Reliability and Validity of a Sentence Intelligibility Measure for Children

Megan M. Hodge & Carrie L. Gotzke
University of Alberta, Edmonton, AB CANADA

CANADIAN LANGUAGE & LITERACY RESEARCH NETWORK

ASHA Convention
November 22, 2008
Chicago, IL

Introduction

Background

- Need for reliable and valid word identification procedures to measure young children's speech intelligibility[1]:
  - Quantifying severity of speech disorder
  - Measuring speech outcomes
- Word identification intelligibility measures based on a conversational speech sample:
  - Have high ecological validity[2,3]
  - Are time consuming
  - Are challenging for children with severe intelligibility deficits
- *Test of Children's Speech Plus (TOCS+ software)* developed to digitally audio record imitated utterances from young children and play these to listeners for word identification[4]
  - Percent words identified correctly = "intelligibility score"
- *TOCS+ Sentence Intelligibility Measure*
  - *TOCS+* software randomly selects items from utterance pools (range 2 – 7 words) to create "unique" 80-word test
  - User specifies maximum item length (number of words) in a test (adjust for child's MLU)

Objectives

- Determine alternate form reliability of *TOCS+ Sentence Intelligibility Measure* for children with and without speech sound disorders (SSD)
- Assess validity of *TOCS+ Sentence Intelligibility* scores by determining their relationship with word identification scores obtained for 100 contiguous words[5,6,7] from a spontaneous speech sample for the same children

Methods

Child Subjects

- N = 72 English-speaking children (18 at each of 4 ages: 3, 4, 5 & 6 yrs)
  All children had receptive language[8], hearing abilities[9], and speech mechanism structure[10] within normal limits.
  In each age group:
  - 9 with typically developing speech
    Obtained scores ≥ 16th %ile on articulation subtest of the *Fluharty Preschool Speech and Language Screening Test (Fluharty-2)*[11] and no history of parent concern or speech-language therapy
  - 9 with speech sound disorders (SSD)
    Identified by referring SLPs and scores < 16th %ile on the *Fluharty-2* articulation subtest.

Recording *TOCS+ Sentence Tests* and Conversational Samples

TOCS+ Sentence Tests

- Two forms administered within two week period
- Recordings made using standard head-mounted mic/pre-amp)
  (Shure WH20 XLP microphone/AudioBuddy Dual Mic Pre-amplifier)
  - Longest utterance in tests:
    - Children with typical speech development:
      - 3 yr. = 4 words; 4 yr. = 5 words; 5 yr. = 6 words; 6 yr. = 7 words
    - Children with SSD
      - Fit with MLU (e.g., for MLU of 4 words, longest utterance = 4 words)

Spontaneous Speech Sample

- 15 minute spontaneous speech sample elicited using interactive play[11] and audio recorded using *TOCS+ Recorder/Player software (TOCS+RP™*[12]) and standard mic and pre-amp
- Starting at 2nd minute of each sample, a 100-word contiguous sample was segmented into utterances following procedures of Shriberg *et al.*[13]
- Each utterance saved as a .wav file

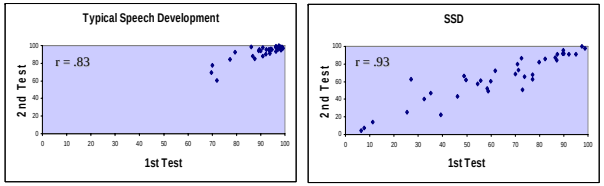
Judging Samples

- Adults with normal hearing, English as a first language, post-secondary education, 18 - 35 years of age
- 3 different listeners judged each child's recordings for each *TOCS+ Sentence Test* and 100-word spontaneous sample
  - Word identification task – maximum 2 presentations per item
  - 72 children x 3 conditions x 3 listeners = 648 listening sessions
    Some listeners participated in more than one session but these were at least a month apart and were never for the same child
- *TOCS+* Intelligibility software ⟶ presented sentence items
- TOCS+ RP™ software ⟶ presented spontaneous utterances
- Listeners instructed to type in the words they heard the child say

Results

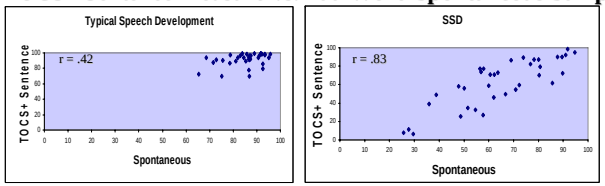
Percentage of words identified correctly, based on mean of 3 listeners' scores = **intelligibility score**

TOCS+ Sentence Measure: Alternative Forms



- Test Time x Group x Age ANOVA
  - No significant difference between test forms
  - Significant group and age effects

TOCS+ Sentence Measure vs. 100-Word Spontaneous Sample



- Sample Type x Group x Age ANOVA:
  - Significant Sample by Group Interaction $F = 10.04$; $p = .002$

Post hoc testing:

- **Typical Speech Development** $F = 21$; $p = .000$
  - Significant difference between sample conditions
- **SSD** $F = .98$; $p = .329$
  - Nonsignificant difference between sample conditions

Conclusions

- **Strong positive correlation between *TOCS+ Sentence* forms for both groups of children**
  - Intelligibility scores did not differ significantly between forms
- **Strong positive correlation between *TOCS+ Sentence* & 100-word spontaneous speech sample scores for children in SSD group**
  - Intelligibility scores *did not differ between sample type for group with SSD* but did for group with typical speech development

TOCS+ Sentence Intelligibility measure shows promise as an efficient, stable and representative measure of ability of young children with SSD to make audio recordings of their speech understandable to unfamiliar listeners.

References

[1]Hodge, M. & Gotzke, C. (2007). Preliminary results of an intelligibility measure for English-speaking children with cleft palate. *Cleft Palate-Craniofacial Journal*, *44*, *163-174.*

[2]Weston, A. & Shriberg, L. (1992). Contextual and linguistic correlates of intelligibility in children with developmental phonological disorders. *Journal of Speech and Hearing Research*, 35, 1316-1332.

[3]Flipsen, P. (2006). Measuring the intelligibility of conversational speech in children. *Clinical Linguistics and Phonetics*, 20(4), 303-312.

[4]Hodge, M. & Daniels, J. (2007). *Test of Children's Speech Plus (TOCS+ Plus) ver. 5.3*. Edmonton, AB: University of Alberta [computer program].

[5]Gordon-Brannan, M. & Hodson, B. W. (2000). Intelligibility/severity measurements of prekindergarten children's speech. *American Journal of Speech-Language Pathology*, *9*, 141-150.

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

[7]Stimley, M. & Hambrecht, G. (1999). Comparison of children's' single word articulation proficiency, single-word intelligibility and conversational speech intelligibility. *Journal of Speech-Language Pathology and Audiology*, 23, 19-23.

[8]Fluharty, N.B. (2001). *Fluharty Preschool Speech and Language Screening Test (Fluharty-2)*. Austin, TX: Pro-Ed Inc.

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

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

[11]Shriberg, L. D. (1986). *PEPPER: Programs to examine phonetic and phonologic evaluation records*. Hillsdale, NJ: Lawrence Erlbaum.

[12]Hodge, M. & Gotzke, C. (2007). *TOCS+ Recorder-Player Software – TOCS+RP™ ver. 2.0*. Edmonton, AB: University of Alberta.

[13]Shriberg, L.D., Kwiatkowski, J. & Rasmussen, C. (1990). *The Prosody-Voice Screening Profile (PVSP): Scoring Forms and Training Materials.* Tucson, AZ: Communication Skill Builders.