Development of a screening tool to identify babies at risk of language delay: A preliminary study

Siddhi Shamkant Pathak & Pallavi Kelkar

Introduction:
Speech and language delays and disorders can pose significant problems for children and their families. Delayed language development has a serious impact on adolescence and adulthood in terms of educational, social and emotional development (Brown, 2014; Levy, 2018). Thus, identifying language delay as early as possible becomes important.

Need for Study:
There are very few tools available for early language screening (Larson, 2016). The available tools for language screening or assessment are either time consuming to administer (Communication DEALL Developmental Checklist; Karanth, 2007), not freely available (Bayley Scales of Infant and Toddler Development Screening Test; Bayley & Aylward, 2005) or do not have an indigenous normative (Receptive Expressive Emergent Language Scale for Kannada speaking children; Madhu et al., 2009).

Some tools (Ages and Stages Questionnaire; Squires & Bricker, 2009) consist of other domains along with speech and language skills which cannot be assessed by an SLP (Singh et al., 2016). In a developing country like India, primary health concerns might receive a higher priority than speech and language delays (Blake et al., 2009); and parents are likely to visit a pediatrician more frequently than they might visit a speech-language pathologist. There is, therefore, the need for a language screening tool with an indigenous normative that is easy to administer for pediatricians, parents and speech language pathologists alike. As the doctor to patient ratio is low in India (Mohan et al., 2017), the tool also needs to be quick to administer.

Aim & Objectives:
To develop a quick and easy to administer language screening tool for babies from 6 to 21 months of age.
In line with the aim, the objectives were:
1) To develop a language screening tool representing semantic, syntactic and pragmatic aspects of language for the age group of 6 to 21 months
2) To investigate reliability and criterion validity of the tool
3) To determine the pass-refer criteria for the tool

Method:
Phase I: Creating an item pool
An item pool of 78 items representing syntax, semantics and pragmatics was created after a review of various sources on development (Amidon & Heistad, 1985; Aylward, 2009; Rabiah, 2012) and classified into five screening levels by the first author, specifically, 6 to 9 months (Level 1), 9 to 12 months (Level 2), 12 to 15 months (Level 3), 15 to 18 months (Level 4) and 18 to 21 months (Level 5). Items in each screening level allowed a three month lag with reference to the minimum age for that level, to account for individual variability in developmental milestones.

Phase II: Validation by judges
The classified item pool was then given to two speech-language pathologists with over ten years of clinical and research experience. Within each level, the judges selected items which they would choose to use for quick screening of language in a clinical setting. Additionally, they independently classified each item as representative of syntactic, semantic or pragmatic abilities. Discrepancies in categorization were discussed between the judges and resolved. Modifications and changes were made based on the feedback and suggestions received from the judges. 42 items were selected and three items were added as suggested by the two experts making the total number of items 45. This was called the Screening Test of Early Language Development - Test Version (STELD-T).

Phase III: Administration of the test version
Age appropriate levels of the STELD-T were administered to 74 babies ranging in age from 6 months to 21 months. All babies had Marathi as their first language. The Receptive Expressive Emergent Language Scale (REELS; Bzoch & League, 1971) was also administered for checking criterion validity. As the parents were educated in English and the REELS was worded in English, the STELD-T was administered to the parents in the English language.

Phase IV: Psychometric evaluation
KR-20 was used for assessment of reliability as the data was dichotomous. Cohen’s Kappa was used for evaluating agreement between REELS and STELD-T. To convert REELS age values to dichotomous data for using the Kappa statistic, an average of the receptive and expressive language age was computed for each baby. This was then compared to the chronological age (corrected age in case of a preterm delivery). An average value within three months of the chronological age was coded as 1, and all other values coded as 0.

Results & Discussion:
Phase I: The item pool represented all three aspects of language. This could facilitate early detection of disorders that might affect pragmatic development; and might be crucial to reducing overall impact of such disorders on a child’s development (Cummings, 2014).

Phase II: Checking the validity of statements by two experienced judges facilitated high face validity of the tool. Cross validation of items through categorization by judges into syntax, semantics and pragmatics ensured that the STELD-T represented all aspects of language development.

Phase III and IV: A KR-20 value of 0.46 for screening level 1 indicated moderate reliability (Stephens et al., 2006). A Kappa coefficient of 0.45 (p=0.037) indicated moderate agreement between the results (Landis & Koch, 1977; Sim & Wright, 2005) of STELD-T and REELS. This result may however be considered acceptable since STELD-T has items pertaining to the pragmatic aspect of language, while REELS focuses...
largely on syntax and semantics.
KR-20 for Level II was 0.85 indicative of high reliability (Taber, 2017). Cohen’s Kappa value was 0.64 (p=0.004) indicating substantial strength of agreement with REELS.
Internal consistency for Level 3 was 0.737 indicating good reliability. Kappa coefficient was 0.43 (p=0.057) suggestive of moderate strength of agreement.
Level 4 had a KR-20 value of 0.60 suggestive of adequate reliability. A Kappa coefficient for of 0.75 (p=0.003) suggested substantial strength of agreement.
For Level 5, internal consistency was 0.92 indicative of high reliability. The Kappa coefficient value was 1.0 (p=0.000) indicating a perfect agreement between STELD-T Level 5 and REELS.
A 70% pass-refer criterion led to decreased sensitivity, while an 80% pass-refer criterion reduced specificity. It also led to a reduction in the Kappa coefficient for Level 5. A 75% criterion was thus finalized as a pass-refer criterion for the STELD-T.

Summary & Conclusion:
Data on prevalence of disabilities in the age range of 0-3 years is scarce owing to the unavailability of efficient tools for screening and developmental surveillance (Fischer et al., 2014). Findings of the present study are a starting point towards changing this scenario. They indicate overall satisfactory reliability and criterion validity for the STELD-T, suggesting that it be used to screen language development in babies from 6 months to 21 months of age in urban areas of Maharashtra.
A low level of item difficulty balanced with a "strict" pass-refer criterion (Westerlund et al., 2006) would ensure that the STELD-T does not overestimate or underestimate the occurrence of language delays. The use of simple language and the nature of items (not dependent on specific tasks), would broaden the scope of utility to include pediatricians or parents and reduce administration time.
Future research would focus on adding to the existing data pool to arrive at a final version of the STELD, and translation to Marathi and other languages, to increase its external validity and utility across India.