The Language Use Inventory for Young Children: A Parent-Report Measure of Pragmatic Language Development for 18- to 47-Month-Old Children

Daniela K. O’Neill
University of Waterloo, Waterloo, Ontario, Canada

Purpose: To demonstrate the internal reliability and discriminative validity of the Language Use Inventory for Young Children (LUI; D. K. O’Neill, 2002), a newly developed parent-report measure designed to assess pragmatic language development in 18–47-month-olds.

Method: To examine internal reliability, the LUI was completed by mail by 177 parents recruited from the University of Waterloo’s Centre for Child Studies database, 175 of whom completed the LUI again within 4 weeks to assess test–retest reliability. To examine discriminative validity, 49 parents of children awaiting assessment at a local speech-language clinic and 49 parents of typically developing children recruited from the Centre for Child Studies database and matched in age and sex to the clinic group completed the LUI.

Results: Alpha values for the subscales of the LUI were at or above acceptable levels (.80–.98), and steady growth in children’s pragmatic language development was demonstrated. The study of discriminant validity revealed sensitivity and specificity levels over 95%.

Conclusions: The LUI’s internal reliability and stability were strongly supported and its sensitivity and specificity in distinguishing between typically developing and language-delayed children exceeded even the most stringent criteria of 90% accuracy.

KEY WORDS: assessment, toddlers, pragmatics, screening, preschool children

Although definitions of pragmatics vary considerably, most definitions focus on our ability to use language effectively and appropriately in social interactions with other people (Bates, 1976). Pragmatics is commonly regarded as the third major component of language ability in addition to knowledge of form (phonology and syntax) and content (semantics) (Ninio & Snow, 1996), although its precise relation to these other aspects is of some debate (Levinson, 1983). This article describes a newly developed parent-report measure of early pragmatic language development, the Language Use Inventory for Young Children (LUI; O’Neill, 2002) and presents evidence of its internal reliability and discriminant validity.

Pragmatic Development in Children Under 4 Years of Age

Researchers studying young children’s pragmatic development have focused on a wide range of topics and ages. Longitudinal studies that have
concentrated on children’s earliest gestural and verbal communicative intents have demonstrated that children begin as early as 9 to 10 months of age to use their gestures and vocalizations for such pragmatic functions as requesting, labeling, answering, greeting, and protesting (Bates, Camaioni, & Volterra, 1975; Dale, 1980). By 2 years of age, as revealed in a longitudinal study investigating the order of emergence of more than 100 communicative acts in mother–child conversation (Ninio & Snow, 1996), children have mastered most of the basic, central communicative uses of speech such as the ability to discuss various topics and negotiate action (e.g., asking \textit{wh}-questions, discussions of the nonpresent, requesting/proposing new activity).

Beyond the age of 2 years, pragmatic abilities studied include children’s mastery of rules of politeness, deictic forms, and indirect forms of speech; the development of conversational skills (e.g., turn taking, topic initiation, contingent responding); the adaptation of utterances to the background knowledge of participants in the conversation; the production of extended discourse and mastery of different styles and registers of speech tailored to specific social roles and social circumstances; and narrative development. Children’s pragmatic language abilities have been demonstrated to rapidly increase and become more sophisticated during the preschool years, with 4-year-olds, for example, tailoring their utterances to knowledge of a listener and to a listener’s age, status, and gender (for review, see Clark, 2003). A chronology of the order of acquisition (e.g., month-by-month) of these more sophisticated pragmatic abilities is, however, not available, given that children in these studies were observed at only a few age points. One aim of the LUI is to provide a more detailed picture of the order of emergence of pragmatic language abilities from 18 to 47 months of age.

\textbf{Need for a Measure to Assess Pragmatic Language Development in Toddlers and Preschoolers}

No standardized test (observational or stand-alone parent report) is currently available that is specifically designed to assess toddlers’ and preschool children’s (i.e., under age 4 years) pragmatic language competence. The need for such a standardized tool is becoming more urgent as clinicians and researchers aim for earlier identification and diagnosis of language and developmental disorders. Governments are also increasingly enacting legislation requiring that children with communication disorders be identified before entering school (Kerr, Guildford, & Bird, 2003; Nuttall, Romero, & Kalesnik, 1999). Moreover, standardized testing is essential, and often legally required, to document a deficit to qualify a child for educational and social services (Paul, 2001).

With respect to existing measures for assessing pragmatic language competence suitable for very young children, the Communication and Symbolic Behavior Scales (CSBS; Wetherby & Prizant, 1993) are regarded as the closest measure to meet this need (Mervis & Robinson, 2005), although the CSBS does not provide an overall score of pragmatic functioning. The CSBS is also quite time-consuming to administer, although a shorter version for use with children 6–24 months of age has recently been developed (i.e., the Communication and Symbolic Behavior Scales Developmental Profile [CSBS DP]; Wetherby & Prizant, 2002). (Two other measures of pragmatics, the Pragmatics Protocol [Prutting & Kirchner, 1983] and the Children’s Communication Checklist–2 [CCC-2; Bishop, 1998, 2003] are designed for children older than age 4. The CCC-2 is a standardized caregiver checklist. The Pragmatics Protocol is clinician-scored based on a conversational sample and is not standardized. A Pragmatics Profile subtest is included in the most recent edition of the Clinical Evaluation of Language Fundamentals–Preschool [CELF-P-2; Semel, Wiig, & Secord, 2004], but only provides cutoff points at 6-month age intervals.)

The few standardized tests available to assess language abilities in children under 5 years of age focus largely on semantic (vocabulary) and grammatical skills and not pragmatics (e.g., the MacArthur Communicative Development Inventories [M-CDI]; Fenson et al., 1993; the Language Development Survey [LDS]; Rescorla, 1989; the Preschool Language Scale–3 [PLS-3]; Zimmerman, Steiner, & Pond, 1992; and the CELF-P-2; Semel et al., 2004; for review, see Hirsh-Pasek, Kohenoff, Newcombe, & de Villiers, 2005). Other tools that have been developed to assess communication more broadly (e.g., Girolometti, 1997; Hadley & Rice, 1993; Haley, Coster, Ludlow, Haltiwanger, & Andrelllos, 1992) have not yet been validated with large samples (McCauley, 2001). This situation has led many researchers to identify an urgent need for reliable, valid, and standardized measures to assess early pragmatic language competence (McCardle, Cooper, & Freund, 2005).

The provision of measures that specifically target pragmatic language development has become more important with the recognition that, for some children, communicative impairment is found primarily at the level of pragmatics rather than vocabulary or grammatical acquisition (e.g., Adams & Bishop, 1989). It is well-recognized that pragmatic language impairment (Bishop, 1998) occurs among individuals with pervasive developmental disorders (Baron-Cohen, 1988), but disproportionate pragmatic difficulties not accompanied by any autistic symptomatology have also been found in other clinical groups (Rice, Warren, & Betz, 2005), including individuals with specific language impairment (Botting & Conti-Ramsden, 1999), hyperlexia (Healy, 1982), fragile X syndrome and Down syndrome (Abbeduto & Murphy, 2004), and neurodevelopmental disorders (Levy, Tennebaum, & Orney, 2000).
Use of Parent Report

With respect to assessing language competence in younger children, and especially pragmatic competence, structured tests have been criticized as providing only a limited picture of the richness and complexity of the child’s communicative behavior and revealing little about the child’s language as it occurs and is used in everyday communication (Leonard, Prutting, Perozzi, & Berkeley, 1978; Owens, 1995). Greater ecological validity has been argued to be possessed by informal or naturalistic assessment methods (Lund & Duchan, 1983).

One answer to concerns about the ecological validity of structured tests has been the use of standardized parent-report measures. Parent report is now a component of many widely used developmental and language screening tools and structured assessments (e.g., the Ages and Stages Questionnaire; Bricker & Squires, 1999; the Receptive-Expressive Emergent Language Test; Bzoch & League, 1971; the Denver Developmental Screening Test; Frankenburg, Dodds, Fandal, Kazuk, & Cohrs, 1975; the CELF-P-2; Semel et al., 2004; the Sequenced Inventory of Communicative Development; Hedrick, Prather, & Tobin, 1984; the LDS; Rescorla, 1989; the Rossetti Infant-Toddler Language Scale; Rossetti, 1990; and the CSBS; Wetherby & Prizant, 1993). In addition, the M-CDI (Fenson et al., 1993), which assesses lexical, gestural, and grammatical development (but not pragmatics) among children 8–30 months of age, relies entirely on parental report. Parent report has been demonstrated to be accurate, valid, and reliable, particularly when assessment is limited to current and emergent behaviors and a recognition format is used (Fenson et al., 1993). As will be shown, the LUI satisfies these conditions. With respect to assessing pragmatic language competence, parent report is especially promising given that parents observe their children’s language in a wide variety of settings, a condition that would be very difficult to mimic using structured testing. Standardized parent-report measures can also provide a cost-effective means of screening and evaluating children.

Development of the Initial Items on the LUI

The LUI (O’Neill, 2002) used in the two studies reported in this article represents the fourth version of the questionnaire formerly known as the Pragmatics Aptitude Test. The 338 items on the original version (O’Neill & Baron-Cohen, 1996) were developed following a review by O’Neill of the literature on typical and nontypical language development. In identifying pragmatic competencies to include on the questionnaire, the approach adopted was not to isolate and classify specific speech acts (e.g., assertives, directives, etc.; Searle, 1969), an approach that has also been found to be too limited by other researchers (e.g., Prutting & Kirchner, 1983, 1987). Rather, the approach was similar to more recent approaches emphasizing “the functioning of language in actual contexts of use” (Verschueren, 1999, p. 9) and the capturing of “discrete, psychologically real types of communicative acts” occurring in children’s interactions with other people (Ninio & Snow, 1996, p. 21).

And most important, the focus was on identifying developments in children’s language use (note that language use is used synonymously with pragmatic competence) influenced by their developing understanding of the mind (O’Neill, 2005)—that is, their understanding of their own and other people’s behaviors, mental states, and differing perspectives (i.e., theory of mind; Astington, Harris, & Olson, 1988). This approach is consistent with views of pragmatics that stress the importance of “understanding intentional human action” (Green, 1989, p. 2) and “understanding the interlocuter’s state of mind” (Ninio & Snow, 1996, p. 191) in communication. Thus, based on the findings of studies investigating children’s developing understanding of mind, items were developed to capture, for example, young children’s ability to direct someone’s attention, to talk or ask about people’s behaviors and mental states, to tease, and to take a listener’s knowledge into account. The firm grounding of the content of the items in established research findings was intended to ensure the LUI’s content validity (Anastasi, 1988) and meets current calls for assessment tools with greater “empirical validity” (Hirsh-Pasek et al., 2005).

Preliminary Studies of the LUI’s Internal Reliability

Two preliminary studies of the LUI’s internal reliability were conducted primarily to shorten the questionnaire from its original 338 items to permit completion within 30 min and develop subscales with acceptable levels of reliability. Corrected-item total calculation (CITC) scores and Cronbach’s coefficient alpha values (Cronbach, 1984) were used to decide whether to retain or delete items. A minimum coefficient level of .3 was generally used in interpreting the CITC scores (Anastasi, 1988). Alpha values over .80 were deemed acceptable in accordance with recommendations of assessment researchers (Salvia & Ysseldyke, 2001).

In the first study, the questionnaire was completed by 183 parents of children between the ages of 13 and 48 months (48% girls, 52% boys; in six 6-month age groups recruited from the University of Waterloo’s Centre for Child Studies database. Item analysis led to the deletion of 132 items. The initial alpha values for the 15 subscales computed on the 206 remaining items...
were encouraging, with 10 subscales demonstrating high internal consistency (as of .80 to .98). To address the lower reliabilities of 5 of the subscales, 26 new items were added (following further review of the literature). Following this first study, the lowest appropriate age for the questionnaire was increased to 18 months from 13 months as parents responded “no” to most items before 18 months.

In a second study, this second version of the questionnaire (O’Neill, 1999) with 232 items was completed by 161 parents of children (52% girls, 48% boys) between the ages of 18 and 47 months recruited as in the first study. Item analysis led to the deletion of a further 55 items. Alpha values computed on the 177 remaining items revealed that the internal consistency of the questionnaire’s 12 scored subscales had improved: 8 subscales had alpha values in the range of .90 to .98, 3 had values of .89 and .80, and one had a value of .74. Three new items were added to the questionnaire, resulting in a total of 180 items. At this time, to increase the ease of use of the questionnaire, two other modifications were made: (a) items were reworded to eliminate any need for reverse scoring, and (b) questions and instructions were reworded based on readability analyses (Gunning, 1952; Kinkaid, Fishburne, Rogers, & Chissom, 1975) to require less than the 8th grade reading ability often advocated for public health information (National Work Group on Literacy and Health, 1998). Two gesture subscales demonstrating a linear decrease in their scores with age (as would be expected if children are replacing these gestures with words) were considered for deletion, but were ultimately retained so that a parent of a child with very low levels of language use would not have to indicate negative responses from the beginning.

The third version of the LUI resulting from these preliminary studies (O’Neill, 2001) was then the subject of two studies to be presented in this article. Study 1 examined its internal and test–retest reliability. Study 2 examined the ability of the LUI to distinguish between children whose language was developing typically and children who subsequently received a diagnosis confirming the presence of language delay.

Study 1

The goal of Study 1 was to evaluate the internal reliability of the LUI and assess its test–retest reliability (i.e., the extent to which scores remain stable across two or more administrations). With respect to test–retest reliability, given the rapidity with which language skills can change over short periods of time and uneven rates of language change, one would not expect high stability. However, within a period of a few weeks, one might expect to see stability. In keeping with time periods used in other similar test–retest reliability studies, parents completed a second questionnaire within 4 weeks of completion of the initial questionnaire.

Method

Materials

The 180-item version of the LUI resulting from the preliminary reliability studies was used. Table 1 provides a description of its main three parts, its 14 subscales, and sample items. A yes or no response was required for 89% of the items, and the remaining 11% presented the options never, rarely, sometimes, or often (the option not anymore was also presented in Gesture Subscale A).

In addition to the LUI, parents provided information pertaining to birth order and the presence of major health or speech problems using a form similar to that used by Fenson et al. (1993). However, an additional new form was developed to assess exposure to other languages. Parents were asked to indicate all adult persons with whom the child regularly interacts, the language(s) spoken by the person, the percentage of time the person speaks one or more languages, and the number of hours per week the child typically spends with the person. Given the substantial percentage of children exposed to other languages at home (12%–15% of children under 4 years according to Statistics Canada, 2001), the inclusion of solely monolingual English speaking children was not merited. Exposure to a second language for less than 20% of the child’s waking hours is generally agreed not to lead to substantial language growth in that second language and such children are not generally considered as bilingual (e.g., Pearson, Fernandez, Lewedeg, & Oller, 1997). Thus, exposure to a second language exceeding 20% of waking hours was established as the threshold for exclusion.

It should also be noted that in households where more than one language is spoken, asking parents to report on the questionnaire as “yes” only the particular items performed in English as opposed to a second language would impose an artificial distinction that would place an unnecessary demand on parents, especially given that it has been documented that bilingual speakers are not always aware of the language they are speaking (Goodz, 1989) and bilingual families often do not adopt a “one-parent one-language” strategy (Pearson et al., 1997). For these reasons, in the information letter accompanying the questionnaire, if a child was exposed to a language(s) other than English, parents were instructed to check an item as “yes” if the child produced the words or utterances asked about “either in English or in the other language(s) he/she was exposed to.”

Participants and Procedure

Time 1. Parents were recruited from the UW Centre for Child Studies database. The questionnaire was mailed...
to 207 parents of children between the ages of 18 and 47 months. This age range was divided into 5 six-month intervals (18–23, 24–29, 30–35, 36–41, and 42–47 months). At the time of the initial call, parents were also asked if they would be willing to complete the questionnaire twice. One hundred ninety-two (93%) of the questionnaires were returned. Fifteen (8%) were dropped for the following reasons: missing data (8), medical reasons (2), second language exposure over 20% of waking hours (2), prematurity exceeding 2 weeks (1), and age older than 47 months (1). This resulted in a final sample size of 177 (see Table 2 for sample characteristics). All the children had acquired

Table 1. Overview of the content of the three major parts and 14 subscales of the Language Use Inventory for Young Children (LUI; O’Neill, 2001, 2002).

<table>
<thead>
<tr>
<th>Part and subscale</th>
<th>Content description of subscale and abbreviated version of sample items</th>
</tr>
</thead>
</table>
| Part 1: How your child communicates with gestures | The child’s imperative use of gestures.  
A: How your child uses gestures to ask for something  
e.g., lifts arms to be carried; points to something wanted  
B: How your child uses gestures to get you to notice something  
e.g., points to something of interest; shows you something of interest |
| Part 2: Your child’s communication with words | The topic and typicality of the child’s first words.  
C: Types of words your child uses  
e.g., use of animal terms; use of “gone”  
D: Your child’s requests for help  
e.g., use of “help”; requests for help with difficult task  
E: Your child’s interests  
e.g., three favourite play activities |
| Part 3: Your child’s longer sentences | The child’s declarative use of words.  
F: How your child uses words to get you to notice something  
e.g., asks you to “look!”; says “guess what?”  
G: Your child’s questions and comments about things  
e.g., asks what something is; asks who something belongs to  
H: Your child’s questions and comments about themselves or other people  
e.g., asks where someone is; says how he/she feels emotionally  
I: Your child’s use of words in activities with others  
e.g., tells another person to stop doing something; asks for a turn  
J: Teasing and your child’s sense of humour  
e.g., says wrong things in a teasing way; tells jokes  
K: Your child’s interest in words and language  
e.g., imitates words heard; asks about the meaning of new words  
L: Your child’s interests when talking  
e.g., three things talked about most  
M: How your child adapts conversation to other people  
e.g., use of “what?”; clarification questions such as “This one?”  
N: How your child is building longer sentences and stories  
e.g., use of “next,” “maybe,” “think,” “tomorrow” |

Note. In the LUI, the specific items in each subscale are preceded by a main question (e.g., for Subscale H: Which of the following things have you heard your child talk about? Following this, a list of items is presented and many of these items are accompanied by an example(s) to aid parents in understanding the question (e.g., Subscale H: Item a: Where someone else is [e.g., Where’s Daddy?; Mommy’s here.]; Item b: What he/she is going to do [e.g., I’m gonna draw a house.] etc.). The sample items provided have been abbreviated given space limitations.
English as their first language, although 13.4% were reported to be exposed to a second language in the home for less than 20% of their waking hours (including French, Spanish, German, Italian, Portuguese, Turkish, Slovenian, Hungarian, Romanian, Slovakian, Croatian, Vietnamese, and Cantonese).

Time 2. Of the final 177 participants, 175 had indicated a willingness to complete the questionnaire twice and, so, when the questionnaire was returned to the lab, a second was mailed with instructions to complete it within 4 weeks of initial completion of the Time 1 questionnaire. One hundred fifty-eight (90%) of these questionnaires were returned. Of these, 17 (11%) were dropped for the following reasons: missing data (5), subsequently being dropped at Time 1 (11), and responding beyond the 4-week time window (1). Thus, a final total of 141 questionnaires were used in the test–retest sample (see Table 2 for sample characteristics).

Scoring. A child’s total score was calculated from the sum of all items in the subscales of Parts 2 and 3 (i.e., it did not include the two gesture scales in Part 1). Responses of yes, sometimes, and often were scored as 1 and never and rarely as 0.

Results and Discussion

Completion Time

Based on parent report, the mean time to complete Parts 1 to 3 of the questionnaire was 27 min, with a range from 22 to 35 min.

Internal Reliability

Item analysis suggested only the reassignment of 1 of the 180 items to a different subscale. Following its reassignment, coefficient alpha values were computed for all three parts of the inventory and all subscales. The resulting alpha values are shown in Table 3 and very strongly confirm the internal consistency and reliability of the LUI. The coefficient alpha values of each of the three parts and 7 of the 12 scored subscales met the clinically most desirable level of .9 or above. The alpha value for Parts 2 and 3 combined, which comprises a child’s total score on the LUI, was .99 (N = 161, SEM = 3.96).

Table 3. Values for Cronbach’s alpha for all parts and subscales of the Language Use Inventory for Young Children (Study 1).

<table>
<thead>
<tr>
<th>Part</th>
<th>Description</th>
<th>Alpha (SEM)</th>
<th># of items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part 1: How your child communicates with gestures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A:</td>
<td>How your child uses gestures to ask for something</td>
<td>.92 (1.14)</td>
<td>11</td>
</tr>
<tr>
<td>B:</td>
<td>How your child uses gestures to get you to notice something</td>
<td>.55 (0.32)</td>
<td>2</td>
</tr>
<tr>
<td>Part 2: Your child’s communication with words</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C:</td>
<td>Types of word’s your child uses</td>
<td>.93 (0.90)</td>
<td>21</td>
</tr>
<tr>
<td>D:</td>
<td>Your child’s requests for help</td>
<td>.87 (0.56)</td>
<td>7</td>
</tr>
<tr>
<td>E:</td>
<td>Your child’s interests</td>
<td>Not scored</td>
<td>2</td>
</tr>
<tr>
<td>Part 3: Your child’s longer sentences</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F:</td>
<td>How your child uses words to get you to notice something</td>
<td>.83 (0.65)</td>
<td>6</td>
</tr>
<tr>
<td>G:</td>
<td>Your child’s questions and comments about things</td>
<td>.91 (0.79)</td>
<td>9</td>
</tr>
<tr>
<td>H:</td>
<td>Your child’s questions and comments about themselves/other people</td>
<td>.98 (1.76)</td>
<td>36</td>
</tr>
<tr>
<td>I:</td>
<td>Your child’s use of words in activities with others</td>
<td>.94 (0.94)</td>
<td>14</td>
</tr>
<tr>
<td>J:</td>
<td>Teasing and your child’s sense of humour</td>
<td>.80 (0.81)</td>
<td>5</td>
</tr>
<tr>
<td>K:</td>
<td>Your child’s interest in words and language</td>
<td>.86 (1.19)</td>
<td>12</td>
</tr>
<tr>
<td>L:</td>
<td>Your child’s interests when talking</td>
<td>Not scored</td>
<td>4</td>
</tr>
<tr>
<td>M:</td>
<td>How your child adapts conversation to other people</td>
<td>.93 (1.26)</td>
<td>15</td>
</tr>
<tr>
<td>N:</td>
<td>How your child is building longer sentences and stories</td>
<td>.98 (1.88)</td>
<td>36</td>
</tr>
</tbody>
</table>
Figure 1. The mean proportion (with 95% confidence interval) of items scored positively on each of the 12 scored subscales shown separately for boys and girls in each of the five age groups.
The alpha values for all remaining subscales except B were all at or above acceptable levels of .80 to .87. Subscale B’s low alpha value (.55) was unexpected given that with the same items in the previous study the alpha value had been .74. One reason for this lack of stability may be the lower number of items in this scale and the fact that children in all age groups scored close to ceiling.

Factor Analysis

An exploratory factor analysis with varimax rotation supported a two-factor solution, with the first factor corresponding to the 10 subscales of Parts 2 and 3 assessing verbal communication (eigenvalue = 8.34) and a second factor corresponding to the 2 gesture subscales in Part 1 (eigenvalue = 1.1). No subscales overlapped in their loadings, and all loadings for the first factor were greater than .78. The two factors accounted, respectively, for 69.5% and 9.2% of the variance explained, for a total of 78.7%. The finding that all the subscales in Parts 2 and 3 loaded onto one factor was not contrary to expectation. As will also be evident from the results of the age analyses and test–retest study to follow, the very rapid growth of these abilities in tandem with each other suggests a fair amount of overlap in the development of these abilities in time. The fact that these subscales load onto one factor with a typically developing group of children also does not preclude the possibility that different results might be found with subpopulations of children with language impairment.

Growth With Age Across Subscales

For each subscale, to examine children’s growth in pragmatic competence with age, a 5 (age groups) × 2 (sex) analysis of variance (ANOVA) was conducted on the summed scores for each subscale. A significant main effect of age \( (p < .05) \) was observed for all subscales. Figure 1 shows the performance of each age group, and sex within each age group, on each subscale. As is evident in this figure, children’s scores increased linearly with age for all subscales in Parts 2 and 3 assessing verbal language use. In contrast, for two gesture subscales in Part 1, children’s scores decreased with age, reflecting children’s replacement of these early gestures with verbal forms.

Results of Tukey’s post hoc tests identified which of the five age groups differed significantly in their
performance on each subscale (see summary in Table 4). The results clearly show the early, rapid growth in pragmatic language competence, especially with respect to the uses of language assessed by the subscales of Parts 2 and 3. For example, as summarized in Table 4, and as can be observed in Figure 1, the 24–29-month age group and all older groups were already showing significantly higher scores on all subscales in Parts 2 and 3 than the youngest age group (18–23 months). In addition, on all subscales in Part 3, the 30–35-month age group and all older age groups had scores significantly higher than the younger two age groups (18–24 and 25–29 months). Finally, only the two oldest age groups (36–42 and 43–47 months) scored significantly higher on the last four subscales of Part 3 (J through N) than all three younger age groups.

Overall, these findings suggest that the goal of the LUI to capture early developments in pragmatic language competence—ideally under age 4 to increase the usefulness of the LUI as a means of screening and evaluating toddlers and preschool children—was met.

### Sex Differences

Whether sex differences in the rate of language development exist is of considerable research and applied interest. Recent reviews suggest the performance of girls is often significantly higher than that of boys on general and specific tests of language development administered before age 7 (Bornstein, Hahn, & Haynes, 2005). In the current study, however, significant sex differences (\( p < .05 \)) were rare and isolated. They were found only for two subscales: F, which assesses declarative use of words, and N, which assesses word use as children are building longer sentences and stories. Moreover, for both these subscales, the Sex × Age Group interaction was not significant, indicating the sex difference was not observed at all age groups. The findings of a much larger ongoing standardization study will determine more accurately whether sex differences exist with respect to children's scores on the subscales of the LUI.

### Intercorrelations Among the Subscales

Table 5 shows the intercorrelations among all 12 scored subscales, controlling for age (in months). Children’s scores on both gesture subscales (A: Imperative Gestures and B: Declarative Gestures) were significantly, but fairly weakly, positively correlated. Scores on Subscale A were weakly and negatively correlated (\( r < .3 \)) with scores on all other subscales. Scores on Subscale B were also largely nonsignificantly and negatively correlated with children’s scores on all other subscales, probably because children scored uniformly high across age groups on this subscale.
This decrease in the use of gestures in relation to an increase in the use of verbal forms is in keeping with the findings of other researchers examining the relation between gestures and early speech. Namely, as other authors have stated, “Gesture paves the way for language development” (Iverson & Goldin-Meadow, 2005, p. 367), and many of the lexical items that children first produce in gesture subsequently appear in their verbal repertoire.

Turning to consideration of the remaining 10 subscales that form Parts 2 and 3 of the questionnaire, it is evident that the intercorrelations among children’s scores on the first six scored subscales (C to I) were higher (rs in range of .617 to .813; M = .709) than the intercorrelations among the last four scored subscales, J to N (rs in range of .236 to .746; M = .545). Overall, the moderate r levels in the range of .32 to .70 for the large majority (78%) of the possible 45 intercorrelations between children’s 10 subscale scores in Parts 2 and 3 of the questionnaire suggest that these subscales are measuring aspects of pragmatic language competence that differ to a certain degree. The high intercorrelation of some of the subscales could have led to their amalgamation. This was not done for two reasons. First, phone interviews with parents suggested that shorter subscales with a greater focus were easier to complete. Second, high intercorrelations found with a typically developing group of children do not preclude the possibility that different results might be found with subpopulations of children with language impairment (indeed, this was found to an extent with the children in the discriminant validity study to be described next).

**Test–Retest Reliability**

Significant growth was observed in children’s scores even within the short 4-week time period. That is, a significant difference in mean score between Time 1 and Time 2 (p < .05) was found for all but three subscales (Subscale A: Imperative Gestures; B: Declarative Gestures; G: Questions/Comments About Things). Nevertheless, test–retest stability and reliability were excellent, with all subscales demonstrating significant Pearson correlations of .85 to .96 (p < .001) between the test and retest scores, except for Subscale B (r = .34). When the correlations were recomputed, controlling for age, the significance results did not change, except that the range of partial correlations for all subscales except B (r = .32) was between .75 and .89 (all ps < .001).

---

**Study 2**

The aim of Study 2 was to determine whether children’s scores on the LUI would distinguish between children with and without language delay. The scores of children whose parents were seeking an initial assessment at a child development center regarding possible language delay were compared with those of a group of children matched for sex and age (in months) drawn from our laboratory database for whom no concern about their language development had been expressed. Thus, this study used a clinical population of children to examine discriminant validity. For this particular clinic, parents seeking an assessment did not need a physician’s referral. Thus, this group of parents might be best categorized as parents who, for one reason or another, felt their child’s language development was not progressing as it should. In the region in which this study was conducted, this child development center is the primary location for speech-language assessments.

**Method**

**Materials**

The materials used were identical to those used in Study 1.

**Procedure and Participants**

During the time period from December 19, 2001, to April 20, 2003, all parents who contacted a local child development center and requested an assessment of their child’s language were asked whether they would be willing to take part in the study by completing the LUI before their child’s initial assessment. Parental consent to obtain information about their child’s assessment and eventual diagnosis was also obtained. The child development center mailed parents the LUI with a postage-paid and pre-addressed envelope to return the questionnaire to the university laboratory (i.e., the completed questionnaire was at no time available to staff members of the child development center). The children of parents who sought an assessment of their language ability at the child development center and took part in the study are referred to as the clinic group.

When the completed questionnaire was received, a parent with a child of the same sex and age in months was contacted from the laboratory’s database and invited to participate in the study by completing the LUI. This age- and sex-matched sample of children drawn from our laboratory database is referred to as the typically developing group.

The clinic group included 49 children (age range: 21–47 months; M = 28.8 months; 40 boys and 9 girls). By design, the typically developing group was matched for sex and age. Questionnaires were initially mailed out to 118 parents and 73 were returned (62% return rate). Twenty-four questionnaires were dropped for the following reasons: prematurity (10), missing data (4), completion date unknown (4), older than 47 months (3), second language exposure beyond 20% of waking hours (2), and
because the child was discharged before assessment (1). Six children were minimally exposed to a second language (less than 6% of the child’s waking hours, including Portuguese, German, Greek, Spanish, or French). The center would not provide further demographic information regarding the families, but it is known that all children resided in two-parent homes. Once a child’s assessment at the clinic was completed and a diagnosis had been given, a member of the clinic staff retrieved from the child’s file the information requested by the laboratory concerning the assessment procedure, the final diagnosis, and any other relevant health information such as the presence of hearing loss. This information was forwarded to the university. Information regarding the final diagnoses of these children is presented next in the Results and Discussion section.

Results and Discussion

Of the 49 clinic children, 43 (88%) were given a diagnosis of expressive language delay. Among these 43 children, secondary diagnoses included receptive language delay \((n = 7)\), receptive and pragmatic language delay \((n = 2)\), articulatory delay \((n = 11)\), and phonological delay \((n = 8)\). Among the remaining 6 of the 49 clinic children, 5 (10%) were given diagnoses of articulation delay and 1 was given a diagnosis of phonological delay. Mild hearing loss was noted for 1 child. None of the 49 children were discharged following assessment. The average time between the initial completion of the LUI and the time of assessment was 2.3 months, with a range from 0 to within 8 months. The mean age of children at the time of assessment was 31.5 months (range: 25 to 52). Of the 49 children, 37% were administered standardized tests during assessment, typically including one or more of the following: the PLS-3, the CELF-P, the M-CDI, the Goldman Fristoe Test of Articulation (Goldman & Fristoe, 1986), the Structured Photographic Articulation Test II (Dawson & Tattersall, 2001), and the Weiss Comprehensive Articulation Test (Weiss, 1980). The remaining children (63%) were assessed using observations conducted by the speech-language pathologist during a session involving free-play and conversation with the child.

As was done in the third reliability study of the LUI, a total score (out of 161) was summed from Parts 2 and 3 of the questionnaire. When the total scores of the clinic group were compared with those of the typically developing group, the difference was strikingly large. Figure 2 shows the scores of the 49 age- and sex-matched child pairs in order of increasing age from left to right. The mean total score of the clinic group was 27.4 \((SD = 23.2)\) compared to 106.5 \((SD = 27.2)\) for the typically developing group. The average difference in scores was 79.1 points \((SE = 5.1)\), a difference that was significant, \(t(96) = 15.47, p < .001\). Moreover, Figure 2 reveals that this difference remained relatively stable across the entire age range from 21 to 47 months.

Figure 3 shows the mean performance for each group over all 12 scored subscales. With the notable exception

---

**Figure 2.** The total scores on the Language Use Inventory for Young Children (LUI), plotted against age (in months), of the 49 age- and sex-matched children in the clinic and typically developing groups.
of their mean performance on the two first Gesture Subscales A and B, the mean performance of the clinic group fell far below that of the typically developing group on all remaining 10 subscales.

Results of a discriminant function analysis using total scores, first conducted with age not taken into account, correctly classified 91.8% of the original 98 cases. Sensitivity (clinic group children identified as such) was 95.9% and specificity (typically developing group children identified as such) was 87.8%. These high rates of correct classification, without even taking age into account, can be understood if one refers back to Figure 2 and notes that using total scores alone, a score of 65 or greater on the LUI was obtained by all but 4 of the youngest typically developing children (under 25 months of age) and only 2 of the clinic children.

A second discriminant function analysis was conducted using total scores corrected for age (i.e., regression was used to predict LUI scores from age and produce a standardized residual score for each child). As expected, this improved the specificity among the youngest typically developing children, and specificity rose to 95.9%. The overall correct classification rate and sensitivity were also 95.9%. The results of both of these analyses strongly support the discriminant validity of the LUI. The LUI exceeds even the most stringent criteria of 90% accuracy suggested by psychometricians (e.g., Plante & Vance, 1994).

Intercorrelations Among the Subscales

Table 6 shows the intercorrelations among all 12 scored subscales controlled for age (in months) from the clinic group. Compared to the findings with the typically developing children in the third reliability study shown in Table 5, the clinic children’s scores on Subscale A (Imperative Gestures) were more strongly (negatively) correlated with their scores on all other subscales ($r$ values of $-.4$–$-.5$) than were typically developing children’s scores on this subscale ($rs < .3$).

The intercorrelations among clinic children’s scores on the 10 subscales that form Parts 2 and 3 of the questionnaire were all significant, but the magnitude of the $r$ values tended to be lower (range: $.31$–$.75$; $M = .54$) than for the typically developing group (range: $.24$–$.84$; $M = .61$). Moreover, compared to the typically developing group, the clinic children’s scores on the later subscales of Part 3 (I to N) showed a less steady increase in the magnitude of their intercorrelation with prior subscales. Overall, therefore, the moderate $r$ levels of many of the intercorrelations suggest that the subscales are measuring aspects of pragmatic language competence that are nonoverlapping to a certain extent, as was suggested by the results among the typically developing children. Nevertheless, the more variable pattern with respect to the strength of the intercorrelations across subscales suggests that for children with language delay, certain pragmatic competencies, such as those tapped by the latter

![Figure 3. The mean performance of the 49 children in the clinic and age- and sex-matched typically developing group on all 12 scored subscales of the LUI.](image-url)
Subscales I to N, may be lagging behind others rather than steadily increasing in tandem with each other. Exactly what the nature of the profile of children with different types of language delay might look like across the subscales of the LUI must await further research more specifically designed to answer this question. It is noteworthy, however, that the types of abilities assessed in Subscales I to N are similar to those found to be problematic in older children judged as impaired in pragmatic competence when assessed using the Pragmatics Protocol and the CCC-2, such as problems with ambiguous references, coherence, inappropriate initiations, and inadequately informative contributions in conversation (Bishop, 2003; Prutting & Kirchner, 1987).

Conclusions

The results of the two studies presented strongly support the internal reliability and discriminant validity of the LUI. Alpha coefficient values for the subscales very strongly confirmed their internal consistency and reliability. Test–retest stability and reliability were also strongly supported. Children’s scores on the subscales showed clear and significant growth occurring between the youngest (18–23 month) group and the oldest (42–47 month) group, meeting a goal to elucidate in more detail pragmatic growth in the less-chartered territory between 24 and 47 months of age. The pace of development in children’s use of language across a wide variety of situations was very rapid. In addition, the discriminant validity of the LUI was strongly supported by its ability to distinguish between children with typically developing language and a clinic sample of children with language delay with sensitivity and specificity levels greater than 95%. Overall, the results further confirm the value of parent report and ease of use of the LUI, with short completion times and few instances of missing data or problems encountered by parents during completion of the questionnaire.

Limitations to this work exist that require further study. Of interest is establishing the concurrent validity of LUI by examining, for example, the relation between children’s performance on the LUI and their performance on available standardized tests that include some aspects of pragmatic competence such as the CSBS (Wetherby & Prizant, 1993) or relevant laboratory-based measures such as storytelling. Such laboratory studies could also examine divergent validity by examining whether children’s scores on the LUI are more strongly correlated with their performance on laboratory measures than on other, more general (i.e., focused more broadly on language development in general and not pragmatics per se) parent-report measures of language development and other standardized tests of language development. Longitudinal studies could examine its predictive relation with children’s performance on measures appropriate for older children such as the Test of Pragmatic Language (Phelps-Terasaki & Phelps-Gunn, 1992). Discriminant validity and the LUI’s sensitivity and specificity could also be investigated with a larger, general (nonclinical) population.

Nevertheless, the results of the studies presented in this article suggest a very good potential for the LUI to serve as a screening measure to detect delay in language use in toddlers and preschoolers. Although the LUI is not standardized at present, a standardization study involving Canadian children is underway.

For clinicians and researchers, the LUI may also provide a measure to investigate and distinguish more systematically differences in areas of impairment (i.e., lexical or grammatical vs. pragmatic) or examine, based on subscale performance, profiles of pragmatic competence for different populations of children displaying language

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>-0.329*</td>
<td>0.421**</td>
<td>0.032</td>
<td>-0.398**</td>
<td>-0.409</td>
<td>0.727**</td>
<td>-0.513**</td>
<td>0.031</td>
<td>0.388**</td>
<td>0.664**</td>
</tr>
<tr>
<td>C</td>
<td>-0.421**</td>
<td>-0.329*</td>
<td>0.513**</td>
<td>0.079</td>
<td>0.747**</td>
<td>0.642**</td>
<td>0.667**</td>
<td>0.537**</td>
<td>0.054</td>
<td>0.645**</td>
</tr>
<tr>
<td>D</td>
<td>-0.398**</td>
<td>-0.409</td>
<td>-0.513**</td>
<td>-0.039</td>
<td>0.436**</td>
<td>0.639**</td>
<td>0.519**</td>
<td>0.755**</td>
<td>0.054</td>
<td>0.372**</td>
</tr>
<tr>
<td>F</td>
<td>-0.513**</td>
<td>0.031</td>
<td>0.436**</td>
<td>0.079</td>
<td>-0.398**</td>
<td>0.717**</td>
<td>0.637**</td>
<td>0.519**</td>
<td>0.054</td>
<td>0.372**</td>
</tr>
<tr>
<td>G</td>
<td>-0.421**</td>
<td>0.031</td>
<td>0.388**</td>
<td>0.079</td>
<td>0.436**</td>
<td>-0.398**</td>
<td>0.717**</td>
<td>0.637**</td>
<td>0.519**</td>
<td>0.054</td>
</tr>
<tr>
<td>H</td>
<td>-0.537**</td>
<td>-0.409</td>
<td>0.645**</td>
<td>-0.039</td>
<td>0.436**</td>
<td>0.639**</td>
<td>0.519**</td>
<td>0.755**</td>
<td>0.054</td>
<td>0.372**</td>
</tr>
<tr>
<td>I</td>
<td>-0.476**</td>
<td>-0.006</td>
<td>0.396**</td>
<td>0.436**</td>
<td>-0.398**</td>
<td>0.747**</td>
<td>0.642**</td>
<td>0.554**</td>
<td>0.054</td>
<td>0.372**</td>
</tr>
<tr>
<td>J</td>
<td>-0.476**</td>
<td>-0.006</td>
<td>0.396**</td>
<td>0.436**</td>
<td>-0.398**</td>
<td>0.747**</td>
<td>0.642**</td>
<td>0.554**</td>
<td>0.054</td>
<td>0.372**</td>
</tr>
<tr>
<td>K</td>
<td>-0.538**</td>
<td>0.013</td>
<td>0.445**</td>
<td>0.497**</td>
<td>-0.398**</td>
<td>0.727**</td>
<td>0.627**</td>
<td>0.639**</td>
<td>0.054</td>
<td>0.372**</td>
</tr>
<tr>
<td>M</td>
<td>-0.538**</td>
<td>0.013</td>
<td>0.445**</td>
<td>0.497**</td>
<td>-0.398**</td>
<td>0.727**</td>
<td>0.627**</td>
<td>0.639**</td>
<td>0.054</td>
<td>0.372**</td>
</tr>
<tr>
<td>N</td>
<td>-0.417**</td>
<td>-0.026</td>
<td>0.421**</td>
<td>0.342**</td>
<td>-0.398**</td>
<td>0.727**</td>
<td>0.642**</td>
<td>0.554**</td>
<td>0.054</td>
<td>0.372**</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01.
impairment. The LUI may also provide a measure to equate children for overall pragmatic language development when investigating environmental or cognitive correlates of language use such as theory of mind reasoning or social competence.

Acknowledgments

This work has been made possible by grants from the Natural Sciences and Engineering Research Council of Canada (1993–1999) and from the Canadian Institutes of Health Research (2000–2005) to the author. Thanks are extended to the parents who participated in these studies; the director and staff of KidsAbility Child Development Centre in Waterloo, Ontario, Canada; and my undergraduate research assistants: Andrina Chen, Jane Topolovec, Mary Bender, Jennifer Pick, Kirsten Weeda, Tanya Kaefer, and Sejal Patel. Portions of the research reported have been presented at the 2001, 2002, and 2003 annual meetings of the Symposium on Research in Child Language Disorders, Madison, WI, and at the 2003 meeting of the Canadian Association of Speech Language Pathologists and Audiologists, St. John’s, NL, Canada. This work is dedicated to my mother, Maria Charlotte O’Neill, for all our conversations—big and small.

References


Contact author: Daniela O’Neill, Department of Psychology, University of Waterloo, 200 University Avenue West, Waterloo, Ontario Canada N2L 3G1. E-mail: doneill@uwwaterloo.ca.