Test-retest reliability of independent phonological measures of 2-year-old speech

Katherine Wittler  
*University of Nebraska at Omaha*

Shari L. DeVeney  
*University of Nebraska at Omaha*, sdeveney@unomaha.edu

Follow this and additional works at: https://digitalcommons.unomaha.edu/spedfacproc

Part of the Phonetics and Phonology Commons, and the Special Education and Teaching Commons

Please take our feedback survey at: https://unomaha.az1.qualtrics.com/jfe/form/SV_8cchtFmpDyGfBLE

**Recommended Citation**

https://digitalcommons.unomaha.edu/spedfacproc/3
Test-Retest Reliability of Independent Phonological Measures of 2-year-old children

Katherine Wittler, B.S. and Shari DeVeneys, Ph.D. CCC-SLP

**Background**

**Informal Assessment Measures**
- Assessment tool not intended for comparison to larger group
- Often used along with standardized, norm-referenced assessments in a comprehensive evaluation of speech-language skills
- E.g., analysis of communication samples, observational

**Independent Phonological Measures**
- Measure speech-sounds without comparison to adult standard
- Used to obtain descriptive information to establish baseline
- E.g., Phonetic inventory (PI); Record of different sounds used, even sounds not produced in appropriate word positions ("bit" for "cat")
- E.g., Word Shape Analysis (WS); Record of sound combination complexity used to form words ("cat" represents CVC shape)

**Test-Retest Reliability**
- Degree to which a measurement is stable over time
- Short-term reliability critical for indications of baseline performance and progress tracking over the course of treatment

**Significance**
- ASHA acknowledged need for evidence-based practices in assessment and treatment (ASHA 2004, 2005) to satisfy eligibility guidelines and provide effective services
- SLPs may assume commonly used informal measures are reliable, but little evidence to support.

**Existing Literature**
- Morris (2009) found test-retest (within one week) instability for PIs with typically developing 18- to 22-month-olds,
- Preston, Ramsele, Oller, Edwards, and Tobin (2011) found that weighted measures of sound accuracy has similar reliability and validity as a relational phonological measure (PCC-R) for a variety of normative and clinical child populations.
- Heilmann, DeBrock, and Riley-Tillman (2013) noted test-retest reliability for other informal measures calculated from communication samples obtained from kindergarten-age children.

**Aim of current study**
- Plot the extension of Morris (2009) work for determining test-retest reliability of independent phonological analyses over time for slightly older child population, 24- to 36-month olds.

**Research Question**
- What is the test-retest reliability (within one week) of the phonetic inventory and word shape analysis measures calculated using intelligible words produced during a 20-minute conversational speech sample for 24- to 36-month old children?

**Method**

- Participants. (n = 3); Ages 29 months to 33 months (M = 30.66, SD = 2.08); mono-lingual
- English speaking
- Identified as typically developing using the following criteria: SS > 85 on Preschool Language Scale 5 (PLS-5); > 25th percentile on MacArthur Bates Communicative Development Inventory (CDI)

**Procedures.** Parent-child dyads participated in two 20-minute play sessions one week apart in a university clinical setting. Toys for each session were randomly assigned and communication samples were obtained during play sessions.

- Sessions were video recorded for later viewing and transcription using the international phonetic alphabet (IPA).
- Transcribers were the 1st author and two trained student research assistants majoring in speech-language pathology. Following procedures used by Morris (2009), initial inter-rater reliability was just above 62%; however, the transcribers re-watched each instance of disagreement up to three times, discussed, and reached agreement. Transcribers agreed on 100% of re-watched instances.
- The 1st author calculated PIs and WS analyses for each participant each session. Inter-rater reliability was established with the faculty advisor who re-analyzed 20% of the data. PI Inter-Rater Reliability was an average of 87% (Range = 85-90%); WS Inter-Rater Reliability was 100%.

**Results**

**Phonetic Inventory.** Initial consonants (productive): P1 was consistent (15 in S1; 14 in S2; P2 and P3 were inconsistent (15;18 and 4;10, respectively)
- Final consonants (productive) - all relatively consistent (14;13, 12;10, 7;7)
- Consonant clusters (productive) – P1 and P2 were inconsistent (16;12, 5;13) P3 was consistent (4;4)

**Word Shape Analysis.** No substantive differences noted across sessions; all produced at least two different words in each of eight target word shape categories (V, CV, CVCC, VC, CVC, CCVC, CVC, CVCC).

**Conclusions and Implications**
- Some support for extension of Morris (2009) findings to older child population. Two of three participants obtained consistent Word-Initial PIs, differences were noted for one participant.

**Selected References**

**Acknowledgements**
- This project was funded through the Fund for Undergraduate Scholarly Experiences (FUSE) grant awarded to the first author. The authors would like to acknowledge the children and their families who participated in this study as well as Ann Walker and Kristin VanVugt for their assistance coding the data set with the UNO Toddler Communication Lab.