sound clarity. In the closed-set task, both identification judgments and distortion ratings contribute to a phonetic accuracy (PA) score. To guide treatment planning, error patterns that contribute to the child's intelligibility deficit can also be identified from an item analysis of listeners' responses on the closed-set task.

The purpose of this study was to determine how well *SIP-CCLP* PA scores predicted *SIP-CCLP* intelligibility scores and intelligibility scores obtained from a spontaneous speech sample for children with and without cleft palate. If *SIP-CCLP* and spontaneous sample intelligibility scores can be predicted with a high degree of confidence from *SIP-CCLP* PA scores, then it would more efficient clinically to have listeners perform only the *SIP-CCLP* closed-set task.

2. METHOD

2.1 Child Participants

Children age 3 years to 6 years, 11 months participated. The mean age of the 12 children with cleft palate was 51.2 mos (SD=11.7 mos) and of the 12 children without cleft palate was 50.1 mos (SD = 11.1 mos). Children with cleft palate had hearing within normal limits, no concomitant physical or cognitive impairments, and receptive and expressive language abilities within normal limits on a standardized measure. Children had a variety of palatal impairments ranging from submucous cleft to bilateral cleft lip and palate. The 12 children without cleft palate all passed a hearing[2] and oral mechanism screening[3], and scored at or above the 16th percentile on a standardized measure of receptive and expressive language and articulation.

Audio recordings were made in a quiet environment directly to a personal computer as digital audio files (SR 48 KHz; QS 16 bits). Recordings were made using a Shure WH20 unidirectional dynamic headset microphone, Audio Buddy Dual Mic Preamplifier and custom software[1]. The *SIP-CCLP* software created a unique order of the stimulus items for each administration. The child was instructed to repeat the name of the picture displayed on the computer monitor after the examiner modeled the word. The software recorded the child's production. A 15-minute spontaneous speech sample was also collected from each child using play scenerios, following Shriberg's[4] procedures. Each utterance in the child's sample was transcribed orthographically. Utterance boundaries were determined using Shriberg's conventions. A section with multiple child utterances longer than two words and few examiner turns was located in the transcript. Within this section, a subsection containing 100 consecutive words[5] and few utterance boundaries was selected for the listening task. During playback, each utterance was presented as an audio-file, in the order of occurrence in the transcript subsection. The transcript served as the key for scoring listeners' responses.

2.2 Listener Judges and Tasks

Seventy-two university students served as listener judges. All listeners had Canadian English as their first language and passed a hearing screening[2]. Three students were randomly assigned to judge each child's *SIP-CCLP* recordings. Three different listeners were randomly assigned to judge each child's spontaneous speech sample. All listening sessions took place in a Madsen OB822 sound booth. Stimuli were presented through a Technics Stereo Integrated Amplifier (model SU-V460) connected to ElectroVoice S-40 compact monitor speakers located in the booth. Listener judges were asked about the comfort level of the playback volume after the practice items for all three tasks (*SIP-CCLP* open- and closed-set, spontaneous sample) and adjustments made as necessary within a range of 55 – 65 dBA.

For the *SIP-CCLP* open-set identification task each listener was instructed to type the word or words heard, using the computer keyboard. If the words were not clear, the listener was instructed to guess. For the *SIP-CCLP* closed-set task listeners were told that they would see two words with a sound or sounds underlined, a "blank" button and a "can't identify" button. They were instructed to focus on what was heard in the underlined position as they heard the child's production of the item. If what was heard corresponded to one of the underlined choices, the judge was instructed to select it. If a different sound(s) was heard in the underlined position, judges were instructed to select the "blank" button and type in what was heard. If the sound(s) heard in the underlined position could not be identified, judges were instructed to select the "can't identify" button. Judges were also instructed to rate what was heard in the underlined position as "clear" or "distorted" if one of the first three buttons (minimal contrast pair or the blank) was selected. The *SIP-CCLP* software randomly generated the order of item presentation for both open- and closed set judging tasks. For the spontaneous speech sample judging task, listeners were told that they would hear a series of utterances varying in word length. They were instructed to write down each word that they heard. Listeners heard each utterance twice if desired.

For the *SIP-CCLP* closed-set task, the responses for each listener judge were examined to determine the number of correct/clear (2 points), correct/distorted (1 point), and incorrect (0 points) items and then the assigned scores for all items were tallied. This sum was converted to a percentage of the total possible score (194 items x 2 = 388). The mean of the three listeners' scores served as the child's PA score. For the *SIP-CCLP* and spontaneous sample intelligibility scores, the number of words identified correctly by each judge was converted to a percentage. The mean of the three judges' scores served as the child's intelligibility score.

3. RESULTS

PA scores for the children with cleft palate (M = 73.4%, SD = 11.6) were significantly lower (t = -4.41, p< .000) than those for the children without cleft palate (M = 89.4%, SD = 5.0). *SIP-CCLP* open-set intelligibility scores for the children with cleft palate (M = 53.3%, SD = 15.9) were significantly lower (t = -4.31, p < .000) than those for the children without cleft palate (M = 76.7%, SD = 10.1). Spontaneous speech open-set intelligibility scores were significantly lower (t = -3.03, p = 0.003) for the children with cleft palate (M = 67.5%, SD = 13.3) than for the children without cleft palate (M = 83.6%, SD = 12.7). As shown in Table 1, Pearson correlation coefficients were high between *SIP-CCLP* PA scores and *SIP-CCLP* intelligibility scores. For the children with cleft palate, correlation coefficients were moderate and similar between *SIP-CCLP* PA and spontaneous speech intelligibility scores (.63) and between *SIP-CCLP* and spontaneous speech intelligibility scores (.62).

Interlistener reliability was examined for the three listeners for each child, for each score type, using the intraclass correlation coefficient (ICC). ICC values for *SIP-CCLP* PA scores and *SIP-CCLP* and spontaneous sample intelligibility scores for the children with cleft palate were .92, .95 and .87 respectively and, for the children without cleft palate, were .91, .96 and .93.

Table 1. Correlation matrix for *SIP-CCLP* phonetic accuracy (PA) and *SIP-CCLP* and spontaneous speech intelligibility scores.

	SIP-CCLP Intelligibility Scores		Spontaneous Speech Intelligibility Scores	
	With CP*	Without CP	With CP	Without CP
SIP-CCLP PA Scores	.93 (p <.000)	.87 (p < .000)	.63 (p = .03)	.48 (p = .113)
SIP-CCLP Intelligibility Scores			.62 (p =.016)	.67 (p = .008)

*CP = Cleft Palate

4. DISCUSSION

ICC values of greater than .9 indicated that listeners achieved acceptable reliability in generating *SIP-CCLP* PA and *SIP-CCLP* intelligibility scores. PA scores over estimated spontaneous speech intelligibility scores on average by 5.9% and 5.8%, and also over-estimated *SIP-CCLP* intelligibility scores on average by 20.1% and 12.7%, respectively, for the children with and without cleft palate. R^2 values (86% and 76%) were high between *PA* and *SIP-CCLP* intelligibility scores for children with and without cleft palate. These results suggest that listeners' responses on the *SIP-CCLP* closed-set judging task can be used to generate PA scores that show a highly predictable relationship to *SIP-CCLP* intelligibility scores. In clinical settings, where time for listeners to perform speech identification tasks is limited, use of the same listener task to estimate an overall intelligibility score, as well as provide error specific information for treatment planning, is an advantage.

REFERENCES

[1]Gotzke, C.L. (2005). *Speech intelligibility probe for children with cleft palate version 3: Assessment of reliability and validity.* Unpublished Master's Thesis, University of Alberta, Edmonton, AB.

[2]American Speech and Hearing Association (ASHA). (1985). Guidelines for identification audiometry. *American Speech and Hearing Association Journal*, 27, 409-453.

[3]Dworkin, J. & Culatta, R. (1996). *Dworkin-Culatta Oral Mechanism Exam – Treatment.* Nicholasville, KY: Edgewood Press.

[4]Shriberg, L. (1986). *Programs to Examine Phonetic and Phonological Evaluation Records (P.E.P.P.E.R.) Version 4.0.* Madison, WI: University of Wisconsin-Madison.

[5]Weiss, C. (1980). *Weiss Comprehensive Test of Articulation.* Hingham, Massachusetts: Teaching Resources Corporation.

ACKNOWLEDGEMENTS

Supported by a grant from the Canadian Language and Literacy Research Network and a graduate scholarship awarded to the first author from the National Science and Engineering Research Council of Canada.