

Testing the Usability of an Arabic Version of TOPL-2 in Measuring Pragmatic Language Impairment in Children and Adolescents with Developmental Dysphasia

Ahmed M. S. Alduais (Corresponding author)

Department of Eng., King Saud University

Riyadh, Kingdom of Saudi Arabia

Tel: 966-551-367-208 E-mail: ibnalduais@yahoo.com

Rasha M. Shoeib

Communication and Swallowing Disorders Unit, King Abdulaziz University Hospital

King Saud University, Riyadh, Kingdom of Saudi Arabia

Tel: 966-559-040-307 E-mail: shoeibr@yahoo.com

Fayza Saleh Al-Hammadi

Department of Eng., King Faisal University

AlAhsa, Kingdom of Saudi Arabia

Tel: 966-505-916-986 E-mail: falhammadi@yahoo.com

Khalid Hassan Al-Malki

ENT Department, King Saud University

Riyadh, Kingdom of Saudi Arabia

Tel: 966-543-300-022 E-mail: kmalky@yahoo.com

Received: March 23, 2012 Accepted: April 16, 2012 Published: June 1, 2012

doi:10.5296/ijl.v4i2.1770

URL: <http://dx.doi.org/10.5296/ijl.v4i2.1770>

Abstract

The purpose of this research was to assess the usability of an Arabic version of TOPL-2 in identifying individuals with PLIs and distinguishing them from those without PLIs.

The method used in this research was that the Arabic TOPL-2 test was administered to twenty-eight individuals (F: 15-M: 13) rigorously diagnosed with DLD (n: 14), SLI (n: 8), HI (n: 2), LD (n: 2) MR (n: 1) and ADHD (n: 1) in the age-range of 6-16 years. For comparison, another twenty-eight participants (F: 14- M: 14) with typical language development in the age-range of 6-16 years also took the same test.

The results indicate that, while the Arabic TOPL-2 is minimally reliable and satisfactorily valid, the internal consistency of the Arabic TOPL-2 is acceptable and its construct validity is generally moderate.

It was concluded that the Arabic TOPL-2 is roughly usable and can distinguish between pragmatically impaired and unimpaired individuals. A small number of the items need to be modified for use with Arab participants.

Keywords: Test usability, Test feasibility, Developmental dysphasia, Pragmatic language impairment, Arabic TOPL-2

1. Introduction

The development of abnormal language skills in children is an uncontrollable phenomenon, but it can be reduced and clinically treated if addressed from the outset of its identification. A basic need for achieving this goal is an assessment tool that can primarily determine the absence or presence of any language impairment(s). Pragmatics, as either a language skill or linguistic component, plays a critical role in the day-to-day communication of human beings in general. When dealing with such a skill, however, a specialized tool is required to scientifically confirm the presence of such a problem. The common tools that are widely used to assess and/or identify pragmatic language impairment(s) in children and adolescents are the following: Clinical Evaluation of Language Fundamentals (CELF-4), Eleanor Semel, Elisabeth H. Wiig, and Wayne A. Secord (PP: Pragmatics Profile subtest & ORS: Observational Rating Scale subtest), (2003), Test of Pragmatic Language (TOPL-2), Diana Phelps-Terasaki & Trisha Phelps-Gunn, (2007), Pragmatic Language Skills Inventory (PLSI), James E. Gilliam & Lynda Miller, (2006), Children's Communication Checklist (CCC-2), D. V. M. Bishop, (2006), Test of Language Development-Intermediate (TOLD-1:3), Donald D. Hammill & Phyllis L. Newcomer, (1988), and Comprehensive Assessment of Spoken Language (CSAL), (Pragmatic Judgment subtest), Elizabeth Carrow-Woolfolk, (1999).

Each of the aforementioned tools has its own characteristics and, more importantly, its own strengths and weaknesses. That being said, only researchers can determine if one tool is superior to another and/or others. Generally speaking, the degree of validity and reliability of any constructed test can provide either benefits or disadvantages to the test. All of the above tests have been proven to be valid and reliable with varying degrees according to test guidelines. The original language of all of the above-mentioned assessment tools is English.

Yet, similar standardized tests for carrying out such assessments in Arabic are, to the author's best knowledge, unavailable. Thus, clinicians, communication disorder consultants, speech pathologists, and other professionals in the field use general Arabic-language tests for the evaluation and assessment of children and adults with PLIs, without acknowledging that each component of language should be assessed and tested using a valid and reliable tool that can accurately determine the absence or existence of a particular language component disorder. The consequence of using a single test for all language components and skills is that neither children nor adolescents with PLIs are given proper clinical treatment, as rehabilitation programs made for such patients are based on partial and superficial evaluations. As a result, these patients are treated neither adequately nor effectively, and their pragmatic language skills remain below acceptable levels. Thus, a major aim of this research is to provide a usable assessment tool that can be used by clinicians and relevant specialists when assessing children with PLIs, in addition to measuring children's and adolescents' pragmatic competence.

This research is the first attempt to produce a usable psycholinguistic marker, namely TOLP-2, for identifying and assessing Arab-speaking children with PLIs. Additionally, the translated test produced for identifying individuals with PLIs—measuring their pragmatic skill and providing feedback and guidelines for the treatment and rehabilitation of children

with this disorder—is not only intended for clinicians and specialists in this area, but can also be used by parents to monitor the development of their child’s language, allowing them to verify whether they are using language in the appropriate context while also indicating whether they exhibit any PLIs.

Pragmatics has been defined in many different ways. Yule, for example, defines it as the linguistic field ‘concerned with the study of meaning as communicated by a speaker (or writer) and interpreted by a listener (or reader)’ (1996, p. 3). In general, persons with a PLI are socially isolated, except from their family members. A major aspect of human language is successful communication between the speaker and hearer; when this communication is disturbed, the major aim of language is lost. Moreover, when a person at a given age feels that they have generally poor communication and feels that others do not understand them, are disinterested in communication with them. Or they are impatient with the person’s weakness in pragmatics, feelings of disappointment set in alongside feelings of isolation from society, (Cummings, 2009 and Kecskes, 2007).

Pragmatic language impairments (PLIs) can actually be a disorder themselves, but they are more commonly a sign of other accompanying syndromes and language disorders. However, the role of the right hemisphere in controlling pragmatic ability lies within the scope of this study and can be discussed under congenital neurological disorders (where there is no brain injury, but there are cell and neuron shortages, clustering, etc.). In the case of this study, both children and adolescents with developmental dysphasia (different types of disorders) exhibited pragmatic language impairment when tested using the Arabic translation of the TOPL-2.

Basically, aphasia can be classified into three types: child aphasia (which is usually called dysphasia to differentiate between adult aphasia and child aphasia), adult aphasia, and aging aphasia, which is more commonly denoted as dementia (Sarno, 1998). Additionally, both aphasia and dysphasia can be further classified into two types: acquired and developmental. The former indicates any type of language dissolution, be it partial or total, resulting from any type of brain injury; the latter represents any type of language disorder that is not due to brain injury. Moreover, other symptoms in the brain such as cell shortages, inherited disease, or natural neurological disorders in some neurons of the brain can also be classified as developmental aphasia (Sarno, 1998 and Ahlsen, 2006). Furthermore, it should be noted that in such incidences, there is an absence of brain injury; therefore, the causes are congenital, that is, the disorder is ‘genetic environment induced prenatal cerebral defect’ (Sarno, 1998, p. 26). In this case, however, the localization of injured brain areas is not common, though the presence of ‘unusual clusterings of cells’ (Obler and Gjerlow, 1999, p. 73) is possible and can be studied. Hence, this study focuses on children and adolescents who suffer from PLI as a sign and/or symptom of any of the many types of developmental dysphasia DD. Possible types of DD, as defined in this study, could include delay language development DLD, specific language impairment SLI, hearing impairment HI, language disability LD, mental retardation MR, and attention deficit hyperactivity disorder ADHD.

As previously mentioned, a certain language may lack instruments and tools for measuring

particular language impairments, as in the case of the Arabic language, which lacks effective tests for identifying PLIs in children and/or adolescents. As a result, researchers need to translate imported tests and verify their usability in the chosen language, while also considering cultural issues. In effect, all of these issues have been examined in studies conducted in other languages, as described below.

A significant study is that of Ketelaars (2010), which starts by constructing a framework for the nature of pragmatic language impairments, that is, pragmatics and types of disorders that are thought to accompany PLIs. This framework also tests the validity and reliability of the Children's Communication Checklist CCC translated into Dutch for administering it to children from the Netherlands. Ketelaars concluded that the results of the Dutch translation version showed 'good internal consistency and reasonable construct validity' in addition to 'a moderate concordance between teacher opinions and Pragmatic Composite cut-off score', (ibid, 2010, p.38).

It seems that the problem of translating instruments for assessing pragmatic language ability in children and adolescents exists for many languages, just as in the case of the Arabic language. For example, in Helland & Heimann's (2007) study, in which the CCC is translated to Norwegian, the authors proposed that pragmatic language impairments among children who are referred to child psychiatric services are more prevalent than in typically developing children. The study showed that the translated version is just as effective and usable as in English, especially for identifying and distinguishing between children with and without PLIs.

In this study, it is proposed that an early-draft Arabic translation of the TOPL-2 could successfully be used to identify children with PLIs and distinguish them from those without PLIs. Consequently, this could lead to the more general claim that rehabilitation programs based on results obtained from TOPL-2 (Arabic translation) for children with PLIs may be more effective and useful than those based on the general evaluation tests currently in use.

2. Method

2.1 Participants

The participants in this study included children and adolescents with PLIs as either pure pragmatic language disorders or accompanied by other linguistic or cognitive disorders. All participants were Saudi nationals and had been diagnosed as having either a cognitive disorder or linguistic disorder and had furthermore been brought to the Communication and Swallowing Disorders Unit, King Abdulaziz University Hospital, College of Medicine, King Saud University, Riyadh, Saudi Arabia, to be treated by an appropriate specialist. The normal children and adolescents that participated were also Saudi nationals and were siblings of the abnormal children who had come to the hospital with their parents.

The researcher followed the *convenience sampling* method for selecting the cases for this study. In other words, twenty eight normal children and adolescents and another twenty eight abnormal children and adolescents were chosen to take the Arabic TOPL-2 test.

The setting of this research was the Communication and Swallowing Disorders Unit (CSDU), Research Chair of Voice, Swallowing, and communication Disorders, King Abdulaziz University Hospital (KAUH), College of Medicine, King Saud University, Riyadh, Kingdom of Saudi Arabia (2010-2011).

It should be noted that this study was submitted to the Research Centre of the King Khalid University Hospital, College of Medicine, King Saud University, Riyadh and was reviewed and approved by the IRB (Institutional Reviewing Board) before being conducted. Table (1) illustrates the characteristics of the participants in this study.

Table 1. Characteristics of participants in the study

Clinical group		Control group
Variable	Characteristics	Characteristics
No. of participants	28	28
Gender	F:15- M: 13	F: 14- M: 14
Age range	6-16 years old	6-16 years old
IQ range (verbal)	52-110	Average
Diagnosis	DLD: 14, SLI: 8, LD: 2, MR: 1, and ADHD: 1	Normal
Native language	Arabic	Arabic
Dialect used	Saudi Arabic dialect	Saudi Arabic dialect
Nationality	Saudis	Saudis

Finally, because the major purpose of this study was to examine the usability of an Arabic version of the TOPL-2 and because a large proportion of the required population was unreachable due to time limitations, the idea of generalizability was not taken into consideration. However, such a produced version of the TOPL-2 could be considered as usable, feasible and effective for identifying children and adolescents with PLIs.

2.2 Measures

A well-known assessment tool was used in this study, namely TOPL-2. The collected data and investigated variables are quantitatively represented for comparative, contrastive and correlational findings among the selected cases. In addition, discrete qualitative variables

rather than continuous ones were used, as PLIs in children and adolescents cannot be definitively measured. A nominal level measurement was used throughout as a measurement scale mainly for measuring pragmatic competency PC. For other variables, such as descriptive ratings of PC and IQ testing, an ordinal level measurement was used instead.

Another secondary measurement included in this study is the psychometric evaluation in order to assess intelligence quotient IQ. To consider whether a child has a clinically treatable linguistic disorder, an IQ test must be administered because the result of this test affects other processes of evaluation. The results for those whose scores were less than the normal IQ level were taken into consideration by the researcher according to the clinician's indications.

The TOPL Test of Pragmatic Language was designed by Diana Phelps-Terasaki and Trisha Phelps-Gunn and was first published in 1992. The test consists of 44 items targeting children and adolescents between the ages of 6 and 12. It was designed as a battery measuring one's ability to use language in social interactions (pragmatics) and measures six sub-components of pragmatics: physical setting, audience, topic, purpose, visual-gestural cues and abstraction. According to the authors of this test, 'it was standardized on a sample of 1,016 children residing in 21 states using gender, residence, race, geographic region and ethnicity as variables' (Phelps-Terasaki & Phelps-Gunn, 2007, p. v). As an evaluation instrument, it consists of a TOPL Booklet, picture Book (in black and white) and an Examiner's Manual Book.

Over the next ten years, it was found that certain changes were needed for this instrumental evaluation battery. As a result, a new version of the TOPL was produced, the TOPL-2, with new additions, modifications and updates that can be summarized as follows:

- 1) Nineteen items for pragmatic evaluation were added;
- 2) Items indicating persuasion as an important sub-component of pragmatics were added;
- 3) A clinical interpretive framework has been included;
- 4) The norms of the test were extended to involve children between the ages of 12 and 18, taking into consideration relationships between behavior and pragmatic language ability;
- 5) The former book of pictures (black and white) was replaced by a colored one;
- 6) The verbal prompts used to draw data from children and adolescents were shortened, taking into consideration cognitive issues such as attention;
- 7) Construct validity, content and other types of validities have been included.

Children and adolescents eligible for this test are those who can "utilize expressive language" (Phelps-Terasaki & Phelps-Gunn, 2007, p. 11), are aged between 6 and 18 years and exhibit any of the following signs: 1) learning disabilities, 2) language delays and/or disorders, 3) reading and comprehension difficulties, and 4) behavioral, attention, emotional and anxiety disorders (summarized from Phelps-Terasaki & Phelps-Gunn, 2007).

There are five major aims or uses of the TOPL-2 test. According to Phelps-Terasaki &

Phelps-Gunn(2007), the aims of the TOPL-2 are the following: 1) to identify individuals with pragmatic language deficits, 2) to determine individual strengths and weaknesses, 3) to document an individual's progress, 4) to measure pragmatic language in research, and 5) to address the needs of specific populations. This research focuses on uses 1, 2 and 3.

Finally, for reasons of reliability and validity, all factors were carefully taken into consideration by the authors of the aforementioned test. The authors 'calculated three types of TOPL-2 reliability: content sampling (Cronbach's coefficient alpha), time sampling (test-retest), and scorer differences (correlation between scorers), (Phelps-Terasaki and Phelps-Gunn, 2007, p. 51). The first coefficient was 91, the second one was estimated at 99, and the third one was found at 98 (ibid, p. 55). It was also stated that for a test such as the TOPL-2, 'reliability coefficients must approximate or exceed .80 in magnitude to be considered minimally reliable; coefficients of .90 or higher are considered most desirable' (ibid, p. 51).

Regarding testing validity, the content-description validity, criterion-prediction validity, and construct-identification validity were all investigated and stated by the authors of the TOPL-2 test. The correlation between IQ scores and TOPL-2 index scores was measured at .52 (a large coefficient), which 'strongly supports the construct-identification validity of the TOPL-2' (ibid, 69). Thus, because this study attempted to produce a usable and feasible Arabic version of the TOPL-2, the reliability and validity issue is addressed in the results section.

2.3 Design

A pre-test post-test complex quasi-experimental design was used in this study. In a notational parallel form, this study can be depicted as follows:

N O X O
N O
R¹ R² O
NO₁ O₂ X-
N O₁ O₂ X-

N= non-equivalent groups

R= randomized group (repeated), that is, R¹ and R²

O= the two measures used

X= Arabic versions of TOPL-2 and PP & ORS subtests from CELF-4

X= non-treatment groups

A non-equivalent normal group was given the translated version of the TOPL-2 in addition to a clinical group. Then, a randomized group took the Arabic TOPL-2 version two times (with one week between each instance). Lastly, non-equivalent normal and abnormal groups took translated versions of the PP & and ORS (CELF-4) subtests and the Arabic version of TOPL-2. The three groups in the last three lines (in the notational depiction of the study) were included for the sake of both reliability and validity.

Because this study adopts numerous features of a non-experimental study approach, internal validity issues are irrelevant and were not taken into consideration.

2.3 Procedure

Data collection: A semi-parallel procedure was employed for choosing the cases; that is, cases were not selected on the basis of the analysis of the previous case (sequential procedure), nor were any of the cases identified in advance (parallel procedure). Instead, a mixed procedure was used in which cases were selected according to drop-in visits to the hospital and their applicability to the objectives of the study.

Authenticity: Consent forms were written to describe the research procedures. In other words, a consent form was given to the parents of the participants, requesting and informing them of data collection as well as the fact that meetings with their children would be used for research purposes only, ensuring a high degree of confidentiality for any data obtained. Another consent form was given to the corresponding staff at King Abdulaziz University Hospital, who in turn approved and agreed to allow archival records and medical reports access to the researcher.

Test administration: The Arabic TOPL-2 was administered in two sessions to the children aged 6-7 and in a single session to the 8-18 age group. However, short pauses and breaks were allowed and given in cases in which the researcher became aware of a critical situation or non-response by the participant.

Time and environment of test administration: The directions recommended by the authors of the original test were followed precisely. The approximate time for administering the test was 1 to 1½ hours. A comfortable table and chair in a small quiet room were all provided for the testing environment. Any distractions or attracting accessories were eliminated in the room to allow the participants to concentrate on the test.

Process performance: A question was first read to the child and/or adolescent, and afterwards, they were shown a picture. All other instructions and directions were followed according to the examiner's manual (Phelps-Terasaki & Phelps-Gunn, 2007, p. 13).

Scoring: A participant was given (1) for a correct response and (0) for an incorrect answer, based on the possible correct and incorrect answers provided by the authors of the original tests.

Preliminary analysis steps: First, the raw score for each participant was determined and recorded in the appendix provided by the authors according to the actual age of the participant. Then, the raw score was converted to a pragmatic language usage index and percentile rank. A descriptive rating for each case was then provided based on the identified index score. The last step involved locating both age and grade equivalents for each case using the relevant recorded raw score.

3. Results

The researchers used the 17th version of SPSS for statistical data analysis. The following

tables and figures illustrate the statistical tools used and the reason(s) for their use.

Table 2 lists both evidenced types of reliability and validity in this study. As shown in the table, the achieved degrees of reliability and validity varied considerably from one type to another. Three types of reliability were measured; two types, the inter-rater and internal reliabilities, were high compared to the test-retest reliability. In addition, five types of reliability were examined to achieve construct validity. As indicated in the table, however, the calculated values of validity were not all positive: face validity is high, whereas content validity is only good. Again, both concurrent validity and convergent validity are low only for the PP subtest when compared to the very low convergent validity of the ORS subtest. The last type of validity was calculated and is shown in the table, also demonstrating low values. In conclusion, it can be assumed that the Arabic TOPL-2 is minimally reliable and satisfactorily valid.

Table 2. Reliability and validity results of the Arabic TOPL-2 and the statistical tools used

Reliability	Statistical tool and result		Validity	Statistical tool and result	
	Tool	Result		Tool	Result
Inter-rater	Pearson	.97-.98	Face	3 raters	High
Test-retest	T-Test	.73	Content	Categories	Good
Internal	Cronbach	.90	Predictive	Uncalculated	
			Concurrent	Pearson	.24 ^{low}
			Convergent	Pearson	.42 ^{pp low} & -.42 ^{ORS}
			Discriminant	Pearson	.50

Table 3 shows the correlation between the Arabic TOPL-2 test items used to measure pragmatic competency and to identify children and adolescents with pragmatic language impairments. Internal consistency results were calculated for both the abnormal group and the normal group. As demonstrated in the table, the internal consistency rates for each pragmatic component varied. For instance, in the case of situational context, it is good in the abnormal group but unacceptable in the normal group. Furthermore, it is acceptable in the discourse context for the abnormal group but excellent for the normal group. Likewise, for the semantic context, it was rated as unacceptable in the abnormal group and good in the normal group. Thus, it can be concluded that the internal consistency of the Arabic TOPL-2 is not poor but nonetheless necessitates further research and modifications.

Table 3. Internal consistency for the Arabic TOPL-2 measures

Population sample (28 normal and 28 abnormal)

TOPL-2 Scale	Cronbach's alpha ^{abnormal}	Cronbach's alpha ^{normal}
Situational context	.83	.016
Discourse context	.77	.95
Semantic context	.75	.80

To ensure that the Arabic translation of the TOPL-2 truly measures what the original TOPL-2 was designed to measure, the construct validity for both the TOPL-2 components and sub-components was statistically calculated. Table 4 presents the calculated values for the pragmatic language components and subcomponents of both the abnormal and normal groups in this study. It is clear that the achieved values for construct validity are extremely high. Construct validity is high for the three pragmatic components, with slight differences between the two groups as well as from one component to another. The highest value of .99 was obtained for the discourse and semantic contexts and also for the purpose subcomponent. The lowest value in the abnormal group, .56, was obtained by the abstractions subcomponent. On the other hand, the lowest value in the normal group, .49, was obtained by the visual-gestural cues subcomponent. More importantly, one subcomponent in the normal group obtained an invalid score (physical context: .066). On the basis of this, one could infer that the construct validity of the Arabic TOPL-2 test is generally moderate and requires some modifications in the translation of test items.

Table 4. Construct validity of the Arabic TOPL-2 version

Correlations between test items and pragmatic language components

TOPL-2 component	Pearson result ^{abnormal}	Pearson result ^{normal}
Situational context	.96	.97
Physical context	.89	.066*
Audience	.95	.98
Discourse context	.97	.99
Topic	.95	.98
Purpose	.99	.98
Semantic context	.99	.99
Visual-gestural cues	.93	.49**
Abstractions	.56**	.98
Pragmatic evaluation	.97	.98

Indicates insignificant values, ** indicates low-level validity, all other values are significant at the 0.01 level.

The major objective of the present study was to test the feasibility and usability of an Arabic translation of the TOPL-2. A minor objective for this study, however, was to assess the ability of the Arabic version to distinguish between children and adolescents who are either pragmatically impaired or unimpaired. The differences between the abnormal group and the normal group are demonstrated in table 5. The mean, standard deviation and range values are generally and significantly higher for the normal group than for the abnormal group. The TOPL-2 test is divided into three components, with each component having either two or three subcomponents. The number of subjects in the two groups is similar: 28 participants, consisting of males and females. The highest mean, SD and range for the components in the abnormal group are 4.7, 6.4, and 26 (semantic component), while the figures for the normal group are 24.4, 8.8, and 26 for the same component. Again, the highest value for the subcomponents was obtained by the physical context for both the abnormal group (6.5, 8.7, and 36) and the normal group (37.9, 14.6, and 41). In contrast, the lowest value for both the pragmatic components and subcomponents was obtained for the abstractions subcomponent of the abnormal group (.17, .54, and 2) and the discourse context component for the normal group (3.03, .74, and 3). In conclusion, the means, standard deviations, and range values for the normal group were generally higher than those of the abnormal group, primarily indicating that there are clear and visible differences between pragmatically impaired and unimpaired participants.

Table 5. Results of PC measures by group (means, standard deviations, and ranges)

TOPL-2 components	Abnormal group				Normal group			
	n	M	SD	range	n	M	SD	range
Situational context	28	2.6	3.6	14	28	15.1	5.2	16
Physical context	28	6.5	8.7	36	28	37.9	14.6	41
Audience	28	3.5	4.5	18	28	22.9	9.1	27
Discourse context	28	1.11	1.3	4	28	3.03	.74	3
Topic	28	1.5	2.4	10	28	12.07	5.1	15
Purpose	28	1.8	2.4	10	28	13.0	5.8	16
Semantic context	28	4.7	6.4	26	28	24.4	8.8	26
V-G cues	28	1.3	1.7	6	28	4.9	1.3	6
Abstractions	28	.17	.54	2	28	4.6	4.04	11
Pragmatic E	28	1.9	2.7	10	28	13.3	4.6	13

Figures 1 and 2 represent descriptive rating levels for both the abnormal and normal group obtained from the Arabic TOPL-2 assessment tool for pragmatic competence performance. It can be seen from these two figures that the descriptive ratings of the abnormal group are generally negative, regardless of their variance, whereas the descriptive ratings for the normal group are generally positive, irrespective of their differences. While more than 35% of the participants in the abnormal group received poor ratings, more than 46% of the normal group was rated above average, and a similar proportion received average ratings. Approximately 7% of the participants in the normal group were rated superior. In comparison, according to the Arabic TOPL-2, none of the normal group participants exhibited a pragmatic language impairment, but the test did identify a number of the abnormal group participants as pragmatically unimpaired (just 7% below the average rate and more than 3% below the above average rate).

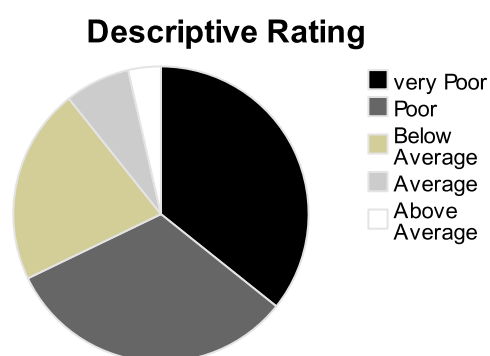


Figure 1. Abnormal Group PC Performance

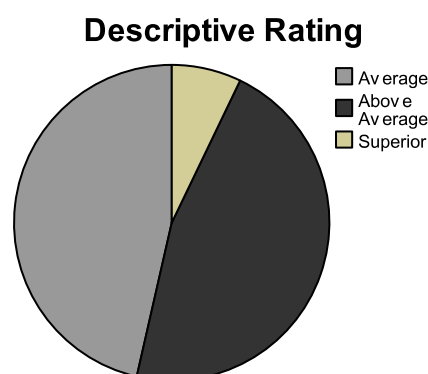


Figure 2. Normal Group PC

Figures 3 and 4 comparatively depict the calculated raw scores obtained by both the abnormal and normal groups using the Arabic TOPL-2 test. It is clear that the participants in the normal group obtained significantly better scores than those in the abnormal group. The highest raw score in the normal group is 41, and the lowest score is 8. This is in contrast to the abnormal group, where the highest score is 20 and the lowest is 0. As shown in figure 3, the number of participants with raw scores between 8 and 9 increased moderately, with a dramatic increase for a raw score of 11. The numbers remain steady at the raw score of 27 and then skyrocket for raw scores of 30 and 32; between the scores of 33 and 41, there is a degree of leveling. The results show peaks at three levels: 11, 30 and 32. On the other hand, in the case of figure 4, the raw scores of the Arabic TOPL-2 show a peak at 0 and then exhibit a noticeable fluctuation between the raw scores of 1 and 5 for the participants in the abnormal group. In the case of raw scores between 6 and 20, a slow decline is noticeable. In summary, the participants in the normal group clearly achieved considerably higher scores than the participants in the abnormal group, supporting the idea that the Arabic TOPL-2 test can distinguish between normal and abnormal individuals in terms of PLIs.

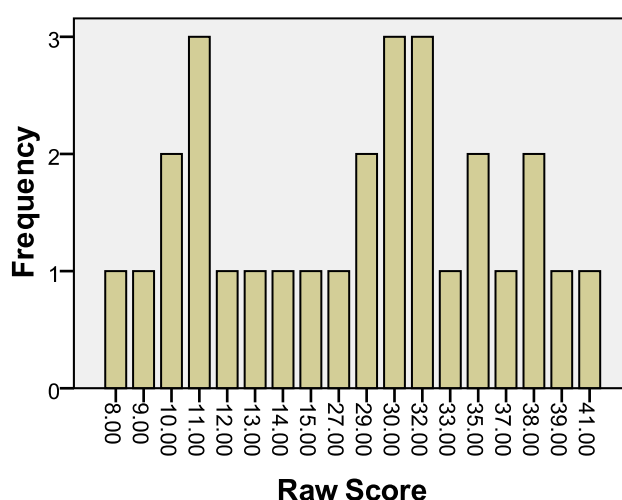


Figure 3. Abnormal TOPL-2 Raw Score-Normal Group

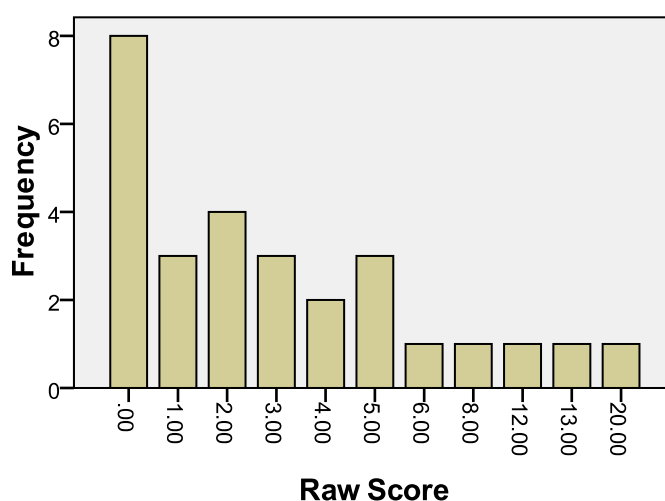


Figure 4. Abnormal TOPL-2 Raw Score Abnormal Group

4. Discussion

The principal goal of this study was to assess the usability and/or feasibility of an Arabic TOPL-2 test version to identify children and adolescents with PLIs and to distinguish them from those without PLIs. As a result, usability and/or feasibility issues in terms of both reliability and validity have been accounted for in detail.

In this study, reliability was documented using different measures to obtain different types of reliability. For inter-rater reliability, for example, the booklets collected from the participants in this study were scored three times over three days. Following the original version of the TOPL-2, a participant is given a score of (1) when an answer is correct and (0) if an answer is incorrect. The inter-rater reliability for scoring the items of the Arabic TOPL-2 was estimated using a Pearson coefficient correlation and ranged between .97 and .98.

For test-retest reliability, twelve subjects took the test twice. The second test was administered one week after the first test. The degree of test-retest reliability of the Arabic TOPL-2 was measured by comparing the summed scores of the two tests using a paired sample T-test. The test-retest reliability of the Arabic TOPL-2 was .73, and the significance was .007, which is significant at the .001 level.

The final area in which reliability was studied in this study is internal consistency reliability, estimated using Cronbach's alpha. The average internal consistency reliability coefficient was .90.

Just as reliability was documented for the measure used in this study, validity was also documented and examined using different methodological techniques. After beginning with translation validity, face validity was also examined in this study. Face validity was achieved by conferring with three experts in the fields of language studies. Two of the experts are professors specializing in English linguistics, and one specializes in Arabic linguistics at the College of Arts of King Saud University, Riyadh, Saudi Arabia. These experts were selected to determine the language skills and/or components that this tool could evaluate and assess, without being informed that it is intended for pragmatics. Indications that were included in the translation, e.g., the title, were all removed to achieve stronger face validity. All of the participating professors, to a great extent, agreed that this tool could only test pragmatics, although two of them also indicated the possibility of some semantic issues.

To achieve as high a validity degree as possible for the measure used in this study, content validity was also assessed by dividing the items of the test into seven categories. Each of these categories assesses a certain component of pragmatic competency PC. Needless to say, the authors, for the purpose of clinical evaluation and implications, had already made this classification. The researchers did their best to maintain this strategy while translating the test into Arabic and also during Arabization. The seven components are as follows: physical context, audience, topic, purpose, visual-gestural cues, abstractions, and pragmatic evaluation. In spite of this, the cultural differences component was inserted covertly during the Arabization process (as will later be discussed in detail).

In addition to translation validity, criterion-related validity was also demonstrated. The predictive validity PV was not calculated, as future predictive indications for atypically developing children and adolescents were not taken into consideration as major goals of this study.

Starting with concurrent validity, the raw scores of the abnormal participants were compared in terms of means and standard deviations (descriptive statistics) and correlations using the Pearson coefficient. The correlation between the two variables raw scores of the clinical and control groups was .24, a very low validity correlation. Despite this, the means and standard deviations of the two different groups were considerably different. For the normal group, the mean was 24.68, while for the clinical group, it was only 3.64. Again, the standard deviation for the first group was 11.50, while that of the second group was 4.69.

The second type of criterion validity studied was convergent validity, which was calculated

using the raw scores of the Arabic TOPL-2. The scores were first correlated with the raw scores of the Arabic PP and then with those of the ORS. The achieved concurrent validity, however, was .42 for the PP and -.42 for the ORS.

The third type of criterion validity addressed was discriminant validity DV, which was calculated for the Arabic TOPL-2 tests. The summed IQ test scores were correlated with the summed scores of the TOPL-2 to investigate the degree of DV (the summed score is the raw score according to the TOPL-2 examiner's manual). After correlating the sum of the achieved scores of the participants for the Arabic TOPL-2 with the IQ test results, the Pearson coefficient correlation was .50. This result indicates that the two variables are moderately correlated, that is, there exists a moderate discriminant validity correlation.

In addition to the standards of reliability and validity, the researchers also considered the issue of translation and Arabization of the TOPL-2. To first ensure that the translation and subsequent Arabization did not affect the original source of the test, a number of procedures were performed. First, the 43 TOPL-2 items along with the 17 items for pragmatic evaluation were literally translated by two different persons, including the first author of this paper. The coincidence and similarity degrees between the two literal translations were theoretically estimated at over 90%. Having achieved this, a back-translation process was performed by the same two translators to determine whether the literal translation would unconsciously lead to any changes in the contents of the original test. The results of the two back-translations were, to a great extent, consistent with one another, as well as with the original text. However, significant differences between the original text and the back-translation were restricted to lexical differences. In other words, while the authors of the original TOPL-2 adhered to the American English variety, the translators of the TOPL-2 used the British English variety instead.

The last step was to perform Arabization on the translated TOPL-2 version. It should be noted that a unified translation was made from the two literal translations, in agreement with the two translators and after some discussion. Despite this, some minor modifications were made in the Arabization process of the TOPL-2 for a number of reasons.

For instance, the items indicating cultural issues, namely proverbs and sayings, were replaced with equivalents from the Arabic language. The numbers of the items corresponding to these expressions are as follows: 20, 26, 29, 40, 42 and 43.

Foreign names were also replaced by Arabian names to make it easier to administer the test to the Arab participants of this study. Specifically, Cindy was replaced by Fatima, Matt by Mohammed, Kate by Aisha, Dan by Ahmed, Scott by Sa'ad, Brad by Badr, Chad by Rashad, Smith by Ali, and finally Bata by Basem.

The most problematic issue was the use of the picture book. In this book, certain items use pictorial explanations featuring girls dressed in a style that is not common in Arabian countries, specifically Saudi Arabia. The participants in this study, particularly children, were not able to recognize the persons in the pictures as girls; they instead considered them as boys, especially those with short hair. The numbers of the items in the picture book related to this

problem are the following: 2, 4, 5, 8, 13, 17, 24, 31 and 33. Similarly, it was impossible to explain to the children—and even the adolescents—how a boy and girl could become friends and socialize, as this occurrence is not common in Arabian countries. For this reason, items that required an illustration of friendship or gatherings of boys and girls and the like were introduced with all friends being either males or females. The clothing style helped in this aspect as trousers, t-shirts, and short hair styles are not at all common among members of Saudi society.

To determine the reliability and validity of the above-calculated data, the results obtained from the normal and abnormal groups are presented in table 5 and figures 1-4. The statistics clearly indicate that the results of the normal group are significantly higher than those of the abnormal group, in terms of means, standard deviations and ranges, as shown in table 5, descriptive ratings, as illustrated in figures 1 and 2, and achieved raw scores, as shown in figures 3 and 4.

On the basis of these results, it can be roughly inferred that the Arabic TOPL-2 can distinguish between pragmatically impaired and unimpaired children and adolescents. The low achieved scores for some cases in the normal group are due to differences between the number of items used and age differences. In other words, in the 6- to 7-year-old age range, a raw score of 8 would correspond to a raw score of approximately 22 for the 8- to 18-year-old age range. Approximately 10% of the participants were identified as pragmatically unimpaired, despite exhibiting slow and poor communication abilities; according to their families, this result could be attributed to two possible reasons. One, their PLI may still be mild or may correspond to a secondary pragmatic impairment (Perkins, 2007 and Perkins, 2000). Second, there are some covert problems in the Arabic TOPL-2 version that could be described as slightly serious.

5. Conclusion

The results of this study partially corresponded to the initial proposals. In other words, the first and primary claim that an early-draft version of the Arabic TOPL-2 could be used to identify and distinguish children and adolescents with PLIs, with considerable and significant differences, was found to be mostly accurate. On the other hand, the second and more general claim that results based on the Arabic TOPL-2 assessment could be used to develop a rehabilitation program for children and adolescents with PLIs was found to be erroneous and was subsequently rejected in this study. A major reason for the failure to achieve such a proposal might lie in the time limitation imposed on this study. However, the early-draft Arabic version of the TOPL-2, which has been preliminarily shown to be feasible and usable, is a significant tool in itself and could greatly contribute to reaching the remaining goals in future work.

This study has clinical implications for speech language pathologists, clinical linguistics, interventionists, phoniatricians and parents. In the case of speech and language pathologists and phoniatricians (in Arabian countries), the results of this study can be compared with results obtained from the general language tests that are currently used to assess children and adolescents in their clinics. As for speech and language pathologists and phoniatricians in

general, the results of the Arabic TOPL-2 can be compared to those of PLIs in other languages. In the case of clinical linguists, the results of this study may help in developing a universal theory of PLIs in all languages based on results and data using the same tool, namely the TOPL-2, serving to assist in identifying children and adolescents with PLIs. Moreover, interventionists can make the most of these results, which could greatly help them in planning rehabilitation and treatment programs for those who have been identified as pragmatically impaired. That is to say, instead of making a general rehabilitation program for all language components, certain language components could receive more attention and intervention than others based on the results of specific assessment tools such as the Arabic TOPL-2 test. Parents can also share such test results with a communication disorders specialist to help their children overcome any linguistic abnormalities that they are exhibiting.

A number of limitations were noted in this study. First, such a study normally requires a much larger number of participants to prove its feasibility; this is especially true in this study's attempt to calculate the validity and reliability of the Arabic TOPL-2. Unfortunately, this limitation seemed inevitable due to the limited number of available cases relating to language disorders, mental disorders or even learning disabilities. Second, the study was limited to patients exhibiting language disorders as symptoms of other primary disorders such as SLI, LD, DLD, etc. Third, the study was also limited to the use of a single assessment tool due to the lack of Arabic batteries that assess only PLIs. These limitations may inhibit the generalizability of the results of this research.

In spite of these limitations, the Arabic TOPL-2 version, which has been proven moderately reliable and satisfactorily valid, can be used for the purposes presented by the researchers in this study. Regardless of whether the test is similarly applied, the issues stated in the discussion as obstacles to the Arabic TOPL-2 version must be taken into consideration, especially the designing of a picture book suitable for Arabian culture in order to minimize the confusion observed during the administration of the Arabic version. Again, the Arabization of names, which was done by the researchers in this study, should be considered; a researcher could either use the previously suggested Arabian names as alternatives or suggest new ones on the basis of the most common names in the country in which the research will be conducted.

Acknowledgement

The first author would like to thank Professor Mohammed Ziad Kebbe for his inspiring ideas and psychological support for the first author to search in this area. The authors would like also to thank the IRB (Institutional Reviewing Board) at King Khalid University Hospital, College of Medicine, KSU, Saudi Arabia, for reviewing the research proposal and issuing the IRB approval letter. Additionally, many thanks are also due to all of the employees at the Research Chair of Voice, Swallowing, and communication Disorders and more specifically at the CSDU (Communication and Swallowing Disorders Unit). Needless to say, many thanks also go to the Department of English Language and Literature, College of Arts, KSU, for providing the first author with the official papers required for data collection and other

procedural issues.

References

- Adams, C. (2002). Practitioner Review: The assessment of language pragmatics. *Journal of Child Psychology and Psychiatry*, 43(8), 973-987. <http://dx.doi.org/10.1111/1469-7610.00226>
- Ahlсен, E. (1995, December). *Gothenburge papers in theoretical linguistics: Pragmatics and aphasia - an activity based approach*. [Online] Available: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.50.255> (October 10, 2010)
- Ahlсен, E. (2006). *Introduction to Neurolinguistics*. Amsterdam, Philadelphia: John Benjamins Publishing Company [electronic version] .
- Alduais, A. M. (2012). Identifying Typical and Atypical Manifestations of Pragmatics Disorder in Arabic: A Multiple-Case-Study on Children with Language Delay Disorders Using Two Pragmatic Language Tests- TOPL-2 & CELF-4. Unpublished MA thesis, King Saud University (KSU), Riyadh, Saudi Arabia.
- Bain, J. L. (2000). Language development in children with attention deficit disorder. Ph.D. dissertation, University of Victoria (Canada), Canada. Retrieved October 13, 2010, from Dissertations & Theses: Full Text.(Publication No. AAT NQ52751).
- Bishop, D. V. M. & Baird, G. (2001). Parent and teacher report of pragmatic aspects of communication: Use of the Children's Communication Checklist in a clinical setting. *Developmental Medicine and Child Neurology*, 43(12), 809-18. Retrieved October 13, 2010, from ProQuest Medical Library. (Document ID: 98646486).
- Burke, L. J. (1993). Pragmatic analysis of communicative behavior in three groups of mentally retarded adults. Ph.D. dissertation, York University (Canada), Canada. Retrieved October 13, 2010, from Dissertations & Theses: Full Text.(Publication No. AAT NN84188).
- Carrow-Woolfolk, E. (1999). *Comprehensive Assessment of Spoken Language (CASL)*. Bloomington: Pearson Assessments, Bloomington, MN.
- Cummings, L. (2009). *Clinical Pragmatics*. Cambridge : Cambridge University Press (CUP).
- Deal, S. (2009) The relationship between pragmatic language skills and behavior. Ed.D. dissertation, Walden University, United States -- Minnesota. Retrieved October 13, 2010, from Dissertations & Theses: Full Text.(Publication No. AAT 3380348).
- Fabbro, F., & Asher, R. E. (Eds.). (1999). *Concise encyclopedia of language pathology*. Amsterdam - Lausanne - New York - Oxford - Sahnnon - Singapore - Tokyo: Elsevier.
- Gilliam, J. E., & Miller, L. (2006). *PLSI: Pragmatic Language Skills Inventory: Examiner's Manual* . Pro-ed Inc.
- Hammill, D. D., & Newcomer, P. L. (1988). *Test of Language Development-Intermediate, 4th Edition (TOLD-I:4)*. Pearson.
- Helland, W. A., & Heimann, M. (2007). Assessment of pragmatic language impairment in

children referred to psychiatric services: A pilot study of the Children's Communication Checklist in a Norwegian sample. *Logopedics Phoniatrics Vocology*, 23-30, Retrieved from: <http://search.epnet.com>, at 01/11/2010 .

Kecskes, I., & Horn, L. R. (Eds.). (2007). *Explorations in Pragmatics: Linguistic, Cognitive and Intercultural Aspects*. Berlin, New York: Mouton de Gruyter.

Ketelaars, M. P. (2010). *The Nature of Pragmatic Language Impairment*. Retrieved 10 15, 2010, from google, Sint Marie Centre: [Online] Available: <http://dare.ubn.kun.nl/bitstream/2066/76512/1/Proefschrift%20Mieke%20Ketelaars.pdf>

Kim, O. H. (1999). Language characteristics and social skills of children with attention deficit hyperactivity disorder. Ed.D. dissertation, Peabody College for Teachers of Vanderbilt University, United States -- Tennessee. Retrieved October 13, 2010, from Dissertations & Theses: Full Text. (Publication No. AAT 9929118).

Kim, O. H., & Kaiser, A. P. (2000). Language Characteristics of Children with ADHD. *Communication Disorders Quarterly*, 154-165.

Lange, H. (2010) Verbal and non-verbal pragmatic deficits in adults with traumatic brain injury: Perspectives of urban speech-language pathologists. M.S. dissertation, University of South Alabama, United States -- Alabama. Retrieved October 13, 2010, from Dissertations & Theses: Full Text. (Publication No. AAT 1484486).

Noveck, I. A., & Sperber, D. (Eds.). (2004). *Experimental Pragmatics*. New York: Palgrave Macmillan Ltd.

O'Neill, D. K. (2007). The Language Use Inventory for Young Children: A Parent-Report Measure of Pragmatic Language Development for 18- to 47-Month-Old Children. *Journal of Speech, Language, and Hearing Research*, 50, 214-228. [Online] Available: <http://search.epnet.com>, at 01/11/2010.

Obler, L. K., & Gjerlow, K. (1999). *Language and the Brain*. Cambridge: Cambridge University Press [electronic version].

Perkins, M. (2007). *Pragmatic Impairment*. Cambridge : Cambridge University Press (CUP).

Perkins, M. R. (2000). The scope of pragmatics disability: A cognitive approach. In N. Muller, *Pragmatics in speech and language pathology: studies in clinical applications* (pp. 7-23). Amsterdam, Philadelphia: John Benjamin Publishing.

Phelps-Terasaki, D., & Phelps-Gunn, T. (2007). *TOPL-2 Test of Pragmatic Language: Examiner's Manual* (2 ed.). Shoal Creek Boulevard, Austin, Texas: Pro-ed Inc.

Philofsky, A. D. (2006). Pragmatic language profiles in autism and Williams syndrome. Ph.D. dissertation, Colorado State University, United States -- Colorado. Retrieved October 13, 2010, from Dissertations & Theses: Full Text. (Publication No. AAT 3246300).

Philofsky, A., Fidler, D. J., & Hepburn, S. (2007). Pragmatic Language Profiles of School-Age Children With Autism Spectrum Disorders and Williams Syndrome. *American*

Journal of Speech-Language Pathology, 368–380. [Online] Available: <http://search.epnet.com>, at 01/11/2010.

Plymale, F. (2008) The language pragmatics of children with spina bifida: A multiple-case study. Psy.D. dissertation, Antioch University New England, United States -- New Hampshire. Retrieved October 13, 2010, from Dissertations & Theses: Full Text. (Publication No. AAT 3368332).

Sarno, M. T. (Ed.). (1998). *Acquired Aphasia* (3 ed.). San Diego, London, Boston, New York, Sydney, Tokyo, Toronto: ACADEMIC PRESS.

Semel, E., Wiig, E. H., & Secord, W. A. (2004). *CELF®-4 Screening Test: Clinical Evaluation of Language Fundamentals 4*. Pearson Education Inc. .

Smith, H. N. (2002). Pragmatic language skills of adolescents with emotional and behavioral disorders. Ph.D. dissertation, Arizona State University, United States -- Arizona. Retrieved October 13, 2010, from Dissertations & Theses: Full Text. (Publication No. AAT 3043831).

Verschueren, J., Ostman, J.-o., Blommert, J., & Bulcaen, C. (2002). *Handbook of Pragmatics*. Amsterdam, Philadelphia: John Benjamins Publishing Company.

Volden, J., Coolican, J., Garon, N., White, J., & Bryson, S. (2009). Brief Report: Pragmatic Language in Autism Spectrum Disorder: Relationships to Measures of Ability and Disability. *Journal of Autism and Developmental Disorders*, 388-393. [Online] Available: <http://search.epnet.com>, at 01/11/2010.

Volden, J., & Phillips, L. (2010). Measuring Pragmatic Language in Speakers With Autism Spectrum Disorders: Comparing the Children's Communication Checklist-2 and the Test of Pragmatic Language. *American Journal of Speech - Language Pathology (Online)*, 19(3), 204-212A. Retrieved October 13, 2010, from ProQuest Medical Library. (Document ID: 2151665871).

Von-Briesen, P. (2007) Pragmatic language skills of adolescents with ADHD. Ph.D. dissertation, Marquette University, United States -- Wisconsin. Retrieved October 13, 2010, from Dissertations & Theses: Full Text. (Publication No. AAT 3263821).

Whelan, C. (2010) Effective stuttering treatment for persons with autism spectrum disorders: A single-subject design study. M.A. dissertation, The George Washington University, United States -- District of Columbia. Retrieved October 13, 2010, from Dissertations & Theses: Full Text. (Publication No. AAT 1475624).

Wolf-Colon, R. (2009) Pragmatic repetition in normal and disordered verbal discourse. Ph.D. dissertation, New York University, United States -- New York. Retrieved October 13, 2010, from Dissertations & Theses: Full Text. (Publication No. AAT 3361993).

Young, E. C., Diehl, J., Morris, D., Hyman, S., & Bennetto, L. (2005). The Use of Two Language Tests to Identify Pragmatic Language Problems in Children With Autism Spectrum Disorders. *Language, Speech & Hearing Services in Schools*, 36(1), 62-72. Retrieved July 2,

2010, from ProQuest Medical Library. (Document ID: 792076791).

Yule, G. (1996). *Oxford Introductions to Language Study: Pragmatics*. (H. G. Widdowson, Ed.) Oxford: Oxford University Press.

Copyright Disclaimer

Copyright reserved by the author(s).

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).