Speech and Language Sciences Section

School of Education, Communication and Language Sciences

Faculty of Humanities, Arts, and Social Sciences

Clinical Assessment of Persian-speaking Children with Language Impairment in Iran: Exploring the Potential of Language Sample Measures

Yalda Kazemi

To be submitted in fulfilment of

the requirements for the degree of

Doctor of Philosophy

April 2013

Declaration of Originality

The material presented in this thesis is the original work of the candidate except as otherwise acknowledged. It has not been submitted previously in part or in whole, for any award at any university, at any other time.

Acknowledgements

Thanks to God for all those unique supports which are beyond the words… just THANKS.

First of all, I would like to express my gratitude to the children and their parents without whose enthusiastic participation this thesis would be impossible.

It is also a pleasure to thank my supervisor, Thomas Klee who was the first person encouraged me to develop a clinical-relevant topic out of immature ideas and then knowledgably elaborated them to more comprehensive concepts. I also owe a deepest gratitude to my kind supervisor, Helen Stringer for always being there as a friend with sincere kindness. Her "looking from another angle" was a big help when everything looked scrambled and her support throughout this journey made it unique for me.

Thank you to Iranian speech therapists for their eager collaboration in different stages, particularly my friends Mahboubeh and Leila who kindly offered help. An extra thank you to Mahboubeh and Laura for helping at the last minutes.

Many thanks to the manager of Kowsar Rehabilitation Centre in Isfahan, Mr. Shariat and his lovely wife, Tina, for providing a space for parts of data collection when I had no access to other facilities, and for their incredible hospitality in between samplings.

Also I would like to show my gratitude to Ann Nockerts and Jon F. Miller for their big effort in developing the Persian-compatible version of SALT software. It was also a big honour for me to be the winner of Jon F. Miller's student travel award for attending and presenting at SRCLD. I also appreciate CRiLLS and School of ECLS for selecting me as a recipient of the postgraduate award.

The last but not least, enormous thanks go to my beloved family, first of all my parents, whose genuine kindness and prays were the base of my success, and to my dearest sister, brothers, and sister-in-law whose kind comments are unforgettable. To my other sister-in-law and my dear aunt who exceptionally assisted with looking after my sons when I was helpless. My family makes the list of helpers long, so, I thank all the family for their sincere patience.

And the major gratitude is for my great love, Mohsen whose favour in this journey is beyond explanation. Thanks for his silent concerns and big patience to provide the best environment for my study and our life. Big kisses for my lovely sons, Parsa and Nami for being my big inspiration to end this journey with enthusiasm, for their unbelievable understanding of why mum cannot build a larger Lego castle, go out for a walk and an ice-cream, and treasure hunting in the park!

Abstract

Access to evidence-based assessment for diagnosing children with primary language impairment (PLI) in Iran is limited. This study aimed to explore diagnostic criteria employed by Iranian speech therapists for defining PLI and examine the diagnostic potential of language sample measures (LSMs) for Persian-speaking children.

Thirty nine speech and language therapists (SLTs) contributed in a qualitative-quantitative study to explore the criteria currently used by Iranian SLTs to assess and diagnose Persian-speaking children with PLI. Personally-defined diagnostic procedures, based on the results of the questionnaires and focus groups were summarised to obtain a general picture of decision-making methods in identifying Iranian children with PLI. The International Classification of Functioning, Disability and Health (ICF) was used as an organising framework for establishing a consensus as to what constitutes a language impairment, since no commonly accepted reference standard currently exists in Iranian clinical practice.

The assessment potential of LSMs in Persian was examined using the framework of diagnostic research and included a pre-accuracy study followed by phase I and II studies. Twenty seven pre-school children with typically-developing language (TDL) and 24 age-matched children with PLI, aged 42 to 54 months, were recruited. Language samples were recorded as each mother played with her child. None of correlations between age and the LSMs were statistically significant in either group of children (pre-accuracy phase). However, a majority of the LSMs could differentiate children at the group level (phase I). Five measures: Grammaticality/Ungrammaticality, Ungrammatical Utterances, MLUw-excluding one-word utterances, and Semantic Errors, provided good diagnostic accuracy when examined at the level of the individual child (phase II).

An ICF-based reference standard for defining PLI in Iranian Pre-school children has been developed to enhance the consensus among Iranian SLTs. It was applied to recruit the children to the DA study, resulting in five LSMs which are clinically able to differentiate between children with and without PLI.

TABLE OF CONTENTS

LIST OF TABLES

TABLE OF FIGURES

CHAPTER 1

INTRODUCTION

Persian-speaking children with language impairment through the lens of evidence-based practice

1.1 Introduction

Due to the growing number of qualified speech and language therapists (SLTs) in Iran, there is an increasing need for more evidence-based resources regarding assessment and intervention which are tailored culturally and linguistically for the Persian language. In 2011 there were over 1000 registered SLT serving over 75 million people in Iran (L. Gholami Tehrani, personal communication, August 2011; Statistical Centre of Iran, 2011). Iranian SLTs suffer from a lack of valid and reliable standardised tools for assessing children's language development and disorders, particularly those based on a clinical framework grounded in evidence-based practice (EBP). Given the relatively new arrival of EBP into the wider speech and language therapy world, it is logical to apply this method in new diagnostic studies of Persian from the outset in order to overcome the problems of traditional methods in child language research which have limited clinical relevance.

A major part of the first chapter comprises information about the linguistic structure of Persian relevant to the current study (section 1.2), followed by an explanation of the notions of child language development and impairment related to the Persian language (section 1.3). To provide an overall picture of the speech and language therapy situation in Iran, this chapter then reviews the historical background of how speech and language disorders are dealt with in Iran, followed by an introduction to the academic system for educating SLTs in Iran (section 1.4). The remainder of the chapter explains the methodological framework of the thesis as well as the main aims of the study (section 1.5).

1.2 Linguistic features of Persian relevant to the current study

Persian is the second member of the category *middle Persian*, the final subdivision of the Indo-Iranian branch of Proto-European languages. Its history goes back more than 700 years BC and it is spoken mostly in Iran and Afghanistan (Dari Persian) and in an archaic form in Tajikistan (Tajik Persian) and the Pamir Mountain region. The Persian language is called *Farsi* by native speakers in Iran (Bahar, 1996; Yarshater, 1989). In this thesis it will be referred to as 'Persian' throughout. This section will discuss those linguistic aspects of Persian relevant to the present research, starting with Persian morpho-syntactic structure.

Morphology is the part of grammar studying word formation by analysing the morpheme, the smallest meaningful or syntactic component of the word, which is

greater than a phoneme and smaller than a word. The morph is the concrete form of a morpheme (whilst morpheme is by definition an abstract concept), of which different varieties are called allomorphs. Allomorphs appear differently according to context or the users' accents. The following examples show an abstract morpheme in Persian, the third person past-tense marker, as realised by four allomorphs (Kalbasi, 2008) (Examples 1-1 and 1-2). Unless otherwise stated, all examples follow the order: Arabic written form *Romanised form* → inflectional ending (English translation).

Example 1-1) a. رفت *ræft*(past verb root of 'go')→ -t (went)
b. خورد *xord* (past verb root of 'eat')→ -d (ate)
c. افتاد *oftad* (past verb root of 'fall')→ -ad (fell)
d. خرید *xærid* (past verb root of 'buy')→ -id (bought)

Example 1-2) a. بشین *be\/in*→ imperative verb marker\present verb root of 'sit'→(sit-imperative in standard Persian accent)
b. بیشین *bi\/in*→ imperative verb marker\present verb root of 'sit'→ (sit-imperative in Isfahani accent, the accent which is studied in the current research).

Persian has two verb roots (or stems): present and past. Present tenses and imperatives are formed from the present root (Example 1-2 and 1-3c) whilst past tenses are formed from the past root (Example 1-1).

1.2.1 Morpheme categories

Morphemes are generally categorised into five types which are defined here.

1.2.1.1 Content / lexical morphemes

These morphemes contain material or lexical meaning and may come independently as simple words. They also include the main morpheme of the verb. 'The lexeme of the word' or 'free morpheme' are other labels for them. In Persian they are the main word classes or parts of speech (i.e. nouns, verbs, adjectives, adverbs, and prepositions) (Meshkato-Dini, 2008). See example 1-3.

Example 1-3 a. کار(noun) *kar*→work
b. خوب(adjective) *xub*→ good
c. نویس (verb root) in می/نویس/م(inflected verb) *nevis*→ to write

Lexemes are an open class of morphemes and can be increased or decreased in number according to social, cultural and economic changes. The process of word formation helps languages to add to their list of lexemes so that new phenomena, creations and

concepts are able to be reflected in any given language through new nouns, adjectives, and verbs. In contrast, the presence of some lexemes deteriorates over time due to the disappearance of the specific event or phenomena referred to, or historical changes in the meaning they convey (Meshkato-Dini, 2008).

1.2.1.2 Functional morphemes

Functional morphemes or words are closed-class, i.e. limited to a known set of words. They do not have material meaning but instead serve specific grammatical roles or relations (Meshkato-Dini, 2008). The Persian functional morphemes are shown in table 1-1.

Type	**Transcription**	**Persian example**	**English meaning**
Pronouns indicate those meanings related to person and quantity	mæn, to, u, ma, ʃoma, iʃan/anha	من، تو، او، ما، شما، ایشان/آنها	I, you, he/she, we, you, they (*anha* for both animate and inanimate subjects, *iʃan* only for animate subjects).
Conjunctions consist of grammatical meanings that relate subordinate or coordinate clauses to the main clauses	væ, ke, æma, zira	وَ، که، اما، زیرا	And, that, but, because
Direct object marker signifies which word or word group is the direct object	ra	را	

Table 1-1 Persian functional morphemes (Meshkato-Dini, 2008)

1.2.1.3 Derivational morphemes/affixes

These types of morphemes play a leading role in novel word formation, outnumbering the inflectional morphemes. They are closer to the word root when making new word categories and change the grammatical category or subcategory of the derived word (Kalbasi, 2008; Meshkato-Dini, 2008). See example 1-4.

Example 1-4

درد(noun) + ناک (adjective-making morpheme) *dærd* + *nak* →pain+ful(adjective)

بر(verbal prefix, not in a derivational role) + داشت (past verb root of 'to possess') *bær* + *daʃt*→ to take

Another specification of derivational morphemes is that they cannot be used with all word roots in the same syntactic category (Kalbasi, 2008; Meshkato-Dini, 2008), mainly due to phonological or pragmatic restrictions (Natel-Khanlari, 1994; R. Nilipour, personal communication, January 2013). For example the derivational morpheme */eʃ/* is used with most present verb roots in Persian to make nouns (e.g. *kuʃ* is the present verb root meaning 'to make an effort', and *kuʃeʃ* is the derived noun of this verb which means effort) but some verb roots (e.g. *bænd* which means shut) do not accept this morpheme because the derived form (e.g. *bændeʃ*) is not in current usage. However, those noncurrent derived words might be used at some point in the future due to the sociolinguistic needs of the community.

1.2.1.4 Inflectional morphemes/affixes

Inflectional morphemes carry certain grammatical concepts and accompany all word classes in the form of prefixes or suffixes. They show specific syntactic relation in the utterance with two main characteristics:

1. They are used with all word roots in the same syntactic category with almost no exception. An example of one exception could be considered with the */an* plural marker which cannot accompany inanimate nouns such as *ketab* (book).
2. They do not change the grammatical category of the word they attach to (Kalbasi, 2008).

The list of inflectional morphemes in Persian is shown in table 1-2.

Type	Transcription	Persian example	Example transcription	English meaning
Plural markers	*/ha* as a general plural marker for all nouns, */an* as a specific plural marker for animate nouns, and their varieties	کتابها، دختران	ketab/*ha*, doxtær/*an*	books, girls
Markers of comparative adjectives and adverbs	*/tær*, */tærin*	بدتر، زیباترین	bæd/*tær*, ziba/*tærin*	worse, the most beautiful
Ordinal markers	*/om*, */omin*	دوم، پنجمین	pænj/*om*, pænj/*omin*	The fifth
Prefixes of verbs in their morpho-syntactic structure	– Potential mood and imperative marker: *be*\, *bo*\, *bi*\; – Past and present progressive marker: *mi*\; – Negative marker: *næ*\, *ne*\	- بروم، برو، بیا - می رود، می رفت - نرو، نمی رود	– *be*\ræv/*æm*, *bo*\ro, *bi*\ya – *mi*\ræv/*æd*, *mi*\ræft – *næ*\ro, *ne**mi*\ræv/*æd*	– go – is/was going – don't go, isn't going
Suffixes of verbs in their morpho-syntactic structure	– all suffixes of present perfect and past perfect if they have not been used in an adjective meaning; apparently past or present participle marker: */e*, */æm*, */i*, */æst*, */im*, */id*, */ænd*, including all inflected forms of past perfect verb, bud: bud/*æm*, bud/*i*, bud/*ø*, bud/*im*, bud/*id*, bud/*ænd* – all present and past verb personal endings: */æm*, */i*, */æst*, */im*, */id*, */ænd*	- دیده ام - دیده بودم - دیدم - دید	– did/*e*/*æm* – did/*e* bud/*æm* – did/*æm* – did/*ø*	– I have seen. – I had seen. – I saw. – He saw.
Causative verb marker	*/an*	رسانید	res/*an*/*id*	He drove (somebody).
Invocatory verb marker	*/a*	کناد	kon/*a*/*d*	

Table 1-2 List of inflectional morphemes/affixes in Persian (Kalbasi, 2008; Meshkato-Dini, 2008)

1.2.1.5 Clitics

Clitics are a class of morphemes which are not independently used and attach to antecedent or subsequent words but unlike the affixes, they are not included in the structure of the word (so the syntactic role of the attached word is not assigned to them). They refer to specific syntactic or structural relations. These relations are described in table 1-3. Those assigned as an antecedent of the word are called proclitics; those that come after a word are enclitics. All Persian clitics are of the latter type (Kalbasi, 2008; Meshkato-Dini, 2008). A complete list of Persian clitics is shown in table 1-3.

Type	Transcription	Persian example	Example transcription	English meaning
Present inflectional forms of *to be*. They are also called *enclitic verbs* or *copulative verbs.*	*/æm, /i, /æst, /im, /id, /ænd*	خوبید. معلمیم.	xub/*id*. moælem/*im*.	You *are* good. We *are* teachers.
Dependent personal pronoun (Inseparable pronouns): a. accusative b. possessive c. Prepositional Phrase (PP) complement	*/æm, /æt, /æʃ, /eman, /etan, /eʃan*	a. بردش. b. دستم c. برایشان	a. bord/*eʃ* b. dæst/*æm* c. bæray/*eʃan*	a. (Somebody) took *him/her*. b. *My* hand c. For *them*
/e-Ezafeh (addition) or genitive sign: a. Nominal or possessive b. Adjective	/e	a. کتابِ پارسا b. کتابِ سودمند	a. ketab/*e* Parsa b. ketab/*e* sudmænd	a. Parsa's book b. Useful book
Indefinite noun marker	*/i*	دختری	doxtær/*i*	*A* girl
Definite noun marker, only in colloquial Persian	*/e*	اسبه	æsb/*e*	*That* horse
Vocative particles	*/a*	خدایا	xoday/*a*	*O'*God

Table 1-3 List of Persian clitics (Kalbasi, 2008; Meshkato-Dini, 2008)

According to the types of clitics mentioned in table 1-3, different syntactic roles can be assigned to clitics in Persian. Full linguistic explanations are beyond the scope of this thesis, however, some of the syntactic roles attributed to several of the clitics mentioned above are as follows: (a) distributionally, enclitics always appear after inflectional morphemes, which acknowledges that they are not a part of the stem they attach to; (b) any exchange in the place of clitics in Persian leads to malformation in syntactic units, as does the simultaneous appearance of enclitics and the words (noun or verb) that they

are representative of, e.g. if a possessive phrase includes possessive enclitics then it cannot also be accompanied by possessive separable pronouns; (c) multiple clitics within 'verbal' or 'noun' categories cannot simultaneously appear in the syntactic units, i.e. possessive enclitics and definite/indefinite markers are incongruent because they both belong to the category of noun enclitics (Eslami & Alizadeh-Lemjiri, 2009; Rasekh-Mahand, 2010).

Considering the above-mentioned features of clitics, it can be concluded that in Persian, enclitics are bound morphemes that are phonologically and morphologically included in the structure of the preceding word but syntactically can be treated as separate words rather than affixes (Carstairs-McCarthy, 1992 pp. 91, 142; Eslami & Alizadeh-Lemjiri, 2009).

1.2.1.6 Verb specification in Persian

Morphologically, the verb in standard Persian refers to a word containing syntactic and morphologic components including the verb root, negation morpheme, passive and causative morphemes, auxiliary verbs and non-verbal components. The verb conveys meaning in terms of tense, aspect, mood, person and number agreement, and transitivity/intransitivity (Mohammad Ebrahimi Jahromi & Haghshenas, 2004). A simple diagram of verb structure in Persian is shown in figure 1-1. Only those verb roots with *tense* + feature are able to be inflected with person and number features. The related morphemes are added to the end of past or present stems as bound suffix morphemes and make a *person and number-inflected* verb (the *person and number agreement* feature) (Mohammad Ebrahimi Jahromi & Haghshenas, 2004).

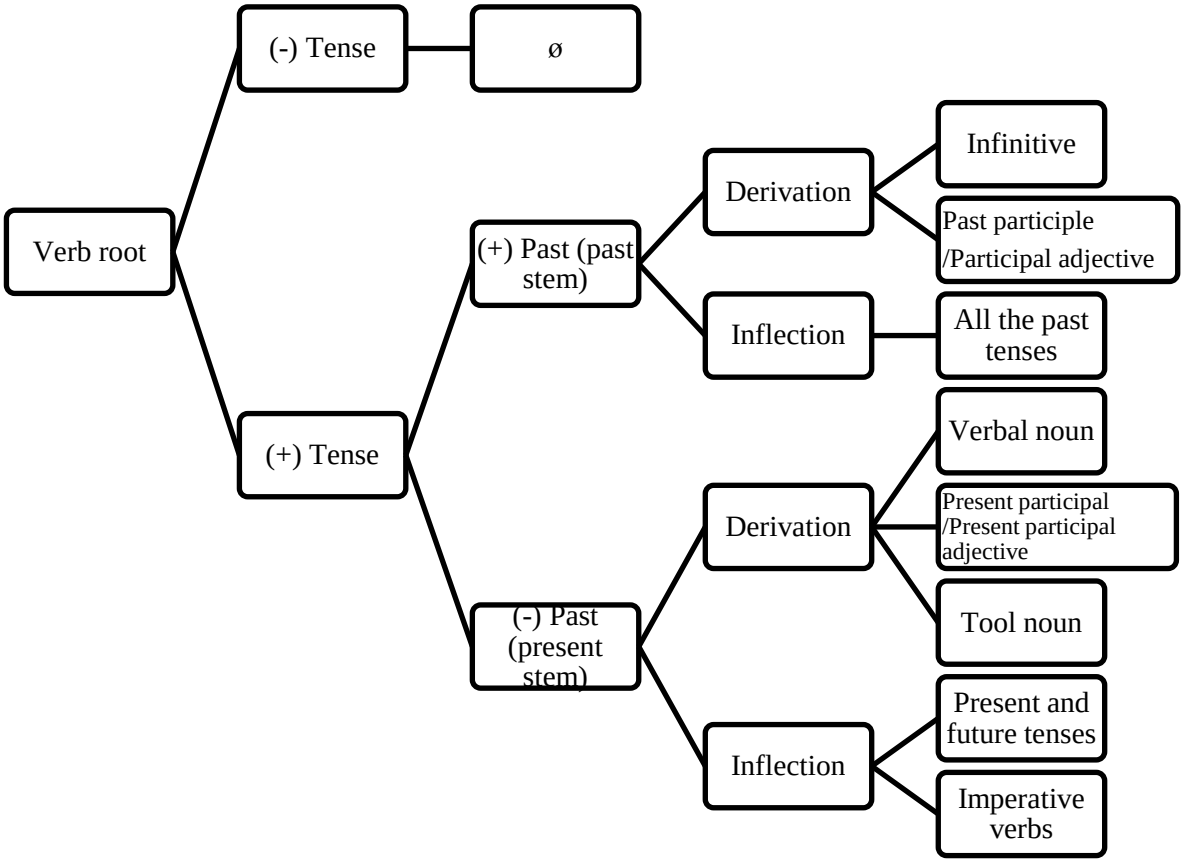


Figure 1-1 Verb root and tense interaction in Persian (Mohammad Ebrahimi Jahromi & Haghshenas, 2004)

In standard Persian, *aspect* can be categorised into two types, either in terms of *ended action*: (a) perfect aspect and (b) imperfect aspect, or in terms of *continuity of the act*: (a) continuous aspect and (b) simple aspect.

Another fundamental feature of the verb in Persian is transitivity/intransitivity. Given a spectrum for transitivity, it can be claimed that some Persian verbs are *more transitive* than others because they have more transitive features in their sentence structures; this is contrary to the traditional Persian grammar in which verbs are either transitive or intransitive or two-phased. In either definition of transitivity in Persian, transitive verbs need objects, sometimes two objects in which case the first object is a Noun Phrase (NP) (noun + *ra* (specific direct object marker in Persian)) and the second is a Prepositional Phrase (PP) complement (non-specific indirect object). This group of verbs is called *two-object* verbs and if either object is missing in a sentence with these verbs, the sentence will be ungrammatical (Mohammad Ebrahimi Jahromi & Haghshenas, 2004).

Negation morphemes in Persian (*ne* or *næ*) usually attach to the beginning of a verb as a bound prefix morpheme and add a negative meaning to the verb; *ne* precedes the prefix *mi*, providing a negative progressive tense, and *næ* attaches to the simple tense (present or past) (Meshkato-Dini, 2008; Mohammad Ebrahimi Jahromi & Haghshenas, 2004).

Auxiliary verbs in Persian are categorised into two types: (a) those without inflectional capability (eg. *bayæd=must*) which are the same for all tenses, persons and numbers, (b) those with inflectional capability (eg. *xastæn=need*) which are inflected according to tense, person and number (Meshkato-Dini, 2008; Mohammad Ebrahimi Jahromi & Haghshenas, 2004). Some of these auxiliary verbs are modals: *bayestæn (=must),* *tævanestæn (=can),* and *ʃodæn (=become).* Three modalities are demonstrated by these modals: epistemic, deontic, and dynamic, with two levels of modalities, obligation and possibility (Akhlaghi, 2008).

Passivisation in Persian is accomplished through the use of three auxiliary verbs (*ʃodæn, gærdidæn,* and *gæftæn,* all with the meaning of *become),* the first of which is the most common. None are used as the main verb but they are inflected after the main verb, which is inflected as a past participle. Third person plural verbs can also be used for passivisation (see example 1-4) (Mohammad Ebrahimi Jahromi & Haghshenas, 2004).

Example 1-4

...می گویند→*mi\guy/ænd*→progressive verb prefix\tell (present root)/third person plural verb marker (they)→ They say ...

1.2.2 Language typology

In this section, Persian will be reviewed according to two existing categories in the linguistic typology of languages: (a) morphological typology, and (b) morpho-syntactic alignment or word order. This categorical view can help explain some Persian-specific features.

1.2.2.1 Morphological typology

From the viewpoint of morphological typology, languages are categorised into two general types: mono-morphemic and poly-morphemic. The words in mono-morphemic languages include only one morpheme and the syntactic associations are specified by word order. Chinese and Vietnamese are two examples of these languages (Kalbasi, 2008). Poly-morphemic languages, conversely, have a word structure consisting of one

or more morphemes. In the classical typology, they can be categorised into three groups: agglutinative, synthetic or inflecting-fusional, and polysynthetic (Dressler, 2010; Kalbasi, 2008). According to this typology, examples of each type are: (a) Turkish, the most obvious example of an agglutinative language, in which the borders between the morphemes are obvious and each one corresponds to one concept; (b) Arabic and Latin, two classical examples of inflecting-fusional languages, in which the morpheme borders are not obvious and the correspondence between morphemes and concepts is not clear; and (c) Yupic and Chukchi, examples of polysynthetic language, in which the border between words and verbs are unclear and many concepts are transmitted by inserting numerous morphemes as one word (Kalbasi, 2008).

Some linguists, however, argue that functionally, languages cannot be classified according to these idealised types (Dressler, 2010; Kalbasi, 2008). They believe that in reality, languages might behave diversely not only in their inflectional morphology and word formation but also in sub-components and sub-modules of their inflectional morphology. These features could be affected by their real behaviour (Dressler, 2010). For instance, Turkish, Persian and English may be typologically distinct because their noun inflections and verb inflections behave differently within the same language (Dressler, 2010; Kalbasi, 2008). We can see that in Romance languages which are derived from Latin, verb and noun morphology is closer to the isolating type, unlike Latin, which is an inflecting-fusional type (Dressler, 2010). Persian, according to the definition above, appears as an agglutinative type in noun morphology and inflecting-fusional in verb morphology (Kalbasi, 2008).

According to Dressler's typological criteria for a language morphological system (Dressler, 2010), Persian morphological specifications are:

a) Morphological richness: Noun morphological markers in Persian generally include case, the direct object (DO) marker *ra* (which is a controversial topic among linguists, Kalbasi, 2008; Karimi, 1996; Meshkato-Dini, 2008), and number, with a plural marker with two forms: */ha*, */an* (Meshkato-Dini, 2008). In additions to these, definition articles include indefinite article */i* which marks unknown, indefinite nouns, and in colloquial Persian, definite article */e* that specifies definite nouns. All other noun morphological features are unmarked. Tense and aspect, in addition to subject-verb agreement or number are marked in verb morphology but there is no grammatical gender marked for verbs, nouns or pronouns (Megerdoomian & MansouriRad, 2000).

The total richness count (the sum of aforementioned markers) for Persian is three for nouns and three for verbs. Dressler considers this type of language as morphologically rich (Dressler, 2010). This leads to Persian approaching an inflecting-fusional language type.

b) Morphotactic transparency: the "morphological pattern" of a given language is transparent if there is a "one-to-one correspondence" between each morpheme category and its representation (Dressler, 2010 p. 112). Example 1-5a shows the formation of a plural noun out of a single noun (*bæʧe*) and a plural marker (*/ha*) with a one-to-one correspondence between the elements. Example 1-5b also indicates an example in verb formation with the same specification between each morpheme and its representation (see example 1-5).

Example 1-5 a) بچه/ها(plural noun)
bæʧe/ha
noun/plural marker
kid/s

b) نگذاشتم (past verb)
næ\gozaʃt/æm
negative morpheme\put (past stem)/first person singular (I)
I did not put.

Persian can be considered transparent since all six noun and verb morphological patterns can be readily identified in their structures. There are some exceptions, however, for example in borrowed nouns from Arabic in which there is no clear distinction between the noun root and the plural marker (see example 1-6). There are Persian equivalences for these plural nouns, though, so they can be excluded from the main Persian morphological system repertoire.

Example 1-6 کتب
kotob
Plural noun
Books

c) Constructional iconicity: this criterion represents the extent to which the morpheme list of a language includes marked forms versus marked meaning. Dressler calls these concepts "morphotactic markedness" and "morpho-semantic markedness", respectively (Dressler, 2010 p.112). Given an example in Persian, plural nouns are all marked (all singular nouns carry */ha* or */an* at the end to construct plural: *bæʧe/ha* = kid/s) in contrast to the unmarked form of a noun

with plural meaning (*dæfater*=notebooks). This feature ascribes Persian approaches to an inflecting-fusional type.

1.2.2.2 Word order

Moghaddam (2001) proposes two hypotheses for the instability of Persian language performance with respect to word order: (a) it might be a historical syntactic change started in Old Persian, continuing through Middle Persian and into modern Persian, from a predomination of SOV as a basic order in old Persian to SVO in spoken modern Persian; (b) it is fundamentally free word-order. She assumes that the first hypothesis can accommodate modern Persian better although there is a strong need for more investigation into Persian word order.

Comrie (1989) also argues that considering a basic word order category for languages with flexible word order is pointless. With this in mind, assigning Persian a specific category of basic word order is very difficult because its behaviour in the spoken mode is different pragmatically and semantically in various contexts, being more likely to behave as "SOV order when the object is phrasal and SVO when the object is a clausal complement" (Moghaddam, 2001 p.17). Moreover, it should be noted that free-word order (or *scrambling* as defined by Karimi, 2003) in Persian is mostly presented in sentences containing direct, specific objects, while the indirect, non-specific objects or counterparts are restricted to preceding relatively attached to the main verb (Karimi, 2003) (see example 1-7).

Example 1-7 من سیبه را با نمک خوردم.
mæn sib/e ra ba næmæk xordæm. (S DO IO V)
pronoun noun/identifier *DO marker* PP complement verb.
I apple/the *ra* with salt ate.
I ate the apple with salt.

In colloquial Persian, however, some forms of scrambling can be found in sentences with a non-specific object. In this case, it is the place of stress that determines the interpretation (Karimi, 2003). The resultant sentence might not be well-formed due to improper word alignment but it is meaningful if it contains proper stress.

On the other hand, free word-order in sentences with a specific object is highly flexible but still depends on the speaker's communicative intent. The main stress on any part of the argument structure can affect word order in the sentence (Karimi, 2003) (see example 1-8). The examples below are based on a given basic word order of a transitive sentence including only a direct, specific object.

Example 1-8 من سیبه را خوردم.

1. *mæn sib/e ra xordæm*. (SOV): declarative sentence with equal stress on all words.
2. *mæn xordæm sib/e ra*. (SVO): stress on either the agent (*mæn*) or the act (*xordæm*)
3. *sib/e ra xordæm mæn*. (OVS): stress on either the object (*sib/e)* or the act (*xordæm*)
4. *sib/e ra mæn xordæm*. (OSV): declarative sentence with equal stress on all words.
5. *xordæm mæn sib/e ra.* (VSO): stress on the act (*xordæm*).
6. *xordæm sib/e ra mæn*. (VOS): stress on the act (*xordæm*).

I ate *the* apple.

Considering the placement of stress, all the examples above are grammatical, although the stress in the last two sentences is restricted to the verbs. Consequently, in Persian, discourse function, showing in speaker's communicative intent, plays an integral role in scrambling.

Although word order is a way of classifying languages, the grammatical relations between the major components of a sentence (subject, object, verb) can also be marked by two grammatical devices: word order and case-marking inflections (Akhtar, 1999). It is hypothesised that languages with a case-marking feature within their morphology are learnt easier than those using word-order specification to show grammatical relations (O'Shannessy, 2011). This feature will be further discussed (section 1.3.8) with regard to child language impairment to seek its potential impact on explaining some grammatical problems observed in primary language impairment (PLI, defined as primary language impairment with no identified origin, which will be fully explained in section 1.5.1).

At this point we turn to a detailed look at the sub-modules of the above-mentioned features in Persian with respect to their role in pathological language, particularly as this is the first time that the vulnerability of Persian inflectional systems or forms in child language is going to be studied systematically. Several of the following sub-modules are used for explaining language impairment in children generally; however, their linguistic connection to Persian has granted their places in the following section.

1.3 Accounts of language development and impairment relevant to Persian-speaking children

Cross-linguistic research on languages other than Persian has shown that there are some factors associated with inflectional morphology which might account for language problems (Penke, 2008). The following lists these factors in association with those accounts which support their role in explaining primary language impairment.

1.3.1 Rich morphology hypothesis (RMH)

Richness and regularity of a morphological system can benefit children with language problems in such a manner that they are able to overcome their limited processing capacity by acquiring the regulations that govern the appearance of morphological features in terms of their frequency (Dressler, 2012). Such factors as "relative frequency, redundancy, regularity, perceptual salience, and pronounceability" are assumed to affect children's language development profiles among different languages so that those languages with less characteristics within each factor are more vulnerable (Leonard, 1998 p. 246). In this way, children with language impairment who speak languages with rich morphological systems and more regular grammar (e.g. Italian, Turkish, and Hebrew) will perform more accurately than those speaking in a less regular language with poor morphological repertoire (e.g. English or French) (Foroodi-Nejad, 2011). This notion is contrary to the feature of Regularity (Penke, 2008) in which non-regular forms are believed to be produced more accurately among children with language impairment. This feature will be discussed shortly.

1.3.2 Morphological complexity

The quantity and nature of the errors in a given language might be affected by the complexity and significance of the inflectional systems within it that are mainly language-specific. These errors are also very sensitive to whether uninflected arrangements are acceptable in a given language (Penke, 2008). Cross-linguistic research has shown that in languages with a rich inventory of morphological forms in each category, the error pattern indicates low omission and more substitution in a given category, e.g. verb affixes in Persian (table 1-2) (Bates, Friederici, & Wulfeck, 1987; Dromi, Leonard, Adam, & Zadunaisky-Ehrlich, 1999; Kunnari et al., 2011; Nilipour & Paradis, 1995).

1.3.3 Type of inflection within different word classes

Researchers assume that different brain activities are responsible for different inflectional types which attribute to particular parts of speech (noun, verb, adjective etc.) or morphological processes (agreement feature, plural making etc.). This differentiation can be found in individuals' *optional selection* of whether to apply them in their language so that some morpho-syntactic inflections are affected whilst the other(s) might remain intact (Clahsen, Bartke, & Göllner, 1997; Laiacona & Caramazza, 2004; Rice, Wexler, & Cleave, 1995). The account of Optional Infinitive to explain children's errors in replacing finite verbs with infinitives hypothesises that this replacement is observed in both children with and without primary language impairment; however, expressions of children with PLI show this feature for an extended period, referred to as *Extended Optional Infinitive* (EOI) (Rice et al., 1995). This account will be further addressed in the discussion section of chapter 4 on the diagnostic accuracy study.

1.3.4 Regularity and rate of appearance

It has been hypothesised that inflectional morphemes can be affected diversely by different language disorders according to the regularity of the morphological process involved or their frequency of appearance (Penke, 2008). Irregular forms are more likely to be correctly produced compared to regular forms since they are stored in the mental lexicon rather than as an abstract concept or symbolic sign of a syntactic role. It has also been assumed that "the frequency effect" is true mostly for regular inflected forms whilst infrequent irregular forms "are stored as fully inflected whole word forms in the mental lexicon" (Penke, 2008, sec. 13.1.4, para. 1).

Frequency of the affixes that appear in different morpho-syntactic structures can be affected by language impairments. The more frequent the inflectional form, the less the impairment can affect it which is explained by the impact of stronger memory tracing for more frequent items in "lexicon storage and access". "Memory traces get stronger with each exposure, making frequently occurring forms easier and quicker to access than infrequent ones" (Penke, 2008, section 13.1.4, para. 1).This can also be explained by the direct activation account that morpho-syntactic strings are not really decoded but stored and recalled as whole units (Rispoli, Hadley, & Holt, 2012).

> "Direct activation is thought to arise from high-frequency word and morpheme combinations in input and use (Bybee, 2006). Both fixed, unanalysed multi-morphemic combinations and limited scope formulae with invariant

> lexical frames are produced by direct activation. For some children, limited scope formulae, such as I want x, can become conversational crutches that they rely on for a sizeable percentage of their spontaneous multi-word production" (Rispoli et al., 2012 p. 1009).

This assumption, nevertheless, has not been tested in languages that have no regularity features in their morpho-syntactic structure (Penke, 2008) or have very few irregular morpho-syntactic constructions, and needs to be considered cautiously in unstudied languages like Persian.

With respect to the current study, Persian-speaking children might show fewer errors in more frequent morphemes as well as those combinations of words and morphemes that are tied together and stored as whole strings (Example 1-9). The former assumption will be tested in this thesis whilst the latter needs a more specific methodology to be scrutinised.

Example 1-9 اینـاش or اینـاهاش or اینسـاش (all with equal meaning)
inaʃ or *inahaʃ* or *inesaʃ*
This is here (Here it is).

This example is transcribed as a whole unit throughout the language sample transcriptions and will be considered as one morpheme but will not be linguistically analysed due to not being in the scope of the current thesis.

1.3.5 Markedness

According to a definition provided by Penke (2008), inflectional affixes are represented in a binary coding system of unmarked (negative) and marked (positive) states (Penke, 2008). The markedness system in the morpho-syntactic structure of a given language might incline individuals with language impairments to replace some complicated inflected forms with unmarked uninflected forms. This replacement is potentially affected by phonological arrangement of the inflected word and the affix, i.e. affixes which are pronounced similarly to the word stems (in terms of place or manner of articulation) are vulnerable to being omitted or replaced by more easily-pronounced forms (Penke, 2008) which is worth further research cross-linguistically.

1.3.6 Pro-drop parameter

In pro-drop or null subject languages (for instance, Spanish, Hungarian, Italian, Turkish, Persian, British Sign Language, Hebrew and so on), the speaker can optionally omit the

pronoun subjects. In some of them (e.g. Spanish) the free inversion of subject and object has also been documented (Gass, 1989). The typological differences related to this parameter have divided languages into pro-drop and non-pro-drop languages. In the former type, the subject does not need to be overtly expressed although it can inflectionally appear as a part of the verb. According to Huang and Li (1996), "the richness of agreement marking of finite verbs" can be thought of as the credible linguistic difference between English-type and Spanish-type languages (Huang & Li, 1996 p. 79). It seems that the rich verbal morphology of pro-drop languages allows speakers to cover their full intended message by including the subject in the inflected verb (e.g. in terms of the person or gender) so the omission of pronouns cannot be considered problematic. With regard to Persian as a pro-drop language, this parameter can be found along with an inflectional agreement feature so that, except for imperative sentences, a clause must include an overt agreement structure if it is pronoun empty. This generalisation is called conditioning of pro-dropping by rich agreement (Neeleman & Szendroi, 2005). Most studies on Iranian language development have shown that children show understanding of agreement in their productions from the very early stages of talking (Jalilevand, Ebrahimipur, & Purqarib, 2012; Meshkato-Dini, 2001; Kazemi et al., 2012). It is worth noting that the study of pro-drop languages like Persian can lead to novel viewpoints of the language acquisition process (Samadi & Perkins, 1998).

1.3.7 Case marking

Case marking can be a way of showing grammatical relations in a given language along with word-order and prepositions (O'Shannessy, 2011; Foroodi-Nejad, 2011). Blake (2001) defines case markers as signposts of how the nouns they deliver relate to the sentence head (verb) in the clause level. In Persian there is only one case marker, DO marker *ra*, and its exact morphological category is controversial (Karimi, 1996; for a list of different definitions of *ra* see Shokouhia & Kipkab, 2003), probably because its role is irregular and sparse (Foroodi-Nejad, 2011 p. 39). The irregularity refers to restricted use of *ra* merely with definite specific direct objects and sparseness signifies its uniqueness in the repertoire of case marking in Persian. Although Karimi considers the Persian overt case marker *ra* as a *(noun) suffix* (Karimi, 1996), it was not found in the categorisation of different suffixes mentioned by other Iranian linguists (Kalbasi, 2008; Meshkato-Dini, 2008). As shown in table 1-1, *ra* is among grammatical morphemes or functional words and this is the main difference between the different linguists' views.

It means that we can assume *ra* is an independent word instead of being a suffix, which will affect quantitative measures of language such as calculating the number of morphemes or words in an utterance.

Regardless of this debate, because Foroodi-Nejad is the only one who investigates *ra* in children with language disorders, referring to her research would be helpful to review the studies on case marking in other languages. In her review of case marking in different languages, Foroodi-Nejad indicates that difficulty in case marking is common across such languages as Hebrew, Turkish, Japanese, and Hungarian, with an exceptional report in German in which both children with and without SLI showed high correct usage of accusative case markers (as stated in Foroodi-Nejad, 2011). Finally, she concludes that the problem children with SLI have in accusative case marking is common across languages, with some exceptions, and that it needs more investigation cross-linguistically. The results of her own study showed that two groups of Persian-speaking children with and without SLI did not have significant difference in terms of percentage of case marking usage (Foroodi-Nejad, 2011). Her study has been further reviewed in chapter 2.

Having all these in mind, we can expect that Persian-speaking children with and without PLI might show differences in using *ra* regardless of its grammatical morphology classification, because of its flexible and irregular appearance. However, it also assists children with language disorders because it signposts the object in a transitive sentence; they might therefore show no problem in transferring their knowledge of word-order through case marking and/or their capability in using *ra* accurately after direct objects. This twofold picture of *ra* will be investigated by examining different errors in the diagnostic accuracy phase of the current study, along with all other grammatical error markers mentioned in this section.

1.3.8 Word order

As mentioned in section 1.3.7, another device for judging word alignment in a sentence is case marking. Judging Persian as having one, and only one, case-marker (*ra* for direct objects), we would expect Persian-speaking children to have the least problem with transitive sentences (which require this marking) with respect to word order. They should show that they have learnt that Persian is a head-first language (subjects come first) and put the verb after the object. However, with flexible word order in Persian transitive sentences, as indicated before, there is a need for more investigation to verify this argument. This position has been challenged by Akhtar (1999), who indicated that

children learn the word order of their language during the early years of their life and it is not affected if they are faced with other types of word order later. Akhtar manipulated word order by creating unknown verbs and including them within word alignments which were new to the children. The results showed that children in their second, third and fourth year of life indicated using one basic word order for all novel transitive structures. Akhtar proposed that two younger groups of children (two and three year-olds) were "in the process of constructing a truly general knowledge of the syntactic significance of word order (which means that all English sentences must employ SVO order)" whereas older children in their fourth year of life "were simply not willing to use the non-SVO structures" (p. 354). Although the younger children were able to use familiar verbs in the right order in sentences, they followed the examiner's probe when producing novel verbs, i.e. if the probe included a different word order, the children repeated that. Hence, a gradual development of word order through the early years of life was suggested, which starts with dependency on specific familiar exemplars of lexical items in younger ages (around the second year of life) until mastery of generalising frequently-heard exemplars of word order to novel situations is achieved (around fourth year of life). Abbot-Smith, Lieven & Tomasello (2001) replicated Akhtar's study and concluded that the trajectory of word order concept in English-speaking children shows a developmental pathway from lexical dependency below age two to a general verb knowledge in children of four to five years of age. In a more recent study (Aljenaie & Farghal, 2009), Kuwaiti children as young as four to five years old also showed a preference for SVO word order whilst six to eight year-olds preferred both SVO and VSO patterns. Word orders in colloquial modern standard Arabic are SVO and VSO. The study concluded that younger children were more semantically-dependent on word-order comprehension, an observation which was also seen in Hungarian children (Babarczy, 2006).

Looking at word order from a psycholinguistic point of view, Hawkins (1983), as cited by Pantcheva (2007), argues that the earlier the lexico-semantic information is placed in an utterance, the more accurately it can be processed. Children's word-order preferences differ from adults', and there are contradictory interpretations of their behaviour. Three principal preferences have been reported as: (a)"new-before-old", (b) "old-before-new" and (c) "no significance" (Narasimhan & Dimroth, 2008 p. 318). 'New' refers to the information or input status relative to the child and shows whether the child prefers to produce new terms before or after old ones. Another account for word order preference in children relates to the accessibility of the referents in the information they are going

to transfer. The distinction between *topic* and *comment* or, in another explanation, *subject* and *object* tries to clarify whether the children prefer to produce *what the utterance is about* prior to *what our prediction of the topic is* or vice versa. Topic or subject is already available to the speaker and listener whilst comment or object is in a predicative condition (Narasimhan & Dimroth, 2008). This may help justify inconsistency in the word order skills of Persian-speaking children, considering that they are speaking in a relatively free word-order which is difficult to analyse according to traditional accounts for word-order problems (such as Akhtar's idea). Regardless of the nature of justification, any research on word-order in Persian-speaking children is entirely new, and this field of study requires extensive investigation with regard to the nature of Persian word order and the problems it causes for children with language impairment.

1.3.9 SLI critical mass and surface hypothesis

A hypothesis by Jones and Conti-Ramsden (1997) suggests that children with SLI do not produce morphologically-appropriate verbs, and in fact do not understand any grammatical category of verbs, unless they have enough exemplars of a certain morpho-syntactic feature of the verbs. This is referred to as *SLI critical mass*. The critical mass hypothesis applies to children with typically developing language (TDL) as well, describing the point when they have access to enough examples of a verbal inflection pattern to recognise, produce and generalise those patterns to a broader range of their grammatical knowledge (Conti-Ramsden & Windfuhr 2002 p.19).

However, from a clinical point of view, it has been discussed that the perception of unstressed words or morphemes in the sentence is difficult for children with PLI. As a result, children with PLI will not be able to produce them due to the restricted access to enough representational examples of those morphological structures in their repertoire of morphemes (Fey, 1986; Weismer & Robertson, 2006). This notion is widely known as the *Surface Hypothesis* (account) which argues that unstressed closed-class morphemes are the main categories that might be affected in this way (Leonard, Caselli, Bortolini, McGregor, & Sabbadini, 1992; Rice, 1994). With regard to Persian, it can be expected that closed-class morphemes or unstressed words and morphemes (e.g. direct object marker *ra*, clitics, or PP complements [non-specific indirect objects]) would be more vulnerable to being omitted by children with PLI.

Another implication of paying attention to language types can be found in Dressler's claim that children learning inflecting-fusional type languages over-generalise

morphological patterns that are prone to grammatical manipulation, after they abstract them in their language, and then they may over-generalise those patterns that are "unproductive but more transparent". He also discusses the idea that "due to the recency effect, ends of words are easier to segment than beginnings which appears to be also the main reason for the general suffixing preference" across languages and considers this as the reason for earlier development of suffixes than prefixes (Dressler, 2010 p.119).
The above-mentioned accounts for morpho-syntactic problems were based on the observation of individuals with aphasia and SLI/PLI in languages other than Persian, particularly in English. Accordingly, they need to be verified in terms of compatibility with Persian linguistic features through performing precise investigations.
The next section will summarise the history of speech therapy and its educational system in Iran with a particular look at child language as well as the referral system.

1.4 Speech therapy in Iran

It is notable that the first Iranian scientific view of the communication mechanism and its disorders dates back to the later part of AD 900 when Avicenna introduced some aspects of today's phonetic theory in his paper called "Treatise on Arabic Phonetics" (Solomon, 2009). Avicenna was an Iranian philosopher and scientist as well as a physician whose two famous books, *Ketab-e Al-shafa* in philosophy, law, mathematics and natural sciences, and *Ketab-e Ghanoun* or *Codes of Laws* in medicine, played a distinguished role in establishing human knowledge at the time. A translation of the latter book into Latin in the 12th century was used as a textbook in many universities throughout Europe (Solomon, 2009; Tschanz, 2003). In his book of phonetics, Avicenna introduces the mechanism of articulation by showing the role of each facial muscle as well as demonstrating the *place* and *manner* of articulation for each Arabic phoneme, which are mostly identical to those of Persian, and the acoustics related to them. This paper could be considered as a revolution in the systematic study of phonetics (Solomon, 2009).
In the later chapters of his book in medicine, Avicenna introduces different communication problems like stuttering, articulation disorders and mutism and suggests some intervention for each of them according to his traditional experience in medicine.

1.4.1 Contemporary speech therapy in Iran

There is no evidence of the study of communication development and disorders between Avicenna and today's Iran, although the Baghcheban school for hearing impaired

children, founded in 1924 by Jabbar Baghcheban (1884-1966), was the first formal academic setting in Iran for the education of children with hearing impairment. He is also known as the father of Iranian sign language, *Baghchehban's sign language* (*Jabbar Baghcheban, Iran's sign language pioneer, remembered*, 2007). Since his death, this school has remained the most well-known school for children with hearing impairment.

What we know as *speech therapy* in Iran was established in 1974 in the College of Rehabilitation and Social Welfare under the supervision of several British scholars. This college trained the first speech therapists to bachelor degree level and soon after, in 1975, the college upgraded to the Faculty of Rehabilitation and Social Welfare (Nilipour, 2002). After Iran's revolution in 1978, the universities closed for a while and when they re-opened, this faculty changed its name to the Faculty of Rehabilitation Sciences and was joined to the Iran University of Medical Sciences. This university was the only university in Iran training professionals in rehabilitation sciences including speech therapy, physiotherapy, occupational therapy, audiology, optometry and management in rehabilitation up to 1992, when the University of Welfare and Rehabilitation Sciences was established and started recruiting students in all the above-mentioned courses as an independent university for rehabilitation sciences. The age of reconstruction and recovery after the Iran-Iraq war (1980-1988) meant that more specialists were needed in rehabilitation including speech therapists. Consequently, the Ministry of Health and Medical Education changed the policies related to rehabilitation, and faculties of rehabilitation were established at other medical universities in big cities around Iran including Tehran, Isfahan, Semnan, Ahvaz, Mashad, Tabriz, Shiraz, Zahedan, and Hamedan. All these faculties have courses in speech therapy at bachelor level. Speech therapy at master's level started to be taught in 1993 in five medical universities. The first PhD students in speech therapy began their studies in 2008 at the University of Welfare and Rehabilitation Sciences in Tehran followed by Tehran University of Medical Sciences. The first speech therapy PhD graduated from the former in January 2012. Throughout these years, however, there have been speech therapists who have moved to other countries pursuing high-quality higher education in their field of professional interest; only one of them returned and established themselves at PhD level in Iran: a speech therapist who graduated from Surrey University in 2008.

1.4.2 Child language study in Iranian universities

Child language is taught in several fields of education in Iran including linguistics, educational psychology, psychology of exceptional children, and development of Persian language; speech therapists also study impairment or disorders in child language. They are trained according to the criteria that the Ministry of Health and Medical Education has approved to meet the special needs of this group of children. Speech therapists have to pass courses in assessment and intervention in speech, language and swallowing disorders in children and adults for four years at undergraduate level. They can choose further education in a two-year course at master's level and then a four-year course at PhD level. Apprenticeship courses are included in all levels but they are essential at undergraduate level to qualify the graduates as a 'therapist' or 'clinician', and all graduates are called speech therapists. None of the other professionals in linguistics and psychology are allowed by the Ministry of Health and Medical Education to intervene in communication disorders. Moreover, speech therapists may register with the Medical Council of Iran, although registration is not compulsory nor a legal requirement. This council has the most comprehensive data bank related to the medical and paramedical professions in Iran and is referred to later in the current project.

1.4.3 Referral to speech therapy Services in Iran

A wide range of people can refer a child with a communication problem to speech therapy. Among the most well-known referral sources are: parents, teachers or care-givers, general practitioners, paediatricians, otolaryngologists and psychologists. Until 2003 speech therapists were supervised by other rehabilitation specialists (e.g. physiotherapists) or physicians (e.g. otolaryngologists). Then, they were permitted to have independent offices for their work after some academic negotiations between the speech therapist and the ministry. The vast majority of Iranian speech therapists now work in private clinics, including specialist rehabilitation centres. Other workplaces for speech therapists include universities and university hospitals, schools for children with special educational needs as well as home visiting.

At times, speech therapists refer their clients to other professionals to ask their advice or to help them decide the proper course of intervention. The increasing attention to establishing clinical judgments based on the best evidence provided by different sources has encouraged Iranian SLTs to grant more rational referrals with the aim of better achievement in child language intervention. Appropriate referrals based on the best

evidence along with concerns which arise from the clients and their parents or family (Dollaghan, 2007) are broadly encouraged in today's academic education.

1.4.4 Insurance services

Two main insurance systems are available in Iran to help people to pay medical costs (1) Medical Services Insurance Organisation and (2) Social Security Organisation. Speech therapy services are partly covered by both and by supplementary private insurance. The main problem with accessing insurance for people with communication disorders is that these disorders are not defined properly for the insurance services so they are unaware of the impact of these impairments on people's lives.

1.5 Methodological framework of the current study and the main aims

The difficulty of accurately assessing the language of young children has always been a concern for researchers. The rapid developmental changes in language during the early years of life, the importance of early intervention and the prevention at first and second levels of communication disorders (i.e., preventing the onset of a disorder and early diagnosis of the disorder, respectively) (American Speech-Language-Hearing Association, 2005) are examples of the complexity involved. Add to this list the incompatibility of using linguistic assessment tools in other languages for children from different cultural and linguistic backgrounds. On the other hand, evidence-based assessment procedures have become more widely used by clinicians and inform researchers to focus on those measures with the best clinical characteristics in their research designs, goals and methodologies. These features should meet clinical needs instead of following the traditional research designs in basic studies and as far as relates to the current study, they should consider a diagnostic point of view in research design and methodology followed by specific statistical frameworks (Dollaghan, 2004).
It can be assumed that in this way, however, using diagnostic accuracy studies (DA) with procedures such as Receiver-Operating Characteristics (ROC) curves (Haynes, Sackett, Guyatt, & Tugwell, 2006) and discriminant function analysis (Gavin, Klee, & Memberino, 1993; Klee, Gavin, & Stokes, 2007) can help to determine the accuracy features and select the best cut-off points for the most promising language sample measure parameters in differentiating language impaired children and normal peers, compared to pre-defined cut-off points in traditional psychometrical procedures.

1.5.1 The organisation of the thesis

The remainder of this chapter will outline the content of each chapter along with the rationale for including any specific framework. Then the detailed aims of the study will be specified.

Chapter 2, 'Research on child language assessment' includes reviews of non-Persian studies and Persian studies relevant to this dissertation. The literature review of the current study will follow the fundamentals of meta-analysis in search of relevant non-Persian studies that comprise similar target population to the current study within the same research design of diagnostic accuracy of language sample measures. Children with primary language impairment (PLI) and their normal developing language (TDL) counterparts are the target population of this study. PLI according to the meaning used in the current study is not a common term among Iranian SLTs. They usually refer to language impairment in children as specific language impairment (SLI) with the exclusionary criteria introduced by Bishop (Bishop, 1992): any disordered language compared to child's chronological age (at least four years old) with no apparent sign or symptom of such associated problems as hearing impairment, cognitive or developmental delay, obvious neurological disorder, motor speech problems, emotional disorders or autistic spectrum disorders. Although in the classic definition of SLI, the child should score more than, for example, 1.5 standard deviations (SDs) below the mean on at least one standardized test, it is not possible at present to apply a benchmark such as this in clinical settings in Iran due to the lack of any standardized language assessment in Persian. The remaining criteria are exclusionary factors in combination with the child's age. SLI is also a relatively new concept in Iranian speech therapy. Its use in academic contexts dates back almost 12 years to when speech therapy started spreading amongst more universities and a new generation of academics started teaching with a new look at the former term *Delayed Speech and Language* (abbreviated *DSL*). DSL included all types of speech and language problems in children with different causes: hearing impairment, mental retardation (a term which is still used in Iran for developmental delay), traumatic brain injury/dysfunction, autism spectrum disorders (ASD), and mutism as well as language disorder with unknown origin (defined as language delay). In some cases SLI was referred to as *childhood or developmental dysphasia* which is still in some use. Primary language impairment (as it is used in this study) is an uncommon term among Iranian SLTs, if in use at all; mainly because it is also relatively new internationally in the field (Law, 2003; Law, Garrett, & Nye, 2003; Tomblin, Zhang, Buckwalter, & O'Brien, 2003) and needs more detailed

interpretation . PLI represents those language specific problems which may be accompanied by deficits in other developmental areas (Thordardottir et al., 2011). It refers to those developmental language problems described by Thordardottir as 'with no definite diagnosis of SLI'. SLI had also been under question for its strict definition which avoided association with other developmental complications; as a result, PLI/SLI are used in the current report to reflect both the Iranian SLTs' use of SLI definition as well as including any uncertainty about potential undetected developmental problems concomitant with language impairment.

The meta-analysis will also include those studies in which at least one language measure derived from language sampling occurs. A full description of language sample measures (LSMs) has been provided in chapter 4 as well as appendix F; however, to provide a brief definition of LSMs, they refer to those measures of language production that quantify a speaker's performance linguistically (either in form or content). Language samples are gathered in any given environment (e.g. natural language produced in a free-play session with mother in the current study) and are fully discussed in chapter 4. The measures of reviewed studies will be defined as they appear in section 2.2 of chapter 2.

The research design of this study employs a paradigm originating in medicine, in which the diagnostic accuracy of clinical measures is examined, with the aim of finding the most predictive LSMs in Persian for differentiating children with PLI from their TDL peers. This method is fully explained towards the end of the current section.

Chapter 3 examines methods used by Iranian SLTs for evaluating child language as determined from a survey and focus groups. While the main part of the current study investigates the diagnostic potential of some quantitative measures derived from language sample analysis, the process of a diagnostic accuracy (DA) study is dependent on specifying impaired and unimpaired children using a valid and reliable assessment or *reference standard* (Sackett & Haynes, 2002). It includes four phases which seek the reliability of an evaluation procedure to differentiate between a population with disorders and their normal peers. This process cannot be completed unless the previously defined diagnostic criteria, the reference standard, has been used in at least one phase of the study (for a complete definition of DA phases, see end of the current section). Mixed qualitative-quantitative research on the criteria which Iranian SLTs usually use to recognize children with language impairment was chosen to gain a reference standard for the later parts of the study, as well as to find out by what means SLTs apply language sampling in their assessment process.

The current study will apply the International Classification of Functioning, Disability and Health (ICF), introduced by the World Health Organisation (WHO), as an organising framework for establishing a consensus as to what constitutes a language impairment, since no commonly accepted reference standard currently exists in Iranian clinical practice. Recently, the ICF has become very popular among researchers and clinicians and it is also going to be incorporated into medical and rehabilitation education in Iran. Considering language development as a system with different components including the child, the environment and other extrinsic factors, provides another category for modelling child language development and disorders (Paul, 2007) in which not only do the children's strengths and weaknesses play a role in language development, but their environment including parents and physical environment may also have a major role in aggravating or improving developmental problems including those related to language. These elements are also involved in the ICF. The impact of environmental factors on communication is inevitable since communication becomes known only within the context in which is carried out. A successful communication intervention should involve the environment in any goal setting, because speech-language therapists are the only educated professionals who study all these factors in depth (Howe, 2008) and consider them in the assessment of child language. Dempsey & Skarakis-Doyle (2010) have shown the interactive relation of different components of ICF in child language impairment, stating that the SLT's role in identifying indirect or nonlinear factors affecting the inefficient use of language by the children must not be underestimated (i.e. in the author's opinion, by only concentrating on the Body Functions and Structures). Addressing other areas is the core property of ICF contribution in its clinical applicability by equally emphasising Activities, Participation and Environmental Factors in the intervention programmes even though their relationship with the other two components might not be linear, which could be seen as the beauty of ICF. A description of the components of ICF is indicated in table 1-4 and figure 1-3 which will be used as a template for integrating the results of the survey on Iranian SLTs.

Components	Description
Body Functions and Structures	Physiological and psychological function of body systems form **Body Functions** and anatomical parts of the body like organs form **Body Structures**. Any deficit or loss in Body Functions and Structures may cause **Impairments**.
Activities and Participation	Physical performance to do a task by an individual forms **Activity** and involvement in everyday life forms **Participation**. Any restriction to do activities causes **Activity Limitations** as well as any difficulty in involvement in everyday life that causes **Participation Restrictions.**
Contextual Factors	Including **Personal Factors** (age, gender, educational and social backgrounds, coping styles, past and present experience, character, and overall behavior), and **Environmental Factors** (such as social attitudes, legal and social structures, architectural features, climate and so on).

Table 1-4 ICF components and related definitions (World Health Organisation, 2002)

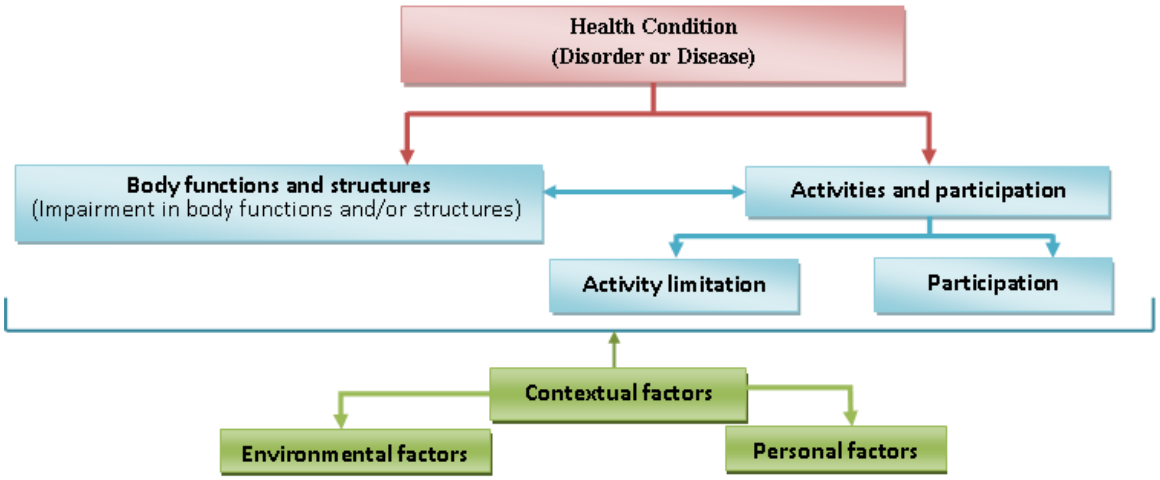


Figure 1-3 Current interaction between different components of ICF (World Health Organisation, 2002)

Chapter 4 is called 'Diagnostic accuracy of some LSMs in Persian'. The necessity of assessment tools on which clinicians can rely when they are making decisions about child language development, and can be certain about their accuracy, is not something that can be ignored. Generally, researchers and clinicians in Iran are likely to encounter situations in which they are not confident that the tests applied are able to identify language impaired children effectively. The problem will be more complicated in the absence of a diagnosed disorder associated with a language problem, mainly known as specific language impairment. Since this term was defined, many issues have arisen with

regard to clinicians' uncertainty about the absence of any kind of accompanying problems with language. As a result researchers have started applying the term primary language impairment unless all criteria for the clinical definition of SLI are met (Thordardottir et al., 2011). This term will help us in this project in two ways: (a) the age range of participants is not within the defined range for SLI (more than four years old) so PLI would be a better description for our population, (b) Although many researchers agree on the definitions for these two terms (SLI and PLI), clinical diagnosis (especially in Iran) suffers from a lack of unified assessment procedure to identify either, and the lack of specific practical definition for these terms in different societies is observed (Thordardottir et al., 2011). This problem appears more crucial in such languages as Persian in which there is no access to language tests. Clinicians must decide merely based upon the interpretation of the observations they make of the child's linguistic behaviour, and associate them to the characteristics of language impairments, while SLTs do not know the accuracy measures of the procedures.

The lack of a reliable and universally acceptable standard for identifying language impairment in children is not something limited to un-studied languages. Even in languages such as English, where many studies have been undertaken to define a reference standard, there is far from a consensus among professionals as to what constitutes language impairment in children (Klee, Wong, Stokes, Fletcher, & Leonard, 2009). The current study uses an EBP approach as a practical framework to try to identify the standards that different studies have set up to find children with LI without known concomitance and differentiate them from typically language developed children. The American Speech-Language-Hearing Association (ASHA) has reported that to make a more cost-effective diagnosis, clinicians would benefit from an EBP approach in their decision making (Klee et al., 2009). The EBP approach, which has been adapted from the term Evidence Based Medicine, has been defined in various ways (for a list of definitions, see Dollaghan, 2007 pp. 1-3), but the aspect that is common among all the definitions is their emphasis on *establishing high-quality research evidence which should be used to guide clinical practice* (Dollaghan, 2007; Klee et al., 2009). The Diagnostic accuracy study is assumed to implement EBP effectively in diagnostic research. Haynes et al. (2006) have introduced a triad for research questions which is also applied in diagnostic accuracy. According to their explanation, the triad of a research question in diagnostic studies includes population, the test under investigation, and the outcome which must be considered in "a valid diagnostic study" (Haynes et al., 2006 pp. 289-290). They have mentioned that the outcome in diagnostic

studies is the *reference*, *gold*, or *criterion* (all synonyms) *standard* which must be defined before conducting the study. A problem will arise when the researcher does not have any standard at the time of the research, similar to the current study. In addition to this triad, they have set up a set of four principles for diagnostic studies called the diagnostic quartet. They believe that studies are accurate only if their data collection is representative of the target population; both the reference standard and diagnostic test are administered to all members of the sample; researchers follow the rules of blinding in interpretation of the results; and finally they are replicable with similar results in a different sample of patients (Haynes et al., 2006).

Usage of diagnostic accuracy studies which have several strong statistical methods for analysing the data can fortify child language impairment diagnostic investigations. These statistical measures are sensitivity, specificity, predictive values and likelihood ratios. The latter two are not dependent on the prevalence of the disorder in the population studied, therefore their role is important in studies with small sample size (Klee et al., 2007). ROC curves are another way of determining the optimal cut-off scores in diagnostic studies (Haynes et al., 2006) and will play a very important role in the current project. These terms have been defined operationally both in section 2.2 of chapter 2 as well as chapter 4. In an attempt to meet the goals of an EBP approach to PLI in Persian-speaking children through the lens of diagnostic accuracy, the four phases of a diagnostic study suggested by Sackett & Haynes (2002) should be followed as described here:

Phase I consists of finding differences between two groups of participants with and without the target problem. Adapting it to the current study, in this phase a group of definitely identified LI children will be compared to a certainly non-LI group of children in terms of LSM scores.

The statistical tests in this phase are simply the tests for comparing the means of the diagnostic test scores.

Phase II reverses the direction of phase I studies, so that the results of the diagnostic test will be used to identify the group to which each child belongs. In the current study, the LSM scores will be analysed blind to recognize the group into which each child can be classified. The same data as in the phase I study can be used in phase II, and the statistical analysis will be based on such outcome measures as sensitivity, specificity and likelihood ratios. These measures are obtainable through ROC analysis or applying the number of impaired and non-impaired cases in a contingency 2×2 table.

Phase III enters another group of participants into the study, a group from the population who are suspected having the impaired condition. To use the example of the current study, in this phase, the same action as in phase II will be performed on the results of children with and without SLI among children in whom it is clinically reasonable to suspect that LI is present. The same outcome measures will be used to analyse the results of this phase and the reference or gold standard will be employed to define children's condition.

Phase IV will be undertaken on the population who receive the diagnostic test during a follow-up intervention programme to find out whether "they fare better than similar patients who are not tested" (Sackett & Haynes, 2002 p. 541).

The current study will involve the first two phases of this process, which will be described more practically in the method section of chapter 4. Recalling the EBP approach to the current study, the following criteria have been used to choose the measures to be studied: (a) general language sample measures which are computable using 'systematic analysis of language transcription' (SALT) (listed in chapter 4) with fair to good accuracy features (Plante & Vance, 1994), according to the results of the meta-analysis of LSM diagnostic accuracy (chapter 2). They must have an equivalent in Persian and be obtainable through a spontaneous language sample collected in the current study. Measures falling outside the age range of the current research have been excluded, as well as those which are impossible to calculate due to the discrepancy between the settings of language sampling and the procedures of the current study (e.g. those which need a structured sampling environment to elicit specific measures such as the emergence of specific morphemes); (b) measures with sensitivity to age, and (c) those in which two experimental groups proved to be meaningfully different via a Phase I DA study.

Chapter 5 is called 'Establishing new measures for evaluating Persian-speaking children with PLI through an evidence-based practice approach: outcomes and suggestions'. The results of the aforementioned studies will be concluded in the final chapter together with the clinical contribution of the outcomes. Theoretical implications will also be discussed within the context of different accounts of PLI, and finally some points of interest will be suggested for future research.

1.5.2 Main aims of the study

This thesis aims to report on the author's attempt to rectify the considerable lack of formal child language assessment methods in Iran by employing well-known methods of test development. Initially, Iranian SLTs would benefit from the results by having access to a set of LSMs that would potentially enhance their confidence to routinely include this method in their assessment repertoire, at least for those who work within the geographical area of the study. A secondary aim would be the significant contribution of the study to cross-linguistic studies on child language assessment, by providing an extra piece of evidence about the much understudied language of Persian with the potential to make a big difference to our judgement about the underlying processes of language impairment. Consequently, the research questions addressed in this study are:

1) What is the current state of child language assessment, particularly in relation to PLI, in Iran? A mixed method framework will be used to examine this question.
2) What are the diagnostic features of a selected set of general language sample measures and Persian-specific measures?
3) What are the most accurate LSMs in Persian with clinical relevance?

A four-phase diagnostic accuracy approach will be applied to answer the second and third research questions.

CHAPTER 2

Research on child language assessment: a meta-analysis

There are few studies considering the assessment of children with PLI who speak Persian as their mother tongue. Hence, this chapter is in two parts: a meta-analysis of non-Persian studies on assessment of children with PLI, followed by a review of Persian studies, classified according to the hierarchy of evidence (Greenhalgh, 2010) and weighed against the criteria of critical appraisals of different types of studies. It ends with an explanation of the clinical procedures available in evaluating child language to justify the method of data collection used in this study.

2.1 The meta-analysis of diagnostic accuracy of language sample measures for identifying primary language impairment in pre-school children

> Systematic review and meta-analysis synthesize external evidence across multiple studies that have been identified and analysed according to explicit and transparent procedures. The emphasis on explicitness and transparency differentiates systematic reviews and meta-analysis from traditional literature reviews, in which the author makes subjective decisions about which information to include and to highlight (Dollaghan, 2007 p. 105).

Meta-analysis statistically quantifies the systematic review results by providing the mean of results of individual studies (Dollaghan, 2007) along with diagnostic accuracy measures such as sensitivity and specificity. Considering that meta-analysis as a research design is incorporated into the evidence-based practice framework, its process begins with a precise research question. The next step is searching through the best and most recent related evidence, followed by a critical appraisal of that evidence (Dollaghan, 2007; Schlosser, Wendt, and Sigafoos, 2007). In addition, it is essential that the target population and controls (clinical conditions), study methods and settings, and outcome measures of the studies included in the meta-analysis are compatible with the research question of the review (or foreground question, FQ, as defined by Dollaghan, 2007). Another point of strength in meta-analysis comes from its powerful statistical methods; it provides the possibility of summarizing the outcomes of the most homogenous studies by combining their small sample sizes, reinforcing the power of the outcomes (Haynes et al., 2006 p. 36).

Since the present study is employing the EBP framework throughout, systematic reviews and meta-analyses were chosen as the frames for reviewing the subject-relevant literature. Most of the diagnostic language sample measures in the field of identifying child language impairment are researcher-made and lack information about diagnostic accuracy. So, “a meta-analysis across the full range of measures can suggest the

measures that appear most promising" (Dollaghan & Horner, 2011 p.1080) and this meta-analysis aims to include those studies with at least one LSM (either through elicitation tasks or language sampling) in their index tests. The combination of heterogeneous studies is unavoidable due to the limited number of studies which share the same index test as well as the variability in bootstrapped index measures (Dollaghan & Horner, 2011).

The repertoire of statistical methods used in meta-analysis includes comparing values of accuracy measures among different studies with similar properties (sample population, study methods, and outcome measures or index test). They consist of sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio and are obtainable by creating a 2 × 2 contingency table of people with and without the target condition (Tables 2-1 and 2-2). In a meta-analysis, all measures are accompanied by a report on a 95% confidence interval (CI) (Dollaghan & Horner, 2011; Glas, Lijmer, Prins, Bonsel, & Bossuyt, 2003).

	Diagnosed by reference standard		
Diagnosed by Index test	Positive	Negative	Total
Positive	True Positive (TP)	False Positive (FP)	Total +
Negative	False Negative (FN)	True Negative (TN)	Total -
Total	Diagnosed +	Diagnosed -	

Table 2-1 The 2 × 2 contingency table of diagnostic accuracy (Adapted from Glas et al., 2003)

The accuracy values (sensitivity, specificity) are judged 'good' if they are between 90 and 100; 'fair' if between 80 and 89; and 'inadequate' if below 80 (Plante & Vance, 1994). LR+ greater than 10 and LR- smaller than .10 are desirable (Dollaghan, 2007) and the narrower the CI, the more reliable the value is (if reported). These measures can also be considered as effect sizes in the meta-analysis of diagnostic accuracy studies (Dollaghan & Horner, 2011 pp.1079-1080).

Diagnostic accuracy measures used in this study	Calculation	Definition
Sensitivity (true positive rate)	TP/(TP + FN)	Proportion of positive test results among people with target condition.
Specificity (true negative rate)	TN/(TN + FP)	Proportion negative test results among people without target condition.
Positive likelihood ratio (LR+)	Sensitivity/(1-Specificity)	The likelihood that a positive test result is found in people with target condition as opposed to people without it.
Negative likelihood ratio (LR-)	(1-Sensitivity)/Specificity	The likelihood that a negative test result is found in people with target condition as opposed to people without it.
Diagnostic Odds Ratio (DOR)	LR+ / LR-	Overall accuracy
Youden's index	Sensitivity + Specificity - 1	Maximum vertical distance between ROC curve and diagonal line, represents the optimal cut-off point.

Table 2-2 A summary of descriptions of diagnostic accuracy measures (Glas et al., 2003; Haynes et al., 2006; Redmond, Thompson, & Goldstein, 2011)**;** TP=True Positive, FP=False Positive, TN=True Negative, FN=False Negative

2.1.1 Foreground question (FQ)

Based on different sources on providing meta-analysis of diagnostic accuracy studies (Dollaghan, 2007, Dollaghan & Horner, 2011; Haynes et al., 2006), the specific aim (or FQ) of the current meta-analysis resulted as the title of this section: 'meta-analysis of diagnostic accuracy of language sample measures for identifying primary language impairment in pre-school children up to five years of age'. So the specific question posed in this meta-analysis is 'what are the most promising LSMs to be employed in identifying PLI in pre-school children?'

2.1.2 Method of searching the literature

A prior assumption would be that few studies relevant to the topic might be found due to the relatively new application of diagnostic accuracy study design in child language impairment accompanied by diagnostic measures such as sensitivity and specificity which are crucial to the methodology of meta-analysis. This assumption would lead to a further one, that a strong conclusion based on pooling the results of different studies might be impossible. However, this meta-analysis can still add to our knowledge of what is missing in this field by providing suggestions for future investigations (Dollaghan & Horner, 2011).

A set of searches through PubMed was designed to find the most recent studies (within the last 10 years) on the diagnostic accuracy of those language sample measures studied in the current project (e.g. mean length of utterance. For a complete list of measures, see chapter 4 or appendix F). The search was filtered for two to five year old children and studies with the following features were excluded:

1. Within a subject unrelated to language domain.
2. Included specified measures as outcome measures or criterion in sampling method only.
3. Any identified clinical groups other than PLI, language delay and SLI including bilinguals, hearing-impaired children etc.
4. Non-English publications unless either they are on Persian population or the required information is stated in abstracts (excluded three studies in German).

No grey literature was searched, only the first keyword (see table 2-3) with the same exclusion criteria used in a search through Google Scholar. As many related studies as possible were gathered following the guidelines suggested by Devillé et al. (2002). The search was finalised on August 21st, 2012 and the results are as follows.

2.1.3 Selection of relevant studies

The search results through PubMed are shown in table 2-3. The results of Google Scholar showed 25 hits with eight eligible studies, four of which were also found through PubMed. A final search through all sources added two more studies.

Key words	Number of hits	Included	Shared with the first search
Diagnostic accuracy + language	21	5	-
Discriminant analysis + language	25	5	1
Sensitivity + grammar	9	1	1
Sensitivity + vocabulary	54	4	1
Sensitivity + mean length of utterance	3	2	2
Sensitivity + language measures	3	2	2

Table 2-3 Search strategy to find the most recent and related studies as of August 10, 2012

One additional study was later found in a book chapter. Three studies from the first search results were excluded due to incompatibility after being thoroughly read. Ultimately, nine studies were included in the analysis, two of which had subsets of either exploratory/confirmatory studies (Simon-Cereijido & Gutierrez-Clellen, 2007) and two of which individually reported index tests in terms of diagnostic accuracies (Conti-Ramsden, 2003; Thordardottir et al., 2011).

2.1.4 Data extraction

The nine studies included are summarised in table 2-4, merging the criteria for appraising diagnostic accuracy studies from Centre of Evidence Based Medicine (*Diagnostic Study Appraisal Worksheet*, 2010) and QUADAS-2 (*QUADAS-2: A quality assessment tool for diagnostic accuracy studies*, 2010). This merging was to make the appraisal items more compatible to the aims of the current study while retaining the fundamentals of meta-analysis of diagnostic accuracy studies. The quality assessment of the eligible studies also included reviewing whether bias concerns have been considered or not. The biases in a meta-analysis study are defined through control procedures as follows: (a) a broader spectrum sample in terms of severity or clinical history assists in reducing *spectrum bias*, (b) *ascertainment bias* is avoidable when the same reference standard is applied to both affected and unaffected clinical groups, (c) to avoid *incorporation bias* and *subjective bias*, the reference standard and index test results should be blind to each other (probably through administration by different examiners) and the results of the index test must not be incorporated to clinical group assignments, respectively (Dollaghan & Horner, 2011). Four studies reported individual diagnostic accuracies for different index tests, and one study (Simon-Cereijido & Gutierrez-Clellen, 2007) had two phases, exploratory and confirmatory, which are shown in individual rows. Consequently, table 2-5 indicates 16 included studies with individually allocated index tests which will be entered into the statistical analysis.

Study	Phases of diagnostic study (DA)	Age range	Representative sample (random or consecutive sampling/spectrum/ size)*	Reference standard	Reference standard applied to all	Reference standard same to all	Reference standard blind to index test	Index test blind to reference standard	Withdrawals explained
Conti-Ramsden, 2003	Discriminant study (phases I&II)	52-70 months	Unclear/Unclear/Small (LI=21, TDL=32)	Referred by specialists and informal assessment by teachers	No	No	Unclear	Unclear	Yes
Klee, Stokes, Wong, Fletcher, & Gavin, 2004; *second study*	Discriminant study (phases I&II)	29 – 61 months	Case-control/Medium/Small (n=45)	SLT's clinical diagnosis prior to the study and two language comprehension test (<1 SD)	Yes	Partly	Unclear	Unclear	No
Simon-Cereijido & Gutierrez-Clellen, 2007	Discriminant study: Exploratory (phase I&II) & Confirmatory phase (phase III)	Average age= 53.7 -55.2 months	Exploratory phase: Unclear/Unclear/Small (n=48) Confirmatory phase: (n=10)	Parent/teacher concern, clinical judgment, and a subtest in Bilingual Spanish and English Assessment (inappropriate to LI diagnosis)	Unclear	Unclear	Unclear	Unclear	No

Study	Phases of diagnostic study (DA)	Age range	Representative sample (random or consecutive sampling/spectrum/ size)*	Reference standard	Reference standard applied to all	Reference standard same to all	Reference standard blind to index test	Index test blind to reference standard	Withdrawals explained
Klee, Gavin, & Stokes, 2007	Discriminant study (phases I&II)	36 – 50 months (for the oldest group, the end limit is unclear and 50 is an estimation)	Unclear/ Medium/Small (n=25)	Sequenced Inventory of Communication Development-Revised (SICD-R), a standardized test of receptive and expressive language, an observation of the child during a free-play with the mother	Yes	Yes	Yes	Unclear	Yes
Heilmann, Miller, & Nockerts, 2010	Discriminant study (phases I&II)	3-5;11 yrs.	Unclear/Medium/Large (n=129)	Refer/receive intervention for being LI, unimpaired children recruited from SALT databases	No	No	Unclear	Unclear	Yes
Wong, Klee, Stokes, Fletcher, & Leonard, 2010	Phase III	49 - 60 months	No/NA/Small (n=29)	SLT's clinical judgment along with Cantonese version of RDLS	Yes	Yes	Yes	Unclear	No

Study	Phases of diagnostic study (DA)	Age range	Representative sample (random or consecutive sampling/spectrum/ size)*	Reference standard	Reference standard applied to all	Reference standard same to all	Reference standard blind to index test	Index test blind to reference standard	Withdrawals explained
Thordardottir et al., 2011	Phase I and II	5 yrs.	Yes/Medium/Moderate (n=92)	SLT's clinical diagnosis prior to the study including exclusionary criteria for PLI	Yes	Partly	Unclear	Unclear	Yes
Eisenberg & Guo, 2012	Phase I&II	3-3;11 yrs.	Unclear/Unclear/Small (n=34)	Parents' concern (using a rating scale) + SPELT-P2	Yes	Yes	Unclear	Unclear	Yes
Gladfelter & Leonard, 2012	Phases I and II	48 – 66 months	Unclear/Unclear/Small (n=55)	Columbia Mental Maturity Scale, SPELT-2, all exclusionary criteria for SLI	Yes	No	Unclear	Unclear	No

Table 2-4 Included diagnostic accuracy studies with a component of language sample measures within the period of 10 years' time (2003-2012)
*US Preventive Services Task Force quality rating criteria for diagnostic accuracy studies: rate 'good' if "uses a credible reference standard; and includes a large number (≥100) of broad-spectrum patients with and without disease (to reduce spectrum bias, see Dollaghan & Horner, 2011)."; 'fair' if "uses reasonable (although not the best) standard; and has a moderate sample size (50–100 subjects) and a 'medium' spectrum of patients."; 'poor' if "uses inappropriate reference standard; ascertains the reference standard in a biased manner; and/or has a very small sample size of very narrow selected spectrum of patients."(Nelson, Nygren, Walker, & Panoscha, 2006).

Study	Number of PLI (mean age in months)	Number of TDL (mean age in months)	Results of accuracy measures	TP	FP	FN	TN
Conti-Ramsden, 2003	21 (37.19)	32 (57)	Past tense marker	17	6	4	26
Klee, Stokes, Wong, Fletcher, & Gavin, 2004; second study	15 (56.40)	Age-matched = 15 (56.87), Language-matched = 15 (35.93)	Combination of age + MLUm +D	15 15	1 0	0 0	14 15
Simon-Cereijido & Gutierrez-Clellen, 2007; Exploratory phase	19 (55.63)	19 (53.95)	Ungrammaticality index	15	0	4	19
			Composite of MLUw + ungrammaticality	15	0	4	19
			Composite of correct use of verb + clitics + articles	15	0	4	19
Simon-Cereijido &Gutierrez-Clellen, 2007; Confirmatory phase	5 (53)	5 (53.40)	Composite of MLUw + ungrammaticality	4	1	1	4
			Composite of correct use of verb + clitics + articles	4	1	1	4
Klee, Gavin, & Stokes, 2007	14 (NR)	11 (NR)	Combination of age + MLUm + D	12	1	2	10
Heilmann, Miller, & Nockerts, 2010	60 (60)	69 (60)	10 LSMs altogether	52	9	8	60
Wong, Klee, Stokes, Fletcher, & Leonard, 2010	15 (55.27)	14 (55.71)	Composite of age + MLUm + D + constant	11	6	4	8
Thordardottir et al., 2011			MLUw at -1SD cut off point MLUm at -1SD cut off point	9 9	22 24	4 4	53 51
Eisenberg & Guo, 2012	17 (41.19)	17 (41.65)	Percent of grammatical utterances Percent of verb tense usage	17 17	2 3	0 0	15 14
Gladfelter & Leonard, 2012	12 (51.58)	15 (51.33)	Finite Verb Morphology Composite (FVMC)	8	2	4	13

Table 2-5 Sample size, Index tests, and 2×2 contingency table of each study; NR=Not reported; Terms' definitions: MLUm=Mean Length of Utterance in morphemes, D=lexical diversity measures by D, Ungrammaticality index=percentage of ungrammatical utterances, MLUw=MLU in words, Finite Verb Morphology Composite (FVMC)=correct use of tense and agreement in a list of verbs (Gladfelter & Leonard, 2012) (The format of this table is adapted from Dollaghan & Horner, 2011 p. 1083); TP=True Positive, FP=False Positive, TN=True Negative, FN=False Negative

2.1.5 Data analysis

The total number of children from eligible studies was 212 typically developed children ranged between 35.93 and 60 months of age, and 178 children with PLI with age range between 37.19 and 60 months. The frequency values of index tests in the contingency table were entered to Meta-DiSc, which is a freely available software specifically designed for the purpose of meta-analysis studies (Zamora, Abraira, Muriel, Khan, & Coomarasamy, 2006). It provides pooled values of sensitivity, specificity, likelihood ratios (LR), diagnostic odds ratios (DOR), as well as symetrical ROC (SROC) across different studies. DOR has less clinical utility unless a meta-regression procedure is used to analyse the data,so will not be applied in this meta-analysis (Zamora et al., 2006). The pooled diagnostic measures are averagely weighted against the power of each study, i.e. its sample size. The pooling method most appropriate for each study is determined by the method of classifying population into unimpaired and impaired: two categories for classification lead to 'dichotomous type' data whereas a spectrum categorisation of impairment (e.g. based on severity) is an indication of 'continuous type' data (Haynes et al., 2006; Zamora et al., 2006). The specific statistical models corresponding to each type of data are out of the scope of the current study (for more information, see Borenstein, Hedges, Higgins, & Rothstein, 2009; as well as the web site of http://www.meta-analysis.com/downloads); however, pooling analysis for the dichotomous data was set to the DerSimonian-Laird random effect model, considering the probability of heterogeneity across the studies involved, with a confidence interval of 95%. Because there are some cells of value zero among the contingency tables, Meta-DiSc was set to add .5 to such cells to enable the measures and CIs to be calculated.

Pooled likelihood ratios: likelihood ratios are assumed to be better diagnostic measures due to their lower sensitivity to sample size (Haynes et al., 2006). For this reason, the pooled LR+ and LR- were calculated and depicted as indicated in forest plot (Figures 2-1 and 2-2). In both diagrams, the red circles show the weight of each index test by the related sample size and blue horizontal lines indicate the confidence interval of each measure. The pooled LRs are demonstrated through red diamonds at the bottom of the plots and the corresponding 95% CI is shown by red horizontal line. The dashed vertical lines depict the same 95% CI of pooled values to provide a better visualisation of LRs across different index tests.

The ideal value for LR+ is greater than 10 and the optimal value of LR- is smaller than .1 (Dollaghan, 2007; Plante& Vance, 1994). This rules out 11 out of 16 index tests in

terms of LR+; however, those five tests with LR+ greater than 10 possess wide CIs with lower bounds falling below 10. The negative LRs of four index tests show ideal values of less than .1; however, similar to LR+, all the upper bound values of tests are greater than .1.

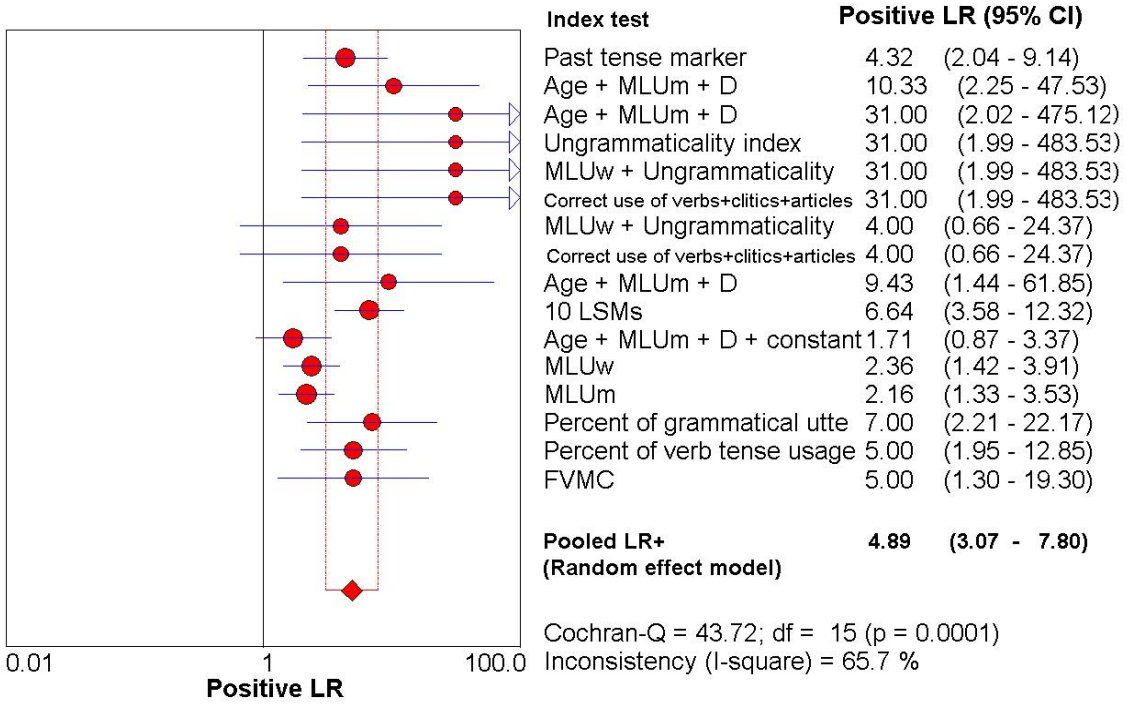


Figure 2-1 Forest plot of pooled LR+ (95% CIs) for each index test (corresponding studies are found in table 2-5)

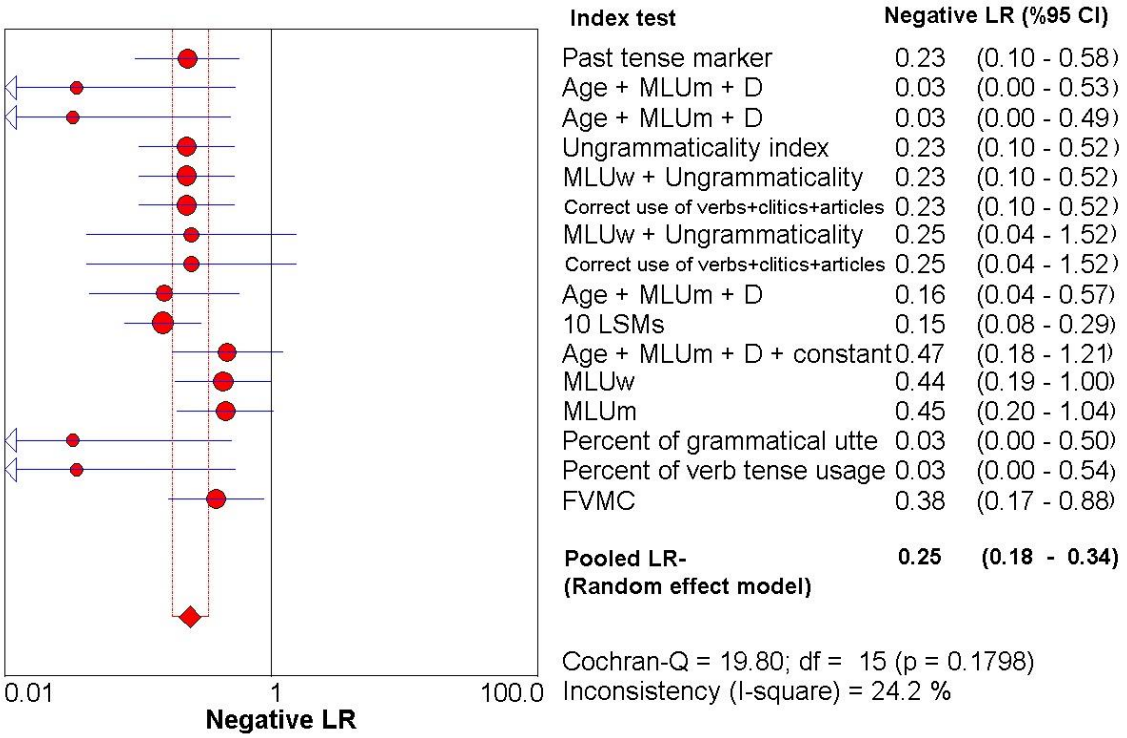


Figure 2-2 Forest plot of pooled LR- (95% CIs) for each index test (corresponding studies are found in table 2-5)

Discussion: As can be seen in tables 2-4 and 2-5, both reference standards and index tests are widely varied across studies along with the sample sizes. They need to be carefully selected and defined in each study with the aim of increasing objectivity. Sampling procedures and inclusion/exclusion criteria also need to be fully addressed along with a clear statement on whether reference standard and index tests are independently administered or not, unlike in several of the eligible studies, which would seriously affect the quality of the evidence. Research design and conduct, severity of target condition, administration of all sorts of tests, and selection of cut points in test interpretation are other sources of difference across studies. All these result in a heterogeneous data set for meta-analysis (Reitsma, 2009) which leads to large confidence intervals in all diagnostic values. With respect to the present meta-analysis, the heterogeneity analysis of both LRs will help to give a better understanding of these outcomes. The relatively high inconsistency of LR+ (I-square) shows that part (65.7%) of variation in LR+ values is due to between-index heterogeneity. This is also supported by the fact that the result of "Cochrane-Q test of the null hypothesis that variations between the results for individual measure (index) are due to chance" is significant (p<.000) (Dollaghan & Horner, 2011 p. 1085; for a full explanation of the statistical terms, also see Zamora et al., 2006). Generally, this means that from a clinical point of view, only four measures including Age+MLUm+D (studied by two studies: Klee et al., 2004, Klee et al., 2007), Ungrammaticality Index (Simon-Cereijido & Gutierrez-Clellen, 2007), MLUw+Ungrammaticality (Simon-Cereijido & Gutierrez-Clellen, 2007), and Correct use of verbs+clitics+articles (Simon-Cereijido & Gutierrez-Clellen, 2007) can be informative (with positive LRs greater than 10 as seen in figure 2-1) in ruling the impaired children as impaired; however, with the wide CIs, the clinician would prefer to use them as supplementary to each other. This picture is different for LR- which shows much less inconsistency (I-square = 24.2%), and Cochrane-Q test of the null hypothesis is strongly insignificant (p>.05). This provides the clinicians with a better basis for a decision in terms of choosing a useful index from those with the least LR- (four measures); however, the wide range of their CIs suggests that they should be interpreted cautiously in terms of ruling out unimpaired children. The overlapping CIs in both LRs is an indication of heterogeneity across measures which means that decisions based on single measure would affect clinical diagnosis. Clinical judgment about children's group membership based on those promising measures is solely about their diagnostic status (Dollaghan & Horner, 2011).

Small sample size also affects CIs around LRs and it can be seen that none of the measures shows the optimal lower bound in LR+ (the largest is 3.58 for '10 LSMs') nor upper bound in LR- (the smallest is .29, again for '10 LSMs'). So, until more informative studies with large enough sample populations are conducted, only a few of the current index tests with sufficiently low LR- show potential in finding children with TDL.

Not all measures examined in the eligible studies were similar, but they all had at least one component of grammatical or semantic measure involved in their index tests. The results fortify the diagnostic accuracy study of the current thesis by providing information on promising measures with optimal LRs irrespective of the imperfect CIs. These measures are classified in table 2-6 based on ideal values of either LR+ or LR- and will be possible index measures for the current diagnostic accuracy study.

LR+	**LR-**
Age + MLUm + D	Age + MLU m + D
Ungrammaticality index	Percent of grammatical utterances
MLUw + Ungrammaticality index	Percent of verb tense usage
Correct use of verbs + clitics + articles	

Table 2-6 Promising language measures in terms of LR values (irrespective of the 95% CI range) which will be entered into the diagnostic accuracy study if they are obtainable through language sampling using Systematic Analysis of Language Transcriptions (SALT)

Limitation: The main limitation of this meta-analysis was the small number of eligible studies which also made the process of blindness impossible. This shows the necessity of conducting such high level evidence-based studies as meta-analysis.

2.2 Persian studies on child language

Quantitatively, research on child language in Iran is not extensive and as will be shown, the majority of studies in this field have descriptive designs. Focusing on the age range of the children in the current study, what follows is a review of Iranian studies on child language, in areas of both development and disorder regardless of their cause, from an EBP point of view. The reason for including all studies in this review is to provide a comprehensive critical summary of the research in this field inside Iran. The results will then be summarised in relation to the main aims of the current thesis. The method

employed was appraising and categorising studies using the levels of evidence suggested by Oxford Centre for Evidence-Based Medicine (Howick et al., 2011).

2.2.1 Method of retrieving studies

The method of retrieving the articles was based on searching through two principal Iranian database websites in which academic publications are recorded, (1) scientific Information Database (SID) (http://www.sid.ir/En/Index.asp) and (2) Iranian Research Institute for Information Science and Technology (IRANDOC) (http://www.irandoc.ac.ir/irandoc-english/irandoc-english.html), as well as MEDLINE and EMBASE as two external databases, CHILDES Forum and Google Scholar search engine. Another source of information was personal communication through email and phone calls. All the academic documents including published articles and unpublished documents (thesis and conference presentations) related to child language studies with the following inclusionary criteria, and were gathered in both Persian and English languages:

1. Studies of Iranian Persian-speaking children including pre-schoolers,
2. Studies of language development or disorder,
3. Include a component of *survey*, *normalising*, *assessment* or *diagnosis* of any part of spoken language skills including *grammar (syntax/morphology)*, *lexicon* or *semantics* as well as similar variables to the current study.

Studies were matched against the criteria of study designs and hierarchy of evidence (Greenhalgh, 2010) (Figure 2-3). Although this hierarchy is for therapy studies, it can be used for all types of studies in clinical settings with slight modifications in appraisal questions to address descriptive or observational studies that form the major research body in Persian speech therapy.

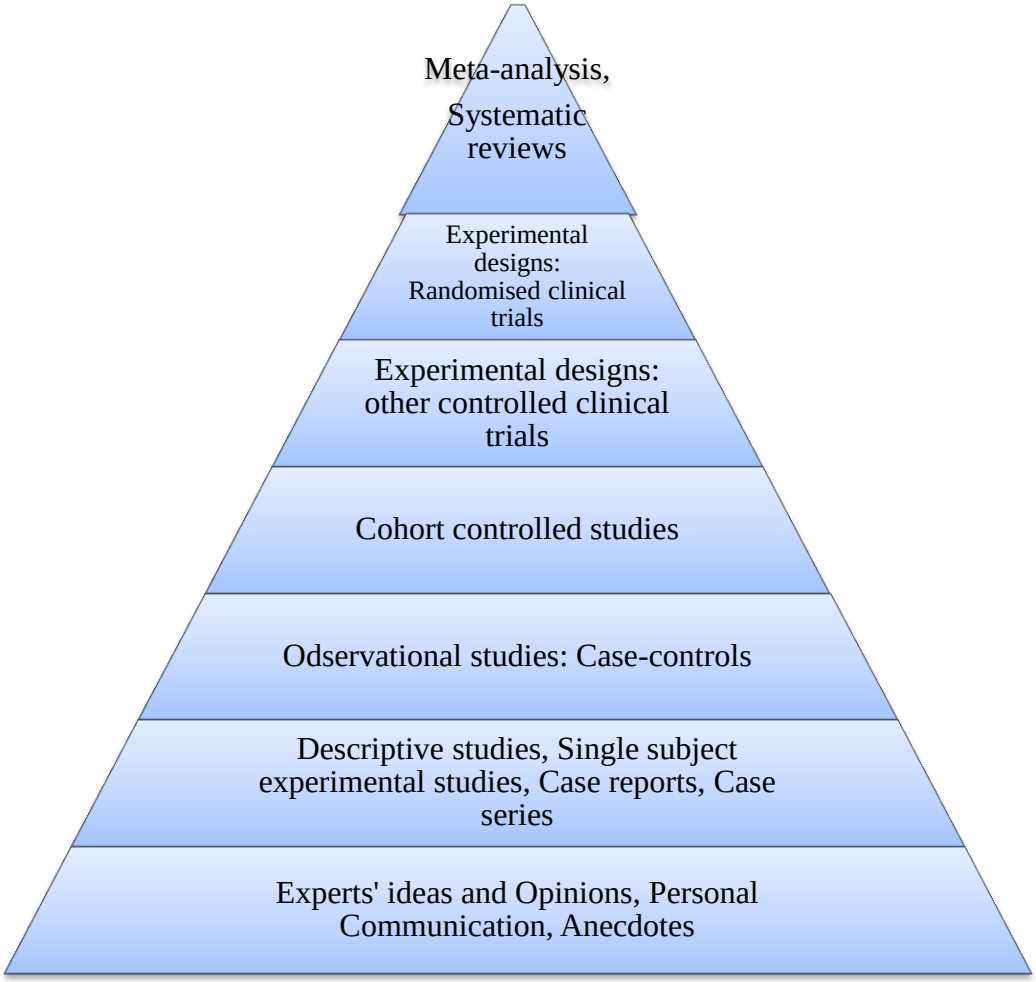


Figure 2-3 Hierarchy of evidence (Greenhalgh, 2010; Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996)

A total of 33 studies met the inclusion criteria specified in this review; the vast majority of them, 27 studies, were descriptive, cross-sectional surveys, representing the lowest level of evidence (Greenhalgh, 2010; *Study Designs*, 2012). Three remaining studies were of observational, cross-sectional analytic or case-control types, one level higher in the hierarchy of evidence.

2.2.2 Psychometric studies of tests in Persian

Thirteen studies were described as *test development* studies that were designed to examine various psychometric features of either translated tests or tests developed for research purposes. Some of them have provided normal scores in large samples and the researchers claim that they can be used as reference (norm-referenced) to find children with low-achievement behaviour relating to the test items, although they are not published as tests, specifically without test manuals. The appraisals of these studies are shown in tables 2-7 and 2-8 using the psychometric criteria introduced by McCauley

and Swisher (1984) in the field of speech and language therapy and most recently employed by Klee et al. (2009). In this appraisal, ten criteria of psychometric properties of these tools were checked against the definitions provided by McCauley and Swisher (1984) as follows:

1. Norming sample should be defined clearly so that the representativeness of the sample is documented by reasonable (a) geographical areas covered, (b) socioeconomic status covered, and (c) "normalcy of subjects in the sample" (p. 38) mentioning the procedure and number of excluded cases. This condition assumes the study as a fully normative study; however, most of the studies reviewed were in the very first stages of developing a test so this condition was used cautiously for all types of studies containing test development. This ensured that sampling in non-norming studies was checked as being representative against the study aims. This is true for all other criteria, too.
2. Sufficient sample size, minimum of 100 cases in each sample group for norming studies (Paul, 2007 p.42).
3. Internal consistency of test structure should be reported in terms of item difficulty or validity or both.
4. The measure of central tendency and variability should be reported for each sample group.
5. Concurrent validity report.
6. Predictive validity report.
7. Test-retest reliability of .90 or higher at .05 significance level or better.
8. Inter-examiner reliability of .90 or higher at .05 significance level or better.
9. A detailed and comprehensive test presentation and scoring system should be provided so that it can be replicated by others.
10. It should be clear who is eligible to do the test and whether there is a need for "specialised training for administrators or scorers".

The resulting appraisal is summarised in tables 2-7 and 2-8.

	Study	Assessment	Abbreviation	Age range	Sample size	Number of psychometric criteria
1	Sayyahi, Soleymani, Mahmoudi Bakhtiyari, & Jalaie, 2011[J]	Non Word Repetition Test		4 – 4;11 (yrs;mths)	30	6
2	Heydari, Torabi Nezhad, Agha Rasouli, & Hoseyni, 2011[J]	Speech Intelligibility Test		3 – 5 (yrs)	100	4
3	Hasanati, Agha Rasouli, Mahmoudi Bakhtiyari, & Kamali, 2011[J]	Sentence Repetition Test		2;6 – 4 (yrs;mths)	72	5
4	Ziatabar Ahmadi, Arani Kashani, Mahmudi Bakhtiari, & Keyhani, 2010b[J]	Rhyme Awareness Task*		5 – 6 (yrs)	100	7
5	Ziatabar Ahmadi, Arani Kashani, Mahmudi Bakhtiari, & Keyhani, 2010a[J]	Tasks for Assessment of First Phoneme of Word*		5 – 6 (yrs)	100	6
6	Kazemi et al., 2008[J]	Persian MacArthur-Bates Communicative Development Inventory	P-MCDI	8 – 16 (mths)	30	4
7	Kazemi et al., 2007[J]	Children's Communication Checklist*	P-CCC	5 – 11 (yrs)	978	5
8	Kazemi & Derakhshandeh, 2007[J]	Oral/Speech Motor Control Protocol*		3 – 6 (yrs)	300	5
9	Kazemi et al., 2012[J]	Mean Length of Utterance	MLU	2;6 – 5;6 (yrs;mths)	171	5
10	Oryadi-Zanjani, Ghorbani, & Keikha, 2006[J]	Mean Length of Utterance*	MLU	2 – 5 (yrs)	580	3
11	Kazemi, Ghasisin, Rezaei, Samadi, & Sharifi, 2007[CP]	Test of Language Development-Intermediate:3	TOLD-I:3	8 – 13 (yrs)	100	5
12	Soleimani & Dastjerdi Kazemi, 2005[J]	Phonological Awareness Test*		4 – 7 (yrs)	203	5
13	Hasanzadeh & Minaei, 2000[LRR]	Test Of Language Development-Primary:3*	TOLD-P:3	4 - 8	1235	6

Table 2-7 Appraisal of Iranian studies on test development for language assessment in Persian; * Specifically mentioned as a norming study; J: Journal, CP: Conference Presentation, LRR: Local research report

Criterion	Number of studies	Studies
1.Sample representative	13 out of 13	All studies
2.Sufficient sample size	5 out of 13	4, 5, 7, 8, 10, 12, 2
3.Internal consistency	11 out of 13	All studies except 9 and 10
4.Measures of central tendency and variability	9 out of 13	2, 3, 4, 7, 8, 9, 10, 11, 13
5.Concurrent validity	1 out of 13	12
6.Predictive validity	0	
7.Test-retest reliability	10 out of 13	All studies except 7, 8, & 10
8.Inter-examiner reliability	4 out of 13	2, 4, 5, 9
9.Test performance instruction	11 out of 13	All studies except 10 and 12
10.Defining the eligibility for test administration	0	

Table 2-8 Psychometric criteria met by each study

None of the above-mentioned studies reported diagnostic measures for the tests or assessments used. Therefore, no judgement about how accurately any of them can identify children with language impairment can be derived from applying these instruments in clinical settings.

2.2.3 Case-control studies

Five analytical observational cross-sectional case-control studies were found (Foroodi-Nejad, 2011; Golpour, Nilipour, & Roshan, 2007; Maleki Shahmahmood, Soleymani, & Faghihzade, 2011; Maleki Shahmahmood, Soleymani, & Jalaei, 2009; Soraya, Bakhtiyari, Badiee, Kazemi, & Soleimani, 2012).Critical appraisal forms from Stanford University were used to assess these studies (*Assessing scientific admissibility and merit of published articles: Critical appraisal form*, 2012) and the results are shown in table 2-9.

Study	Hypothesis clearly stated	Population	Random selection of controls	Number considered for enrolment vs. enrolled	Age	Inclusion/exclusion criteria stated	Clear and same diagnostic criteria for both groups	What studied	Data collection valid and reliable	Statistical analysis explained	Results
Golpour et al., 2007[J]	Yes	10 cases (severe-profound hearing impaired) ~ 10 controls (age matched)	No	Only enrolled reported	4 – 5 yrs.	Yes	Same and clear	Total utterances, lexical and grammatical words, total words, Type-Token Ratio, Mean Length of Utterance in words	Insufficiently described	*t*-test	Significant difference in all measures except TTR in both occasion of free speech and descriptive language
Maleki Shahmahmood, Soleymani, & Jalaei, 2009[J]	Stated not clearly	12 cases (SLI) ~12 controls (language-age matched)	Yes	Control: 40~12 Case: 15~12	Control: 4;1 (±2) yrs., Case: 5;7 (±6) yrs.	Yes	Diagnosis overlaps outcome measures	Test of Language Development-Primary (Farsi version), Percentage of some grammatical morphemes and words	Insufficiently described	Mann-Whitney U, t-test, not specifically-explained which one is used for which measure	- Meaningful difference between two groups in some sub-tests of TOLD-P): Conjunction words, Oral words, Imitation, Spoken quotient, Organization quotient, Semantic quotient. - No meaningful difference in percentage of some grammatical morphemes and words.

Study	Hypothesis clearly stated	Population	Random selection of controls	Number considered for enrolment vs. enrolled	Age	Inclusion/exclusion criteria stated	Clear and same diagnostic criteria for both groups	What studied	Data collection valid and reliable	Statistical analysis explained	Results
Maleki Shahmahmood, Soleymani, & Faghihzade, 2011 [J]	Stated not clearly	13 cases (SLI) ~13 controls (age matched)	Yes	Only enrolled reported	Control: 67 (±6.8) mths.; Case: 67 (±6.9) mths.	Yes	Diagnosis overlaps outcome measures	Test of Language Development-Primary (Farsi version), Mean length of Utterance in morphemes, Percentage of some grammatical morphemes and words	Yes for TOLD-P, Insufficiently described for other measures	Mann-Whitney U, t-test, not specifically-explained which one is used for which measure	- Meaningful difference between two groups in some sub-tests of TOLD-P): Conjunction words, Oral words, Imitation, Spoken quotient, Organization quotient, Semantic quotient. - Meaningful difference between two groups in MLU-m, lexical and grammatical words, lexical to grammatical words ratio

Study	Hypothesis clearly stated	Population	Random selection of controls	Number considered for enrolment vs. enrolled	Age	Inclusion/exclusion criteria stated	Clear and same diagnostic criteria for both groups	What studied	Data collection valid and reliable	Statistical analysis explained	Results
Foroodi-Nejad, 2011[D]	Yes	9 cases (SLI) ~16 controls (age matched)	Not explained	Only enrolled reported	Control: 69 (±9) mths., Cases: 67 (±13) mths.	Yes	Same but not clearly-explained	ENNI's macro-structure and micro-structure (Schneider, Dubé, & Hayward, 2005), Case marking in Persian, using clitics, agreement and tense use	Yes	*t*-test, Mann-Whitney U	- Meaningful difference between two groups in ENNI measures; - Significant difference in the percentage of correct use of case marker *ra*; - Meaningful difference in percent of clitics usage; - No difference inpercentage of correct agreement; - Meaningful difference inmean proportion of *mi*\ usage.
Soraya et al., 2012[J]	Yes	42 cases (prematurely-born) ~ 42 controls (age matched)	No	Only enrolled reported	All: 18-36 mths.	Yes	Same and clear	MacArthur-Bates P-MCDI (Toddler form) vocabulary section	Yes	Two-way ANOVA	Meaningful difference in vocabulary size between two groups

Table 2-9 Summary and assessment of case-control studies on Iranian children with and without language impairment; J: Journal, D: Dissertation

Similar to non-Persian studies, the small sample size is a disadvantage of the above-mentioned studies. Besides that, in all but one, the number of children who were not enrolled in the study is not explained. Also, only two studies reported a random sampling and others had either no clear report of sampling procedure (two studies) or did not recruit the population randomly.

Three studies had a clearly-stated hypothesis and the remaining two with an unclear hypothesis are those in which diagnosis overlapped with the outcome measures which would affect the validity of both. Four studies employed language sample measures either structured or informal; however, none of them submitted a sufficient description of administration procedure. In some studies (e.g. Golpour et al., 2007), it was observed that the operational definition of the measures was not compatible with well-known definitions which caused a big problem in validity appraisal of these studies.

Apart from Foroodi-Nejad's study, no other study reported controlling statistics for the efficacy of the results such as confidence interval or effect size; however, they were calculated by the author if sufficient data was available for computing. Table 2-10 shows the relevant 95% CI and effect size for those studies with a group of children with PLI.

Study	**Effect size (Cohen's *d*)**	**95% CI**
Maleki Shahmahmood, Soleymani, & Jalaei, 2009 **Test:** TOLD-P		
Semantic quotient	-1.15	-2.01, -.28
Organisation quotient	-1.52	-2.43, -.61
Spoken language quotient	-.93	-1.77, -.09
Sentence imitation	-3.04	-4.22, -1.87
Oral vocabulary	-2.93	-4.09, -1.78
Relational vocabulary	-1.13	-1.99, -.26
Maleki Shahmahmood, Soleymani, & Faghihzade, 2011 **Test:** free speech language sampling		
MLUm	-1.65	-2.54, -.76
Percentage of content words	1.5	.63, 2.37
Percentage of grammatical words	-1.5	-2.37, -.63
Grammatical word to content word ratio	-1.6	-2.48, -.71
Foroodi-Nejad, 2011 **Test:** structured elicitation task		
Percentage of correct use of case marking (*ra*)	-2.23	-3.25, -1.21
Percentage of clitics usage	-1.19	-2.07, -.31
Correct percentage of agreement		
Mean proportion of *mi*\ usage	-4.23	-5.66, -2.81

Table 2-10 The calculated 95% CIs and effect sizes for Iranian case-control studies; effect sizes of .2 or less is considered small, around .5 are medium, and those equal or greater than .8 are large (Cohen, 1988); TOLD-P=Test of Language Development-Primary, MLUm=Mean length of utterance in morphemes.

All the effect sizes of different measures are large (greater than .8) which documents the large differences between children with and without PLI in terms of language measures. Two studies by Maleki Shahmahmood, Soleymani, & Faghihzade (2011) and Foroodi-Nejad (2011) will be of particular reference later when examining the difference between the groups of children with and without PLI in the present study.

2.2.4 Descriptive studies

The studies with no comparison groups and no intention for test development were categorised as descriptive case studies, which included 15 studies. They were either cross-sectional or longitudinal and describe normal or impaired language development in Iranian children mostly with typically development. A review of them has been shown in table 2-11 using criteria from 'assessing scientific admissibility and merit of published articles, critical appraisal form, sections P-R' (*Assessing scientific admissibility and merit of published articles: Critical appraisal form*, 2012).

Study	Cross-sectional (CS) vs. Longitudinal (L)	Population	Age	Inclusion/exclusion criteria clearly stated	Random sampling	What studied	Data collection	Validity and reliability stated	Sources of bias stated	Statistical analysis explained	Results
Fahim[CB]	L	1	7 – 34 mths.	No	No	Speech and language development in early years of life	Parents reports and transcription, voice recording	No	No	NA	A comprehensive description of different developmental stages in early years of life including vocabulary, semantic relations and grammatical categories and morphology.
Jalilevand, Ebrahimipur, &Purqarib, 2012[J]	L	2	12 – 36 mths.	Yes	No	What and when are the question words perceived in Persian?	Video-recorded free speech, 120 minute per months from 12 to 36 mts old.	Yes	No	Descriptive statistics	In Persian, questioning emerges by changing in intonation and follows the same pattern as development of question words in English.
MahmoudiBakhtiyari, Soraya, Badiee, Kazemi, &Soleimani, 2012 [J]	CS	42	18 – 36 mths.	Yes	No	Expressive lexicon size	Persian-MCDI Vocabulary section	Cronbach's alpha reported along with professional consensus in a pilot study	No	Two-way ANOVA	Nouns were the largest categories among all three age groups. The expressive lexicon increased by age.

Study	Cross-sectional (CS) vs. Longitudinal (L)	Population	Age	Inclusion/exclusion criteria clearly stated	Random sampling	What studied	Data collection	Validity and reliability stated	Sources of bias stated	Statistical analysis explained	Results
Mehdipour, Shirazi, & Nematzadeh[CB]	L	21	18 – 24 mths.	Yes	No	Vocabulary count and type; sentence development	P-MCDI, parents and caregivers' reports	No	No	Descriptive statistics, development by age comparisons using t-test	Vocabulary increases by age. Nouns are the most frequent category. No sex difference in terms of expressive vocabulary size. High percentage of declarative sentences followed by imperatives and questions.
Meshkato-Dini, 2001[J]	L	2	NC	No	No	Grammar and vocabulary development	On-line transcription. Universal grammar as an analytic framework applied.	No	No	NA	An explanation for different grammatical features based on universal grammar account
Meshkato-Dini, 2004[J]	L	2	23 – 42 mths.	No	No	Emerging sequence of inflectional affixes and morphemes	On-line transcription.	No	No	No	Emerging morphology within these ages has been demonstrated through several tables.
Modarres Zadeh, 2010[J]	CS	60	2.5 yr.	Yes	Yes	Expressive vocabulary	Persian picture naming test	Cronbach's alpha and split-half reliability reported.	No	Descriptive statistics	Not contingent with the study question: the test is capable of showing the development of lexicon in this age range.

Study	Cross-sectional (CS) vs. Longitudinal (L)	Population	Age	Inclusion/exclusion criteria clearly stated	Random sampling	What studied	Data collection	Validity and reliability stated	Sources of bias stated	Statistical analysis explained	Results
Naderi & Seifenaraghi, 1993[CB]	L	180	0-2 yrs.	Not efficien tly	No	Speech development in Farsi	Behavioural observation and note-taking, parents' interview	No	No	Descriptive statistics	A comprehensive quantitative report on all variables.
Oryadi-Zanjani & Ghorbani, 2005[J]	CS	90	4 – 5 yrs.	Yes	No	MLUm, mean length of 5 longest utterances, Verb count, Number of relative clauses, Number of words, Speech rate	Conversation and descriptive speech sampling with no data on the length and the number of utterances	No	No	ANOVA, *t*-test	A comparative study on the indexes in different accents.
Pouladi & Khoddam, 2002[CB]	CS	60	4 – 5 yrs.	NC	Yes	MLU	50 utterances within two contexts: conversation and picture description	NC	NC	NC	- For 4-4;6 yrs: MLUw =3.5, MLUm=7.09 - For 4;6-5 yrs: MLUw=4, MLUm= 7.5
Rahmany, Marefat, & Kidd, 2010[J]	CS	51	2 – 7 yrs.	Yes	No	Understanding relative clauses in Persian	Picture selection task	No	No	ANOVA-Error analysis	More problems observed in processing of non-canonical word order sentences.
Rezaei, Shavaki, Arshi, & Keyhani, 2011[J]	CS	110	2.5 – 4 yrs.	Yes	Yes	Receptive and expressive vocabulary	Persian picture naming test	Cronbach's alpha and split-half reliability reported.	No	ANOVA, independent *t*-test, correlation coefficient	Picture pointing and naming tasks can differentiate between ages 2.5 and 4 years old. No sex difference was seen. Expression and reception develop in parallel.

Study	Cross-sectional (CS) vs. Longitudinal (L)	Population	Age	Inclusion/exclusion criteria clearly stated	Random sampling	What studied	Data collection	Validity and reliability stated	Sources of bias stated	Statistical analysis explained	Results
Samadi & Perkins, 1998[J]	L	3	1;8 – 3;4 yrs.	No	No	Developing Persian-LARSP	Persian-LARSP checklist	NR	NR	CHAT system and LARSP procedure	Reporting language grammar and vocabulary following stages of LARSP

Table 2-11 Review on descriptive studies on Iranian children with and without language impairment; J: Journal, D: Dissertation, LRR: Local research report, CB: Conference booklet, NC: Not clear, NR: not reported

The importance of descriptive studies in current speech therapy in Iran is their capability for increasing awareness of the nature of Persian language development in children. The majority of descriptive studies were about the specifications of Persian language development in Iranian children, with children as young as newborns recruited in sampling. There are cross-sectional studies among this group with large sample sizes (more than 100) that aimed to provide some comparisons across different Persian linguistic features, which would be beneficial in decision making about what to look for at which age group in future studies.

2.2.5 Procedures in assessment of child language

Collecting data to assist clinicians with the process of diagnosing language disorders in children has been a primary concern in the history of communication disorders. From Brown's very first attempts to formulate a comprehensive description of language development in early childhood (Brown, 1973) to the most recent standardised language tests for children, all researcher-clinicians have been seeking the most clinically reliable methods of screening, assessment and follow up for child language conditions. In general, there are two major procedures to evaluate language behaviour in children:

1. **Formal assessment:** Structured tests are the best known and most reliable type of formal language assessments. They need to be administered in predetermined environments following step by step instructions with no deviation from the conventions defined in the test manual. They also need to be administered by well-trained examiners with adequate time allocated (Paul, 2007). They provide the users with several pages of reports on scores for children with normal language development (normative data). They have been formed upon the assumption that if children can behave within the spectrum of normal language development under test conditions, they will be able to show their ultimate language performance in natural settings where there is no pressure on them to communicate and their language competence will be presented with minimum struggle (Paul, 2007).

   The two main types of structured or formal assessments are *norm-referenced* and *criterion-referenced*. Norm-referenced tests compare children's language behaviour against their normal peers, mainly based on chronological age, and tell you if a given child is delayed or not. Criterion-referenced tests provide the user with a set of signs which are clinically meaningful in terms of being symptoms of language disorders. Then the clinicians can compare their clients to

the average scores of those specific signs of disorders to say whether there is any problem or not. They are, in fact, indications of a child's pass/fail situation (Paul, 2007).

2. **Informal assessment:** Some groups of children might be misclassified through the use of standardised language tests. Standardised tests can be unreliable for very young children (Allen, Bliss, & Timmons, 1981), i.e. under 2;6-3 years old due to their short attention span and /or for children with complicated clinical conditions or even typically developing children with high impulsivity. This is when an alternative method, in which children's individual circumstances are considered in language evaluation, may lead the clinician to a more reliable assessment result (Eisenberg, Guo, & Germezia, 2012). Informal assessment procedures contribute to clinicians' judgements in language assessment and include those contextualised assessment procedures that take clients' and their relatives' circumstances into consideration. Language sample analysis (LSA) is the most popular procedure within this framework, and is growing continuously in use (Heilmann et al.; 2010). LSA is a procedure to document functional use of language in children, particularly those in early years of life through pre-school years (Eisenberg et al., 2012; Heilmann et al., 2010). Compared to norm-referenced language tests, some LSA have been shown to be more reliable in young children (Eisenberg et al., 2012). Language sampling is completely context-dependent and if no other limitation is imposed, it is obtainable through parents' personal recordings of their own children. This method, however, needs to be controlled and validated by applying some rules in the process of the collection of language samples, e.g. situation, interactant, recording, language context, transcription conventions etc.(Eisenberg, Fersko, & Lundgren, 2001). Table 2-12 shows concerns when eliciting language sample analysis as the assessment procedure for child language. If they are managed carefully with regard to the goals of research or clinical intervention, a representative sample would result (Tomasello & Stahl, 2004). What follows are the essentials for providing a valid method of sampling which will produce a more reliable outcome for decision-making.

Sample size: The length of the sample is important because of its direct influence on the calculation of measures. It should be either a *predetermined time* or *utterance limit*. Although some studies have shown that sample size does not affect the subsequent diagnosis, this still needs investigation under an EBP framework. Hutchins et al. suggest that by increasing the number of language samples or participants involved, however, a standardized language sample is possible (Hutchins, Brannick, Bryant, & Silliman, 2005). Tomasello suggests that the more samples (irrespective of session length), the better; however, the disadvantage of this for clinicians and researchers is increased transcription time. For (orthographically) transcribing of a one hour speech sample, the estimation is 10 to 20 hours. An alternative for researchers is to use pre-transcribed samples provided in such transcribing systems as CHILDES or SALT which obviously removes the need for any transcribing time (Tomasello & Stahl, 2004).
Setting: The setting should be chosen to be compatible to the research question, goals of assessment and child's condition (Tomasello & Stahl, 2004). It can be at the child's home, clinic, school room or play room. The most important thing is to consider the environmental background distracters for the children and the necessity to have a clear sound recording.
Participants: Potentially anybody can be the interactant with the child. The most studied participants have been parents, especially mothers, clinicians and sometimes peers. The studies in this domain have yielded different results; so it is difficult to determine one single interactant as the best (Eisenberg et al., 2001; Haynes, Purcell, & Haynes, 1979).
Instructions given to interactants: are important in that they will directly affect the language output on both sides of the communication. Many studies suggest using a child-directed strategy in sampling with directions from the interactant, which allows the children to have goal-directed inputs and have longer and sometimes more complicated outputs (Girolametto & Weitzman, 2002).
Activity: Which is the best activity to elicit nouns, verbs, complex sentences and syntax, broader range of lexicon, variety of speech acts etc. is the core question when the researcher or clinician is deciding on the activity in which the communicators are going to be involved. To meet this variety of aims, different activities can be designed and this element of language sampling is one which can be modified in an attempt to make the LSA approach standardization. The most frequently used activities include free play (with the advantage of potentially providing a long sample), conversations, narratives in telling/re-telling stories or favourite experiences or self-talk during play. Sometimes, researchers have designed other researcher-made activities to find the closest activity to standardisation (Bornstein, Painter, & Park, 2002; Eisenberg et al., 2001; Hutchins et al., 2005; Southwood & Russell, 2004; Wagner, Nettelbladt, Sahlen, & Nilholm, 2000). If the activities are structured, the ultimate method could be classified as a formal assessment.
Materials: it is advised that materials used in sampling are adapted to the participants' developmental age (Eisenberg et al., 2001; Paul, 2007). This includes providing some unfamiliar and a wide variety of toys in a context of here-and-now themes for younger and shy children and more decontextualised themes (there-and-then) with more familiar toys that help the older children to get detached from a type of here-and-now experience with new toys to more fantasy contents (Eisenberg et al., 2001; Wanska, Bedrosian, & Pohlman, 1986). Miller (1981) suggests those toys which encourage children to get involved more in different activities (e.g. Eisenberg et al., 2001). Materials have also been shown to have an effect on the linguistic-related variables of the study and equally some non-linguistic variables such as expressing temporal concepts. More manipulatable toys and free activities help the younger children to communicate more verbally (Klein, Moses, & Jean-Baptiste, 2010).

Table 2-12 Key elements in selecting an LSA procedure in child language assessment (the elements adapted from Eisenberg et al., 2001)

In the view of the *descriptive-developmental model* introduced by Naremore (1980 as stated in Paul, 2007), also called *communication-language approach* (Lahey, 1988), a *descriptive* view of language leads us to obtain a representative language sample which can be comprehensively analysed in terms of different aspects of child linguistic capabilities (Paul, 2007). Language sample analysis, as reviewed above, provides us with the method best suited to explain different linguistic features of child language (Eisenberg et al., 2001) which is obtainable in both structured and natural environments. LSA can be performed through several procedures such as calculating MLU, LARSP, Index of Productive Syntax (IPSyn) (in the domain of grammar), Number of Different Words (NDW) and many other measures of interest to researchers, some of which may be combined to boost the clinical diagnostic accuracy of measures; a property which makes this procedure attractive.

Several of the aforementioned procedures with fair to good diagnostic features confirmed by the meta-analysis of previous studies (see Table 2), as well as those capable of differentiating between children with TDL and with PLI in the phase I DA of the current study, will be observed in the phase II DA in searching for their diagnostic accuracy in Persian. A survey of Iranian speech and language therapists will support the phases of the DA studies so as to demonstrate Iranian SLTs' current assumptions about the assessment of PLI, which have been considered as the reference standard for the DA studies. The next chapter explains the process of finding a reference standard from Iranian SLTs' child assessment procedures.

CHAPTER 3

Survey of methods used by Iranian speech and language therapists to evaluate child language

3.1 Introduction

Due to the limited access to reliable assessment tools in Iran, speech therapists use different, personally-developed assessment methods to identify language impaired children. This causes problems both for intervention and research purposes due to the subjectivity of the assessment procedures and bias in clinical judgement. This type of implicit, individualized-type decision making can be seen in the profession of speech and language therapy as a whole; several researchers (Lof, 2011; Lyons et al., 2008; Roulstone, 2001; Thordardottir et al. 2011) have attempted to raise awareness about the need for change to a more explicit condition in which the therapists have access to more objective and reliable sources of materials in their clinical decision making. Moreover, clinicians have shown that sometimes there is a gap between what have been developed as child language tests and what they internally experience in their own work (Lyons et al., 2008), i.e. children's general behaviour in clinical settings contradicts the background criteria essential for administrating a structured language test. This contradiction makes it impossible to make a reasonable judgement about a child's language behaviour.

Clinical judgement and agreement have been considered as an acceptable reference in clinical studies, including both diagnosis and intervention, when the sources of assessment are restricted (Joffe & Pring, 2008; Kemp & Klee, 1997; Lyons et al., 2008; Peña, Reséndiz, & Gillam, 2007; Roulstone, 2001). Among the first attempts to study clinical judgment in speech and language therapy was Allen et al.'s study in which they evaluated the agreement between SLTs' clinical judgment and the results of three standardised language tests in 182 pre-school children (Allen, Bliss, & Timmons, 1981). They claimed that clinical judgment can serve "as a possible defence against test bias" (p. 66) compared to test results, which might penalise the child in the presence of factors like behavioural difficulties. They emphasized, though, that neither clinical judgment nor test results should be considered separately since both of them include sources of bias. For the former, this bias comes from SLTs' "internal norms derived from their experiences with language behaviour" (Allen et al., 1981 p.68).

In a similar study (Aram, Morris, & Hall, 1993), the congruence between clinical diagnoses based on SLTs' judgement compared with a measure based on discrepancy criteria for SLI/PLI showed similar results in "potential mismatch" (p.588) between these two approaches in identifying children with SLI/PLI. Different professionals (speech-language pathologists, psychologists, neurologists, paediatricians, and

psychiatrists with expertise in speech and language) providing intervention for children with SLI/PLI were asked to refer 252 children (ages ranging from 3;0 to 5;11 years;months) for a further evaluation of the children's language abilities (to determine SLI/PLI). The results indicated that the range of congruence between clinical judgement and a set of discrepancy criteria was wide, ranging from 20% to 71.4% (based on alternative language measures of *Illinois Test of Psycholinguistic Abilities* - Auditory Assessment, and MLU age, respectively) and the authors finally raised the issue of providing an operational definition for SLI/PLI as well as considering a "dual-criteria definition of SLI" (p.589) to bring clinical and research definitions of SLI/PLI closer together. Implementing a further source of assistance in diagnosing language problems, Glascoe (1991) reported that parents' concern correctly identified 72% of children with speech-language problems from 157 children. These two studies also included parents' concerns about children's speech and language development in their clinical judgement measures and emphasised that more investigation would help improve the role of parents in detecting child language problems (Aram et al., 1993; Glascoe, 1991). Bishop and McDonald also conducted a similar study including parents' reports about their children's language behaviour through Children's Communication Checklist-2 (CCC-2) into the clinical diagnosis and found that parents' opinion as "complementary information" would benefit clinical diagnosis in combination with test results (Bishop & McDonald, 2009 p. 600).

Lyons et al (2008) asked Irish SLTs to describe their assessment procedure when identifying children suspected of having SLI/PLI. Questionnaire interviewing and focus groups were used to collect data in a mixed-method study emphasising the complementary roles of quantitative and qualitative methods despite their different origins in terms of ontology and epistemology (Lyons et al., 2008). Questionnaire interviews of 349 SLTs addressed the quantitative phase and focus group data collection of 10 SLTs served the qualitative phase. The results showed that this qualitative-quantitative mixed method was successful in describing Irish SLTs' diagnostic procedures in finding children with SLI. Irish SLTs were found to base their clinical decision making on three assessment factors: inclusionary criteria, exclusionary criteria and qualitative markers. The most frequently reported factors included morphological problems, word order problems, word finding difficulties, and difficulty with relational concepts (Lyons et al., 2008 p. 433).

3.2 Specific aims of the survey

This survey explores the case definitions and diagnostic criteria employed by SLTs working in Iran for defining childhood language impairment. One aim of this is to describe how children with language difficulties are assessed in Iran; another aim is to inform the development of a clinically realistic reference standard that can be used in a study to be reported subsequently, relating to the diagnostic accuracy of selected language sample measures. The specific questions are:

1. How do Iranian SLTs assess and identify children with PLI?
2. To what extent do Iranian SLTs use language sampling and language sample measures (LSMs)?

3.3 Method

A mixed method, qualitative-quantitative approach was chosen to address the research questions because the nature of the case was new, un-researched and context-dependent in Iran. Furthermore, the numerical data from the quantitative part would fortify the results of the qualitative data and help modify the final model of language assessment generated, by maximising the number of participants involved. As the researcher intended not to interfere with the SLTs' opinions, the qualitative method seemed to suit this purpose, too. Mixed methods triangulate data collection by utilising qualitative and quantitative procedures and have been reported as being appropriate to approach new subjects and new communities and to gain a better picture of what is happening within the context with respect to the variables of the study (Creswell, Shope, Clark, & Green, 2006; Damico & Simmons-Mackie, 2003; Lyons et al., 2008; Marshall, Goldbart, & Phillips, 2007).

Two methods of data collection, questionnaire interviewing and focus groups, were applied to address data triangulation by looking at the issue from different angles to describe it as comprehensively and completely as possible. Although triangulation, methodologically, contains a concurrent analysis (Roulstone, 1997), a similar sequential data analysis as described in Roulstone (1997) was performed due to the similarity with the nature of the data collection in the present survey. Roulstone's schematic study methodology has been included in the diagram of participant recruitment to show how each part of the process supplies the other parts (Figure 3-1).

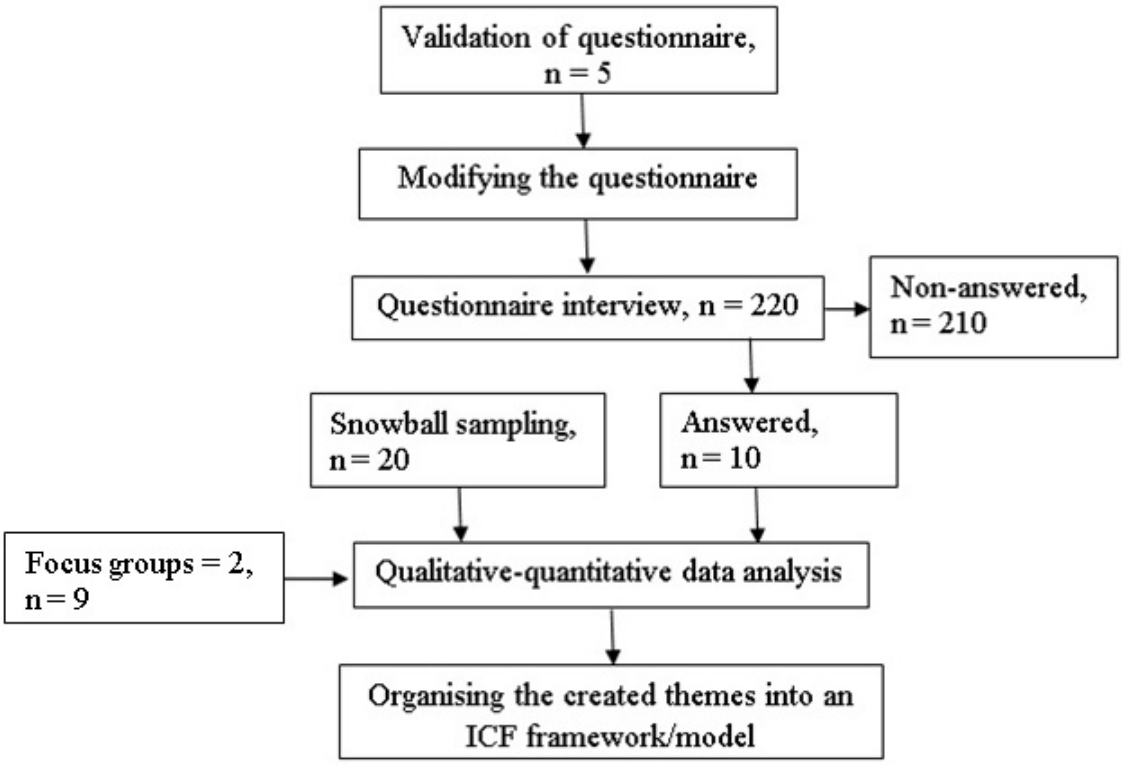


Figure 3-1 Schematic of recruiting SLTs to participate in the survey

3.4 Questionnaire interview

No published studies of Iranian SLTs were found in terms of their views of communication problems, nor involvement in improving intervention policies specifically for Iran. Questionnaire interviewing is a well-known procedure in qualitative studies which is suitable for approaching unstudied subjects in unstudied populations like Iranian speech therapists. This method was therefore selected to pool Iranian SLTs' opinions about their caseload management with regard to child language disorders.

The specific objectives of this survey were:

1. Defining the demographic characteristics of Iranian SLTs including age, gender, socioeconomic status (SES), and academic level (Appendix A-2, questions 1 to 6)
2. Defining the caseload status of communication disorders of Iranian SLTs, including size of caseload, referral status, and diagnosis status (Appendix A-2, questions 7 to 10).

3. Defining the assessment areas for identifying child language impairment and their importance/priority in the assessment process (Appendix A-2, questions 11 and 12).
4. Defining the status of formal and informal language assessment methods, including the extent to which Iranian SLTs use standardised assessments as well as language sample analysis (Appendix A-2, questions 13 to 23).
5. Defining therapist-specified assessment procedures in the diagnosis of child language impairment (Appendix A-2, questions 24 to 26).

3.4.1 Questionnaire design

The initial draft of the questionnaire was designed to address the main goals of the survey, including questions on what and how Iranian SLTs consider when they plan to assess a child suspected of language impairment with no associated problems, PLI. The content was an organised list of those areas that Iranian SLTs usually study in university and was mainly based on the researcher's experience in teaching related courses. Demographic data was targeted in several introductory questions including gender, age, socioeconomic status (SES) of the region they serve as well as caseload and the referrals they have. The SLTs were asked to mention if they use standardised/formal assessments, what and why, as well as using informal assessment, particularly in the format of language sample analysis along with mentioning which type of analysis and why. Some questions were adapted from Kemp and Klee's questionnaire (1997), taking into account the working context of Iranian clinicians.

After reviewing and modifying the first draft, five Iranian SLTs with postgraduate degrees (all with MAs and one with a PhD, and all being lecturers at different universities), were sent a second draft and asked to point out any ambiguity with regard to the structure and meaning of the questions. The final questionnaire was prepared after reviewing and taking account of their comments.

The questionnaire aimed to be brief and not take more than 20 minutes of the respondent's time in an attempt to encourage busy SLTs to complete and return the form. It included a cover letter summarising the survey goals at the beginning and assuring the respondents of anonymity apart from their age and gender as well as stating that in case of any uncertainty, they would be able to contact the researcher through the contact details provided. The questionnaire started with 25 closed-response questions, including multiple-choice, ranking, selecting the most appropriate items from a list, and short answer. It ended with three open-ended questions asking the SLTs to outline any

specific features with which they decide a child is language impaired, anything which was missed in the list of items assessed, and their own personal assessment procedure in the child language domain (The questionnaire is in Appendix A).

3.4.2 Participants and procedures

The target population was all Iranian SLTs registered with the Iran Medical Council (IRIMC). Statistics showed that 911 SLTs were registered in May 2009. Of those, 220 were recruited to the study during the summer of 2009. After getting permission from Isfahan University of Medical Sciences and travelling to Tehran, IRIMC was asked to provide a list of SLTs working throughout Iran. According to the organization's rules, researchers are not allowed to have access to the personal data of professionals, hence the random sampling of the SLT population was done at the IRIMC and 220 out of 911 registered SLTs (by summer 2009) were selected on site using a website producing random numbers for studies (http://www.randomizer.org/form.htm). IRIMC did not permit a larger number of SLTs to be sampled. Since the IRIMC regulations also limited access to SLT's personal data (i.e. name, address and phone numbers), the researcher was asked to provide the required number of questionnaires, envelopes and pre-paid returned envelopes to the organization, who in turn then sent them to respondents.

Because of this, the process of following up the recruited people to achieve the predicted better response rate was not possible. In an attempt to improve the outcome, the time limit for receiving the completed questionnaires was set to be open-ended. A total of only 10 completed questionnaires were received within six months (4.5%) and 36 questionnaires were returned undelivered (16.4%). Consequently, snowball sampling was performed by contacting SLTs using email addresses available to the researcher. SLTs were requested to hand the questionnaire to other SLTs; the final number of completed questionnaires recruited from snowball sampling was 20, giving a total of 30 questionnaires to analyse.

3.4.3 Results

Results are described corresponding to the above-mentioned objectives.

1. Demographic specification of Iranian SLTs

The demographic specifications of the sample population are indicated in tables 3-1 and 3-2.

Age in years (n=29)			Gender (n=30)	
Mean	**SD**	**Range**	**Female**	**Male**
31;5	6	24 (24-48)	26 (86.7%)	4 (13.3%)

Table 3-1 Demographic data of sampled Iranian SLTs.

Length of clinical work in years (n=30)			Academic level (n= 30)		SES of workplace (n= 30)		
Mean	**SD**	**Range**	**BSc.**	**MSc.**	**Below average**	**Average**	**Above average**
7;9	5;13	18 (2-20)	14 (46.7%)	16 (53.3%)	3 (10%)	18(60%)	9 (30%)

Table 3-2 Length of work in years, academic level and Socio-Economic Status (SES) of workplace in a sample population of Iranian SLTs.

As seen in table 3-1, females outnumbered males by 26 (86.7%) to four (13.3%) which is expected due to the higher acceptance rate of females compared to males in Iranian universities on courses in speech therapy.

Table 3-2 shows that clinical experience had an average length of seven years and nine months (SD=5;13) and ranged between two years and 20 years which shows an acceptable range of experience among the population. If the minimum student apprenticeship, i.e. two years, is added to this period of time, it can be expected that the participants with the lowest period of clinical work would have gained an acceptable number of years of experience in managing intervention. Almost half of the participants had a BSc. degree (46.7%); a greater number of them graduated with a MSc. (53.3%).

SES is operationally defined as SLTs' judgement about the condition of their workplace, socially and economically. Sixty per cent of SLTs evaluated the SES of their workplace as being average which is not able to be verified by external sources; however, it is expected based on the researcher's personal experience and is compatible with studies in non-Iranian populations (Kemp & Klee, 1997).

2. The caseload status of communication disorders with Iranian SLTs.

The SLTs' caseload statistics showed that pre-school children (between three and six years old) were the largest group referred to SLTs (M=31.25), followed by school-aged children (older than six years old) (M=16.35) and infants and toddlers (younger than three years old) (M=8.96). This was the reason for choosing the age range of children to participate in the subsequent diagnostic accuracy (DA) study of children between 3;6 and 4;6 (Table 3-3).

Number of cases younger than 3 years old (n=27)			Number of cases between 3 and 6 years old (n=28)			Number of cases between 6 and 12 years old (n=28)		
Mean	**SD**	**Range**	**Mean**	**SD**	**Range**	**Mean**	**SD**	**Range**
9	10	36 (0-36)	31	56	300 (0-300)	16	22	100 (0-100)

Table 3-3 The size of SLTs' caseload within 3 months ending on the date of responding

The main reason for referrals to SLTs was late-talking with average ranking score of 4.41 out of five, then pronunciation problems with average rating of four out of five followed by language difficulty (3.67 out of five), stuttering (3.58 out of five), communication problems (2.5 out of five), memory problems (two out of five) and voice problems (1.32 out of five) (Table 3-4).

Referral reason (n=respondents)	Mean	Mode	Range
Late-talking (n=27)	4.41	5	2 (3-5)
Pronunciation problems (n=27)	4	5	3 (2-5)
Language difficulty (n=27)	3.67	5	4 (1-5)
Stuttering (n=26)	3.58	3	4 (1-5)
Communication problems (n=26)	2.5	1	4 (1-5)
Memory problems (n=25)	2	1	4 (1-5)
Voice problems (n=22)	1.32	1	4 (1-5)

Table 3-4 The reasons for referrals to SLTs in preschool age range, ranked from 1=the least referrals to 5=the most referrals

A confirmation question showed that the percentage of referrals due to late-talking and pronunciation problems were 37.42% and 34.88%, respectively. This percentage was 30.16% for language difficulties followed by stuttering (25.56%), communication problems (17.47%), memory problems (10.1%) and voice problems (3.7%) (table 3-5). As observed, the percentage of referrals confirms SLTs' ranking of referrals.

Percentage of referral reason (n=respondents)	Mean (SD)	Minimum	Maximum
Late-talking (n=26)	37.42(20.8)	10	85
Pronunciation problems (n=25)	34.88(21.2)	8	90
Language difficulty (n=26)	30.61(22.8)	2	80
Stuttering (n=25)	25.56(14.1)	10	50
Communication problems (n=23)	17.47(17.8)	0	60
Memory problems (n=22)	10.09(15.3)	0	50
Voice problems (n=20)	3.7(4.8)	0	20

Table 3-5 The percentage of referrals due to each reason

Table 3-6 shows the percentage of actual diagnosis by SLTs out of the referrals with corresponding age which means that if children are referred to SLTs as late-talkers, approximately 30-31% of them would be actually diagnosed by the SLT as 'late-bloomers' or 'language impaired (LI) due to developmental delay' at average ages of 2.60 and 3.54 years old, respectively. Twenty seven per cent would be diagnosed as having 'articulation disorders' at age of 3.92 years old, followed by 'LI due to hearing problems' (23.25%, mean diagnostic age=2.77 years) and 'LI due to autism spectrum disorders (ASD)' (12.44%, mean diagnostic age=3.15). The least diagnosed group is 'LI due to non-specific reasons (or PLI/SLI)' by 16.03% of actual diagnosis by SLTs at age of 4.28 years old.

Diagnosis (n=respondents)	**Mean (SD)**	**Minimum**	**Maximum**
Language disorder due to mental retardation (n=28)	31 (23.4)	0	90
Age in years (n=14)	3.5 (2.4)	1	10
Late-blooming or late-talking (n=29)	30 (19.1)	0	85
Age in years (n=17)	2.6 (0.7)	1	4
Articulation disorder (n=28)	27 (21.8)	0	85
Age in years(n=19)	3.9 (1)	3	6
Language disorder due to hearing impairment (n=27)	23 (20.3)	0	80
Age in years (n=13)	2.8 (1.5)	1	6
Language disorder without specific reason (SLI) (n=26)	16 (14.9)	0	50
Age in years (n=16)	4.3 (2)	3	10
Language disorder due to ASD (n=27)	12 (14)	0	50
Age in years (n=13)	3.1 (0.8)	2	5

Table 3-6 Percentage of actual diagnosis by SLTs out of the referrals with corresponded age; ASD=Autism spectrum disorder

3. What SLTs include in their assessment to help them identify child language impairment with respect to importance and priority.

Question 11 asks which areas of assessment SLTs address in their evaluation procedures to help them in identifying language impairment. The 'procedures' by which SLTs usually evaluate each area of assessment were also investigated by this question. It seems that practitioners' conceptualisation of their personal assessment process has led them, in some cases, to select a procedure which is incongruous to the nature of the corresponding area of assessment. For instance eight SLTs have chosen parent interview as the procedure of assessment of pure tone screening which is not a sensible approach (Table 3-8). They might have conceptualised that the results of a pure tone audiometry can be retrieved by asking parents (or by observing the report from audiologist or by informal assessment by an audiologist). A further reason for observing

this kind of inconsistency might be a validity bias in questionnaire designing insofar as the labels are misleading or simply the long list of items has led the respondent not to pay careful attention to the content. Tables 3-7 to 3-10 show a descriptive picture of their responses. As can be seen, all the 40 assessment areas mentioned in the questionnaire were selected by the SLTs as areas that they would evaluate if they suspected a child to be language impaired.

Child History	Number (Per cent)	Child History	Number (Per cent)	Child history	Number (Per cent)
Gender	25(83.2%)	***Bilingualism***	26(86.7%)	Checklist	3 (10%)
Parent report or interview	11 (36.7%)	Parent report or interview	26 (86.7%)	Informal assessment	4 (13.3%)
Observation	16 (53.3%)	Observation	9 (30%)	Colleagues' judgment	1 (3.3%)
Family history	28 (93.2%)	Checklists	3 (10%)		
Parent report or interview	28 (93.2%)	Informal assessment	2 (6.7%)	***Social interaction with parents and peers***	30(100%)
Checklist	2 (6.7%)	Other procedures	1 (3.3%)	Observation	21 (70%)
Medical history	28(93.2%)	***Language development***	29(96.7%)	Checklist	2 (6.7%)
Parent report or interview	28 (93.2%)	Parent report or interview	29 (96.7%)	Informal assessment	7 (23.3%)
Observation	2 (6.7%)	Observation	10 (33.3%)	Colleagues' judgment	2 (6.7%)
Checklist	4 (13.3%)	Checklist	4 (13.3%)	Other procedures	1 (3.3%)
Informal assessment	2 (6.7%)	Informal assessment	7 (23.3%)	***History of attending nursery and duration***	24 (79.9%)
Colleagues' judgment	3 (10%)	***Pretend play***	27(89.9%)	Parent report or interview	25(83.3%)
Other procedures	1 (3.3%)	Parent report or interview	22 (73.3%)	Observation	3 (10%)
History of Otitis Media	27 (89.9%)	Observation	20 (73.3%)	Checklist	1 (3.3%)
Parent report or interview	25 (83.3%)	Informal assessment	14 (46.7%)	Colleagues' judgment	1 (3.3%)
Observation	1 (3.3%)	Colleagues' judgment	1 (3.3%)	***Parents' educational level***	25(83.3%)
Checklist	3 (10%)	Other procedures	1 (3.3%)	Parent report or interview	26(86.7%)
Informal assessment	1 (3.3%)	***Quality and quantity of Language stimulation in the environment***	30(100%)	Checklist	3 (10%)
Standardized tests	1 (3.3%)	Parent report or interview	29 (96.7%)		
Colleagues' judgment	7 (23.3%)	Observation	7 (23.3%)		

Table 3-7 The areas Iranian SLTs usually consider in **case histories** to assess language impairments in pre-school children with a breakdown of the procedures they undertake in their assessment. Respondents could select more than one response for each item.

	Number (Per cent)		Number (Per cent)		Number (Per cent)
Hearing Status	**28 (93.3%)**	**Cognition**	**29 (96.6%)**	**Neurological status**	**28 (93.2%)**
Pure tone screening		***IQ estimation***		***Memory assessment***	
Parent report or interview	8 (26.7%)	Parent report or interview	6 (20%)	Parent report or interview	7 (23.3%)
Observation	2 (6.7%)	Observation	6 (20%)	Observation	6 (20%)
Checklist	3 (10%)	Checklist	2 (6.7%)	Checklist	3 (10%)
Informal assessment	7 (23.3%)	Informal assessment	6 (20%)	Informal assessment	22 (73.3%)
Standardized tests	4 (13.3%)	Standardized tests	9 (30%)	Standardized tests	5 (16.7%)
Colleagues' judgment	16 (53.3%)	Colleagues' judgment	15 (50%)	Colleagues' judgment	2 (6.7%)
Whispering test		***Play assessment***		***Neurologist's referral letter***	
Parent report or interview	1 (3.3%)	Parent report or interview	13 (43.3%)	Parent report or interview	10 (33.3%)
Observation	1 (3.3%)	Observation	20 (66.7%)	Observation	7 (23.3%)
Checklist	2 (6.7%)	Checklist	3 (10%)	Checklist	3 (10%)
Informal assessment	5 (16.7%)	Informal assessment	16 (53.3%)	Informal assessment	5 (16.7%)
Colleagues' judgment	10 (33.3%)	Colleagues' judgment	7 (23.3%)	Standardized tests	1 (3.3%)
PTA (Threshold test)		***Painting assessment***		Colleagues' judgment	13 (43.3%)
Parent report or interview	3 (10%)	Parent report or interview	10 (33.3%)	***Test of motor skills (fine and gross movements)***	
Informal assessment	1 (3.3%)	Observation	14 (46.7%)	Parent report or interview	12 (40%)
Standardized tests	7 (23.3%)	Checklist	1 (3.3%)	Observation	18 (60%)
Colleagues' judgment	14 (46.7%)	Informal assessment	17 (56.7%)	Checklist	5 (16.7%)
SRT		Standardized tests	6 (20%)	Informal assessment	20 (66.7%)
Parent report or interview	2 (6.7%)	Colleagues' judgment	8 (26.7%)	Standardized tests	5 (16.7%)
Standardized tests	5 (16.7%)			Colleagues' judgment	8 (26.7%)
Colleagues' judgment	13 (43.3%)				

Table 3-8 The areas Iranian SLTs usually consider in **hearing status, cognition and neurological status** to assess language impairments in pre-school children with a breakdown of the procedures they undertake in their assessment. Respondents could select more than one response for each item.

	Number (Per cent)		Number (Per cent)		Number (Per cent)
Oro-motor development	**26 (86.6%)**	**Pre-verbal skills**	**28(93.2%)**	***Semantic relations***	28(93.2%)
Parent report or interview	13 (43.3%)	Parent report or interview	13 (43.3%)	Parent report or interview	11 (36.7%)
Observation	16 (53.3%)	Observation	15 (50%)	Observation	6 (20%)
Checklist	4 (13.3%)	Checklist	7 (23.3%)	Checklist	7 (23.3%)
Informal assessment	15 (50%)	Informal assessment	25 (83.3%)	Informal assessment	21 (70%)
Standardized tests	6 (20%)	Standardized tests	3 (10%)	Standardized tests	4 (13.3%)
Colleagues' judgment	2 (6.7%)	Colleagues' judgment	1 (3.3%)	***Syntax***	28(93.2%)
Other procedures	1 (3.3%)			Parent report or interview	8 (26.7%)
Language processing	**26(86.6%)**	**Receptive language**	**30 (100%)**	Observation	6 (20%)
Parent report or interview	8 (26.7%)	***Phonological awareness***	26	Checklist	6 (20%)
Observation	8 (26.7%)	Parent report or interview	6 (20%)	Informal assessment	20 (66.7%)
Checklist	6 (20%)	Observation	7 (23.3%)	Standardized tests	5 (16.7%)
Informal assessment	23 (76.7%)	Checklist	4 (13.3%)	***Morphology***	27 (89.9%)
Standardized tests	3 (10%)	Informal assessment	18 (60%)	Parent report or interview	7 (23.3%)
Colleagues' judgment	1 (3.3%)	Standardized tests	6 (20%)	Observation	5 (16.7%)
Non-word repetition tasks	**26(86.6%)**	Colleagues' judgment	1 (3.3%)	Checklist	7 (23.3%)
Parent report or interview	5 (16.7%)	***Vocabulary***	29 (96.6%)	Informal assessment	19 (63.3%)
Observation	8 (26.7%)	Parent report or interview	16 (53.3%)	Standardized tests	5 (16.7%)
Checklist	4 (13.3%)	Observation	8 (26.7%)	***Conversational rules***	27(89.9%)
Informal assessment	20 (66.7%)	Checklist	10 (33.3%)	Parent report or interview	8 (26.7%)
Standardized tests	4 (13.3%)	Informal assessment	24 (80%)	Observation	5 (16.7%)
Colleagues' judgment	1 (3.3%)	Standardized tests	4 (13.3%)	Checklist	7 (23.3%)
				Informal assessment	20 (66.7%)
				Standardized tests	3 (10%)

Table 3-9 The areas Iranian SLTs usually consider in **receptive language and areas associated with communication** to assess language impairments in pre-school children with a breakdown of the procedures they undertake in their assessment. Respondents could select more than one response for each item.

	Number (Per cent)		Number (Per cent)		Number (Per cent)
Expressive language	**30 (100%)**	***Vocabulary***	30 (100%)	***Conversational rules***	28(93.2%)
Joint attention	26 (86.6%)	Parent report or interview	16 (53.3%)	Parent report or interview	7 (23.3%)
Parent report or interview	8 (26.7%)	Observation	7 (23.3%)	Observation	10 (33.3%)
Observation	18 (60%)	Checklist	8 (26.7%)	Checklist	4 (13.3%)
Checklist	2 (6.7%)	Informal assessment	21 (70%)	Informal assessment	21 (70%)
Informal assessment	17 (56.7%)	Standardized tests	4 (13.3%)	Standardized tests	2 (6.7%)
Standardized tests	1 (3.3%)	Other procedures	1 (3.3%)	***Narrative and reasoning skills***	30 (100%)
Colleagues' judgment	1 (3.3%)	***Semantic relations***	28 (93.2%)	Parent report or interview	9 (30%)
Use of gestures	25 (83.2%)	Parent report or interview	7 (23.3%)	Observation	9 (30%)
Parent report or interview	9 (30%)	Observation	7 (23.3%)	Checklist	6 (20%)
Observation	19 (63.3%)	Checklist	6 (20%)	Informal assessment	25 (83.3%)
Checklist	2 (6.7%)	Informal assessment	21 (70%)	Standardized tests	2 (6.7%)
Informal assessment	15 (50%)	Standardized tests	4 (13.3%)	Other procedures	1 (3.3%)
Standardized tests	1 (3.3%)	***Syntax***	29 (96.6%)		
Phonetic inventory	27 (89.9%)	Parent report or interview	7 (23.3%)		
Parent report or interview	13 (43.3%)	Observation	5 (16.7%)		
Observation	6 (20%)	Checklist	4 (13.3%)		
Checklist	7 (23.3%)	Informal assessment	23 (76.7%)		
Informal assessment	18 (60%)	Standardized tests	5 (16.7%)		
Standardized tests	4 (13.3%)	***Morphology***	28(93.2%)		
		Parent report or interview	5 (16.7%)		
		Observation	5 (16.7%)		
		Checklist	3 (10%)		
		Informal assessment	22 (73.3%)		
		Standardized tests	3 (10%)		

Table 3-10 The areas Iranian SLTs usually consider in **expressive language** to assess language impairments in pre-school children with a breakdown of the procedures they undertake in their assessment Respondents could select more than one response for each item.

SLTs were also asked to rank these areas on a five-point Likert scale (one, very low importance; five, very high). The majority of SLTs ranked three areas as being not a high priority in their assessment: History of Otitis Media, Bilingualism, and Parents' education. These areas were ranked less than three on the five-point Likert scale, on average. Eighteen other areas received a priority of high and very high, on average, i.e. three or more on the Likert scale. They are as follows ordered by their mean score: Receptive vocabulary, Pragmatics, Syntax and Expressive vocabulary, Morphology, Cognition status and Semantic relations, Language processing, Diagnostic therapy, Hearing test, Neurology status, Memory test, Non-word repetition, Pre-verbal skills, Play assessment, Language stimulation, Social interaction, Oro-motor tests, and Family history (Table 3-11). Figure 3-2 visualises this ordering by mean of ranks and the corresponding confidence interval.

Area of assessment	**Number of respondents**	**Minimum**	**Maximum**	**Mean rank (SD)**
Receptive-expressive vocabulary	27	3	5	4.20 (.79)
Pragmatic skills	27	2	5	4.19 (.87)
Expressive vocabulary	25	2	5	4.18 (.88)
Receptive-expressive syntax	28	2	5	4.18 (.82)
Receptive-expressive morphology	28	2	5	4.14 (.87)
Receptive-expressive semantics	27	2	5	4.02 (.97)
Cognition status	27	1	5	4.02 (.96)
Language Processing	26	2	5	4.00 (.98)
Diagnostic intervention	13	2	5	3.61 (.77)
Hearing test	27	1	5	3.56 (1.31)
Neurological status	26	1	5	3.52 (1.28)
Memory assessment	27	1	5	3.50 (.97)
Non-word repetition	24	1	5	3.42 (1.06)
Preverbal skills	27	1	5	3.39 (1.23)
Play assessment	27	1	5	3.37 (1.02)
Language stimulation in environment	26	1	5	3.23 (1.21)
Social interaction	28	1	5	3.16 (1.31)
Oro-motor development	27	1	5	3.15 (1.23)
Family history of similar problem	27	1	5	3.09 (1.27)
Parents education	26	1	5	2.38 (1.13)
Bilingualism	26	1	4	2.10 (1.07)
History of Otitis Media	27	1	4	2.04 (.90)

Table 3-11 Areas of assessment ordered by the priority mean rank scored by SLTs

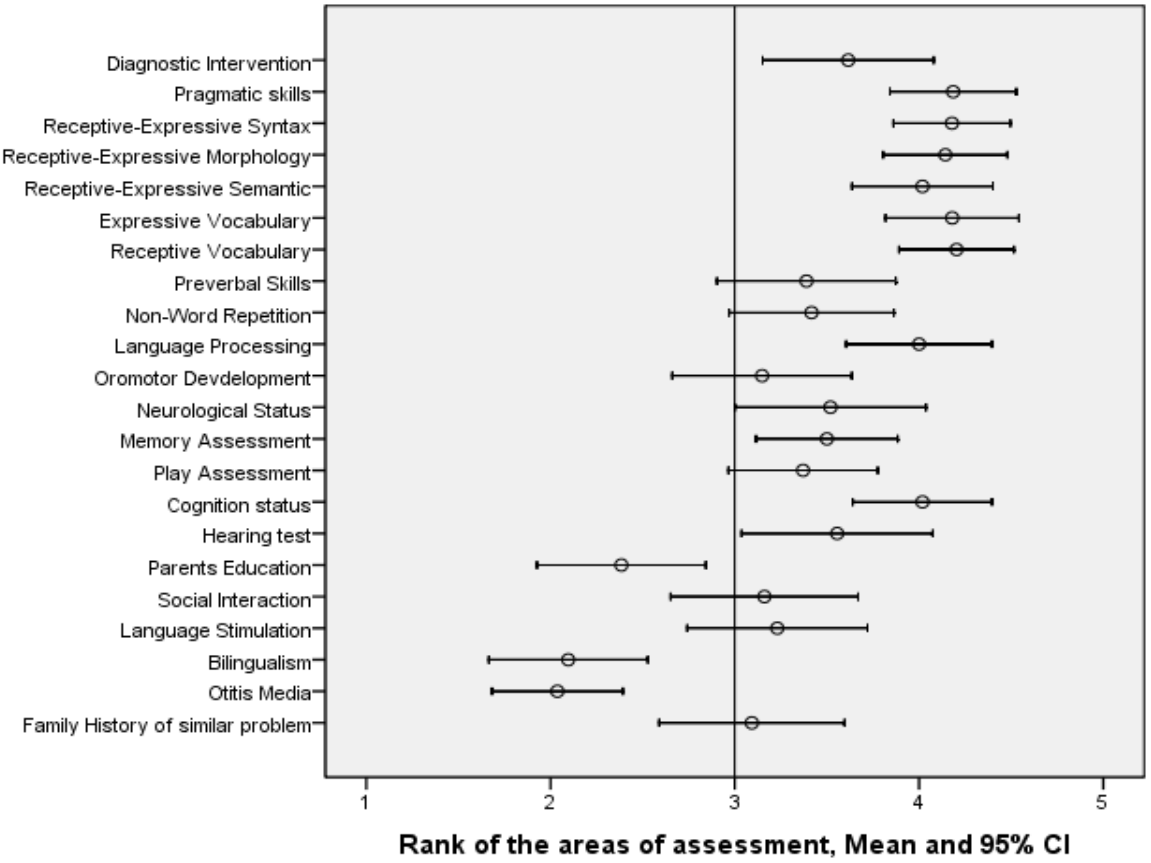


Figure 3-2 Areas ranked by Iranian SLTs as the most helpful in identifying children with PLI, ordered by mean of ranks and 95% confidence intervals shown in horizontal axis

4. Status of formal and informal language assessment.

The majority of Iranian SLTs (n=23, 76.6%) do not use standardised assessments in their clinical assessment plans; versus 13.3% of them (n=7) who reported using them (Figure 3-3).

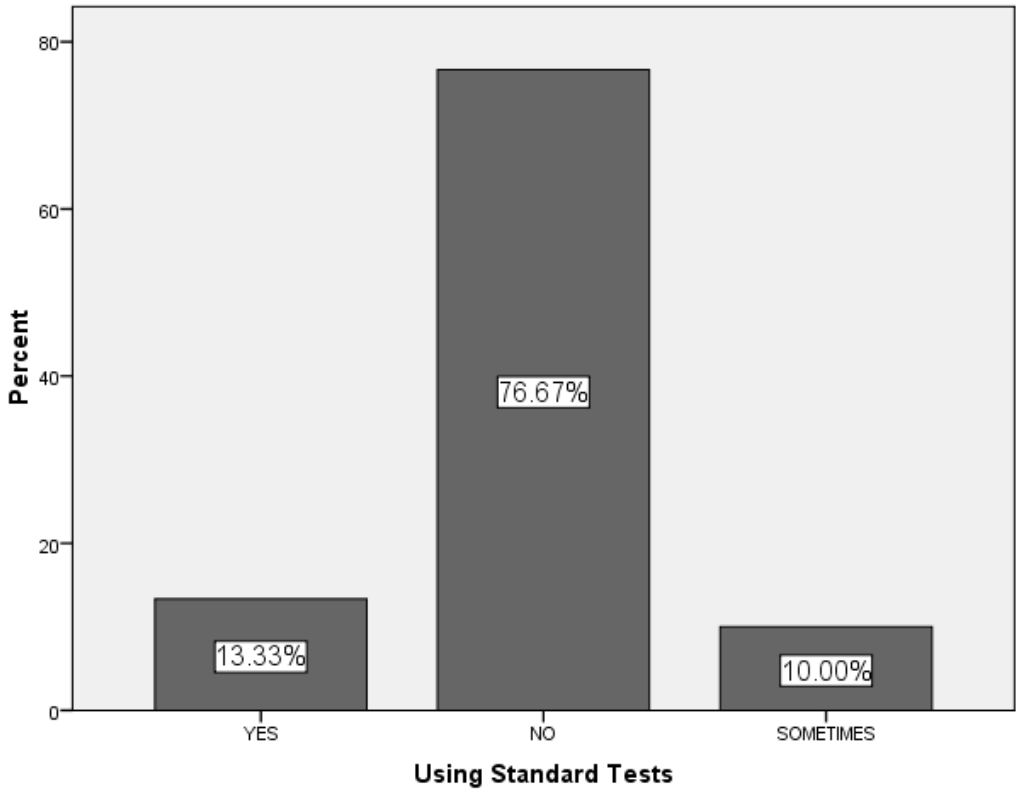


Figure 3-3 Percentage of Iranian SLTs who reported using standardised tests (n=30).

Nine SLTs chose the reason for using standardised tests as ‘Differential Diagnosis + Diagnosis’ (n=4, 13.3%), ‘Only Diagnosis’ (n=2, 6.67%) followed by ‘Screening’, ‘Screening + Differential Diagnosis + Diagnosis’ and ‘all options’ (n=1 for each category, 3.3%) (Figure 3-4).

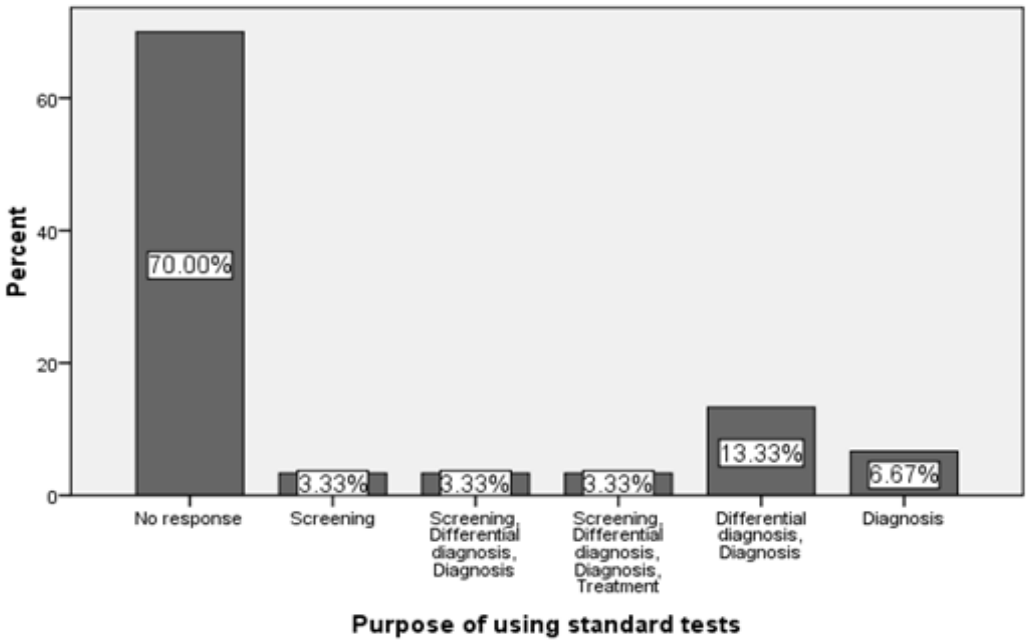


Figure 3-4 Purposes of using standardised assessment by Iranian SLTs (n=30).

Among the reasons for not using standardised tests in their assessments, 60% of SLTs indicated the lack of availability of such tests in Iran (Figure 3-5).

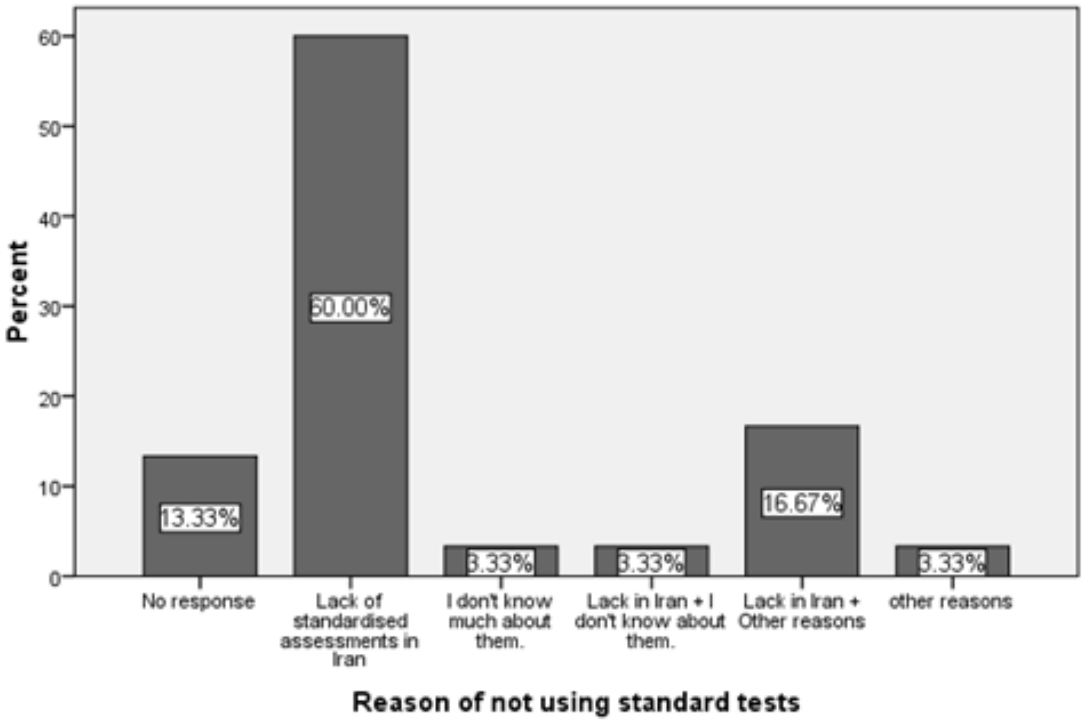


Figure 3-5 Reasons for not using standardised assessment by Iranian SLTs (n=30).

Eighty per cent of Iranian SLTs sampled (n=24) reported using language sample analysis to assess children in their evaluation plans. If we add those who sometimes use LSA, this increases to 83.33% (n=25) (Figure 3-6).

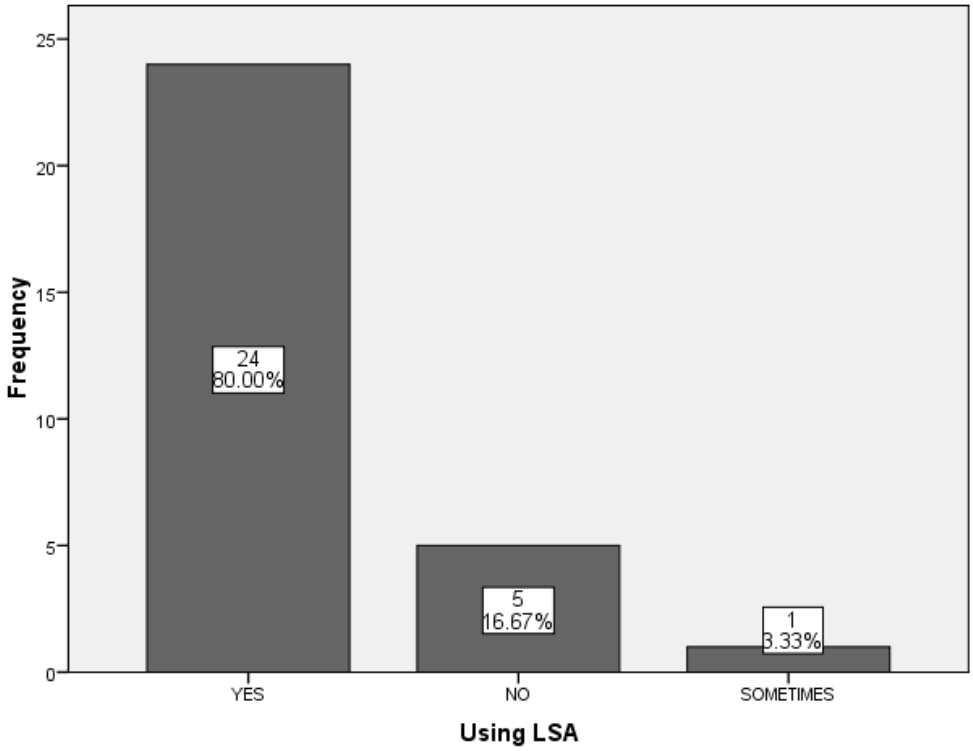


Figure 3-6 Status of using language sample analysis in assessment by Iranian SLTs (n=30).

5. The status of language sample analysis and its related measures

Providing an operational definition, language sampling for Iranian SLTs is a general term they use to refer to all types of collecting samples of child's language, including those procedures mentioned in the question. Imitation tasks elicit language samples by providing a model for the child to repeat, e.g. the simplest way is: 'say what I say'. Elicitation tasks are known as "elicited production" (Paul, 2007 p.50) in which unlike imitation tasks, the child is directed to the target form or content by providing some indirect cues (Paul, 2007). In natural language sampling, child's spontaneous language production is collected when she is communicating in a real communicative environment with real communicative partners. In question 17, SLTs are asked to specify their elicitation procedure in collecting language samples. Forty per cent (n=12) of SLTs employ a combination of 'imitation tasks + elicitation tasks + natural language sampling' as their method in collecting language samples, followed by 23.33% (n=7) who use the combination of 'elicitation tasks + natural sampling', and 13.33% (n=4) who use only 'elicitation tasks'. Less than 10% of SLTs use either 'imitation tasks + elicitation tasks' or merely 'natural language sampling' (n=2 for each category, 6.67%) (Figure 3-7).

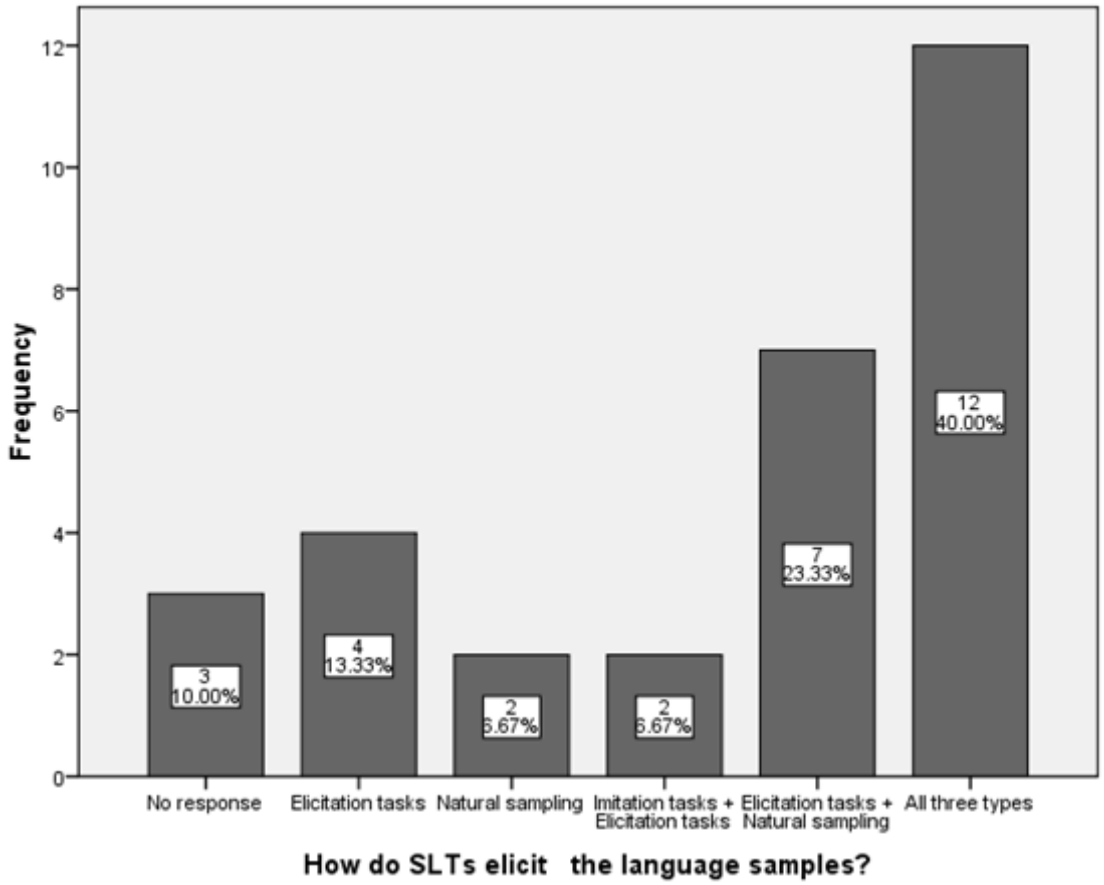


Figure 3-7 Elicitation procedures used by Iranian SLTs for language sampling (n=30).

A total of 43.33% of SLTs (n=13) believe that natural language sampling is the best procedure to elicit child language for diagnostic purposes, followed by 36.67% (n=11) who believe that elicitation tasks are more helpful in this case (Figure 3-8).

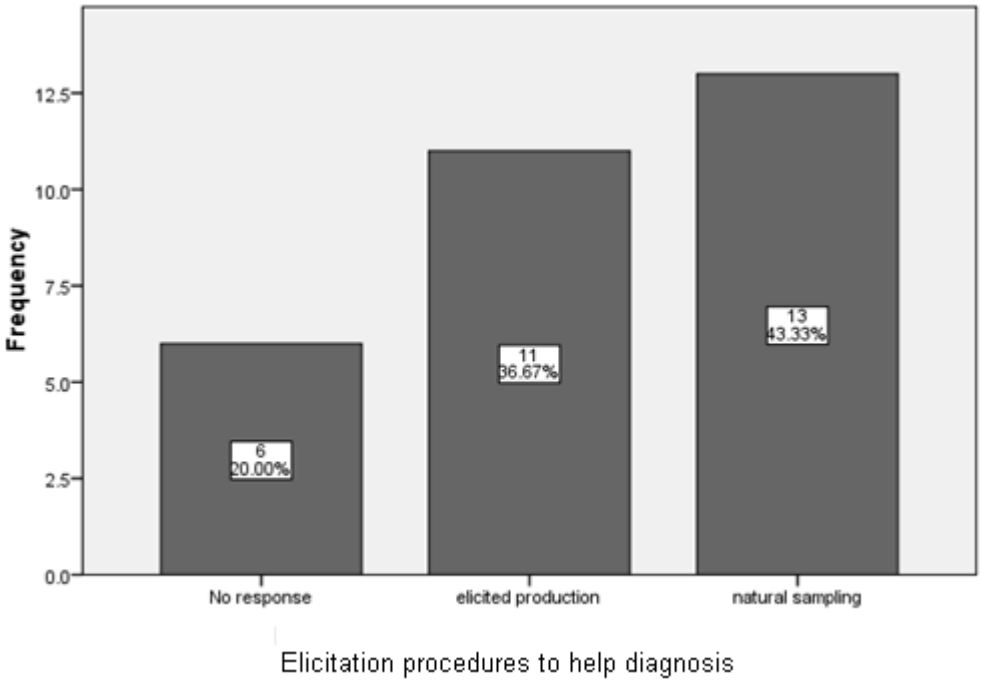


Figure 3-8 The best elicitation tasks from the Iranian SLTs' point of view (n=30).

Language sample length is usually based either on time or number of utterances. For those SLTs who reported basing language samples on time, the average length of language samples reported was 14.08 (SD=10.05) minutes. For SLTs who reported basing language samples on utterance length, the average was 81.36 (SD=141.1). Seventeen SLTs reported that they analyse between three and 60 samples per year (*M*=27.23, SD=19.44) (Table 3-12).

	Mean (SD)	**Minimum**	**Maximum**
Language Sample length (in minutes) (n=26)	14.08 (10.04)	1	45
Language Sample length (in utterances) (n=11)	81.36 (141.1)	10	500
Number of LSA per year (n=17)	27.2 (19.44)	3	60

Table 3-12 The preferences for language sampling by time limit or by number of utterances along with the average number of language samples per year analysed by Iranian SLTs.

The most frequent method of recording is ‘audio recording’ (n=6) followed by those who use three methods of ‘audio recording, real-time transcription and parents’ diary’ (n=5), then combination of ‘audio recording and parents’ diary’ (n=4). A complete illustration of the findings for question 19 is shown in Figure 3-9.

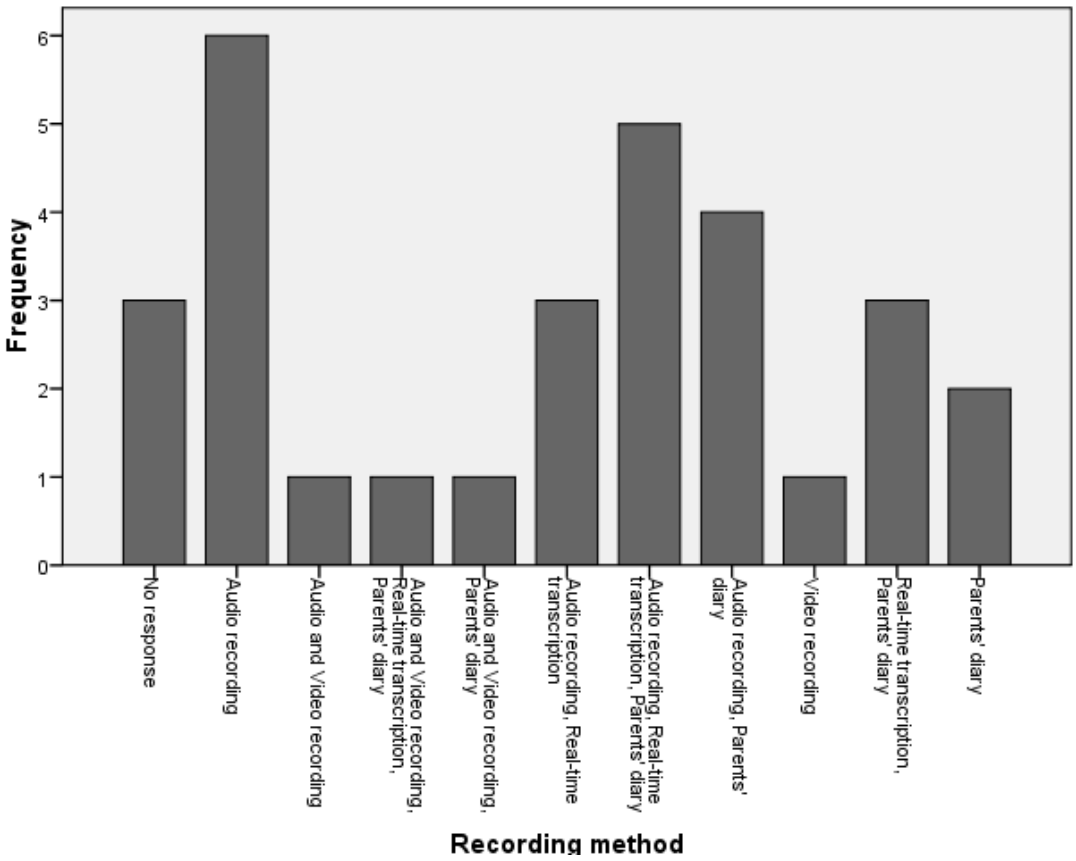


Figure 3-9 Methods of recording chosen by Iranian SLTs (n=30)

Twenty per cent of Iranian SLTs sampled reported that although they apply language sampling in their assessment plans, they don't use any specific language scale in their language analysis (n=6). The other 21 SLTs, however, apply at least one type of scale in their analysis. The combination of MLU and lexical diversity is employed by 16.67% of SLTs (n=5) followed by merely MLU (n=4, 13.33%) with other combinations forming only 6.67% per cent of the responses. One SLT also reported using other scales in language analysis (Figure 3-10).

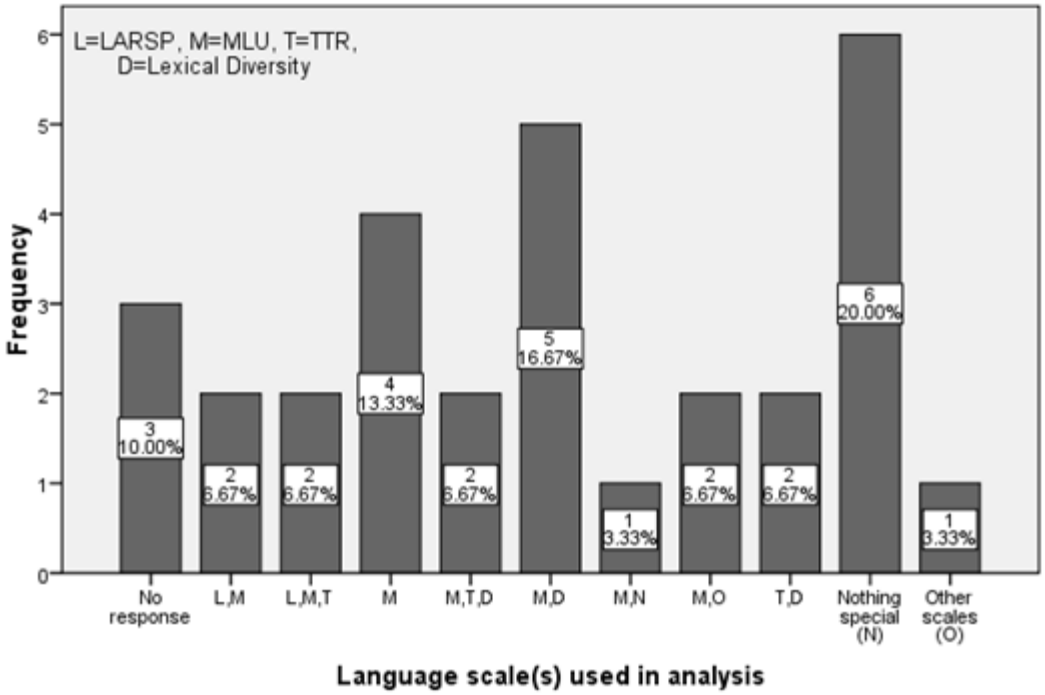


Figure 3-10 Language scales used by Iranian SLTs (n=30)

The most frequent reasons for including LSA in the assessment are diagnosis and treatment with a frequency of 12 respondents (40%). All other reasons fall between 3.33% and 13.33% of the responses (Figure 3-11).

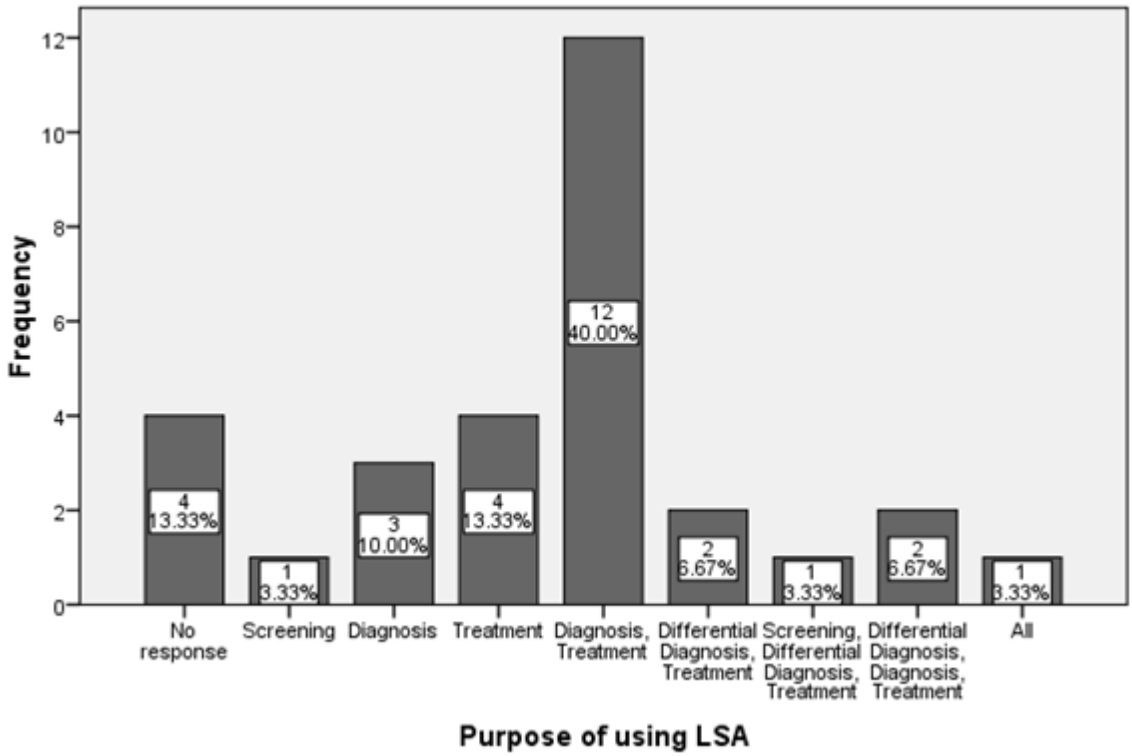


Figure 3-11 Purposes of applying LSA by Iranian SLTs (n=30).

Recall from figure 3-6, five SLTs reported that they do not use LSA. They were, therefore, asked to give their reason for not using LSA as shown in table 3-12. According to this table, eight more SLTs also replied to this question which makes the number of respondents up to 13. This will exclude 17 SLTs as non-respondent due to the fact that they all use LSA in language assessment. Twenty per cent (n=6) of the whole sample (n=30) (equal to 46.15% of those who do not use LSA in their assessments), indicated that the main reason is lack of time. All other categories have one respondent each with the following combinations: no training, lack of expertise, lack of either hardware or software, and a combination of all as well as financial constraints (Figure 3-12).

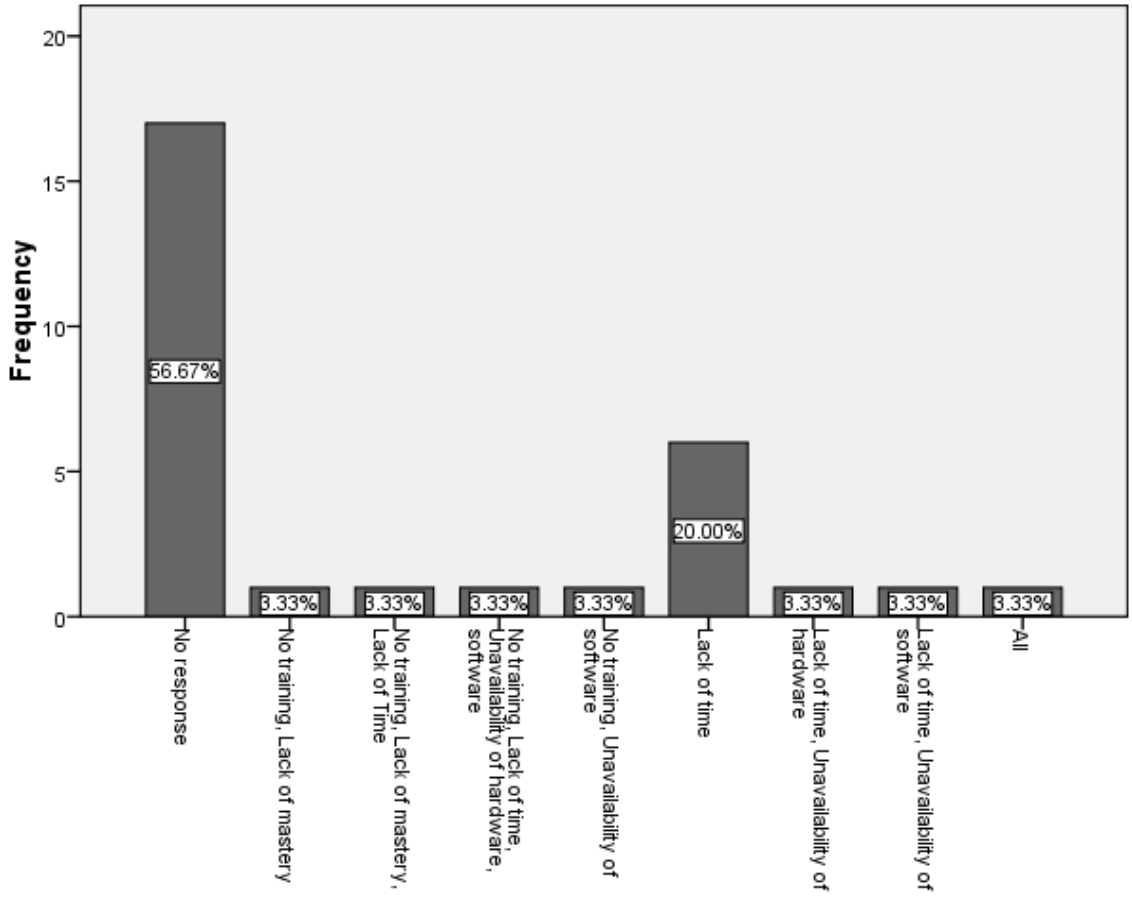


Figure 3-12 Reasons for not using LSA in assessment by Iranian SLTs (n=30).

6. Therapist-specified assessment procedure

Questions 24 to 27 were open-ended and the qualitative data analysis of these questions and the focus groups will be explained in the discussion section; however, a quantitative report of themes elicited from the answers to the last three question is given here to attain a quantitative picture of the results and to justify the prioritising the themes

selected for developing the assessment model. This report follows Roulstone's procedure in quantification of qualitative data (Roulstone, 2011), an alternative definition for content analysis that has been expanded by Hsieh and Shannon (2005) which is the data analysis framework for the current survey. According to Hsieh and Shannon's review of the emerging method of content analysis, it is defined and "used primarily as a quantitative research method, with text data coded into explicit categories and then described using statistics" (p.1278). Text data can be retrieved through different data collection procedures such as open-ended questions, interviews, focus groups, and observation (Hsieh, & Shannon, 2005), so this method will be used in analysing data collected in next part of this survey which is focus groups. As a results, "qualitative content analysis is defined as a research method for the subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns" (p. 1278).

Following the conventional content analysis, no predefined themes were chosen for categorising the concepts which emerged from the SLTs. So, after thoroughly reading and re-reading the answers to the open-ended questions (questions 24 to 26) as well as discussions in focus groups, a long list of initial codes was derived for each question which represented the key concepts and will be reported shortly when reporting each question and results of focus groups, individually. The concepts were then re-coded within sets of more general themes (or codes) which shared similar contents, and re-categorised according to the potential underlying relations among them (Hsieh & Shannon, 2005). The emergent themes were prioritised by considering the frequency of SLTs' agreement with each concept. A simple count (Roulstone, 2001 p. 335) was the optimal way of collecting and analysing the data quantitatively. The global themes, along with the relevant frequency data, shape the output of the current study that systematically combines quantitative and qualitative methods in search of underlying patterns in the unstudied domain of evaluating primary language impairment in Iran.

Question 24 had 21 respondents which shows a 70% response rate. The first row of content analysis identified 59 symptoms (initial codes) named by SLTs as what leads them to a diagnosis of SLI/PLI linguistically. Examples are higher usage of nouns compared to other word classes, shorter sentences, omission of different parts of speech such as prepositions or verbs or conjunctions, telegraphic speech, higher usage of imperatives, incoherence in narratives and topic maintenance, problems with verb inflections (agreement) in terms of tense, number and mode and aspect, problems with communication rules like turn-taking etc. Interestingly, their answers did not merely

include linguistic signs but also they sometimes considered the accompanying signs like family history, auditory processing, memory problems and so on, with linguistic ones to give a more comprehensive picture of SLI/PLI specifications. Moreover, none of these categories earned a majority of the responses. The highest rate was for "omission of grammatical morphemes" with seven responses followed by "phonological problems" with six responses. A more in-depth review of the categories, on the other hand, depicted the similarities amongst them which led to the decision that they could be organised within two main themes of *language-dependent symptoms* and *associated symptoms*. They were called symptoms due to the fact that in some cases, controversial aspects of the same feature were seen so that within the same category some SLTs believed that there should not be a sign of obvious problems whilst others recalled several signs of problems (e.g. pragmatic problems vs. no pragmatic problems). Consequently, these different definitions were interpreted and categorised under one uniting theme, or global theme, to show that SLTs take all these symptoms into consideration in their definitions of SLI/PLI, whether they are problematic or not. Global themes are shown in Table 3-13 with corresponding frequencies observed.

Symptoms	**Frequency***
Language-dependent symptoms	
Phonological	13
Syntactical	23
Morphological	24
Morpho-syntax	8
Semantic	8
Pragmatic	8
Semantic-pragmatic	4
Associated symptoms	
Auditory processing	4
Memory problems	4
Repetition problems	6
Hearing, cognition, neurological status	4
Family history	1

Table 3-13 Global themes raised by Iranian SLTs with respect to their definition of SLI/PLI along with their superior themes; *Number of SLTs who mentioned each symptom in their answers.

In response to question 25, only two SLTs answered that they consider intelligibility and type of communication (verbal vs. nonverbal) in addition to the areas mentioned in the questionnaire.

Question 26 contained a variety of answers which descriptively explained SLTs' own assessment procedures used to identify language impairment in pre-school children. Assessment procedure descriptions from the 27 SLTs who responded to the question were fully read and coded according to the categories which emerged while reading through the transcriptions. In some cases, it was necessary to split or merge the codes. Marshall calls this process of coding "Constant Comparative Coding" (Marshall et al., 2007 p. 540). A total of 47 categories were identified with a frequency range from one for 'several assessment areas' to 15 for 'interviewing parents' and 14 for 'language sample analysis'. The frequency of each category based on the number of SLTs who mentioned it in their answer was counted and re-categorised under 17 areas of assessments. In merging the initial categories into more global themes, an attempt was made not to lose even a single individual category since the mere citation of a category was valuable enough to be considered in the final model of Iranian type of child language assessment. The importance stems from the researcher's intention to introduce a preliminary reference standard for further research, making every single mention of value.

The 17 global themes (within 12 ranked themes) have been demonstrated in table 3-14 with corresponding frequencies observed. The relationships between codes and their configuration were then reviewed in accordance with the codes derived from focus groups to organise the initial model of the assessment and to "reflect the investigation of a process happening within a context rather than a static one-off experimental situation, moving towards the development of a theory, rather than starting with theory-driven concepts" (Roulstone, 1997 p. 302). This sort of text-driven data (open-ended questions and focus groups) is the foundation of content analysis which is defined as "systematic classification process of coding and identifying themes or patterns" (Hsieh & Shannon, 2005 p.1278), a type of bottom-up inference.

As shown in table 3-14, a total of 24 SLTs consider 'parents' communication patterns' in their own assessment protocol. This is because any small verbal or non-verbal communication from the parents can affect child language development or cause a delay if they do not show willingness to communicate with their child or mention that they have no time to interact with their child verbally and non-verbally. In the second place

are ‘child’s overall health’, seeking referrals to different professionals including audiologist, psychologist, and neurologist, as well as a ‘comprehensive assessment of expressive language through language sample analysis’, in their own definition which might include any length of language sample collected in any setting and with any interactant of SLT’s choice, and contain any type of either quantitative measures, e.g. MLU or number of different words (NDW), or qualitative ones, e.g. semantic relations between words or type of relative clauses. ‘Interviewing parents’ which specifically includes ‘bilingualism’ comes in the third place, followed by ‘audio-visual skills’, particularly memory as well as ‘verbal imitation processing’, which is defined as imitation tasks including repeating words and sentences with increasing length and within different interval times. In fifth place is ‘assessment of language comprehension or receptive language’ in any type of language tasks, and then ‘vocabulary expression’ which is considered separately from assessment of expressive language because it exclusively includes child’s expressive vocabulary by asking him/her to name individual pictured nouns. Phonological assessment through articulation tests, specifically named ‘phonetic test’, and ‘preverbal skills’ (e.g. joint attention) come in sixth place followed by ‘grammatical assessment’ in seventh place. Grammatical assessment was named separately even by those SLTs who had previously mentioned that they would comprehensively assess expressive and receptive language. This happens with other linguistic aspects like phonology, semantics and pragmatics, too and probably shows the importance of individual domains in assessment.

‘Motor skills and motor coordination’, assessed by several specific gross and fine motor tasks (e.g. balanced walking on a line or inside a series of coloured circles, sequentially pointing to nose and head etc.), as well as ‘personal professional judgement about child’s age-appropriate language development’ come in eighth place. Due to unavailability of approved language development milestones for Iranian children, this judgement is very subjective and completely relies on SLT’s personal observation and experience. ‘Play assessment’ by defining the type of play (sensory-motor stage, pre-symbolic, symbolic etc.) and judging its appropriateness to age comes in ninth place followed by ‘consulting with colleagues about the child’s clinical condition’. ‘Diagnostic intervention’ is in eleventh place which is a term for a period of monitoring child’s language improvement by providing professional counselling or one month intensive language therapy (equal to four standard therapy session, each takes 30 to 45 minutes) with the presence of parents. The last place is for ‘pragmatic assessment’ and

'speech rate' which is assessed by the number of words per minute, again with no reference to normal range.

Rank	Areas of assessment	Frequency*
1	1. Parents' communication patterns (verbal or non-verbal)	24
2	2. Medical history including referrals for hearing, cognitive and neurological status	23
2	3. Comprehensive assessment of expressive language through language sample analysis	23
3	4. Interviewing parents and bilingualism	22
4	5. Audio-visual skills, including memory and verbal imitation processing	16
5	6. Assessment of language comprehension (vocabulary, sentences, and narrative)	10
5	7. Vocabulary expression	10
6	8. Phonological assessment	7
6	9. Preverbal skills	7
7	10. Grammatical assessment	6
8	11. Motor skills and motor coordination	5
8	12. Comparing child's development to normal development (developmental milestones) by personal professional judgment	5
9	13. Play assessment	4
10	14. Consultation with colleagues to seek the clinical judgment agreement	3
11	15. Diagnostic intervention	2
12	16. Pragmatic assessment	1
12	17. Speech rate	1

Table 3-14 Areas of assessment that Iranian SLTs consider in their personally-designed evaluation of pre-school child language;*Number of times each item is mentioned in SLTs' responses

These themes will be combined with emergent themes from focus groups to organise the final framework for child language assessment in Iran, specifically targeted at PLI, and consequently will be used as the reference standard in phases of the diagnostic accuracy study.

3.5 Focus groups

Two focus groups were organised concurrently with analysing the open-ended questions to clarify the unclear ideas as thoroughly as possible.

3.5.1 Participants and Procedure

Because recruiting a random sample of SLTs for the focus group was not possible, an attempt to reduce bias was tried by inviting SLTs who worked in different areas of Isfahan city. The SLTs who contributed in this part were nine qualified clinical speech therapists from different parts of Isfahan city. Their demographic specifications are shown in figure 3-13.

The SLTs' answers to the aforementioned questions of the questionnaire were summarized and two main questions were designed to be given to the discussion groups including:

- Which areas do you usually assess to identify a child with no concomitant problem as being language impaired *or* PLI?
- What is your specific procedure of assessment in the field of language impairment with no explicit sign of associated problems *or* PLI?

The contributors were encouraged to involve themselves in the discussion as much as possible by the coordinator (main researcher). The coordinator tried not to interfere in the discussion unless to help maintain the topic or to engage the members who were silent. Her probes followed the questioning techniques used by Roulstone (2001) as follows:

- Laddering: "to explore the super-ordinate, more abstract hierarchy of a construct by asking questions such as 'why is that important to you, what kind of client would that be'" (p. 333);
- Pyramiding: "leads the respondent down the hierarchy to subordinate constructs. Typical questions might be 'can you give me an example of that? How would you identify one of those?'" (p. 334).

In addition to audio-recording the discussion groups, note-taking was another main method of data collection in this part of study to reduce the impact of missing data.

3.5.2 Results

The sex and academic level of SLTs participating in focus groups are summarised in figure 3-13. All the participants had more than three years of clinical experience.

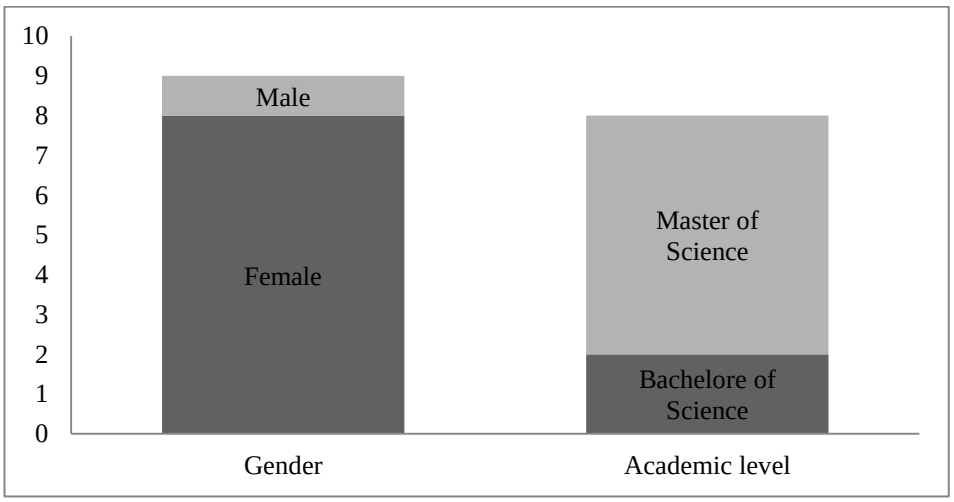


Figure 3-13 Description of SLTs involved in discussion groups

A categorization of the results of the discussion groups showed that the majority of SLTs (n=7) use diagnostic treatment or diagnostic intervention in their assessment procedure to identify children with SLI/PLI. Given that this item had not been included in the questionnaire survey, it was decided to ask as many available SLTs as possible to answer the question: 'Do you apply 'diagnostic intervention' as a procedure to identify SLI/PLI?' and if so, 'how much does it help you to diagnose a child as SLI/PLI? Rate it from 1=very low to 5=very high'. These questions were asked of those SLTs who replied to the aforementioned questionnaire by email (n=20). Thirteen SLTs responded to this question, all with a positive response to use of diagnostic intervention and the ranking ranged from two to five (*M*= 3.61, Mode=4). Results of this item have been integrated into table 3-11 and figure 3-2.

The qualitative data analysis follows the same method discussed in analysing the open-ended questions of the questionnaire survey (section 2.2.3, part 6). Several concepts came up about the assessment areas including the age at diagnosis, the complications around standardized tests, as well as diagnostic intervention and the procedure of carrying out the assessments. The themes brought up with regard to the areas that each individual SLT considers in assessment of language impairment with no concomitant problem *or* PLI, were derived from the groups' discussion. They were coded with no pre-defined coding system to meet the criteria of being open to every novel category. The categories, afterward, were reviewed to merge and organise into fewer themes and the themes produced were compared against themes extracted from the open-ended questions of the questionnaire interviews, resulting in four main areas of consideration in assessment procedure: (a) interviewing parents, siblings, and nursery teachers; (b) referrals to reliable professionals and colleagues; (c) overall child observation; and (d) child language assessment through either parents' records (child's voice-recording, parents' diaries or checklists) or direct assessment (including observation of language behaviour, semi-structured assessments, and natural language sampling).The detailed assessment areas were matched to the results of the questionnaire interview and the final model was developed based on the International Classification of Functioning, Disability and Health (ICF) framework of World Health Organisation (WHO) (2002) (Figure 3-14).

3.6 Discussion

The first and foremost aim of this survey was to develop an Iranian-specific reference standard for child language assessment by exploring and explaining Iranian SLTs

individual knowledge, because a suitable instrument was unavailable to use in the diagnostic accuracy study (Creswell et al., 2006).The two qualitative methods employed helped to "generate a model that is grounded in the viewpoints of the participants" and "subsequently tested or refined using quantitative methods" (Creswell et al., 2006 p. 3). In responding to the open-ended questions, SLTs answered what they think *should be* observed in a child with SLI/PLI as well as what they do based on what *was observed* in real clinical settings: two different processes called "prescriptive knowledge" and "descriptive knowledge", respectively by Ellis (1992, as stated in Roulstone, 1997 p. 302), and it has been emphasised that it is necessary to study what exactly happens in a therapy session with an SLT who needs to turn her/his theoretical/prescriptive knowledge into practice.

The data, accordingly, were analysed qualitatively and an integrated assessment method was established based on the most agreed assessment domain and procedure raised in both questionnaire responses and focus groups. The practical scientific framework for classifying the impairments globally developed by the WHO was also used to classify the criteria that Iranian SLTs are applying in their diagnostic procedures. The quantitative data analysis of the main themes integrated into the model provides the reader with information on how numbers can concretely reinforce the abstract results. The whole compilation will make the results of this study more practical for users inside Iran (Figure 3-14).

In an attempt to relate the global themes and their sub-themes to a functional framework (Figure 3-14), some of the initial codes were restored as the most important examples of what Iranian SLTs believe should be included in an assessment. They are going to be embedded into the ICF framework along with all the superior themes mentioned in tables 3-13 and 3-14.

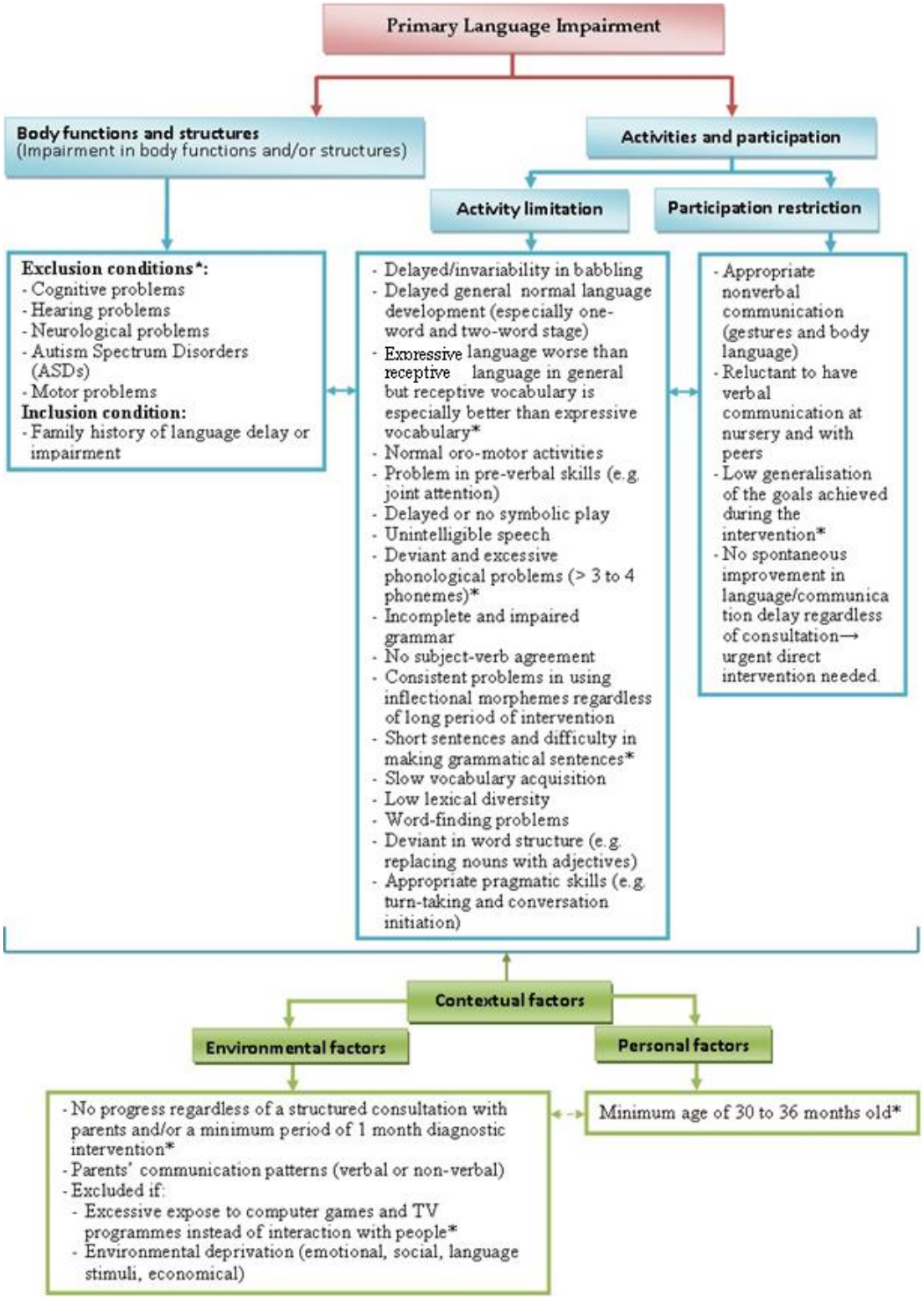


Figure 3-14 International Classification of Functioning, Disability and Health (ICF) framework adapted for the results of the survey on Iranian SLTs; * Items emphasised as very important by SLTs in focus groups.

The systemic model chosen to review and analyse the results of this survey enables SLTs to look at language impairment as an internal factor, comprising the child's personal strengths and weaknesses, surrounded by a variety of external factors including the communication environment, setting and communicators' factors (Paul, 2007). The ICF was found to be the most advanced and integrated example of a systemic model in such a way that it tries to show all the child's capabilities in an interaction with her communication environment and linguistic domains. The ICF can be interpreted as a practical manifestation of a systemic model. The conceptual framework embedded in the ICF was found to be the most suitable one to describe how Iranian SLTs consider different factors influential to the child language from different sources, personal factors, the child's ability in social communication, the child's environment (people and settings), in their clinical assessment to gain a broader picture of the child's language condition. This "conceptualisation of language impairment" is a major advantage of the ICF and may provide a coherent interaction between "empirical evidence" and "clinical observation" (Dempsey & Skarakis-Doyle, 2010 pp. 425, 426), a prominent feature of this framework that is most compatible with the entity of the current survey on Iranian SLTs' child language assessment.

Dempsey's tutorial includes a comprehensive description of an ICF application in child language sciences with an excellent comparison among studies as the supporting evidence for the conclusion (Dempsey & Skarakis-Doyle, 2010). Integrating the results of the current study with what Dempsey & Skarakis-Doyle mentioned in the tutorial shows that what Iranian SLTs consider in their assessment of child language includes parts of the ICF framework, although they may be unaware of the title of the framework employed. Their descriptive knowledge about PLI and what they actually do in an assessment session demonstrates their knowledge that PLI in children is not the only way to account for their communication problems. They try to cover many suspected areas of impairment in their assessment because they believe that PLI in children is not solely affected by linguistic deficits but that there are several interfering variables involved. What is seen in the child's language behaviour, by their definition, is the result of mutual inter-connections of Body Functions and Structure (represented by including/excluding criteria), Activities (embodied in such signs as the expressive language level compared to receptive language level, phonological problems, difficulty in grammatical formulation etc.), Participation (generalisability of the targeted features in the therapy etc.), and Contextual Factors –Personal Factors (minimum age of 30 to 36 months old) and Environmental Factors (such as no progression after a period of

diagnostic intervention, parents' interaction patterns, and the amount of communication by people vs. electronic games etc.). If they wanted to look at children with PLI only from an etiological point of view, such as a medical model of language impairment (Paul, 2007), they would have not included the areas other than Body Functions and Structure, or at most Personal Factors, in their assessment procedures, whilst in fact it is obvious that all factors have equal values in SLTs' clinical decision making. ICF represents this well, both comprehensively and cohesively (Dempsey & Skarakis-Doyle, 2010; Washington, 2007). The power of the ICF to systematise this array in a clinically practical way may support researchers to undertake evidence-based studies whose frameworks systematically follow the concepts embedded in the ICF, to align their study results to real clinical conditions, and to tie research to (strictly stated) immediate practice.

Despite limitations in this study which will be discussed shortly, the viewpoint taken here is that the resulting framework is capable of being applied in clinical studies to compare a participant's unique profile against this framework to find out any potential matching and/or divergence between corresponded factors. The researcher, then, would be able to map out one or several representations for both normal language and various language impairments. It might enable the field of child language science to introduce a novel arrangement of symptoms for defining normality and/or abnormality or create an alternative categorization of language impairment according to a systemic model. The global point of view through the lens of ICF would also collect scattered personally defined assessments around a widely accepted framework and create a "core assessment and evaluation approach" (Aram et al., 1993 p. 586) by restricting the occurrence of examiner bias.

3.7 Limitations and suggestions

While the poor return rate and snowball sampling seriously affected the generalisability of the findings, the process was out of the researcher's control and the results could be replicated if the IRIMC's regulations continue to be the same. Roulstone argues that the principle of a qualitative study lies in answering the question: 'what is worth paying attention to?' versus the replicability in quantitative studies (Roulstone, 1997). The challenge of the current study was to meet this principle so as to provide the results within the ICF framework, which is increasingly being considered as a valuable coherent clinical framework by speech and language therapy professionals (Dempsey & Skarakis-Doyle, 2010).

A parallel analysis of homogeneity and heterogeneity of results can also lead us to a conclusion about how much Iranian SLTs agree with respect to what and how child language should be examined in a search for language impairment in pre-school children. Both consensus and disagreement in SLTs' initial assessment decisions reported by Roulstone (2001) suggests that "consensus within the profession is no guarantee of efficacy" and that looking for differences based on evidence (more research) would be "an important part of the process of advancing our consensual knowledge base" (Roulstone, 2001 pp. 346-347). A systematic operationalised definition of terms embedded in this framework benefits the clinicians (and other related professionals) in so far as they would find a higher congruence between what they experience in real clinical diagnosis and the clinical competence of the proposed ICF framework (Aram et al., 1993). This should be a direction of research in the future. The proposed preliminary ICF-based reference standard for identifying pre-school children with PLI in Iran would benefit future research by providing an integrated approach of theory and practice in intervention. Besides, it will support the next phases of both the current and upcoming diagnostic accuracy studies, being the sole reference standard, although preliminary, to enhance Iranian SLTs' knowledge about the role of evidence-based practice in speech-language therapy.

CHAPTER 4

The diagnostic accuracy of language sample measures in Persian

4.1 Language sample analysis as a diagnostic device

Finding a reliable clinical assessment tool to identify primary language impairment (PLI) in Iranian pre-school children was the core aim of the third part of this project. Language sampling and analysis can accommodate almost all types of research questions about communication skills in children as well as the people who interact with them, making it a good tool for studying both sides of an interaction with the child (Rowe, 2011).

The evaluation procedure closest to a naturalistic communication environment is language sampling. As will be discussed shortly, of all clinical evaluation procedures, language sample analysis has been found to be the most appropriate to use in a clinical setting (Rowe, 2011). Natural language sampling was the first choice for collecting language in the current study due to its ecological validity and being "direct, objective and reliable" (Bornstein, Painter, & Park, 2002 p. 688). Moreover, the majority of Iranian SLTs (83%) were in favour of using LSA in their evaluation of child language, which is similar to the results of other studies (Hux, Morris-Friehe, & Sanger, 1993; Kemp & Klee, 1997). In a later study by Loeb and colleagues, however, 93% of SLTs reported using LSA in assessment (as cited in Eisenberg et al., 2001 p. 323) despite many SLTs' reluctance to use LSA because it is time-consuming (See results of chapter 2) (Heilmann, 2010; Kemp & Klee, 1997). Compared to standardised language tests, quantitative LSA is a more sensitive measure of language impairment in young children (Eisenberg, Guo, & Germezia, 2012). It has been documented in many studies that due to the assessment transparency between LSA and natural communication contexts, its flexible nature would also help to increase the efficiency of clinical management (e.g. Costanza-Smith, 2010; Heilmann et al., 2010; Hewitt et al., 2005). Moreover, in the current study, natural language sampling was generally the most frequently-reported elicitation procedure used by Iranian SLTs (n=25, 83%) and at least 40% of SLTs (n=12) reported that they employed natural language sampling during clinical assessment and diagnosis in combination with other elicitation tasks.

The initial inclusion criteria for LSA measures to be examined in this study were those reported as being sensitive to changes in age and/or having good diagnostic accuracy in English (chapter 2 on meta-analysis of the non-Persian studies). Heilmann and colleagues report that the diagnostic accuracy of some language sample measures was based on those measures with reported sensitivity to age and language status in previous studies (Heilmann et al., 2010); accordingly, some measures in the current study were selected from their study, including mean length of utterance in morphemes (MLUm),

number of different words (NDW), word errors, and utterance errors. Other measures were also chosen because of their importance in Persian, including four measures adapted from Foroodi-Nejad's study, case marker *ra*, object clitic, subject-verb agreement, and present tense marker *mi*, as well as those with feasible grammatical and semantic properties in finding PLI. The above-mentioned Persian-specific measures had been shown to be able to differentiate Persian-speaking children with and without PLI (Foroodi-Nejad, 2011) along with the measure of MLUm in Maleki Shahmahmood and colleagues' study (2011). They, subsequently, were examined against the age correlation in a real sample of Persian-speaking children (pre-accuracy study) to find out the most suitable measures for the current study.

Given that the three phases of the DA study share participants and procedures, only their results will be explained separately. Hence what follows is organised as participants, the procedure and then the results as reported in pre-accuracy, phase I and phase II studies.

4.2 Participants

The sampling was done in Isfahan city located in central Iran (Figure 4-1) and included two groups: children with typically developing language and children with primary language impairment, as will be explained in 4.2.1 and 4.2.2.

Figure 4-1 A map of Iran showing the geographical location of Isfahan city (Retrieved 15/12.2012 from http://shiraz.airport.ir/HomePage.aspx?TabID=5765&Site=shiraz.airport&Lang=fa-IR)

4.2.1 Children with Typically Developing Language (TDL)

After getting permission from the State Welfare Organisation (Isfahan branch) to see a list of nurseries in Isfahan, 10 nurseries were randomly selected from 338 registered nurseries in Isfahan and adjacent satellite towns. They were asked to provide a list of children between 42 and 54 months old. The initial aim was to randomly sample five or six children from the registered children; however, some of the nurseries had less than five children in this age range. Consequently, all the children that fell within the age range in those nurseries were sampled and the difference was balanced from another nursery. A total number of 55 children with typically developing language between 42 and 54 months old were randomly selected in this way. In the initial screening, the child's age and medical history, including hearing, neurological, mental and physical health, were checked using their medical documents stored in nurseries as well as asking teachers. The information was checked again when the parents attended the language sampling session. Concerns or problems with the child's communication were also obtained from parents and teachers. In the case of any concern, the child would be assessed using a routine Iranian SLT assessment and observation by the main researcher. PLI diagnosis was based on researcher's professional judgement through a routine language assessment, mainly clinician-made (researcher as clinician), which included parts of the reference standard introduced as ICF framework in chapter 3. Two children showed signs of PLI in this stage, leaving 53 as being TDL. Other children who did not participate were those whose parents were either reluctant to join the study or did not attend the appointment. Those two children were replaced by the children from the same nurseries to bring the sample back to 55 children (Figure 4-2). The above-mentioned criteria were considered as the reference standard due to the fact that there is no gold standard assessment for identifying language impairment in Iran. In fact, a professional clinical judgement is considered as the reference standard. Despite the subjective bias that this procedure had, different components of the procedure were verified by examining through carefully-adapted research methods to fulfil the lack of gold standard, both qualitatively and quantitatively (see the results of chapter 3). Moreover, including clinical judgment in clinical diagnosis of PLI in the absence of a gold standard was shown to be widespread in other studies (Dollaghan & Horner, 2011; Thordardottir et al., 2011).

4.2.2 Children with Primary Language Impairment (PLI)

Iranian SLTs working at six university-dependent clinics, and also 24 SLTs who worked independently in Isfahan city, were asked to refer clients who met the inclusion criteria within a period of five months (August to December 2009). Although they were requested to select those children between 36 and 60 months of age whose expressive language was longer than one word and intelligible enough to be recorded and analysed, there were still children among referrals who did not meet the criteria, probably due to the SLTs' inattention to the specific inclusion criteria or because they were simply interested in assisting with the research. The wider age range was chosen to allow the replacement of children who did not meet the criteria with ones who did, where possible. Also because some measures of interest (mean length of utterance excluding one-word utterances) required excluding utterances shorter than two words, children's samples at the one-word stage of expressive language would be completely uninformative for this study. Certainly, they will be of value in studies on the nature of form and content of one word utterances in this range of age. The whole process of language transcription and analysis was also dependent on accessing an intelligible language sample. The source of unintelligibility in children's speech might be variable and sometimes interfere with grammatical development (Estigarribia, Martin, & Roberts, 2012), so needed to be controlled.

Upon calling the parents, the researcher asked whether the child met all criteria. Some children were undetectable due to changing their contact details (n= 25), some parents were unwilling to participate (n= 5) and some stated that their child no longer needed speech therapy or had been discharged from therapy (n= 4). Speech of some children was not intelligible enough to be included (n= 5, two children had been diagnosed as severe phonological disorder) or was at one-word stage (n=9) or both (n=1). Thirty three children were outside the age range of the study. Two children were bilingual and two others did not meet the appointment. In sum, 24 out of 110 children referred were recruited to the study as children with PLI. The classification of two children changed from the TDL to PLI group after receiving the reference standard assessment by the researcher resulted in a total of 26 children with PLI. Of those, however, two children were withdrawn; one due to parents' unwillingness and one after checking language samples which showed insufficient sample size (less than 20 minutes). Finally, the language samples of 24 children with PLI were entered to the analysis (Figure 4-2).

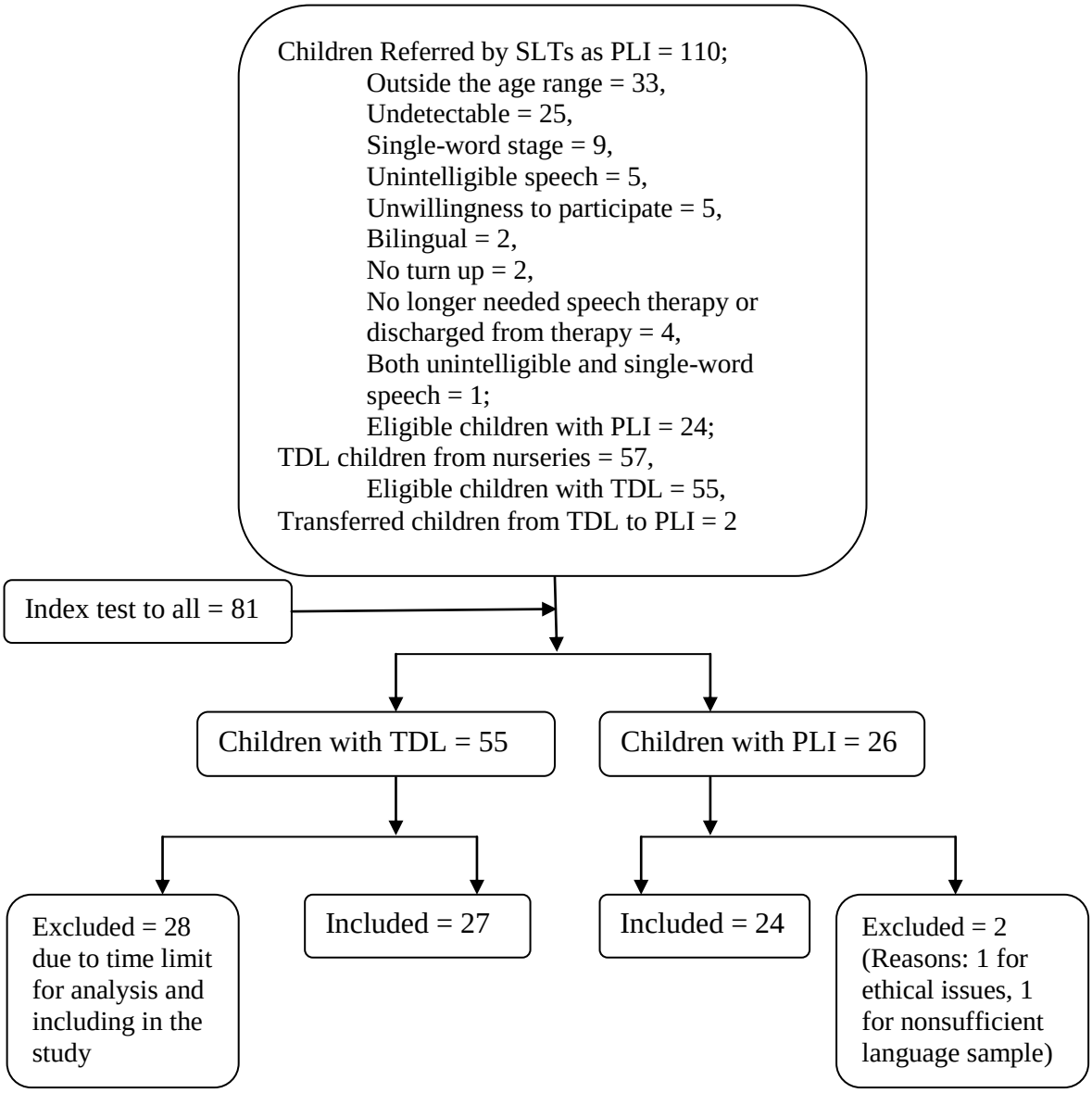


Figure 4-2 Flow diagram of referred and recruited children to the study

4.3 Procedure

The mothers were asked to make an appointment with the researcher to record a play session with their child. The recordings were carried out in three sites: Navab medical centre (speech therapy section), located in central Isfahan city, Al-Zahra grand hospital (speech therapy section), located towards the south of Isfahan, and Kowsar private rehabilitation and physiotherapy centre, located towards the east of Isfahan. The first two sites belong to Isfahan University of Medical Sciences and the last one is a private rehabilitation centre. The distribution of the locations was so that the parents were able to choose among them; the closer to their home, the more convenient. It was attempted

to make the place as quiet as possible, however, noise was inevitable in some cases (e.g. plumbing repairs above the room etc.). However, this made the recording conditions similar to the actual situations in which Iranian SLTs work with children.

An information sheet was given to the parents and they were asked to read and sign a consent form (Appendices B and C). At this stage, some of the parents were asked to withdraw from the study (see above) and consequently other children were recruited. They were also provided with the opportunity of talking to the main researcher about their child's development as a 'thank you' for taking part in the study. The children were also given sweets after the session to show the researcher's gratitude. They were asked to allocate up to one hour of their time for the sampling with no external pressures to leave the session early. Demographic data along with medical history, bilingualism, status of child hobbies, child-care, and family socioeconomic data were requested using the basic information form (Appendix D).

After interviewing the parents to confirm the child's developmental health, the mothers were invited to attend a free play session as the main interactant with their children. Such factors as background noise, other people interference and unexpected events might affect the quality of recording. Also the impact of speaker, context and topic should be considered while recording a child for a spontaneous language sample (Hoff, 2011; Pan, 2011). Having mothers serve in the role of the interlocutor during the language sample has been shown to be a factor in increasing children's language output and maximising the value of some language measures (Bornstein et al., 2002) in addition to providing a more familiar communication environment in which young children can behave naturalistically. Pragmatically, conversation activities have a facilitating effect on the length and complexity of language output (Eisenberg et al., 2001; Haynes, Purcell, & Haynes, 1979) and it is supposed that during free-play interaction with mothers, children's 'spontaneous' language productivity would be enhanced compared to structured contexts in which sometimes 'communication', by its true meaning, would not happen at all (Evans & Craig, 1992). Free-play was also believed to be developmentally closer to the young age of the sample population so that they would be more motivated to be involved in the communication interaction and show significantly more complex structures in their production (Klein, Moses, & Jean-Baptiste, 2010). Whereas story retelling and narrative activities have been revealed to elicit more complex morpho-syntactic utterances from older children (mean age 5;4) (Southwood & Russell, 2004), Klein et al. (2010) showed that free-play involves more complexity in content with the young children (mean age = either 2;8 or 3;4

years;months) suggesting that cognition and linguistic performance are interacting developmentally. They concluded that, developmentally, complexity in semantic relationships would be endorsed best in free-play activities with manipulated objects among young children. On the other hand, the lexical richness might be affected by the place of language sampling although the samples are collected through free play. Children might express more diverse vocabulary when the play is with mothers at home where they are more familiar with the environment and feel more relaxed. Free-play, on the other hand, is considered a time-consuming task if the researcher or clinician wants to elicit a representative language sample with adequate samples of different grammatical structures; however, in anticipation of providing evidence for its inappropriateness with young children, it is assumed as the best method for collecting language samples from children younger than 4;6 years of age.

Mothers and children were, then, directed to the sampling room furnished with a carpet, a children's table and two chairs. A doll's house with two dolls, two vehicles and a set of animals along with a picture book were used to elicit the language. Although the mothers were asked to tell the story from the book to their child and ask for retelling, some of them did not do so during the 20 minutes of the sampling time. As a result, the first twenty minutes of the free play session were considered to be analysed, whether story-telling was included or not. The toys were aimed to be age-appropriate and meet such factors as being diverse enough to cover both here-and-now (topics with immediate presence in terms of place and time that happens in the immediate environment) as well as there-and-then topics (which contain information about non-present environment) to result in as much variety of language structure as possible (Eisenberg et al., 2001; Paul & Norbury, 2012; Wanska, Bedrosian, & Pohlman, 1986). Figure 4-3 shows the materials used along with the furnished room provided for the sampling in one of the centres.

Figure 4-3 Materials and furniture in the sampling room

The mother was directed to start playing with her child on the child's choice of toys with emphasis on making the conversation as exciting as possible for their child to encourage her/him to participate in conversation. She was also requested to use open-ended questions to elicit more expressive language (i.e. using the *who, how, why* questions instead of *yes/no's*) and mostly follow the child's directed games and talk. The mother was also asked to encourage her child to talk more than herself about toy related themes (i.e. a doll house with two dolls, a bus, a broken car, a set of animals and a jungle play mat) and not to ask for automated and serial speech (e.g. counting or singing) very much.

Language samples were recorded using an Olympus WS-311 digital voice recorder or an Edirol R-09 digital voice recorder positioned as close as possible to where the child was playing. All utterances in the 20-minute session were orthographically transcribed. Decisions about utterance segmentation and morphemes and word roots were made using their definitions in Persian grammar and semantic literature (Kalbasi, 2001; Karimi, 2003; Mahootian, 1997; Meshkato-Dini, 2008; Nilipour & Raghibdoust, 2001) as well as those compatible criteria in English, particularly by looking at the Systematic Analysis of Language Transcripts (SALT) system which was the main analyser of the data. Transcription challenges led the researcher to look for software which specifically takes the rich morphological structure of Persian into account in its conventions. This richness, as shown in chapter 1, is largely situated in the inflectional morpheme inventory.

Two well-known pieces of software in the field of child language, especially in English are (a) CLAN (MacWhinney, 1994) and (b) SALT (Miller & Iglesias, 2008, 2012). A brief comparison of the two programs is in table 4-1.

Software	PROS	CONS
SALT	1.Uncomplicated method of calculating language measures; 2. Short training period.	1.Not freeware; 2. No Persian studies; 3. Prefixes not addressed in SALT conventions and analysis (one of the most prominent morphological features of Persian).
CLAN	1.Freeware; 2.Broadly-applied in studies; 3. Two Persian databases downloadable from CHILDES website.	1.Complex set of instructions (command lines); 2. Lengthy process, perhaps more applicable in research than clinical practice; 3. Unpublished Persian transcription rules.

Table 4-1 Comparison between two types of language analysis software in terms of applying to Persian (adapted from Kazemi, Nockerts, Klee, Stringer, & Miller, 2012)

The diagnostic study needed software with high clinical applicability, comprising the least complicated conventions as well as quick analysis and computation of language sample measures. Although SALT met the desired criteria compared to CLAN, it did not have a convention for transcribing Persian prefix morphemes which led to them being missed in analysis. Consequently, the developers of SALT, Ann Nockerts and Jon F. Miller, agreed to adapt SALT so as to accommodate prefix marking. The new version of SALT accommodates the morphological complexity of Persian and similar languages (Miller & Iglesias, 2012). The Persian Transcription Conventions Protocol (PTCP) introduced for diagnostic study, also provides a compatible range of transcription conventions for utterance segmentation and counting Persian morphemes (See table 4-2). Quantitatively, the major difference can be directly seen in NDW and indirectly in the number of different bound morphemes which may be derived from the list of bound morphemes. Qualitatively, the main difference is observed in outputs in the "word root table" and "bound morphemes table". As an example, the large number of potential combinations of verb formulations in Persian, with three prefixes and three suffixes, is demonstrated in the following one-word sentence (Example 4-1) which contains five morphemes. The reason that it is considered as one-word is that only the verb stem can stand alone as a single morpheme and all other affixes are bound morphemes, although some of them are written independently (e.g. می) from the verb root in script:

Example 4-1

نمی خونمش.
ne\mi\xun/am/esh.
Negative marker\progressive marker\verb root/first singular verb marker/objective clitic.
not\-ing\read/I/it.
I am not reading it.

A comparison between two examples of SALT transcriptions, before (Example 4-2a) and after (Example 4-2b), using the revised version, can be found in the following example (each underlined unit counts one word):

Example 4-2

نه، نمی خواد بخوابه.
می خواستن سوار بشن ماشینشون را اینطوری.
اینطوری روشن می کنن بعد صداش اینطوری می شه {داقوداقو}.

a) C *na, ne/mi/xa/d be/xab/d.*
C *mi/xast/an savar be/sho/an mashin/eshun ra intor/i2.*
C *intor/i2 roshan mi/kon/an bad seda/esh intor/i2 mi/sho/d {daqodaqo}.*
MLUw=5.33; MLUm=11.33; NDW=11; NTW=16
Number of different bound morphemes=11

b) C *na, ne\mi\xa/d be\xab/d.*
C *mi\xast/an savar be\sho/an mashin/eshun ra intor/i2.*
C *intor/i2 roshan mi\kon/an bad seda/esh intor/i2 mi\sho/d {daqodaqo}.*
MLUw=5.33; MLUm=11.33; NDW=13; NTW=16
Number of different bound morphemes=9

No, he/she is not going to sleep.
They wanted to get on their car like this.
They start it like this then its sound is like this {daqodaqo}.

In the original, un-adapted version of SALT, some of the word roots such as *sho, kon* would be counted as bound morphemes and some of the bound morphemes like *ne\, be* were ignored and included in NDW computation whilst, obviously, their grammatical categories are entirely different (Kazemi et al., 2012).

Utterance conventions

1 Included utterances:

1.1. Transcribing 20 minutes consecutive intelligible utterances

1.1.1. Fully transcribed utterances (Brown,1973)

1.1.2. Portions of utterances, entered in parentheses to indicate doubtful transcription, are used (Brown,1973)

1.2. The criteria to consider a part of speech as an utterance are as follows with the priority order:

1.2.1. Utterances that carry connectives are one utterance (Fletcher & Garman, 1988; Klee 1992) including coordinating conjunctions (and, but, so, and then, then) unless there is:

1.2.1.1. Ellipsis of an element in the first clause and/or

1.2.1.2. There is an anaphoric relation to an antecedent phrase or lexical item in the first clause (Fletcher & Garman, 1988; Klee 1992);

1.2.1.3. All utterances with subordinating conjunctions (such as after, before, but, if, when, that, then, because) will be considered as one utterance (Rice, Redmond, Hoffman, 2006).

1.2.2. Terminal intonation contour, rising or falling (Miller, 1981)

1.2.3. Pauses of greater than two seconds (Miller, 1981)

1.3. Immediate imitations of adult utterances, exact self-repetitions and identical utterances will be included as well as one-word utterances.

(Note: Some of the utterances might be excluded for specific analysis reasons; however, they are transcribed and might be analysed according to the certain goals of analysis. The excluded utterances in the current study are:

- Totally unintelligible or partially intelligible utterances.
- Counting sequences, and phrase social responses without evidence of productivity in the rest of the sample;
- Single-word utterances will be excluded to create a new series of measures in the analysis phase.)

Morphemes

1. All inflected morphemes should be separated from the word stem using either backward slash \ -for prefixes- or forward slash / -for suffixes.
2. Content or lexical morpheme; lexeme of the word; free morpheme:
   2.1. Include material or content meaning;
   2.2. Five groups: nouns, verbs, adjectives, adverbs and prepositions; (e.g. man, book, home, good, fast, still ...) (Kalbasi, 2008; Meshkato-Dini, 2008)
3. Functional morphemes; as described in table 1-1.
4. Bound morphemes (inflectional); as described in table 1-2;
5. Clitics; as described in table 1-3.
6. Exceptions:
   6.1. Block phrases such as */inehash/ /inahash/ /inesash/* etc. compute as one morpheme (See also example 1-9).
   6.2. Calculate compound nouns as one morpheme, like */baqvahʃ/*, */ʃahrebazi/, /kenardarya/* as well as compound specific nouns including those with titles like */mamanbozorg/, /aqajun/, /xalemaryam/* etc.

(Continued in next page)

7. Normalizing:
   7.1. Phonetic and phonemic errors
   7.2. Dialectal differences (including the different dialects in producing the prefix marker of imperative verbs */be/bi/bu/* in Isfahani accent: *bi\ʃin = be\ʃin* = sit)
   7.3. Do not normalize morphological errors unless they show a type of phonetic errors: e.g. *mi\ʃin* (=sit) should not be normalised as it shows an error in using prefix *mi* instead of *be* in an imperative verb.

SALT adapted rules for Persian

1. All terminal punctuation and the conventions related to intelligibility; incomplete words and utterances; omitted items; hesitations, repetitions, interruptions, pauses, and reformulations; neologism, symbolic sounds and voices; and automatic speech and counting will be treated according to SALT conventions.
2. Persian-specific codes:
   Transcribing & Coding in colloquial spoken Persian is far different from the formal spoken or written Persian making it more complicated to transcribe. The following codes have been set up to meet the Iranian SLTs' need for a standard transcribing manual to facilitate more robust methodology in research on language sample analysis, particularly using SALT.
   - /e (3rd person verb marker): *mi\xor/e*.
   - /e1 (past or present participle marker): *oftad/e1 ru zamin*.
   - /e2 (e ezafe): *sar/e2 kar*
   - /e3 (emphasis on a definite noun): *un arusak/e3 ra be\de*.
   - /i (single 2nd person present tense) *che goft/i?*
   - /i1 (emphasise on a definite attribution of a noun): *kar/e2 dorost/i2 mi\kon/d.*
   - /i2 (indefinite noun marker): *yek bache/i2 bud/ø.*
   - *chi →che*
   - All the definite pronouns will be linked to the previous and next part of all combinations; */chetor/, /harche/, /inqadr/, /unha/* etc.
   - e (single 3rdperson present tense of "to be") → *ast*
   - /in (plural 2nd person present tense of "to be")→ */in1*
   - /n (plural 3rd person present tense of "to be") → */an*
   - /am (=possessive pronoun) → */am1: kif/am1*
   - /e (=3rd person verb marker) → */d*
   - *ya* (=or) → *ya1*
   - Verbs: *\g/ → \gu/, \sh/ →\sho/, \r/ → \ro/, \d/ → \de/, \zasht/ → \gozasht/*
   - *o, a* (=and) → *va*
   - *o, a* (=direct object marker) → *ra*
   - *bad* (=bad) → *bad1*
   - *chera* (with the meaning: yes) → *chera1*
   - *ye* (=indefinite noun marker) → *yek*
   - *do* (=two) → *do1*
   - All the possessive pronouns are transcribed in the complete written format: */am1, /et, /esh, /emun, /etun, /eshun*

Table 4-2 The Persian Transcription Conventions Protocol (PTCP)

Adapted transcriptions were analysed using *SALT-2012 research version* to explore the language measures as well as errors (Tables 4-4 and 4-5). The grammatical and semantic sub-categories of SLTs' answers that are computable using LSA were also included in the list, i.e. incomplete sentences, short sentences, wrong subject/verb agreement, problem in verbal inflectional morphemes, missing */e-Ezafeh* (addition or genitive sign), missing verbs, missing prepositions, missing conjunctions, missing objectives, difficulty in sentence formulation, low lexical diversity, ambiguous sentences, inappropriate responses to questions. They are of particular interest because grammatical and semantic aspects of expressive language are the focus of the current stage of the study, and Iranian SLTs had mentioned them as the areas of concern in the language samples of children with PLI (see the results of chapter 3).

The measures have been classified into two categories: General LSMs and Persian-specific measures. Within each there are two subcategories of measures and errors (Tables 4-3 to 4-6).

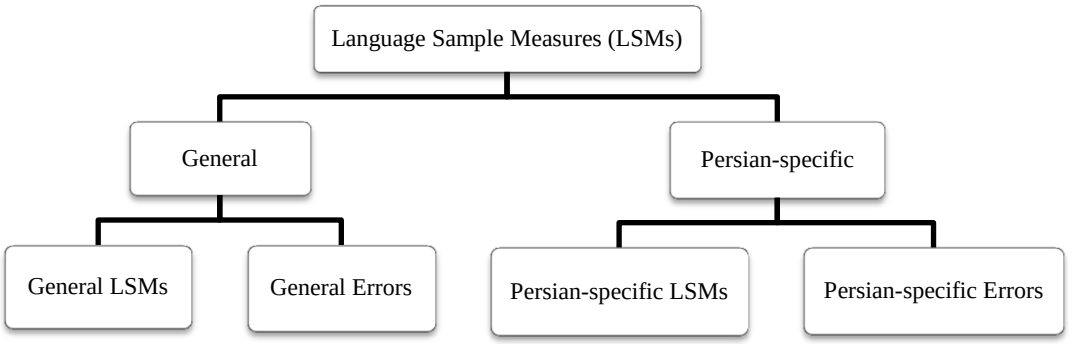


Table 4-3 Measures with subcategories

General Measures*	Acronym
1. Number of Total Complete and Intelligible (C&I) Utterances	NTU
2. Mean Length of Utterances C&I in morphemes	MLUm
3. Mean Length of Utterances C&I in words	MLUw
4. Number of Total Words C&I	NTW
5. Number of Different Words C&I	NDW
6. Total Number of One-Word Utterances	TNOU
7. Mean Length of Utterances in morphemes – excluding one-morpheme utterances	MLUm-exc
8. Mean Length of Utterances in words – excluding one-word utterances	MLUw-exc
9. Total number of verbal morphemes	TNVM
10. Percentage of intelligible utterances	Intelligibility
General Errors	**Acronym**
1. Number of clitics errors	Clitic errors
2. Number of verb inflectional errors (Finite Verb Morphology)	VIE
3. Number of semantic errors	Semantic errors
4. Total number of errors	
5. Total number of grammatical utterances	Grammatical utterances
6. Total number of ungrammatical utterances	Ungrammatical utterances
7. Percentage of grammaticality	Grammaticality
8. Percentage of ungrammaticality	Ungrammaticality
9. Number of missing verb markers	Missing verb marker
10. Number of missing prepositions	missing prepositions
11. Number of missing conjunctions	missing conjunctions
12. Number of missing verbs	missing verbs
13. Number of wrong agreement	wrong agreement
14. Number of wrong word order	wrong word order
15. Number of nonsense strings of words	nonsense strings of words
16. Number of wrong responses to questions	wrong responses

Table 4-4 General LSMs and Errors; *Definitions of all measures in this table are provided in appendix F.

Persian-specific Measures (definition of morphemes are in chapter 1)	**Acronym**
1. Total number of plural marker */ha*	plural marker */ha*
2. Total number of direct object (DO) marker *ra*	DO marker *ra*
3. Total number of prefixes	
4. Total number of progressive verb marker *mi*	progressive verb marker *mi*
Persian-specific Errors (definition of morphemes are in chapter 1)	
1. Total number of missing */ha* (Plural noun marker)	missing */ha*
2. Total number of missing *ra* (direct object marker)	missing *ra*
3. Total number of missing */e Ezafeh* (addition or "genitive sign")	missing */e Ezafeh*
4. Total number of missing *mi* (Progressive marker)	missing *mi*
5. Total number of missing objective clitic	missing objective clitic
6. Total number of missing possessive clitic	missing possessive clitic
All Persian-specific errors (definition of morphemes are in chapter 1, see also table 4-2 for codes)*	
1. Total number of missing *ra* (direct object marker)	
2. Total number of wrong usage of *ra* (direct object marker)	
3. Total number of missing objective clitic	
4. Total number of wrong usage of objective clitic	
5. Total number of missing verb marker	
6. Total number of wrong agreement	
7. Total number of missing *mi* (Progressive marker)	
8. Total number of wrong usage of *mi* (Progressive marker)	
9. Total number of word order error	
10. Total number of missing preposition	
11. Total number of wrong usage of preposition	
12. Total number of missing verb	
13. Total number of wrong usage of verb in terms of meaning	
14. Total number of nonsense string of words (see appendix F)	
15. Total number of missing *i1* (emphasise on a definite attribution of a noun)	
16. Total number of wrong *i1*(emphasise on a definite attribution of a noun)	
17. Total number of missing *i2* (indefinite noun marker)	
18. Total number of wrong *i2* (indefinite noun marker)	
19. Total number of missing *be-eltezami* (Potential mood verb marker)	
20. Total number of wrong usage of *be-eltezami* (Potential mood verb marker)	
21. Total number of missing demonstrative pronoun	
22. Total number of wrong usage of demonstrative pronoun	
23. Total number of missing PP Complement (non-specific object)	
24. Total number of wrong usage of PP Complement (non-specific object)	
25. Total number of missing /*ha* (Plural marker)	
26. Total number of wrong /*ha* (Plural marker)	
27. Total number of missing possessive clitic	
28. Total number of wrong usage of possessive clitic	
29. Total number of missing *e1* (past or present participle marker)	
30. Total number of wrong usage of *e1* (past or present participle marker)	
31. Total number of missing */e-Ezafeh* (addition or "genitive sign")	
32. Total number of wrong usage of */e-Ezafeh*(addition or "genitive sign")	
33. Total number of missing *e3* (emphasis on a definite noun)	
34. Total number of wrong usage of *e3* (emphasis on a definite noun)	
35. Total number of wrong verb tense	
36. Total number of wrong verb mode	
37. Total number of wrong verb root	
38. Total number of missing *bo/ bi/ be/* imperative	
39. Total number of wrong usage of *bo/ bi/ be/* imperative	
40. Total number of missing *ne/* (negative verb marker)	
41. Total number of wrong usage of *ne/* (negative verb marker)	
42. Total number of wrong responses to questions	
43. Total number of missing causative verb marker	
44. Total number of wrong causative verb marker	
45. Total number of missing conjunction	
46. Total number of wrong questions	

Table 4-5 Persian-specific Measures and Errors. *For those measures with no definition for their morphemes (either in the table or in chapter1) or code (in table 4-2), a definition has been provided or the error is defined in appendix F.

All General LSMs were computable directly from the SALT report of complete and intelligible utterances of the first speaker (the child). For one-word utterances, the command *Save as Separate Transcript* from *Explore* was set to the *Utterance Length* of one word for the first speaker. The new transcripts were then analysed using the command *Standard Measures* of the *Analysis*. The Total Number of Verbal Morphemes was calculated from summing up the verbal morphemes reported through the command *Bound Morpheme Table* of the *Analysis*.

Additionally, each transcript was coded according to the type of the error observed in the language sample listed under the title All Persian-specific Errors (Table 4-5). Some of the errors under the name of General Errors were computed from the combination of several errors as follows:

a) Clitic Errors: Sum of all missing Persian clitics listed in section 1.2.1.5, table 1-3, including Missing Objective Clitic, Missing Possessive Clitic, and Missing */e-Ezafeh* (addition or "genitive sign");
b) Verb Inflectional Errors (or Finite Verb Morphology Composite as described by Gladfelter & Leonard, 2012): Sum of all missing verbal inflections listed in table 1-2, including Missing *mi* (Progressive marker), Missing Verb Markers, Wrong Agreement, Missing Potential Mood Verb Marker, and Missing Imperative Mood Verb Marker.
c) Semantic Errors: Sum of those units that via omission or wrong use affect the meaning of the utterance, including Missing Verbs, Wrong Responses, Missing Prepositions, and Nonsense String of Words
d) Total Errors is the sum of all errors listed under the heading All Persian-specific Errors in table 4-5.
e) Total Number of Grammatical Utterances includes those utterances without any error code in the transcript and its percentage is called Grammaticality (As called by Eisenberg & Guo, 2012).
f) Total Number of Ungrammatical Utterances is the result of deduction of total grammatical utterances from total complete and intelligible utterances and the percentage is called Ungrammaticality (As called by Simon-Cereijido & Gutierrez-Clellen, 2007).

4.3.1 Reliability

The content validity of transcribing conventions was frequently checked against utterance and morpheme criteria mentioned in textbooks and articles (e.g. Brown, 1973;

Fletcher & Garman, 1988; Kalbasi, 2008; Klee 1992; Meshkato-Dini, 2008; Miller, 1981; Miller & Iglesias, 2008; Rice et al. 2006) as well as by consulting with supervisors and Iranian linguists to ensure that the criteria encompassed the most agreed characteristics from different sources. Subsequently, as a pilot study, at the beginning of the transcription phase, four language samples were transcribed three times and checked with respect to the utterance and morpheme conventions.

It was not possible to measure inter-rater reliability of the transcriptions because there was no access to a skilled Persian transcriber. In lieu of that, intra-rater reliability based on 10 language samples from the PLI group was calculated. Ten language samples were listened to three times each on two occasions with a time interval of four months, and main LSMs including MLUw, MLUm, NDW, and NTW were calculated. This showed high correlation between the two transcript occasions for the main measures of LSA (see table 4-6). As normal distribution was not assumed due to the small sample size, the non-parametric correlation was calculated using Spearman's rho at the .01 level (2-tailed).

Measure	***r***	***p***
MLUw	.879	.001
MLUm	.976	.000
NDW	.988	.000
NTW	.960	.000

Table 4-6 Intra-rater reliability for four main LSMs based on 10 language samples; MLUw=Mean length of utterance in words, MLUm=Mean length of utterance in morphemes, NDW=Number of different words, NTW=Number of total words.

Moreover, the recordings of all children were listened to three times if there was any doubt about transcription in terms of intelligibility. Then all doubtful utterances or words were transcribed using the SALT convention (X character) to show the unintelligibility of the items and they were excluded from the final SALT analysis which only included Complete and Intelligible Utterances (Miller & Iglesias, 2008).

Reliability was calculated only for main LSMs and did not include error codes. It would be of value if either kind of reliability check, inter-rater or intra-rater, was performed on transcriptions in terms of allocating different codes (as defined in appendix F) to the errors observed, so that the definitions of various codes can also be examined.

4.4 Results

4.4.1 Participants' demographic and background specifications

As shown in figure 4-4, boys with PLI outnumbered the girls, which is not unexpected. The gender difference, however, was not significant between two conditions (Chi-square (1) = 3.843, p>.05). Generally, more boys participated in the study than girls.

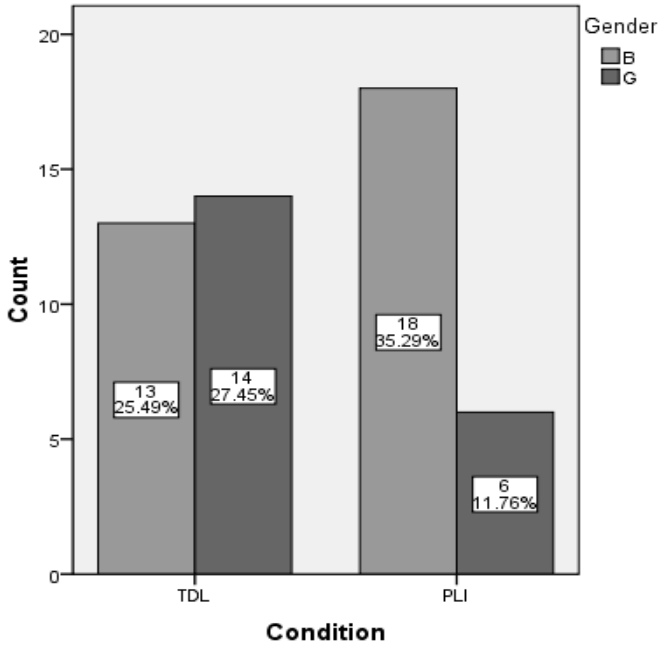


Figure 4-4 Participating children by gender and condition; B=Boy, G=Girl, TDL=children with typically developing language, PLI=children with Primary language impairment

Participants' average age in months was 48.92 (SD=3.56) and 47.33 (SD=4.15) in the groups of children without and with PLI, respectively with no significant difference (t (49) = 1.47, p>.05) (Table 4-7). The groups did not differ in terms of being born pre-term (n=51, Pearson Chi-Square (1) = .054, p>.05) or birth weight (n=51, t (49) = 1.55, p>.05). Exposure frequency to either a foreign language (n=23, t (1.24) = -1.75, p>.05) or television viewing (n=34, t (32) = -1.04, p>.05) also did not show any significant difference between two groups (Table 4-7).

Condition	Age		Birth weight		Exposure to a foreign language (hours/day)		Exposure to TV programs (hours/day)	
	n	*M* (SD)	n	*M* (SD)	n	*M* (SD)	n	*M* (SD)
Children with TDL	27	48.9(3.5)	27	3218(454)	12	2.8(1.2)	18	2.6(1.7)
Children with PLI	24	47.3(4.1)	24	3024(433)	11	8.3(1.4)	16	3.5(2.8)

Table 4-7 Participants' features with no significant difference between children with and without PLI; N=number of respondents, M=mean, SD=standard deviation; PLI=Primary language impairment, TDL=Typically developing language

On the other hand, children with TDL produced their first word significantly earlier than their PLI peers (n= 42, $t(22.48) = -2.88$, $p<.01$) and the age of 1.2 months old sounds too early for producing a meaningful word by the child. It might be due to misinterpretation of one or two parents about the first meaningful word so that they probably considered cooing or babbling as first word. Age of producing two-word phrases was on average 13 months ahead compared to children with PLI (n=43, $t(32.93) = -5.29$, $p<.000$) and they have been looked after by people other than parents for more hours than children with PLI (i.e. day-care, grandparents or other relatives. Note that no non-Persian-speaking carer was reported in response to this question) (n= 38, $t(36) = 2.44$, $p<.05$). Parents' education in the group of children with TDL was significantly higher than their PLI counterparts (father's: n= 48, $t(46) = 3.15$, $p<.01$); mother's: n=49, $t(47) = 2.76$, $p<.01$) (Table 4-8).

Condition	First word (mts)		2-word phrases (mts)		Being taken care by non-parents (hrs/w)		Father education (yrs)		Mother education (yrs)	
	N	M (SD)	N	M (SD)	N	M (SD)	N	M (SD)	N	M (SD)
Children with TDL	24	1.2(3.6)	21	16.7(5.6)	26	27.1(8.8)	27	14.3(2.6)	27	14(2.6)
Children with PLI	18	15.9(7.9)	22	29.9(1.1)	12	19.4(9.3)	21	11.3(4)	22	11.6(3.6)

Table 4-8 Participants' features with significant difference ($p<.05$) between children with and without PLI; N=number of respondents, M=mean, SD=standard deviation; PLI=Primary language impairment, TDL=Typically developing language

A total of 31 children were first-born (61%), and 12 children in both groups were born preterm with a range of one to eight weeks (*M*= 2.75, SD = 2). Ten children had a history of Otitis Media (OM) and 13 had a history of hospitalisation due to preterm complications (n=4), severe flu (n=2), disorders like hyperactivity (n=2), head accidents (n=2) or convulsions (n=3) with no significant effect on development based on the paediatrician's judgement, and their parents did not state any concern about these conditions. With respect to having a family history of behaviour problems, neurological disorder, language impairment or hearing impairment, 25 children had no history versus 15 children with one relative with a history of mentioned problems. A total number of 24 children in both groups were exposed to a foreign language which showed no significant difference between two groups (Tables 4-9 and 4-10).

	Birth rank			Preterm birth		History of OM		History of hospitalization or specific disability		Family history of language-related problems*		
Condition	**1**	**2**	**3**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**	**Do not know**
Children with TDL	17	8	2	6	21	5	22	3	24	3	24	0
Children with PLI	14	9	1	6	18	5	19	10	14	12	11	1
Total	31	17	3	12	39	10	41	13	38	15	25	1

Table 4-9 Medical history among two sample groups; *Such as behaviour problems, neurological disorder, language impairment or hearing impairment; TDL=Typically Developing Language, PLI=Primary Language Impairment; OM=Otitis Media.

	Exposure to other language		Source of other language				
Condition	**Yes**	**No**	**DVDs**	**Satellite channels**	**Parents**	**DVDs + Satellite**	**Parents + Satellite**
Children with TDL	13	14	2	5	1	4	1
Children with PLI	11	13	2	6	1	2	0
Total	24	27	4	11	2	6	1

Table 4-10 Conditions of foreign language experience among two sample groups; TDL=Typically Developing Language, PLI=Primary Language Impairment.

4.4.2 Descriptive features of LSMs

In this section, the descriptive statistics for the language sample measures are divided into two categories as explained in tables 4-4 and 4-5. The average scores of each LSM are reported for each experimental group along with their standard deviation (SD), standard error (SE), and range (Tables 4-11 to 4-14).

	Children with TDL			**Children with PLI**		
General LSMs	***M* (SD)**	**Range**	**SE**	***M* (SD)**	**Range**	**SE**
NTU	184.5 (33.4)	132	6.4	204.7 (57.4)	271	11.7
MLUm	3.83 (.64)	2.35	.12	2.55 (.71)	2.55	.14
MLUw	2.69 (.39)	1.47	.07	1.88 (.40)	1.54	.08
NTW	493 (100)	398	19.2	390 (145)	540	29.6
NDW	151 (256)	95	4.94	108 (321)	119	6.55
TNOU	65 (19.1)	72	3.69	101.8 (38.4)	159	7.84
MLUm-exc	5.20 (.72)	2.86	.13	3.68 (.73)	2.53	.15
MLUw-exc	3.58 (.43)	1.82	.08	2.71 (.32)	1.43	.06
TNVM	73 (23)	85	4.5	55 (36)	117	7.4
Intelligibility	90 (5)	18	1	90 (5.4)	19	1.1

Table 4-11 Descriptive statistics of General LSMs for each group; LSM=Language Sample Measures, TDL=Typically Developing Language, PLI=Primary Language Impairment; M=Mean, SD=Standard Deviation, SE=Standard Error of Mean, NTU=Number of Total Utterances, MLUm=Mean Length of Utterance in morphemes, MLUw= Mean Length of Utterance in words, NTW=Number of Total Words, NDW=Number of Different Words, TNOU=Total Number of One-word Utterances, MLUm-exc.= Mean Length of Utterance in morphemes-excluding one-word utterances, MLUw-exc.= Mean Length of Utterance in words-excluding one-word utterances, TNVM=Total Number of Verbal Morphemes.

General Errors	Children with TDL			Children with PLI		
	***M* (SD)**	**Range**	**SE**	***M* (SD)**	**Range**	**SE**
Number of clitics errors	.18 (.39)	1	.08	1.96 (2.51)	8	.51
Number of verb inflectional errors (Finite Verb Morphology)	1.44 (1.67)	7	.32	3.87 (3.05)	12	.62
Number of semantic errors	.92 (.91)	3	.17	11.8 (9.32)	33	1.9
Total number of errors	7 (4.1)	13	.8	26.3 (13)	49	2.6
Total number of grammatical utterances	177 (32.5)	128	6.2	174 (51)	248	1.4
Total number of ungrammatical utterances	7.2 (3.6)	14	.7	31 (15.5)	55	3.2
Percentage of grammaticality	96 (1.9)	6.7	.36	85 (6.7)	26	1.36
Percentage of ungrammaticality	4 (2)	6.7	.36	15 (7)	26	1.36
Number of missing verb markers	.22 (.7)	3	.13	1.29 (1.8)	8	.37
Number of missing prepositions	.2 (.4)	1	.07	1.42 (2.6)	11	.53
Number of missing conjunctions	.15 (.7)	4	.15	.04 (.2)	1	.04
Number of missing verbs	.41 (.7)	2	.14	3.6 (5.6)	28	1.14
Number of wrong agreement	1.15 (1.2)	5	.24	2.08 (1.6)	6	.33
Number of wrong word order	.56 (.8)	2	.15	1.04 (1.6)	8	.33
Number of nonsense strings of words	.30 (.6)	2	.11	3 (3)	11	.61
Number of wrong responses to questions	.04 (.2)	1	.03	3.8 (5.9)	24	1.20

Table 4-12 Descriptive statistics of General Errors for each group; TDL=Typically Developing Language, PLI=Primary Language Impairment; M=Mean, SD=Standard Deviation, SE=Standard Error of Mean.

	Children with TDL			Children with PLI		
Persian-specific Measures	***M* (SD)**	**Range**	**SE**	**M(SD)**	**Range**	**SE**
Total number of plural marker */ha*	6.7 (4.5)	15	.9	3.1 (3.5)	13	.7
Total number of direct object marker *ra*	19.3 (8.2)	32	1.6	13.5 (10)	34	2.0
Total number of prefixes	67.3 (21)	93	4.0	55.5 (37)	133	7.5
Total number of progressive verb marker *mi*	32.7 (12)	40	2.3	22.1 (22)	83	4.4
Persian-specific Errors						
Total number of missing */ha*	.26 (.8)	4	.15	0	0	0
Total number of missing *ra*	.56 (.8)	3	.15	2.54 (3.7)	12	.75
Total number of missing */e Ezafeh*	.15 (.3)	1	.07	.83 (1.7)	7	.35
Total number of missing *mi*	.07 (.2)	1	.05	.50 (.8)	3	.17
Total number of missing objective clitic	0	0	0	.29 (.75)	3	.15
Total number of missing possessive clitic	.04 (.19)	1	.03	.62 (1.09)	3	.22

Table 4-13 Descriptive statistics of Persian-specific Measures and Errors for each group; TDL=Typically Developing Language, PLI=Primary Language Impairment; M=Mean, SD=Standard Deviation, SE=Standard Error of Mean.

	Children with TDL			Children with PLI		
Persian-specific errors	***M* (SD)**	**Range**	**SE**	***M* (SD)**	**Range**	**SE**
missing *ra* (direct object marker)	.26 (.8)	4	.15	0	0	0
wrong *ra* (direct object marker)	.26 (.44)	1	.86	.37 (.71)	2	.14
missing objective clitic	0	0	0	.29 (.75)	3	.15
wrong objective clitic	.07 (.26)	1	.05	.17 (.63)	3	.13
missing verb marker	.22 (.7)	3	.13	1.29 (1.8)	8	.37
wrong agreement	1.15 (1.2)	5	.24	2.08 (1.6)	6	.33
missing *mi* (Progressive marker)	.07 (.2)	1	.05	.50 (.8)	3	.17
wrong *mi* (Progressive marker)	0	0	0	.17 (.48)	2	.1
word order error	.56 (.8)	2	.15	1.04 (1.6)	8	.33
missing preposition	.2 (.4)	1	.07	1.42 (2.6)	11	.53
wrong preposition	.26 (.52)	2	.1	.58 (1.5)	6	.31
missing verb	.41 (.7)	2	.14	3.6 (5.6)	28	1.14
wrong verb	.33 (.62)	2	.12	.46 (.78)	3	.16
nonsense string of words	.30 (.6)	2	.11	3 (3)	11	.61
missing *i1* (indefinite noun marker)	.04 (.19)	1	.03	.08 (.41)	2	.08
wrong *i1* (indefinite noun marker)	.04 (.19)	1	.03	.04 (.10)	1	.04
missing *i2*	.07 (.26)	1	.05	0	0	0
wrong *i2*	.07 (.26)	1	.05	.08 (.28)	1	.06
missing *be-eltezami* (Potential mood verb marker)	.11 (.42)	2	.08	.37 (.64)	2	.13
wrong *be-eltezami* (Potential mood verb marker)	.11 (.32)	1	.06	.29 (1.42)	7	.29
missing Demonstrative pronoun	.07 (.26)	1	.05	.12 (.49)	2	.09
wrong Demonstrative pronoun	.22 (.50)	2	.09	.12 (.34)	1	.07
missing PP Complement	0	0	0	.04 (.20)	1	.04
wrong PP Complement	.07 (.26)	1	.05	.08 (.28)	1	.06 Table continued

	Children with TDL			Children with PLI		
Persian-specific errors	***M* (SD)**	**Range**	**SE**	***M* (SD)**	**Range**	**SE**
missing */ha* (Plural marker)	.26 (.8)	4	.15	0	0	0
wrong */ha* (Plural marker)	.15 (.45)	2	.09	.17 (.48)	2	.1
missing possessive clitic	.04 (.19)	1	.03	.62 (1.09)	3	.22
wrong possessive clitic	.33 (.62)	2	.12	.50 (.93)	3	.19
missing *e1* (past or present participle marker)	.4 (.19)	1	.03	0	0	0
wrong *e1* (past or present participle marker)	0	0	0	.04 (.20)	1	.04
missing */e-Ezafeh* (addition or "genitive sign")	.15 (.3)	1	.07	.83 (1.7)	7	.35
wrong */e-Ezafeh* (addition or "genitive sign")	.07 (.26)	1	.05	.04 (.20)	1	.04
missing *e3* (emphasis on a definite noun)	0	0	0	.21 (.51)	2	.1
wrong *e3* (emphasis on a definite noun)	0	0	0	.13 (.34)	1	.07
wrong verb tense	.19 (.48)	2	.09	.46 (.72)	2	.14
wrong verb mode	.04 (.19)	1	.03	.04 (.20)	1	.04
wrong verb root	.26 (.44)	1	.08	.13 (.45)	2	.09
missing */bo/bi/be* imperative	.04 (.19)	1	.03	.04 (.20)	1	.04
wrong */bo/bi/be* imperative	0	0	0	.08 (.28)	1	.06
missing */ne* (negative verb marker)	.04 (.19)	1	.03	.17 (.63)	3	.13
wrong */ne* (negative verb marker)	0	0	0	.04 (.20)	1	.04
wrong response	.04 (.2)	1	.03	3.8 (5.9)	24	1.20
missing causative verb marker	0	0	0	.04 (.20)	1	.04
wrong causative verb marker	0	0	0	.04 (.20)	1	.04
missing conjunction	.15 (.7)	4	.15	.04 (.2)	1	.04
wrong question	0	0	0	.21 (.83)	4	.17

Table 4-14 Descriptive statistics of all Persian-specific errors for each group; TDL=Typically Developing Language, PLI=Primary Language Impairment; M=Mean, SD=Standard Deviation, SE=Standard Error of Mean.

4.4.3 Pre-accuracy study of LSMs in Persian

In the pre-accuracy study, the correlation between selected LSMs and age was investigated at the group level to find out how they correlate to age. Accordingly, those age-correlated measures could be entered into the next level of the study which is the first phase of diagnostic accuracy (Sackett & Haynes, 2002). It has been shown that measures related to age development grow consistently with age, and they can be assumed to have good capability to differentiate between TDL and LI (Gavin et al., 1993; Klee, 1992; Klee et al., 2007; Moyle, Ellis Weismer, Evans, & Lindstrom, 2007; Sahakian & Snyder, 2012).

The distribution of each measure was checked to determine the most appropriate test for examining the association between age and the measures. Since the sample size in each group of children with and without PLI was less than 50, the results of a Shapiro-Wilk test of normality were considered to judge normal distribution of each measure within the two groups. Normal distribution of General Measures in both groups was assumed because the p-value of Shapiro-Wilk's test for all of them was greater than .05. This result directed the next step of testing the association between age and the measures: to apply a Pearson correlation test for all measures. The results of exploring the association between age and General Measures in both groups of children have been demonstrated in table 4-15.

	Correlation coefficient (*p*-value)	
General LSMs	**Children with TDL, n=27**	**Children with PLI, n=24**
NTU	-.184 (.35)	.027 (.90)
MLUm	.265 (.18)	.294 (.16)
MLUw	.212 (.28)	.291(.16)
NTW	-.044 (.82)	.155 (.46)
NDW	.050 (.80)	.226 (.28)
TNOU	-.145 (.47)	-.198 (.35)
MLUm-exc	.251(.20)	.206 (.33)
MLUw-exc	.185 (.35)	.220 (.30)
TNVM	.161(.42)	.228 (.28)
Intelligibility (%)	-.104 (.60)	-.263 (.21)

Table 4-15 Correlation between age and General LSMs; LSM=Language Sample Measures, TDL=Typically Developing Language, PLI=Primary Language Impairment, NTU=Number of Total Utterances, MLUm=Mean Length of Utterance in morphemes, MLUw= Mean Length of Utterance in words, NTW=Number of Total Words, NDW=Number of Different Words, TNOU=Total Number of One-word Utterances, MLUm-exc.= Mean Length of Utterance in morphemes-excluding one-word utterances, MLUw-exc.= Mean Length of Utterance in words-excluding one-word utterances, TNVM=Total Number of Verbal Morphemes.

In General Errors, however, the normal distribution could not be assumed for any of the error measures except for Total Grammatical Utterances and Total Ungrammatical Utterances in children with TDL. Normal distribution was seen in the children with PLI for Verb Inflectional Errors; Total Errors; Total Grammatical Utterances and Total Ungrammatical Utterances; Grammaticality and Ungrammaticality, and Wrong Agreement but not for the others. The association between age and General Errors, therefore, is reported for two groups by keeping this in mind (Table 4-16).

	Correlation coefficient (*p*-value)	
General Errors	**Children with TDL, n=27**	**Children with PLI, n=24**
Clitics Errors	-.141 (.48)	.096 (.66)
Verb Inflectional Errors	-.150 (.45)	-.171 (.42)
Semantic Errors	-.059 (.76)	-.080 (.71)
Total Errors	-.261 (.18)	-.059 (.78)
Grammatical Utterances	-.155 (.43)	.033 (.87)
Ungrammatical Utterances	-.311(.11)	-.010 (.96)
Percentage of Grammaticality	.285 (.14)	.003 (.98)
Percentage of Ungrammaticality	-.285 (.14)	-.003 (.98)
Missing verb markers	-.305 (.12)	-.218 (.30)
Missing prepositions	-.055 (.78)	-.288 (.17)
Missing conjunctions	-.253 (.20)	.213 (.31)
Missing verbs	-.251 (.20)	-.117 (.58)
Wrong agreement	-.038 (.85)	-.138 (.51)
Wrong word order	-.450* (.01)	-.190 (.37)
Nonsense string of words	.116 (.56)	.024 (.91)
Wrong responses	.304 (.12)	-.229 (.28)

Table 4-16 Correlation between age and General Errors; * Correlation is significant at the .05 level (2-tailed) or less; TDL=Typically Developing Language, PLI=Primary Language Impairment.

Following the same procedure, a test of normality for Persian-specific measures in both groups did not show normal distribution for the majority of measures except for Total Number of *ra*, Total Number of Prefixes, and Total Number of *mi* for the TDL group, and Total Number of Prefixes for the children with PLI. The correlation between age and Persian-specific Measures is reported in view of this result (Table 4-17).

	Correlation coefficient, r (p-value)	
Persian-specific Measures	**Children with TDL, n=27**	**Children with PLI, n=24**
Total number of plural marker */ha*	.543** (.003)	-.113 (.60)
Total number of direct object marker *ra*	-.097 (.62)	.131 (.54)
Total number of prefixes	.151 (.45)	.219 (.30)
Total number of progressive verb marker *mi*	.204 (.30)	.258 (.22)
Persian-specific Errors		
Missing */ha* (Plural noun marker)	.111 (.58)	0
Wrong usage of *ra*	-.207 (.3)	.104 (.63)
Missing *ra*(direct object marker)	.058 (.77)	.472*(.02)
Missing */e Ezafeh*	-.208 (.29)	.168 (.43)
Missing *mi* (Progressive marker)	.036 (.85)	-.116 (.59)
Missing objective clitic	0	-.328 (.11)
Missing possessive clitic	.101 (.61)	.079 (.71)

Table 4-17 Correlation between age and Persian-specific Measures and Errors; * Correlation is significant at the .05 level (2-tailed); ** Correlation is significant at the .01 level (2-tailed); TDL=Typically Developing Language, PLI=Primary Language Impairment.

The correlation coefficients indicate weak or negligible associations between age and the majority of measures. The only measures correlated to age are Total Number of Plural Marker */ha* ($r(27) = .54$, $p<.01$) and Total Number of Wrong Word Order ($r(27) = -.45$, $p<.01$) in the children with TDL, and Total Number of Missing *ra* (direct object marker) ($r(24) = .47$, $p<.05$) in the children with PLI. In the best case, just less than 30% of variation in these measures can be explained by age (r^2 = 29%, 20% and 22%, respectively) which is interpreted as low (Taylor, 1990). Founded on the results of the pre-accuracy study, the only age-correlated measures were Total Number of Plural Marker */ha*, Wrong Word Order, and Missing *ra*. This dissociation between age and other measures is interpretable by remembering the small sample size in both groups, i.e. 27 in the children with TDL and 24 in the children with PLI, as well as the small range of ages. This affects the correlation coefficient to the extent that less variation in

the groups can be seen in terms of measures in association with age. In fact, it is not due to the dissociation *property* of the measures but due to the narrow range of age within each group as well as small sample size. So, in future studies, a larger sample size with wider age range might document a significant correlation between age and the language measures (Bland & Altman, 2011). Keeping this in mind, measures were entered into the next phase of study according to either possessing good diagnostic accuracy in previous studies or being considered important in diagnosis of PLI by Iranian SLTs in the survey study. Regardless of being classified into these categories, though, some Persian-specific measures were examined because they were exclusively Persian-specific.

4.4.4 Phase I Diagnostic Accuracy

Given the principle of a phase I diagnostic accuracy study (Sackett & Haynes, 2002), the main aim here was determining if differences existed at the group level between children with and without PLI with respect to LSMs.

Because the majority of measures in both groups were normally distributed, the statistical test used for examining group differences was an independent-sample *t*-test. Those measures with at least one group with non-normal distribution were examined by a Mann-Whitney U test. The analysis included the measures examined in the pre-accuracy study.

The review of table 4-18 leads to the conclusion that the mean scores of the majority of General Measures in the children with TDL were significantly different from their PLI counterparts except Number of Total Utterances and Intelligibility. Children with TDL, on average, showed higher MLUm, MLUw, NTW, NDW, MLUm-exc., MLUw-exc., and TNVM than children with PLI; however, the number of one-word utterances in their language samples was fewer than children with PLI. Although one-word utterances were excluded from the main analysis with both groups, number of one-word utterances was still considered as an informative LSM in analysis of both phase I and II.

	Children with TDL, n=27		Children with PLI, n=24				
General LSMs	***M* (SD)**	**Range**	***M* (SD)**	**Range**	***t* (df)**	***p*-value**	**Effect size (95% CI)†**
NTU	184.5 (33.4)	132	204.7 (57.4)	271	-1.561 (49)‡	.125	.43 (-.13, .99)
MLUm	3.83 (.64)	2.35	2.55 (.71)	2.55	6.755 (49)***	.000	-1.87 (-2.53, -1.21)
MLUw	2.69 (.39)	1.47	1.88 (.40)	1.54	7.232 (49)***	.000	-2.02 (-2.70, -1.35)
NTW	493 (100)	398	390 (145)	540	2.962 (49)**	.005	-.82 (-1.40, -.25)
NDW	151 (256)	95	108 (321)	119	5.367 (49)***	.000	-.15 (-.70, .40)
TNOU	65 (19.1)	72	101.8 (38.4)	159	-4.241(32.9)***	.000	1.22 (.62, 1.82)
MLUm-exc	5.20 (.72)	2.86	3.68 (.73)	2.53	7.381(49)***	.000	-2.07 (-2.75, -1.38)
MLUw-exc	3.58 (.43)	1.82	2.71 (.32)	1.43	7.975 (49)***	.000	-2.24 (-2.94, -1.54)
TNVM	73 (23)	85	55 (36)	117	2.136 (38.3)*	.039	-.59 (-1.16, -.03)
Intelligibility (%)	90 (5)	18	90 (5.4)	19	-.387 (49)‡	.700	.00 (-.55, .55)

Table 4-18 Mean (SD) and mean comparison between two groups in terms of General LSMs; *Difference is significant at the .05 level (2-tailed). ** Difference is significant at the .01 level (2-tailed). *** Difference is significant at the .000 level (2-tailed); †Cohen's *d:* effect sizes of .2 or less are considered small, around .5 are medium, and those equal to or greater than .8 are large (Cohen, 1988); ‡Normal distribution of the measure assumed so that t-test is used for testing mean differences; LSM=Language Sample Measures, TDL=Typically Developing Language, PLI=Primary Language Impairment, NTU=Number of Total Utterances, MLUm=Mean length of Utterance in morphemes, MLUw= Mean length of Utterance in words, NTW=Number of Total Words, NDW=Number of Different Words, TNOU=Total Number of One-word Utterances, MLUm-exc.= Mean length of Utterance in morphemes-excluding one-word utterances, MLUw-exc.= Mean length of Utterance in words-excluding one-word utterances, TNVM=Total Number of verbal Morphemes.

Similar results were observed in General Errors in which three measures including Total Number of Grammatical Utterances, Missing Conjunctions, and Wrong Word Order were unable to differentiate between the two groups of children with TDL and PLI (Tables 4-19). According to the results shown in table 4-19, Children with PLI had significantly higher errors than children with TDL in Clitic Errors, Verb Inflectional Errors, Semantic Errors, Total Errors, Total Number of Ungrammatical Utterances, Percent of Ungrammaticality, Missing verb markers, Missing prepositions, Missing verbs, Wrong Agreement, Nonsense String of Words, and Wrong Responses. The percentage of Grammaticality in children with PLI, however, was significantly lower than their TDL counterparts.

General Errors	Children with TDL, n=27 *M* (SD)	Range	Children with PLI, n=24 *M* (SD)	Range	Test quantity (df)	p-value	Effect size (95% CI) †
Clitics Errors	.18 (.39)	1	1.96 (2.51)	8	170*	.001	1.00 (.42, 1.59)
Verb Inflectional Errors (Finite verb Morphology)	1.44 (1.67)	7	3.87 (3.05)	12	156.5*	.001	.99 (.41, 1.57)
Semantic Errors	.92 (.91)	3	11.8 (9.32)	33	34*	.000	1.67 (1.03, 2.31)
Total Errors	7 (4.1)	13	26.3 (13)	49	47.5*	.000	2.02 (1.35, 2.70)
Total Number of Grammatical Utterances	177 (32.5)	128	174 (51)	248	.28 (49) ‡	.781	-.07 (-.62, .48)
Total Number of Ungrammatical Utterances	7.2 (3.6)	14	31 (15.5)	55	-7.24 (25.1)* ‡	.000	2.14 (1.45, 2.83)
Percentage of Grammaticality	96 (1.9)	6.7	85 (6.7)	26	19*	.000	-2.26 (-2.96, -1.56)
Percentage of Ungrammaticality	4 (2)	6.7	15 (7)	26	19*	.000	2.16 (1.47, 2.85)
Missing verb markers	.22 (.7)	3	1.29 (1.8)	8	184**	.001	.79 (.22, 1.36)
Missing prepositions	.2 (.4)	1	1.42 (2.6)	11	234*	.034	.67 (.10, 1.23)
Missing conjunctions	.15 (.7)	4	.04 (.2)	1	323	.955	-.21 (-.76, .35)
Missing verbs	.41 (.7)	2	3.6 (5.6)	28	96.5***	.000	.81 (.24, 1.38)
Wrong agreement	1.15 (1.2)	5	2.08 (1.6)	6	210*	.027	.65 (.09, 1.22)
Wrong word order	.56 (.8)	2	1.04 (1.6)	8	256.5	.163	.38 (-.17, .94)
Nonsense string of words	.30 (.6)	2	3 (3)	11	88***	.000	1.27 (.66, 1.87)
Wrong responses	.04 (.2)	1	3.8 (5.9)	24	127***	.000	.92 (.34, 1.49)

Table 4-19 Mean (SD) and mean comparison between two groups in terms of General Errors; *Difference is significant at the .05 level (2-tailed). ** Difference is significant at the .01 level (2-tailed). *** Difference is significant at the .000 level (2-tailed); †Cohen's *d:* effect sizes of .2 or less is considered small, around .5 are medium, and those equal or greater than .8 are large (Cohen, 1988); ‡Normal distribution of the measure assumed so that t-test is used for testing mean differences; TDL=Typically Developing Language, PLI=Primary Language Impairment.

With regard to Persian-specific measures, six out of eight measures differentiated between the two groups of children (Table 4-20). On average, children with TDL expressed significantly higher total number of plural marker */ha,* direct object marker *ra*, prefixes, and progressive verb marker *mi*. However, the missing rate of *ra* (direct object marker), */e Ezafeh* (addition or "genitive sign"), *mi* (Progressive marker), objective clitic, and possessive clitic in the language samples of children with TDL were significantly less than children with PLI. Affixes were calculated in different measures; some of them were composites and have been defined in appendix F.

The majority of effect sizes indicate that the size of difference between observed averages of each measure in two groups was remarkable. In terms of Cohen's interpretation of his effect index, effect sizes of .2 or less are considered small, around .5 are medium and those equal to or greater than .8 are large (Cohen, 1988). Nevertheless, simply relying on this prescribed explanation of the magnitude of difference is not advised as a large effect size does not necessarily mean that a measure is better than the one with small effect size. The measures should be assayed in terms of clinical applicability (Durlak, 2009) which is the job of phase II of the diagnostic accuracy stage in the current study.

Persian-specific Measures	Children with TDL, n=27		Children with PLI, n=24		Test quantity(df)	*p*-value	Effect size (95% CI)†
	***M* (SD)**	**Range**	***M* (SD)**	**Range**			
Total number of plural marker */ha*	6.7 (4.5)	15	3.1 (3.5)	13	150**	.001	-.87 (-1.45, -.30)
Total number of direct object marker *ra*	19.3 (8.2)	32	13.5 (10)	34	195*	.015	-.63 (-1.19, -.07)
Total number of prefixes	67.3 (21)	93	55.5 (37)	133	1.38 (35.3)‡	.17	-.39 (-.95, .16)
Total number of progressive verb marker *mi*	32.7 (12)	40	22.1 (22)	83	179.5**	.006	-.60 (-1.16, -.04)
Persian-specific Errors							
Missing */ha* (Plural noun marker)	.26 (.8)	4	0	0	276	.052	-.44 (-1.00, .12)
Wrong usage of *ra*	.26 (.44)	1	.37 (.71)	2	316.5	.852	-.19 (-.74, .36)
Missing *ra* (direct object marker)	.56 (.8)	3	2.54 (3.7)	12	219*	.033	.75 (.18, 1.32)
Missing */e Ezafeh*(addition or "genitive sign")	.15 (.3)	1	.83 (1.7)	7	244.5*	.049	.56 (.00, 1.13)
Missing *mi* (Progressive marker)	.07 (.2)	1	.50 (.8)	3	237*	.018	.75 (.18, 1.32)
Missing objective clitic	0	0	.29 (.75)	3	270*	.029	-
Missing possessive clitic	.04 (.19)	1	.62 (1.09)	3	239*	.011	-.58 (-1.01, -.15)

Table 4-20 Mean (SD) and mean comparison between two groups in terms of Persian-specific Measures and Errors. *Difference is significant at the .05 level (2-tailed). ** Difference is significant at the .01 level (2-tailed). *** Difference is significant at the .000 level (2-tailed); †Cohen's *d:* effect sizes of .2 or less is considered small, around .5 are medium, and those equal or greater than .8 are large (Cohen, 1988); ‡Normal distribution of the measure assumed so that t-test is used for testing mean differences; TDL=Typically Developing Language, PLI=Primary Language Impairment.

4.4.5 Phase II Diagnostic Accuracy

The phase II study examined the differences in LSMs to find out whether each one could be allocated to either a TDL or PLI child without being aware of the child's initial grouping.

Up to this point, the measures have been screened for finding which ones are most likely to correctly identify children with and without PLI. The criteria applied in the screening of the most capable measures were as follows:

1) Covered by the best evidence in previous studies;
2) Mentioned by Iranian SLTs as benchmarks in their own assessment procedures (although their procedures in calculating different LSMs are different from the conventions applied in the current study, and Iranian SLTs raised the issue that sometimes they record language samples with no purpose of calculating any specific LSM);
3) Being sensitive to age;
4) Capability of distinguishing between children with and without PLI (phase I DA). This criterion was considered in selecting the measures to enter into the phase II DA. The measures selected for the phase II study are shown in table 4-21. Calculation of the composite measures (e.g. semantic errors, grammaticality etc.) has been explained in appendix F.

General LSMs	General Errors	Persian-specific Measures and Errors
1) MLUm	1) Clitics Errors	1) Total number of plural marker */ha*
2) MLUw	2) Verb Inflectional Errors (Finite verb Morphology)	2) Total number of direct object marker *ra*
3) NTW	3) Semantic Errors	3) Total number of progressive verb marker *mi*
4) NDW	4) Total Errors	4) Missing *ra* (direct object marker)
5) TNOU	5) Total Number of Ungrammatical Utterances	5) Missing */e Ezafeh* (addition or "genitive sign")
6) MLUm-exc	6) Percentage of Grammaticality	6) Missing *mi* (Progressive marker)
7) MLUw-exc	7) Percentage of Ungrammaticality	7) Missing objective clitic
8) TNVM	8) Missing verb markers	8) Missing possessive clitic
	9) Missing prepositions	
	10) Missing verbs	
	11) Wrong agreement	
	12) Nonsense string of words	
	13) Wrong responses	

Table 4-21 LSMs to be analysed in a phase II diagnostic accuracy study; LSM=Language Sample Measures, NTU=Number of Total Utterances, MLUm=Mean Length of Utterance in morphemes, MLUw= Mean Length of Utterance in words, NTW=Number of Total Words, NDW=Number of Different Words, TNOU=Total Number of One-word Utterances, MLUm-exc.= Mean Length of Utterance in morphemes-excluding one-word utterances, MLUw-exc.= Mean Length of Utterance in words-excluding one-word utterances, TNVM=Total Number of Verbal Morphemes.

To find out the diagnostic accuracy of the aforementioned LSMs, they were examined using the Receiver Operative Characteristics curves command of SPSS-19. ROC is designed to portray the "true cases" of impairment against "false positives" and ROC curves are mainly employed to explore the best cut-off points for a given measure or test score which result in the best measures of diagnosis, i.e. sensitivity (S_n), specificity (S_p) and likelihood ratios (LR). In a 2×2 contingency table, as shown in table 4-23, true positive (TP) shows the number of cases with the target condition who were also diagnosed as affected by the index test, whereas false positive (FP) shows the frequency of cases without the target condition who are diagnosed as affected by the index test. True negative (TN) shows the frequency of cases who do not have the target condition and were also diagnosed as unaffected by the index test, whilst false negative (FN) shows the number of cases with the target condition but diagnosed as unaffected by the index test. The cut-offs derived from ROC statistics show the cut-off point in a measure which is the point of reference for deciding a person's health condition in that given

measure; any higher or lower score than this point would be interpreted as either affected or unaffected based on the definition of the index test. ROC curves "nicely display the trade-offs of using one or more cut-offs for the test" (Haynes et al., 2006 p. 284) by plotting S_n by $1\text{-}S_p$. Two of most referable properties of ROC curves are as follows:

1) The size of the largest difference between S_n and $1\text{-}S_p$, defined as Youden's Index, statistically indicated by *J*. Its statistical definition is "the farthest point from the line of equality (diagonal line)" (*ROC Curve*, 2012 p. 7). So, the equation for *J* is "maximum (Sensitivity + Specificity – 1)" which is "the maximum vertical distance between the ROC curve and the diagonal line" (Redmond et al., 2011 p. 108) . The resultant S_n and $1\text{-}S_p$ coordinates are the best ones amongst other points and its corresponding numerical value is the optimal cut-off of the measure (Redmond et al., 2011).
2) The Area Under Curve (AUC) which is another property of the relationship between S_n and $1\text{-}S_p$ shows how far the curve is from the lower right corner of the ROC curve box in terms of area. A measure is best if its AUC is the largest, i.e. very close to the upper left corner of the graph (Haynes et al., 2006). When comparing the measures, those with the largest AUC would score better and are more capable in distinguishing between children with and without PLI. Haynes at al. (2006) consider an AUC of .8 and higher as "a reasonable powerful" value (p. 351). The respective ROC curves of LSMs are shown in figures 4-5 to 4-33.

AUC and corresponding standard error, sensitivity, specificity, LR+ and LR-, and Diagnostic Odds Ratio (DOR, overall accuracy) of each LSM are calculated and indicated in tables 4-24, 4-26, and 4-28. The related interpretation of each diagnostic accuracy value is as follows (See also table 4-22):

1) The closer the AUC to one, the better. The best diagnostic test provides an AUC of 1. An AUC of .50 indicates that the test has no potential in distinguishing the population with and without the target condition. Moreover, if the confidence interval of an AUC contains the .50 area, it means that the test might not be a suitable one for the purpose of diagnosis (Hanley & McNeil, 1982; Zweig & Campbell, 1993).

2) A low standard error of a given AUC is an indication of the probable accuracy of the AUC of a measure in the current sample compared to the sample population.
3) The optimal cut-off points are the indications of the best standard deviation of a given measure. Every single new case of the suspected condition's score can be compared to this cut-off point and judged in terms of possessing the target condition.
4) Sensitivity and specificity are defined as the accuracy measures of a given test in identifying cases with and without the expected condition, respectively. The best measures hold a S_n and S_p of 1 which is very rare in reality; however, the closest number to one is optimal (Haynes et al., 2006; Dollaghan, 2007). Accuracy values (sensitivity, specificity) are judged "good" if between 90 and 100; "fair" if between 80 and 89; and "inadequate" if below 80 (Plante & Vance, 1994). It is also desirable that their CIs do not include the poor values as described.
5) The vulnerability of S_n and S_p to the sample size and base rate of the scores in the target population "limits their usefulness as accuracy metrics, especially when they are derived from samples in which the base rate is low" (Dollaghan, 2007 p. 93). LRs, however, are less affected by base rate effect and more preferable in populations in which base rates are located in the two extremes of the spectrum of the target score (Dollaghan, 2007). "The likelihood ratio is the likelihood that a given test result would be expected in a patient with the target disorder compared to the likelihood that the same result would be expected in a patient without the target disorder" (*Likelihood Ratios*, 2012, para. 1). LR+ greater than 10 and LR- smaller than .10 are desirable and indicate greater confidence in interpreting that a given score comes from a person with the target condition (Dollaghan, 2007).
6) Overall accuracy or diagnostic odds ratio (DOR) is another measure arrived at by dividing LR+ by LR- (Haynes et al., 2006). This accuracy measure has been claimed as the only unified statistic for accurately directly comparing diagnostic tests (Glas et al., 2003). Higher DORs show the better overall capability of the test in classifying cases as with and without target condition.

Diagnostic accuracy measures used in this study	Calculation	Definition
Sensitivity (true positive rate)	TP/(TP + FN)	Proportion of positive test results among people with target condition.
Specificity (true negative rate)	TN/(TN + FP)	Proportion negative test results among people without target condition.
Positive likelihood ratio (LR+)	Sensitivity/(1-Specificity)	The likelihood that a positive test result is found in people with target condition as opposed to people without it.
Negative likelihood ratio (LR-)	(1-Sensitivity)/Specificity	The likelihood that a negative test result is found in people with target condition as opposed to people without it.
Diagnostic Odds Ratio (DOR)	LR+ / LR-	Overall accuracy
Youden's index	Sensitivity + Specificity - 1	Maximum vertical distance between ROC curve and diagonal line, represents the optimal cut-off point.

Table 4-22 A summary of descriptions of diagnostic accuracy measures (Glas et al., 2003; Haynes et al., 2006; Redmond, Thompson, & Goldstein, 2011)**;** TP=True Positive, FP=False Positive, TN=True Negative, FN=False Negative

In order to compute the CIs and LRs as well as DOR, the children were re-classified using the optimal cut-offs of each LSM. Thirty one LSMs were individually recoded into different variables in SPSS-19 and the resultant counts of identified children with TDL and with PLI were entered into the rows of Crosstabs command. Consequently, 2×2 contingency tables were created for each measure (Tables 4-23, 4-25, 4-27) and were then employed to calculate 95% confidence intervals (CI) for sensitivity, specificity, and LRs using an online statistical calculator available at the Centre for Evidence-Based Medicine (http://ktclearinghouse.ca/cebm/toolbox/statscalc). DORs and their CIs also were calculated using the Excel Worksheet called 'CI Calculator' available at http://vl.academicdirect.org/applied_statistics/binomial_distribution/ref/CIcalculator.xls.

		Reference standard	
Index test: General LSM		**Diagnosed PLI**	**Diagnosed TDL**
MLUm	**PLI**	22 (TP)	6 (FP)
	TDL	2 (FN)	21 (TN)
MLUw	**PLI**	22	6
	TDL	2	21
NTW	**PLI**	11	1
	TDL	13	26
NDW	**PLI**	19	6
	TDL	5	21
TNOU	**PLI**	18	4
	TDL	6	23
MLUm-exc.	**PLI**	16	0
	TDL	8	27
MLUw-exc.	**PLI**	20	0
	TDL	4	27
TNVM	**PLI**	11	2
	TDL	13	25

Table 4-23 Correspondence between general language sample measures (LSMs) and clinical diagnoses in terms of number of children; TDL=Typically Developing Language, PLI=Primary Language Impairment, LSM=Language Sample Measures, MLUm=Mean length of Utterance in morphemes, MLUw= Mean length of Utterance in words, NTW=Number of Total Words, NDW=Number of Different Words, TNOU=Total Number of One-word Utterances, MLUm-exc.= Mean length of Utterance in morphemes-excluding one-word utterances, MLUw-exc.= Mean length of Utterance in words-excluding one-word utterances, TNVM=Total Number of verbal Morphemes; TP=True Positive, FP=False Positive, TN=True Negative, FN=False Negative.

General LSMs	AUC (95% CI)	SE	Optimal cut-off	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)	DOR (95% CI)
MLUm	.906 (.828, .984)	.040	3.39	.92 (.74, .98)	.78 (.59, .89)	4.12(2.0, 8.4)	.11(.03, .4)	38.5 (6.9,212.5)
MLUw	.928 (.863, .993)	.033	2.37	.92 (.74, .98)	.78 (.59, .89)	4.12(2.0, 8.4)	.11(.03, .4)	38.5 (6.9,212.5)
NTW	.721 (.577, .864)	.073	359	.46 (.28, .65)	.96 (.82, .99)	12.37(1.7,88.9)	.56(.4, .8)	22 (2.6,189.4)
NDW	.848 (.743, .953)	.053	132	.79 (.59, .91)	.78 (.59, .89)	3.56(1.7, 7.4)	.27(.1, .6)	13.3 (3.5,50.7)
TNOU	.822 (.700, .943)	.062	83	.75 (.55, .88)	.85 (.67, .94)	5.06(2, 12.9)	.29(.1, .6)	17.25(4.2,70.5)
MLUm-exc.	.923 (.855, .991)	.035	4.08	.66 (.46, .81)	.98 (.85, .99)	36.96(2.3, 585)	.35(.2, .6)	106.8(5.8,1973)
MLUw-exc.	.950 (.895, 1.00)	.028	2.96	.82 (.63, .92)	.98 (.85, .99)	45.92(2.9, 720)	.18(.08, .4)	250.5(12.7,4919)
TNVM	.648 (.488, .808)	.082	42.5	.46 (.28, .65)	.93 (.77, .98)	6.19(1.5, 25.1)	.58(.4, .8)	10.57(2,55)

Table 3-24 Diagnostic values (with 95% CIs) of General LSMs with best cut-off points; LSM=Language Sample Measures, MLUm=Mean length of Utterance in morphemes, MLUw= Mean length of Utterance in words, NTW=Number of Total Words, NDW=Number of Different Words, TNOU=Total Number of One-word Utterances, MLUm-exc.= Mean length of Utterance in morphemes-excluding one-word utterances, MLUw-exc.= Mean length of Utterance in words-excluding one-word utterances, TNVM=Total Number of verbal Morphemes; AUC=Area under Curve, SE=Standard Error, LR=Likelihood Ratio, DOR=Diagnostic Odds Ratio, CI=Confidence Interval.

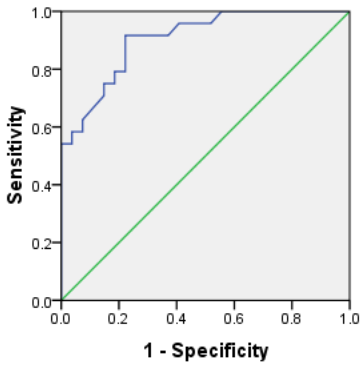


Figure 4-5 ROC curve of MLUm

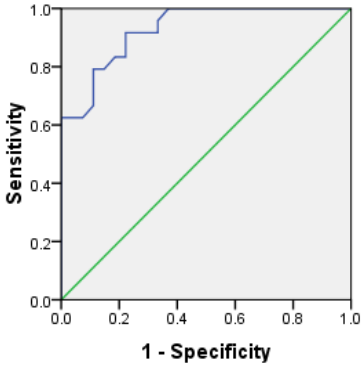


Figure 4-6 ROC curve of MLUw

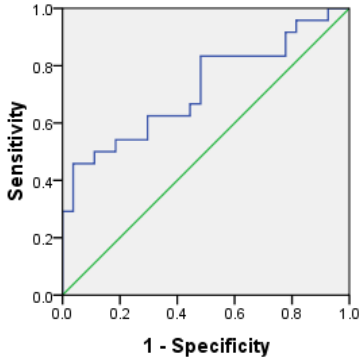


Figure 4-7 ROC curve of NTW

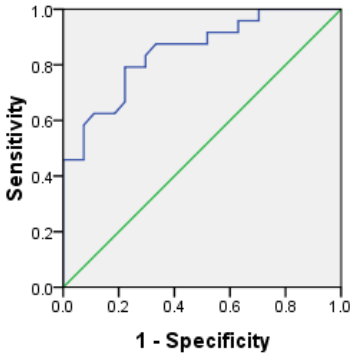


Figure 4-8 ROC curve of NDW

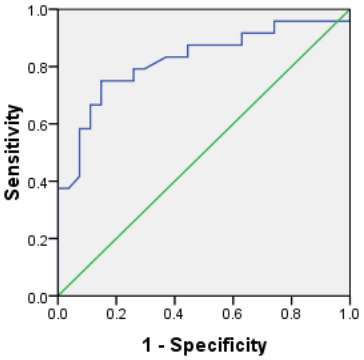


Figure 4-9 ROC curve of TNOU

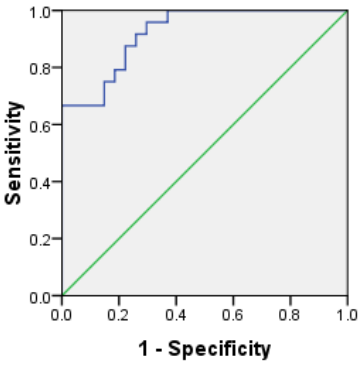


Figure 4-10 ROC curve of MLUm-exc.

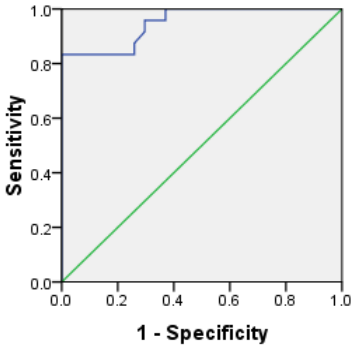


Figure 4-11 ROC curve of MLUw-exc

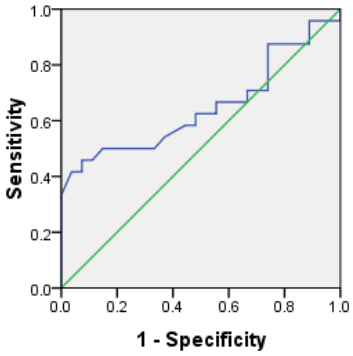


Figure 4-12 ROC curve of TNVM

Index test: General Errors		Reference standard	
		Diagnosed PLI	**Diagnosed TDL**
Clitics Errors	**PLI**	10	0
	TDL	14	27
Verb Inflectional Errors (Finite verb Morphology)	**PLI**	12	2
	TDL	12	25
Semantic Errors	**PLI**	22	1
	TDL	2	26
Total Errors	**PLI**	18	0
	TDL	6	27
Ungrammatical Utterances	**PLI**	21	1
	TDL	3	26
Percentage of Grammaticality	**PLI**	24	4
	TDL	0	23
Percentage of Ungrammaticality	**PLI**	24	4
	TDL	0	23
Missing verb markers	**PLI**	13	3
	TDL	11	24
Missing Prepositions	**PLI**	6	0
	TDL	18	27
Missing Verbs	**PLI**	20	7
	TDL	4	20
Wrong agreement	**PLI**	8	2
	TDL	16	25
Nonsense string of words	**PLI**	20	6
	TDL	4	21
Wrong Response	**PLI**	15	1
	TDL	9	26

Table 4-25 Correspondence between General Errors and clinical diagnoses in terms of number of children; TDL=Typically Developing Language, PLI=Primary Language Impairment.

General Errors	AUC (95% CI)	SE	Optimal cut-off	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)	DOR (95% CI)
Clitics Errors	.738 (.595, .880)	.073	1.50	.42 (.25, .61)	.98 (.85, 1)	23.5(1.4,381)	.59(.4,.8)	39.83(2.2,729)
Verb Inflectional Errors	.758 (.625, .892)	.068	3.50	.50 (.31, .69)	.93 (.77, .98)	6.75(1.7,27.1)	.54(.3, .8)	12.5(2.4,65)
Semantic Errors	.948 (.874, 1.02)	.038	2.50	.92 (.74, .98)	.96 (.82, .99)	24.75(3.6,170)	.09(.02,.3)	286.0(24.2,3370)
Total Errors	.927 (.847, 1.00)	.041	15.5	.74 (.54, .87)	.98 (.85, 1)	41.44(2.6,653)	.26(.1, .5)	156.5(8.3,2950)
Ungrammatical Utterances	.958 (.906, 1.01)	.027	13.50	.87 (.69, .96)	.96 (.82, .99)	23.62(3.4,162.6)	.13(.04,.3)	182.0(17.6,1880)
Percentage of Grammaticality	.971 (.934, 1.00)	.019	94.25	.98 (.83, 1)	.84 (.66, .92)	6.1(2.6,14.2)	.02(.00,.3)	255.9(13,5018)
Percentage of Ungrammaticality	.971 (.934, 1.00)	.019	5.70	.98 (.83, 1)	.84 (.66, .92)	6.1(2.6,14.2)	.02(.00,.3)	255.9(13,5018)
Missing verb markers	.716 (.570, .862)	.074	.50	.54 (.35, .72)	.89 (.72, .96)	4.87(1.6,15.1)	.52(.3, .8)	9.45(2.2,40.1)
Missing prepositions	.639 (.483, .794)	.079	1.50	.26 (.13, .45)	.98 (.85, 1)	14.86(.9,250)	.75(.6, .9)	19.86(1.05,375)
Missing verbs	.851 (.741, .961)	.056	.50	.83 (.64, .93)	.74 (.55, .87)	3.21(1.6,6.2)	.22(.09,.6)	14.27(3.6,56.5)
Wrong agreement	.676 (.527, .825)	.076	2.50	.33 (.18, .53)	.93 (.77, .98)	4.5(1.06,19.1)	.72(.5,1)	6.25(1.2,33.2)
Nonsense string of words	.864 (.758, .971)	.054	.50	.83 (.64, .93)	.78 (.59, .89)	3.75(1.8,7.8)	.21(.09,.5)	17.5(4.3,71.4)
Wrong responses	.804 (.674, .934)	.066	.50	.62 (.43, .79)	.96 (.82, .99)	16.87(2.4,118.4)	.39(.2, .6)	43.33(5,376)

Table 4-26 Diagnostic values (with 95% CIs) of General Errors with best cut-off points; AUC=Area under Curve, SE=Standard Error, LR=Likelihood Ratio, DOR=Diagnostic Odds Ratio, CI=Confidence Interval.

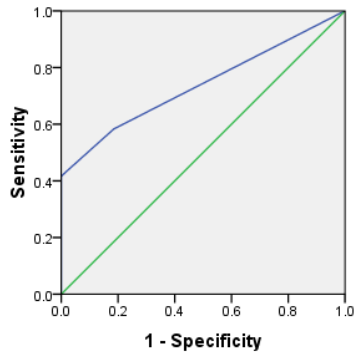


Figure 4-13 ROC curve of Clitic Errors

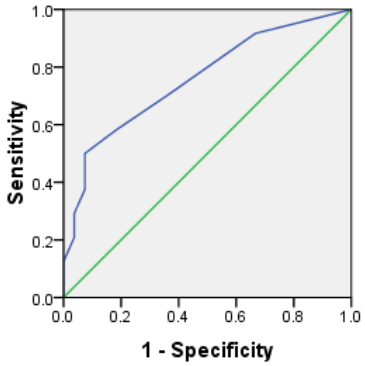


Figure 4-14 ROC curve of Verb Inflectional Errors

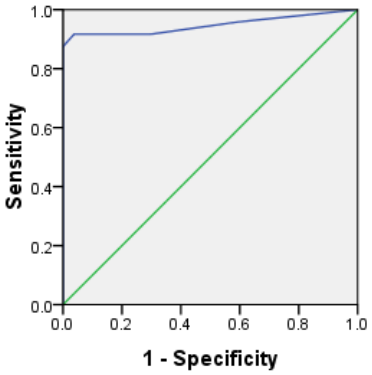


Figure 4-15 ROC curve of Semantic Errors

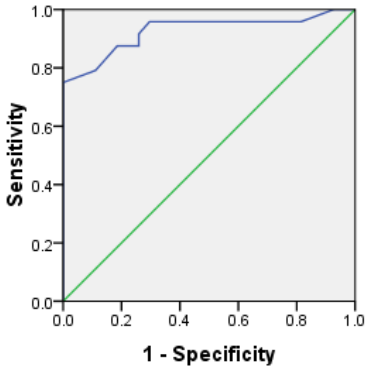


Figure 4-16 ROC curve of Total Errors

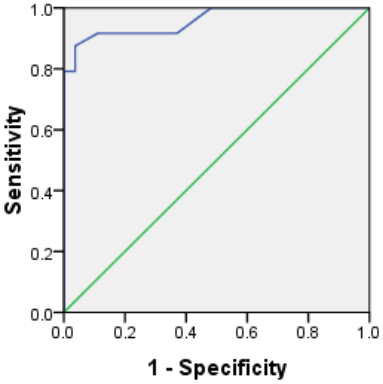


Figure 4-17 ROC curve of Total Ungrammatical Utterances

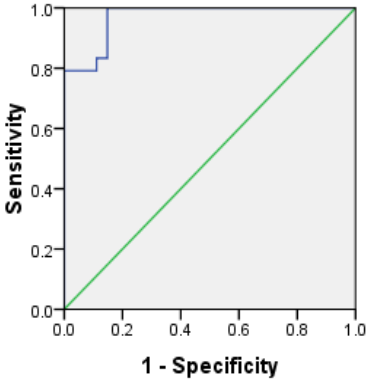


Figure 4-18 ROC curve of Grammaticality

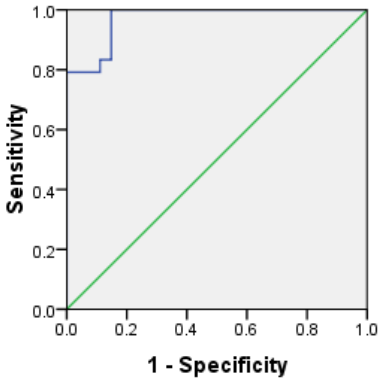


Figure 4-19 ROC curve of Ungrammaticality

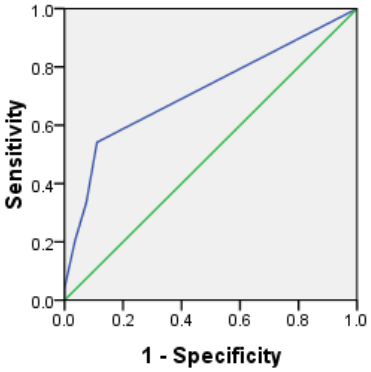


Figure 4-20 ROC curve of Missing Verb Markers

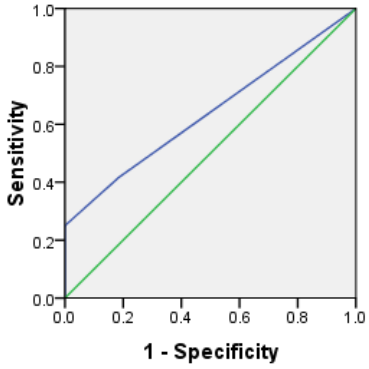


Figure4-21 ROC curve of Missing Prepositions

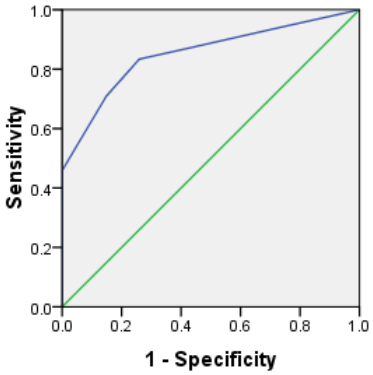


Figure 4-22 ROC curve of Missing Verbs

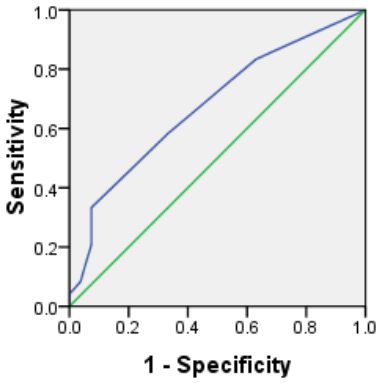


Figure 4-23 ROC curve of Wrong Agreement

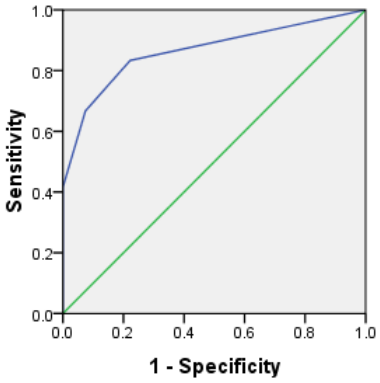


Figure 4-24 ROC curve of Nonsense String of Words

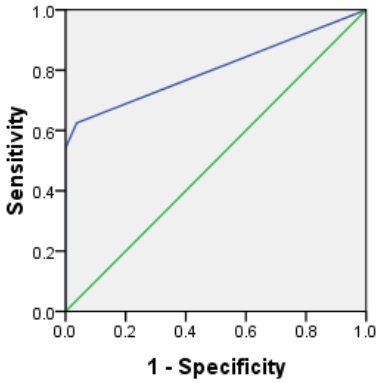


Figure 4-25 ROC curve of Wrong Response

Index test: Persian-specific Measures and Errors		Reference standard	
		Diagnosed PLI	**Diagnosed TDL**
Total number of plural marker */ha*	**PLI**	16	2
	TDL	8	25
Total number of direct object marker *ra*	**PLI**	12	4
	TDL	12	23
Total number of progressive verb marker *mi*	**PLI**	13	0
	TDL	11	27
Missing *ra*	**PLI**	9	3
	TDL	15	24
Missing */e Ezafeh*	**PLI**	9	4
	TDL	15	23
Missing *mi*	**PLI**	8	2
	TDL	16	25
Missing objective clitic	**PLI**	4	0
	TDL	20	27
Missing possessive clitic	**PLI**	7	1
	TDL	17	26

Table 4-27 Correspondence between Persian-specific Measures and Errors, and clinical diagnoses in terms of number of children; TDL=Typically Developing Language, PLI=Primary Language Impairment.

Persian-specific Measures and Errors	**AUC (95% CI)**	**SE**	**Optimal cut-off**	**Sensitivity (95% CI)**	**Specificity (95% CI)**	**LR+ (95% CI)**	**LR- (95% CI)**	**DOR (95% CI)**
Total number of plural marker */ha*	.769 (.632, .905)	.070	2.50	.68 (.47, .82)	.93 (.77, .98)	9(2.3, 35.1)	.36(.2, .6)	25.0(4.7, 133)
Total number of direct object marker *ra*	.699 (.550, .848)	.076	11.50	.50 (.31, .69)	.85 (.67, .94)	3.37(1.2, 9.1)	.59(.4, .9)	5.75(1.5, 21.7)
Total number of progressive verb marker *mi*	.723 (.570, .876)	.078	13	.54 (.35, .72)	.98 (.85, 1)	30.24(1.9,483)	.47(.3, .7)	64.56(3.5,1180)
Missing *ra*	.662 (.510, .814)	.078	1.50	.37 (.21, .57)	.89 (.72, .96)	3.37(1, 11)	.70(.5, .9)	4.8(1.1, 20.6)
Missing */e Ezafeh*	.623 (.466, .779)	.080	.50	.37 (.21. .57)	.85 (.67, .94)	2.53(.9, 7.1)	.73(.5, 1)	3.45(.9, 13.2)
Missing *mi*	.634 (.478, .790)	.080	.50	.33 (.18, .53)	.93 (.77, .98)	4.5(1, 19.1)	.72(.5, 1)	6.25(1.2, 33.2)
Missing objective clitic	.583 (.424, .743)	.081	.50	.18 (.07, .37)	.98 (.85, 1)	10.28(.6, 181)	.83(.7, 1)	12.37(.63, 243)
Missing possessive clitic	.631 (.475, .788)	.080	.50	.29 (15, .49)	.96 (.82, .99)	7.87(1.0, 59.5)	.74(.6, 1)	10.71(1.2, 95)

Table 4-28 Diagnostic values (with 95% CIs) of Persian-specific Measures and Errors with best cut-off points; AUC=Area under Curve, SE=Standard Error, LR=Likelihood Ratio, DOR=Diagnostic Odds Ratio, CI=Confidence Interval.

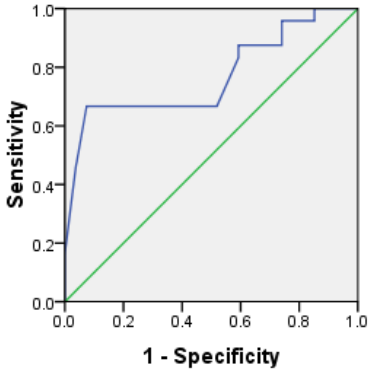


Figure 4-26 ROC curve of *ha*

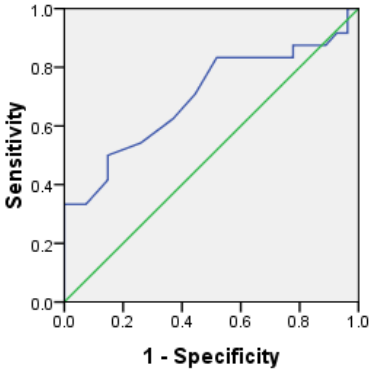


Figure 4-27 ROC curve of *ra*

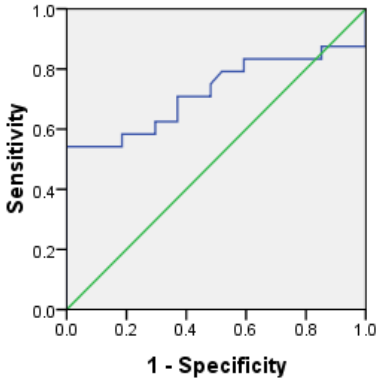


Figure 4-28 ROC curve of *mi*

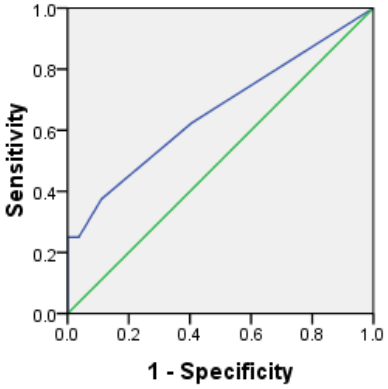


Figure 4-29 ROC curve of Missing *ra*

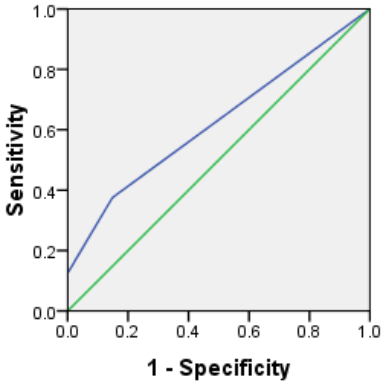


Figure 4-30 ROC of Missing */e-Ezafeh*

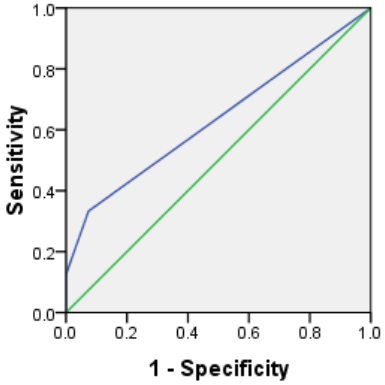


Figure 4-31ROC curve of Missing *mi*

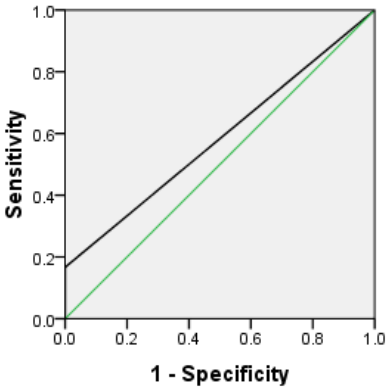


Figure 4-32 ROC of Missing Objective Clitics

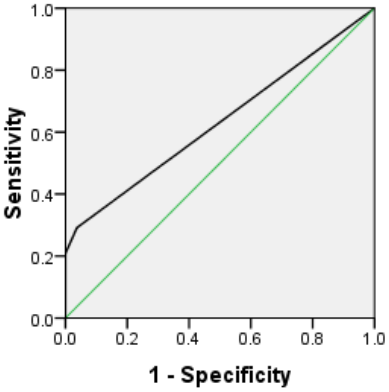


Figure 4-33 ROC of Missing Possessive Clitics

4.5 Discussion

The current research provides new evidence regarding the clinical utility of language sample measures in diagnosing PLI in Iranian Persian-speaking children. This series of studies started with a pre-accuracy study of the association between age and language sample measures, in order to supply the further phases of the study with measures developmentally sensitive to age. The results of the Persian study did not support the association between age and LSMs due to the fact that small sample size and age range have strong impacts on the results of a correlation study (Bland & Altman, 2011) and the only way of improving the results is broadening the age range with a large enough sample size in each age group.

Although making a comparison between studies with different inclusion/exclusion criteria and definition of language impairment with no specific reason is not advised unless a large number of similarities are found with the sample population (Miller & Fletcher, 2005), it is plausible that studies will strengthen each other in terms of what to include as possible variables for analysis in studies with similar methodology. Having said that and regardless of the pre-accuracy outcomes, however, it was assumed that those LSMs with the ability to diagnose children with primary language impairment would also be able to do the same in Persian, as there were similar studies on the diagnostic accuracy of those LSMs in languages other than English. The studies most related in terms of methodology were named in the meta-analysis chapter (chapter 2).

Phase I of Sackett and Haynes' progression of studies is aimed at finding measures with the potential to differentiate Persian-speaking children with and without PLI. Children with and without PLI were statistically different on 29 of 36 language sample measures at the group level. They were categorised into two main categories of General and Persian-specific. The measures in the General category included: MLU in morphemes and words, Number of Total Words, Number of Different Words, Number of One-word utterances, MLU in morphemes and words-excluding One-word Utterances, and Number of Total Verbal Morphemes. General Errors also included Clitics Errors, Verb Inflectional Errors (Finite Verb Morphology), Semantic Errors, Total Errors, Total Number of Ungrammatical Utterances, Grammaticality, Ungrammaticality, Missing verb markers, Missing prepositions, Missing verbs, Wrong agreement, Nonsense string of words, and Wrong responses.

Persian-specific measures were: Total number of plural marker */ha*, Total number of direct object marker *ra*, Total number of progressive verb marker *mi*, Missing *ra* (direct

object marker), Missing */e-Ezafeh* (addition or genitive sign), Missing *mi* (Progressive marker), Missing objective clitics, and Missing possessive clitics.

The measures, subsequently, were analysed in the form of a diagnostic accuracy study to further investigate their clinical diagnostic competence. The diagnostic accuracy study of these LSMs showed that some of them held promising values of AUC, sensitivity and specificity, and likelihood ratios. They are shown in table 4-29, ordered by AUC. Although it has been claimed that DOR is the "single indicator of test performance", AUC has also been suggested as an alternative to DOR (Glas et al., 2003) but both need to be accompanied by LRs and CIs in order to fully evaluate the most promising measures. Some measures with high DORs do not hold promising AUCs, and in some cases the lower end of their CI approximates the diagonal line which means that the results of that measure should be considered with caution. Moreover, interpreting AUC should be looked at in parallel with sensitivity and specificity, with the intention that these features are central in clinical decision making. High sensitivity might be more important when including children with PLI and high specificity is more desirable when identifying children without PLI. The application of sensitivity, however, is preferred by clinicians due to its essential role in identifying clients with impairment out of all the clients referred to the clinics (Klee et al., 2007).

Language Sample Measures	AUC (95% CI)*	Sensitivity (95% CI)**	Specificity (95% CI)**	LR+ (95% CI)†	LR- (95% CI)†	DOR (95% CI)††
1. Grammaticality	.971 (.934, 1.00)	.98 (.83, 1)	.84 (.66, .92)	6.1(2.6, 14.2)	.02(.00, .3)	255.9(13, 5018)
2. Ungrammaticality	.971 (.934, 1.00)	.98 (.83, 1)	.84 (.66, .92)	6.1(2.6, 14.2)	.02(.00, .3)	255.9(13, 5018)
3. Ungrammatical Utterances	.958 (.906, 1.01)	.87 (.69, .96)	.96 (.82, .99)	23.62(3.4, 162.6)	.13(.04, .3)	182.0(17.6, 1880)
4. MLUw-exc	.950 (.895, 1.00)	.82 (.63, .92)	.98 (.85, .99)	45.92(2.9, 720)	.18(.08, .4)	250.5(12.7, 4919)
5. Semantic Errors	.948 (.874, 1.02)	.92 (.74, .98)	.96 (.82, .99)	24.75(3.6, 170)	.09(.02, .3)	286.0(24.2, 3370)
6. MLUw	.928 (.863, .993)	.92 (.74, .98)	.78 (.59, .89)	4.12(2.0, 8.4)	.11(.03, .4)	38.5 (6.9, 212.5)
7. Total Errors	.927 (.847, 1.00)	.74 (.54, .87)	.98 (.85, 1)	41.44(2.6, 653)	.26(.1, .5)	156.5(8.3, 2950)
8. MLUm-exc	.923 (.855, .991)	.66 (.46, .81)	.98 (.85, .99)	36.96(2.3, 585)	.35(.2, .6)	106.8(5.8, 1973)
9. MLUm	.906 (.828, .984)	.92 (.74, .98)	.78 (.59, .89)	4.12(2.0, 8.4)	.11(.03, .4)	38.5 (6.9, 212.5)
10. Nonsense string of words	.864 (.758, .971)	.83 (.64, .93)	.78 (.59, .89)	3.75(1.8, 7.8)	.21(.09, .5)	17.5(4.3, 71.4)
11. Missing verbs	.851 (.741, .961)	.83 (.64, .93)	.74 (.55, .87)	3.21(1.6, 6.2)	.22(.09, .6)	14.27(3.6, 56.5)
12. NDW	.848 (.743, .953)	.79 (.59, .91)	.78 (.59, .89)	3.56(1.7, 7.4)	.27(.1, .6)	13.3 (3.5, 50.7)
13. TNOU	.822 (.700, .943)	.75 (.55, .88)	.85 (.67, .94)	5.06(2, 12.9)	.29(.1, .6)	17.25(4.2, 70.5)
14. Wrong responses	.804 (.674, .934)	.62 (.43, .79)	.96 (.82, .99)	16.87(2.4, 118.4)	.39(.2, .6)	43.33(5, 376)
15. Total */ha*	.769 (.632, .905)	.68 (.47, .82)	.93 (.77, .98)	9(2.3, 35.1)	.36(.2, .6)	25.0(4.7, 133)
16. Verb Inflectional Errors	.758 (.625, .892)	.50 (.31, .69)	.93 (.77, .98)	6.75(1.7, 27.1)	.54(.3, .8)	12.5(2.4, 65)
17. Clitics Errors	.738 (.595, .880)	.42 (.25, .61)	.98 (.85, 1)	23.5(1.4, 381)	.59(.4, .8)	39.83(2.2, 729)
18. Total *mi*	.723 (.570, .876)	.54 (.35, .72)	.98 (.85, 1)	30.24(1.9, 483)	.47(.3, .7)	64.56(3.5, 1180)
19. NTW	.721 (.577, .864)	.46 (.28, .65)	.96 (.82, .99)	12.37(1.7, 88.9)	.56(.4, .8)	22 (2.6, 189.4)
20. Missing verb markers	.716 (.570, .862)	.54 (.35, .72)	.89 (.72, .96)	4.87(1.6, 15.1)	.52(.3, .8)	9.45(2.2, 40.1)
21. Total *ra*	.699 (.550, .848)	.50 (.31, .69)	.85 (.67, .94)	3.37(1.2, 9.1)	.59(.4, .9)	5.75(1.5, 21.7)
22. Wrong agreement	.676 (.527, .825)	.33 (.18, .53)	.93 (.77, .98)	4.5(1.06, 19.1)	.72(.5, 1)	6.25(1.2, 33.2)
23. Missing *ra*	.662 (.510, .814)	.37 (.21, .57)	.89 (.72, .96)	3.37(1, 11)	.70(.5, .9)	4.8(1.1, 20.6)
24. TNVM	.648 (.488, .808)	.46 (.28, .65)	.93 (.77, .98)	6.19(1.5, 25.1)	.58(.4, .8)	10.57(2, 55)
25. Missing prepositions	.639 (.483, .794)	.26 (.13, .45)	.98 (.85, 1)	14.86(.9, 250)	.75(.6, .9)	19.86(1.05, 375)
26. Missing *mi*	.634 (.478, .790)	.33 (.18, .53)	.93 (.77, .98)	4.5(1, 19.1)	.72(.5, 1)	6.25(1.2, 33.2)
27. Missing possessive clitic	.631 (.475, .788)	.29 (15, .49)	.96 (.82, .99)	7.87(1.0, 59.5)	.74(.6, 1)	10.71(1.2, 95)
28. Missing */e Ezafeh*	.623 (.466, .779)	.37 (.21. .57)	.85 (.67, .94)	2.53(.9, 7.1)	.73(.5, 1)	3.45(.9, 13.2)
29. Missing objective clitic	.583 (.424, .743)	.18 (.07, .37)	.98 (.85, 1)	10.28(.6, 181)	.83(.7, 1)	12.37(.63, 243)

Table 4-29 Diagnostic values (with 95% CIs) of language sample measures with best cut-off points, ordered by Area under Curve (AUC); *The closer the AUC to 1 indicates the better diagnostic competency. **Accuracy values (sensitivity, specificity) are judged "good" if between .9 and 1; "fair" if between .8 and .89; and "inadequate" if below .8. *** Higher DOR shows better overall diagnostic competence of the measure. LR+ greater than 10 and LR- smaller than .10 are desirable; AUC=Area under Curve, SE=Standard Error, LR=Likelihood Ratio, DOR=Diagnostic Odds Ratio, CI=Confidence Interval.; Abbreviation of measures are in table 4-21.

The results shown in table 4-29 would lead us to conclude that although the first five measures with high AUC and good sensitivity and specificity are suitable for diagnostic purposes, not all five hold desirable LRs and CIs; the values which have been claimed by some researchers as more preferred measures compared to other diagnostic features of a test (Dollaghan, 2008 p.87). The very low negative LR for the first two measures (LR-=.02) indicates that a negative score in either Grammaticality (higher scores than cut-off 94.25 percent) or Ungrammaticality (lower scores than 5.70 percent) indexes are very unlikely to have come from a child with PLI, whereas this is not true for the corresponding positive LRs. Their wide CIs, though, include the unacceptable defined range of LRs (Less than 10 for LR+ and greater than .10 for LR-). This picture is the same for the next two measures, Number of Ungrammatical Utterances and MLUw Excluding One-word Utterances, with an exception in LR+ whose high values indicate that a positive score in either one is very likely to have come from a child with TDL (LR+=23.62 and 45.92, respectively). Negative LRs do not show suitable performance. Both LRs, again, show wide CI range. Interestingly, the only measure which holds a high level in all diagnostic accuracy features, regardless of large CIs, is Semantic Errors (AUC (95% CI) = .948 (.874, 1.02), Sensitivity = .92, Specificity = .96, LR+ = 24.75, LR- = .09, DOR = 286). One reason could be that this measure is the sum of all semantic errors so the cumulative effect of all these measures would cause the high level of accuracy (see appendix F for definitions of errors).

The CIs of the best LRs of the five measures need to be considered by users in deciding how precisely the measures would act in differentiating between children with and without PLI (Dollaghan, 2008). If the related 95% CI also fits within the optimal range of either positive or negative LRs, the user can be sure of its good diagnostic performance (Dollaghan, 2008). To some researchers, likelihood ratios are considered as better measures of accuracy than sensitivity and specificity since they are less affected by the prevalence of the target condition (Dollaghan, 2008). They can also be used by clinicians to estimate the post-test probability of a given test or measure if the clinicians aim to assess a given child using those measures. The post-test probability shows the percentage of test accuracy if used to confirm the target condition in new clients. Pre-test probability with the equation $P(D+) = D+ / (D+ + D-)$ is .47 for Grammaticality and Ungrammaticality, .41 for Ungrammatical Utterances, .39 for MLUw-exc., and .43 for Semantic Errors, "where D+ indicates the number of patients with target disorder, D- indicates the number of patients without target disorder, and P(D+) is the probability of the target disorder" in each measure (*Pre-test Probability,*

2012; para. 4). The related frequency of cases for each measure is shown in tables 4-23 and 4-25. The post-test probabilities of the first five measures are shown in table 4-30. The resultant post-test probabilities show that for example if a child with the suspected condition of PLI is examined by the measure of Semantic Errors and her result falls above the cut-off point, the probability of being PLI would be around 96% using the online nomogram for calculating post-test probabilities. So, the clinician can be 96% certain that the child has PLI if her/his test result is positive (greater than the cut-off).

Language Sample Measure	Post-test probability% (approximately)
1. Grammaticality	87%
2. Ungrammaticality	87%
3. Ungrammatical Utterances	95%
4. MLUw-exc	97%
5. Semantic Errors	96%

Table 4-30 Post-test probability of the first five language sample measures

4.5.1 Linguistic classification of measures

Classifying the measures according to the linguistic domains of grammar and semantics shows that Grammatical Measures outnumber Semantic Measures with 22 of 29 measures (Table 4-31). The remaining seven Semantic Measures, however, include the best diagnostic measure of Semantic Errors, which supersedes all other measures as already explained.

4.5.1.1 How informative are grammatical language sample measures?

Of the three measures of grammatical production (Grammaticality, Ungrammaticality, and Total Number of Ungrammatical Utterances), Grammaticality and Ungrammaticality were shown in the meta-analysis to be good indexes in identifying language impairment, alone or in combination with other measures (Eisenberg & Guo, 2012). Simon-Cereijido and Gutierrez-Clellen (2007) demonstrated that the combination of Ungrammaticality and MLUw showed fair to good sensitivity and specificity along with another composite score of Correct Use of Verbs + Clitics + Articles for young Spanish-speaking children (3;11 – 5;1 years old), sensitivity of 79% and specificity of 100% in the exploratory study (equal to phase II DA) and 80% for both sensitivity and specificity in the confirmatory study, which was consistent with previous research on older Spanish-speaking children (5;0 – 7;11 years old). The combinations formed after the researchers did not find any specific grammatical

measure with the capability to differentiate between children with and without language impairment. The measures, therefore, were combined and showed the above results. The researchers concluded that those particular features of Spanish grammar are not delayed in Spanish-speaking children.

The Grammaticality percentage was shown between 65% and 100% for English-speaking typically-developed children between 2;6 and 7;8 years of age (n=41) (Dunn, Flax, Sliwinski, & Aram (1996) as stated in Eisenberg et al., 2012). In Westerveld and Gillon's study (2010), 77% of 5-year-olds had a Grammaticality percentage of 85 and over (as stated in Eisenberg et al., 2012). Eisenberg and Guo (2012) explored the diagnostic accuracy of Grammaticality Percentage in 3-year-olds with and without language impairment using a picture description task. The results showed a sensitivity of 88% and specificity of 100% for this measure with LR+ of less than 10. The diagnostic values of three measures of grammaticality in the current study are very close to previous studies; Grammaticality and Ungrammaticality show good sensitivity and fair specificity (98% and 84%, respectively) with a very low negative LR (.02). Ungrammatical Utterances also possesses fair sensitivity of 87% and good specificity of 96% with a very high positive LR (23.62). Neither of the above two studies, however, reported the CI of the measures which leads decision making about the results to be taken cautiously.

Amid the General Grammatical LSMs, MLUw-excluding One-word Utterances holds the best AUC of .950, fair sensitivity of .82 and good specificity of .98. Its positive LR is also very high, at 45.92, and the negative LR is very close to the optimal point (.1), at .18 (Table 4-31). The post-test probability is calculated at 97% which shows a very good estimation of inclusion of new cases of PLI. The main version of MLUw resulting from all complete and intelligible utterances in the transcript, on the other hand, has good sensitivity, at .92, and an unacceptable specificity, at .78 without good LR+. However, LR- is extremely close to the optimal point (.1) at .11. Again, all the CIs are large. MLUw excluding one-word morpheme utterances was introduced by Klee and Fitzgerald (1985) as 'mean syntactic length' (MSL) with the aim of reducing the "noise" within the sample. They observed that an average of 31% of the utterances produced by children between two and three years of age were one-word utterances and in further studies, one-word utterances accounted for 34% to 50% of the utterances in the samples of normal language and SLI samples respectively (Klee, 1991). Other studies showed the same results (for a review, see Klee, 1992) and suggested finding the clinical applicability of a new measure (named Mean Syntactic Length (MSL) by Klee

& Fitzgerald, 1985) by withdrawing single-morpheme and single-word utterances from the analysis to reduce the effect of "pragmatic constraints of the conversation" (Klee, 1991 p. 325) and non-informative one-word utterances in terms of grammar. Recall from chapter 1 that single-word utterances in Persian might include more than one morpheme and they may be of interest for being investigated individually; however, the current results are congruent with other studies when one-word utterances are excluded from the sample resulting in higher MLUs. This event will be more discussed in chapter 5.

Grammatical LSMs	AUC (95% CI)*	Sensitivity (95% CI)**	Specificity (95% CI)**	LR+ (95% CI)†	LR- (95% CI)†	DOR (95% CI)††
1. Grammaticality	.971 (.934, 1.00)	.98 (.83, 1)	.84 (.66, .92)	6.1(2.6, 14.2)	.02(.00, .3)	255.9(13, 5018)
2. Ungrammaticality	.971 (.934, 1.00)	.98 (.83, 1)	.84 (.66, .92)	6.1(2.6, 14.2)	.02(.00, .3)	255.9(13, 5018)
3. Ungrammatical Utterances	.958 (.906, 1.01)	.87 (.69, .96)	.96 (.82, .99)	23.62(3.4, 162.6)	.13(.04, .3)	182.0(17.6, 1880)
4. MLUw-exc	.950 (.895, 1.00)	.82 (.63, .92)	.98 (.85, .99)	45.92(2.9, 720)	.18(.08, .4)	250.5(12.7, 4919)
5. MLUw	.928 (.863, .993)	.92 (.74, .98)	.78 (.59, .89)	4.12(2.0, 8.4)	.11(.03, .4)	38.5 (6.9, 212.5)
6. Total Errors	.927 (.847, 1.00)	.74 (.54, .87)	.98 (.85, 1)	41.44(2.6, 653)	.26(.1, .5)	156.5(8.3, 2950)
7. MLUm-exc	.923 (.855, .991)	.66 (.46, .81)	.98 (.85, .99)	36.96(2.3, 585)	.35(.2, .6)	106.8(5.8, 1973)
8. MLUm	.906 (.828, .984)	.92 (.74, .98)	.78 (.59, .89)	4.12(2.0, 8.4)	.11(.03, .4)	38.5 (6.9, 212.5)
9. TNOU	.822 (.700, .943)	.75 (.55, .88)	.85 (.67, .94)	5.06(2, 12.9)	.29(.1, .6)	17.25(4.2, 70.5)
10. Total */ha*	.769 (.632, .905)	.68 (.47, .82)	.93 (.77, .98)	9(2.3, 35.1)	.36(.2, .6)	25.0(4.7, 133)
11. Verb Inflectional Errors	.758 (.625, .892)	.50 (.31, .69)	.93 (.77, .98)	6.75(1.7, 27.1)	.54(.3, .8)	12.5(2.4, 65)
12. Clitics Errors	.738 (.595, .880)	.42 (.25, .61)	.98 (.85, 1)	23.5(1.4, 381)	.59(.4, .8)	39.83(2.2, 729)
13. Total *mi*	.723 (.570, .876)	.54 (.35, .72)	.98 (.85, 1)	30.24(1.9, 483)	.47(.3, .7)	64.56(3.5, 1180)
14. Missing verb markers	.716 (.570, .862)	.54 (.35, .72)	.89 (.72, .96)	4.87(1.6, 15.1)	.52(.3, .8)	9.45(2.2, 40.1)
15. Total *ra*	.699 (.550, .848)	.50 (.31, .69)	.85 (.67, .94)	3.37(1.2, 9.1)	.59(.4, .9)	5.75(1.5, 21.7)
16. Wrong agreement	.676 (.527, .825)	.33 (.18, .53)	.93 (.77, .98)	4.5(1.06, 19.1)	.72(.5, 1)	6.25(1.2, 33.2)
17. Missing *ra*	.662 (.510, .814)	.37 (.21, .57)	.89 (.72, .96)	3.37(1, 11)	.70(.5, .9)	4.8(1.1, 20.6)
18. TNVM	.648 (.488, .808)	.46 (.28, .65)	.93 (.77, .98)	6.19(1.5, 25.1)	.58(.4, .8)	10.57(2, 55)
19. Missing *mi*	.634 (.478, .790)	.33 (.18, .53)	.93 (.77, .98)	4.5(1, 19.1)	.72(.5, 1)	6.25(1.2, 33.2)
20. Missing possessive clitic	.631 (.475, .788)	.29 (15, .49)	.96 (.82, .99)	7.87(1.0, 59.5)	.74(.6, 1)	10.71(1.2, 95)
21. Missing */e Ezafeh*	.623 (.466, .779)	.37 (.21. .57)	.85 (.67, .94)	2.53(.9, 7.1)	.73(.5, 1)	3.45(.9, 13.2)
22. Missing objective clitic	.583 (.424, .743)	.18 (.07, .37)	.98 (.85, 1)	10.28(.6, 181)	.83(.7, 1)	12.37(.63, 243)
Semantic Language Sample Measures						
23. Semantic Errors	.948 (.874, 1.02)	.92 (.74, .98)	.96 (.82, .99)	24.75(3.6, 170)	.09(.02, .3)	286.0(24.2, 3370)
24. Nonsense string of words	.864 (.758, .971)	.83 (.64, .93)	.78 (.59, .89)	3.75(1.8, 7.8)	.21(.09, .5)	17.5(4.3, 71.4)
25. Missing verbs	.851 (.741, .961)	.83 (.64, .93)	.74 (.55, .87)	3.21(1.6, 6.2)	.22(.09, .6)	14.27(3.6, 56.5)
26. NDW	.848 (.743, .953)	.79 (.59, .91)	.78 (.59, .89)	3.56(1.7, 7.4)	.27(.1, .6)	13.3 (3.5, 50.7)
27. Wrong responses	.804 (.674, .934)	.62 (.43, .79)	.96 (.82, .99)	16.87(2.4, 118.4)	.39(.2, .6)	43.33(5, 376)
28. NTW	.721 (.577, .864)	.46 (.28, .65)	.96 (.82, .99)	12.37(1.7, 88.9)	.56(.4, .8)	22 (2.6, 189.4)
29. Missing prepositions	.639 (.483, .794)	.26 (.13, .45)	.98 (.85, 1)	14.86(.9, 250)	.75(.6, .9)	19.86(1.05, 375)

Table 4-31 Diagnostic values (95% CIs) of language sample measures with best cut-off points, ordered by Area under Curve (AUC) and categorised by two linguistic domains of Grammatical and Semantic; Light shade indicates that the measure is categorised as General and dark shades represent Persian-specific measures; Abbreviation of measures are in table 4-21; See table 4-29 for complete interpretation of accuracy values.

MLU was shown to have good sensitivity and specificity of 94.7% and 89.5% respectively in Bedore and Leonard's study of 38 children with and without SLI between 3;7 and 5;9 years of age (Bedore & Leonard, 1998); however, Klee and colleagues computed marginally acceptable LR+ of 9.0 (95% CI=2.42, 33.53) and acceptable LR- of .06 (95% CI=.01, .40) for Bedore and Leonard's study (Klee et al, 2007). This study was not retrieved in the search for meta-analysis due to its date of publication.

MLU in either morphemes or words has been shown to increase accuracy of diagnostic measures when combined with other measures of grammar or syntax (see chapter 2 for a meta-analysis of related diagnostic studies; see also Klee et al., 2007). Klee et al. evaluated the discriminant function of the combination of Age + MLUm + D in identifying children with SLI, which resulted in 44 of 45 (97.8%) children being correctly classified as PLI or TDL (Klee et al., 2004). They examined whether this composite measure would be useful for classifying Cantonese-speaking children into groups of PLI and TDL. A follow-up study about the accuracy of the same discriminant function was later reported using a new sample of children, and the diagnostic values for four year-olds were reported as follows: S_n = 97% (76%, 100%); S_p = 91% (68%, 98%); LR+ = 10.33 (2.25, 47.53); LR- = .03 (0, .53) (as stated in Klee et al., 2007). They attributed the large CIs to the small sample size which is similar to the current study. Their results for the 95% CI were replicated in a similar study finding the discriminant function of the composite of Age + MLU + D in English as follows: S_n = 86% (60%, 96%); S_p = 91% (62%, 100%); LR+ = 9.43 (1.44, 61.85) ; LR- = .16 (.04, .58). The figures of all these results together led the authors to emphasise that any decision based on the wide CIs should be accompanied by caution; no decision is preferred to an uncertain one. Also they recommended that more thorough language assessment would be essential for two year-olds due to the fact that their CIs were much larger than other age groups (Klee et al., 2007). This is true for all measures reported in the current study or similar studies with large CIs even though the current study still holds a significant place among the collection of research on Iranian child language. The results on Cantonese-speaking children (2004) were replicated in a larger independent sample of 49-60 month old children which showed different results from the first study. The diagnostic values reported were as follows: sensitivity of 73.3% (48%, 89%); specificity of 57.1% (33%, 79%); LR+ of 1.71 (.87, 3.37); and LR- of .47 (.18, 1.21). The results were ascribed to the differences between the averages of measures in two studies as well as the behaviours of different groups in formal tests

which were used as reference standards. The authors proposed that this clinical composite would not be clinically useful in the identification of Cantonese-speaking children with SLI and that future research should consider whether any other measure of language sample or language processing with promising results in the diagnosis of English SLI would behave the same in Cantonese. The final suggestion was to examine the diagnostic accuracy of sentence imitation as the current single processing measure in Cantonese with the ability to differentiate TDL from PLI (Wong et al., 2010).

Accuracy measures of MLUw and MLUm in a study on French-speaking 5-year-old children with and without PLI in Quebec showed their highest measures were at -1SD as follows: for MLUw, sensitivity was 40%, specificity was 85%, LR+ was 2.92, and LR-was .67. For MLUm, sensitivity was 36%, specificity was 87%, LR+ was 2.68, and LR- was .74 (Thordardottir et al., 2011).The CIs, again, were not reported. As can be seen, the only acceptable value is of specificity which can be judged as fair (Plante and Vance, 1994).

One important thing, however, should be taken into consideration when applying the results of MLU (and probably other general language sample measures): that this measure is not representative of the complexity of the child's linguistic performance, nor an index of the child's "overall linguistic health" (Klee, 1991 p.327). It only measures the average utterance length of the child's whole communicative output to depict how she has behaved in her language production compared to her peers in a particular situation.

In the present study, children with PLI had significantly higher Clitic Errors than children with TDL. Accuracy measures show that this measure is not efficient in identifying children with PLI (sensitivity = .42) but it performs well in identifying unimpaired cases (specificity = .98). Contrary to positive LR with acceptable value of 23.5, negative LR does not hold acceptable metric (.59), in addition to the fact that all computed CIs are large. This result suggests that if total number of Clitic Errors in a given child's language sample is higher than the cut-off point of 2.50 this measure is only 42% accurate in diagnosing the child as PLI, versus 98% efficient in identifying a child with a total Clitic Errors of less than 2.50 as TDL. When dividing Clitic Errors into its sub-components, children with TDL did not miss any objective clitics in their language sample whereas children with PLI had an average of .29 (SD=.75) missing objective clitics, which was significantly higher. Measure of Missing Objective Clitics, however, does not hold informative diagnostic accuracies except for specificity of .98 and a positive LR of 10.28. Persian-speaking children with PLI also used significantly

fewer objective clitics in their responses to a structured task compared to their TDL peers (Foroodi-Nejad, 2011). The percent of objective clitic use in the responses of children with PLI was 36% (SD=19) compared to their TDL peers with 55% (SD=14) which showed a large effect size of *d*=-1.13. The results of the present study could be viewed as complementing Foroodi-Nejad's because it considers the errors in clitic usage. The role of the clitic in grammar systems has been in debate among linguists insofar as allocating a known underlying syntactic structure to them is very difficult (Simon-Cereijido & Gutierrez-Clellen, 2007; also see chapter 2). It was also observed that different categories attributed to clitics in Persian sources and their components varied among different Persian linguists (Kalbasi, 2008; Meshkato-Dini, 2008). This controversy makes the explanation of the results difficult with respect to the type of errors seen in Clitics Errors. However, generally, similar to Spanish-speaking children who showed both omission and substitution in their spontaneous language (Simon-Cereijido& Gutierrez-Clellen, 2007), Persian-speaking children indicated both omissions and wrong use of clitics in their language samples.

In another of Wexler's accounts of language impairment (the extended unique checking constraint [UCC] account of LI; see Wexler, 2003) the optionality issue was extended to clitics to document clitic omissions observed in Spanish-speaking children (as cited in Simon-Cereijido & Gutierrez-Clellen, 2007). It would probably be appropriate to employ this account in explanation of other linguistic difficulties in the population of children with PLI when their errors are omissions. For the current study, also, the majority of error indexes with the capability of distinguishing children with PLI from their TDL counterparts were *missing-type* errors or *omissions* so that they can simply be accounted for by Wexler's optionality: Persian-speaking children with PLI choose whether to use or not use the target inflectional morpheme, either the finite verb inflectional morpheme or other types of inflectional morphemes in the structure of Persian. Individual clitic types with the capability of differentiation between two groups will be discussed later, when Persian-specific measures are discussed.

Verb Inflectional Errors (VIE) is the next General Grammatical measure in table 4-31 where children with PLI score significantly higher than their TDL peers, which also represents their higher error rates in VIE subcomponents as follows: Missing Verb Marker, Wrong Agreement, and Missing *mi*\. The results of a discriminant analysis study of Finite Verb Morphology Composite (FVMC) showed that, in general, the accuracy of FVMC was fair to good in discriminating LI in conversational tasks (Reported sensitivity and specificity were 84% and 100%, respectively) (Bedore&

Leonard (1998) as reported in Klee et al., 2007 and Eisenberg et al., 2012). FVMC is defined as the percentage of correct use of the copula and auxiliary *be*, third person singular –*s*, and past tense –ed. VIE accompanied with another measure of verb inflection, Total Number of Verbal Inflections (TNVM), represents the child's production of verb morphology. Children with PLI scored significantly less in TNVM compared to children with TDL, which confirms their higher error rate in VIE. The sensitivities of both measures of TNVM and VIE, however, were not acceptable (46% and 50%, respectively) whilst the specificities were classified as good (both 93%) leaving them 24th and 16th among other 29 measures in terms of AUC (Table 4-29). According to this result, if a given child produces more than 42.5 (cut-off point) counts of verbal morphemes in her language sample or less than 3.50 errors in verbal inflections, it can be concluded that TNVM and VIE are 93% accurate in identifying this child as TDL.

The subcomponents of VIE did not show better results in accuracy: the best sensitivity was only 54% for Missing Verb Markers and the best specificity was 93% for Wrong Agreement and Missing *mi*\. LRs of all individual and composites were not acceptable and again the calculated CIs were wide. A more recent study on FVMC which added auxiliary *do, does*, and *did* to the list, showed 100% sensitivity and specificity for distinguishing children with and without SLI between the ages of 4;0 and 4;6 years and both LRs met the optimal criteria (Gladfelter & Leonard, 2012). The CIs related to this age range were not provided. Another measure of tense marking in the current study is Total *mi*\ with differentiation capability in identifying PLI which is consistent with Foroodi-Nejad's study. She reported that Persian children with primary language impairment produced significantly less tense markers (morpheme *mi*\) in their speech, with a large effect size of -2.12 (Foroodi-Nejad, 2011). Contrary to the present study, the agreement feature in Foroodi-Nejad's study showed that children of both groups did not differ in correct use of this feature and both showed a high level of proficiency in using the agreement feature in their language samples. Agreement feature in the current study was examined through the frequency of wrong agreement and showed that children with PLI behaved significantly less accurately than their TDL counterparts with a medium effect size of d=.65. However, the Wrong Agreement measure does not provide satisfactory diagnostic accuracies except for a specificity of 93%.

Direct Object (DO) marking was investigated through observing three measures within children's language productions: Total Number of DO Marker *ra*, Missing *ra*, and Wrong Usage of *ra*. The first two measures showed differentiating competence whilst

Wrong Use of *ra* did not (table 4-20). The insignificant difference in Wrong Use of *ra* is inconsistent with the MRH account, in which more substitution errors are observed in languages with rich morphological systems, rather than omissions. The inconsistent language behaviour of children with PLI in these three measures of one grammatical event is compatible with our expectation (chapter 1) that *ra* specification in Persian probably makes Persian-speaking children with PLI display an unstable picture in *ra*-related error analysis. The sparse and irregular appearance of *ra* had been claimed to be problematic for children with PLI as it was documented in two errors containing *ra*. In addition, the significantly higher Missing *ra* and lower Total Number of DO Marker *ra* in children with PLI can also be explained by the optionality component of 'Extended Optional Infinitive' (EOI) account. The extended optional infinitive account implies that children with PLI show longer inconsistency in using the proper morphemes of verb tense and agreement and a lack of awareness of arbitrary application of these morphemes in specific sentence contexts (Rice et al., 1995). This lack of awareness of compulsory grammatical structures of verbs, generally called 'Optional Infinitive' (OI), is assumed normal and related to "a biologically-based principle" connected to the maturation age of three years (Gladfelter& Leonard, 2012 p. 4). Exposure to more grammatical representations of the affected structures in conversation, however, will help the children to correct the forms and make them comparable to adults (Klee et al., 2007; Leonard, 2009). This notion could be extended to other morphological structures and connect the observed missing errors to the optionality in using grammatical markers. It is also probable that EOI can be modified by dropping 'Infinitive' (as was done by Wexler in creating UCC) and leaving it as 'Extended Optionality' (EO). In PLI, problems in underlying grammar or linguistic knowledge have been attributed as the reasons for EOI (Klee et al, 2007). EOI has also been documented in languages other than English including Dutch, French, German, and Swedish (Klee et al., 2007). In Foroodi-Nejad's study, children with PLI were significantly less accurate compared to children with TDL in the correct use of case marker *ra*, with a large effect size of d= -1.97 (Foroodi-Nejad, 2011). The results of *ra*-related measures in the current study are consistent with Foroodi-Nejad's research; Persian-speaking children with PLI on average produced fewer *ra* in their language samples and also had higher rates of Missing *ra* compared to the children with TDL. Both measures hold a medium effect size of d= -.63 and d= .75 for Total Number of *ra* and Missing *ra*, respectively. The diagnostic accuracy of these measures, however, indicates that they are inefficient in diagnosing PLI (sensitivity of 50% for Total *ra* and 37% for Missing *ra*) and shows

only fair performance in identifying TDL (specificity of 85% for Total *ra* and 89% for Missing *ra*). Additionally, none of the LRs meet an acceptable level of diagnostic accuracy.

The plural noun marker has been shown to have a very low error rate in children with TDL (Eisenberg et al., 2012). Children with PLI also did not show significant difference compared to children with TDL in producing noun plural marker --*s* independently from other markers (Conti-Ramsden, 2003), even when it is included in a noun grammatical morpheme composite with other markers, for children with average age of 7;9 years old (Moyle et al., 2011), suggesting that "not all grammatical markers are problematic for children with language impairment" (Conti-Ramsden, 2003 p.551). Rice and Oetting (1993) documented that plural marking in nouns was "robust" with "high level of accuracy" in the productions of children with PLI (Rice & Oetting, 1993 p.1255). Rice & Oetting also found that plural marking was context-dependent, i.e. 'quantifier + noun' context allocated less plural marking than 'determiner + noun'. They, finally, came up with the notion that the appearance of the plural in quantifier contexts in the speech of children with PLI was optional and therefore needed further investigation (Rice & Oetting, 1993). Later, Leonard's Surface Hypothesis (1998) was used to justify this observation by explaining that acoustically, plural –*s* duration is longer because it comes at the end of noun or phrase. Consequently it will provide children with PLI with better reception of a more durable sustained pronunciation of -s which repairs the limited processing capacity up to a level represented in the similar production of -*s* to their unimpaired counterparts (Marinis, 2011). Persian plural noun marker /*ha*, in contrast, appeared on fewer occasions in expressions of children with PLI at the significant level of .001 and its specificity was classified good, 93%. Other diagnostic values, however, showed unacceptable sensitivity of 63% and LRs outside acceptable values. Since there is no research available on the acoustic specifications of Persian spoken in childhood, judging Leonard's Surface Hypothesis with respect to Persian language acoustics would be inadequate. The answer would have to be revealed by undertaking more specifically-designed research and considering the relationship between plural noun marker and subject-verb agreement.

The last Persian-specific grammatical measure represents the language behaviour of Persian-speaking children in producing the most unique feature of Persian, */e-Ezafeh*. This feature signifies the addition or genitive sign of nominal, possessive or adjectival relationship of two words. Extensive investigation of the literature has revealed that no clinical Persian language study has addressed this before, nor has an equivalent in

English; although English *of* or *'s* can be considered its counterpart. Persian children with PLI produced significantly less */e-Ezafeh* than their TDL peers with marginal significance and effect size. Its accuracy measures also do not reach more than fair in specificity (85%) with poor sensitivity of 37% and unacceptable LRs. Due to its unique place in Persian grammar and the fact that majority of Iranian SLTs include */e-Ezafeh* in their interventional plans, it is highly recommended that future research explore its diagnostic value through more specifically-designed tasks within a wider age range.
A general example of diagnostic behaviour of Persian-specific measures documents that if a given child produces the count of each measure of plural marker more than 2.50, DO marker more than 11.50, and progressive marker more than 13, the measures are 93%, 85%, and 98% capable of diagnosing the child as TDL, respectively.
Generally speaking, weak diagnostic behaviour of some Persian-specific measures compared to some general or composite measures could be partly attributed to the fact that they individually represent one single morphological phenomenon rather than encompassing several aspects of grammar.

4.5.1.2 How informative are semantic language sample measures?

Semantic LSMs include measures capable of distinguishing between the two groups. The first measure is Semantic Errors with a large effect size in differentiating between children with and without PLI. Despite wide CIs, all the related accuracy measures of Semantic Errors were within the good and acceptable range compared to other measures, with an AUC of .948, sensitivity of .92, specificity of .96, LR+ of 24.75, LR- of .09, and DOR of 275. As seen in table 4-31, Semantic Errors has the best accuracy values of all Semantic LSMs. A probable explanation of the high Semantic Errors in the language samples of children with PLI could be attributed to their problem with multi-argument structures and their vulnerability in processing verb structures with more than two arguments due to the nature of their limited processing capacity (LPC) (see Simon-Cereijido & Gutierrez-Clellen, 2007 for a review). The verb structure in Persian is recognised to have at least eight syntactic complements (or arguments) and with more than 4000 Persian verbs, the repertoire of the probable combinations of complements looks huge (Rasooli, Moloodi, Kouhestani, & Minaei-Bidgoli, 2011). The only Persian work documenting this linguistic feature is from the Dedegan Research Group that has provided the first dataset on obligatory and optional syntactic complements called "Valency Lexicon for Persian Verbs" ("Valency Lexicon for Persian Verbs," 2012). It is not within the scope of the current study to investigate this linguistic aspect in detail;

however, considering the enormous size of this repertoire it would show how big this argument structure would appear, to be learned by Persian-speaking children with PLI. The language processing of this repertoire would seem problematic for children with PLI and it might show the reason for their observed difficulties in missing other lexical items regardless of their grammatical role. Children with PLI would find it easier to overlook a word rather than to formulate it inflectionally in order to demonstrate the essential syntactic categories. Add to this the notion of 'Lexicalised Inflected Verbs' (will be discussed shortly), as well as the fact that the Persian verb system has changed over time and it is not possible to construct simple verbs out of nouns or adjectives. This attrition has forced Iranian linguists to substitute compound verbs to fulfil the expectation of society to access new verbs for new concepts (Bateni, 1989). Clearly, the first impact of this type of word formation is to increase the number of compound verbs in the form of prepositional verbs or auxiliary verbs. Increasing the verbs' syntactic complements (semantically speaking, arguments) will burden child's memory capacity and result in difficulty pursuing appropriate words. This overload leads to word-finding difficulties and makes the child choose the easiest way of eliminating struggling: omission.

In addition to being included in Ungrammaticality, Missing Verbs and Missing Prepositions were also considered as Semantic LSMs and in the calculation of Semantic Errors. As explained in chapter 1, verb structure in Persian includes several affixes; both number and tense are marked in Persian, but not gender. These affixes attach to the verb root and sometimes a single word transfers the information regarding the subject (as it is pro-drop, discussed in chapter 1), object (through objective clitics), number, and tense (verbal markers and verb enclitics as shown in tables 1-2 and 1-3) (example 4-2a). This example is also an independent sentence if the objective clitic has a referee in the previous utterances. There are also compound verbs and prepositional verbs which should be added to this list as shown in example 4-2b.

Example 4-2

a) می خوندیمش.
mi\xund/im/esh.
Progressive marker\past verb root/plural first person verb marker/objective clitic.
were -ing\read/we/it.
We were reading it.

b) بر می داردمون.

bar mi\dar/ad/emun. (SALT transcription format: *mi\bardar/ad/emun*.)

Verbal particle progressive marker\present verb root/single third person verb marker/plural first person objective clitic.

Verb particle *were -ing\take/she or he/us*

She/He is collecting us.

It seems that Persian-speaking children do not abstract the string of morphemes individually but perceive them as a single unit. It might be the case that a given inflected verb structure in Persian is learned as a whole semantic item by children with language impairment, particularly if the acoustic features of individual morpheme are not salient enough to be perceived within the context, which is explainable by Leonard's Surface Hypothesis (Marinis, 2011). Consequently, children with PLI would treat the inflected form of the verb as one lexical unit. This feature has been referred to as *direct activation* by Rispoli et al. (Rispoli et al., 2012) inspired by Bybee's explanation that the more frequent groups of verbs and morphemes in input and use will prevent children from considering the verb infinitive (or root) separate from the verb tense or person marker (in brief, verb inflections).This will populate the lexicon with a repertoire of *lexicalised inflected verbs* (LIV) in addition to those non-verb and content words, so that they might be prone to make more omission errors rather than substitutions. Wexler's modified OI theory for languages like Spanish, with a dominant characteristic being null subject, has been rephrased by Simon-Cereijido & Gutierrez-Clellen as "lexical verbs in Spanish are not predicted to be substituted by infinitives" (Simon-Cereijido & Gutierrez-Clellen, 2007 p. 319). The aforementioned notion of LIV accompanied by Direct Activation (as follows) is likely to be seen in Persian children without language difficulty, as well; nevertheless, this is, once again, another uncharted area of child language in Iran which needs to be addressed before making any further assumptions.

The direct activation account could be regarded as the reason for distinctive performance between children with TDL and children with PLI in terms of such Semantic Errors as Nonsense String of Words as well as Wrong Responses that reflect deficits in transferring the appropriate meaning to the listener. This also includes phrases or strings of words with clear signs of the child's difficulty to the extent that she has no idea of what the sequence of words means and has produced it as a single linguistic unit. This claim, along with others including constraints in processing

capacity as well as a socio-pragmatic explanation, however, needs to be tested through conducting scientifically-designed research.

Language sample measures of lexical diversity considered in the current study were differently-calculated variants of NDW and NTW. English-speaking children with PLI show the first signs of difficulty with the late emergence of vocabulary development. This later vocabulary development shows a small lexicon and is accompanied by slow acquisition rate as well as difficulty in learning new words in an everyday language environment (Mainela-Arnold, Evans, & Coady, 2010; Sheng & McGregor, 2010; Watkins, Kelly, Harbers, & Hollis, 1995). The natural language sample measures of lexical development, such as NTW and NDW have been shown to be sensitive in distinguishing children with PLI from TDL peers. Similar results have been repeatedly reported on the differentiating capability of these traditional measures of lexical development for children with and without PLI. NTW and NDW of samples of 50 utterances showed the differential diagnosis between two groups of children between the ages of 24 and 50 months with PLI and their normal counterparts in Klee's study (Klee, 1992). Watkins et al. showed that children with PLI produced fewer different words (NDW) than their age-matched peers in 50, 100, and 200 word sample length (Watkins et al., 1995). Leonard and colleagues (1997) also showed significantly higher NDWs in the normal group compared to the PLI group in samples of 100 utterances (Leonard, Miller, & Gerber, 1999). In all these studies, children with PLI presented lower scores than their TDL counterparts matched for chronological age. In some cases, similar to MLU, they performed similarly to language-matched children with TDL or to younger children (Owen & Leonard, 2002). Miller's suggestion in matching children on their NDW would now seem reasonable because it would result in a more compatible population to compare (Miller, 1991). In a similar way to MLU, as discussed, NTW and NDW should be considered for what they are representing: measures of expressive vocabulary size. Although other factors such as word finding problems or limited processing capacity are involved in the difference in function between children with TDL and children with PLI (Watkins et al., 1995), it is not the role of NTW or NDW to reflect this. The aforementioned factors need to be separately studied with appropriately selected methodologies and the latter measures should be seen or used as indicators of overall expressive vocabulary size. They can be used in matching participants as well as documenting individual changes in vocabulary size during intervention (Watkins et al., 1995). There is another assumption that longer samples in terms of utterances would have an inflation effect on such lexical measures as NTW and NDW. Consequently, it

would be inevitable to assume that children with higher MLUs who produce more words directly affect the values of NTW and NDW in the normal group compared to PLI group (Klee, 1992; Owen & Leonard, 2002). In searching for a solution to this assumption, Klee suggested the application of a standard number of words instead of utterance length or sample time (Klee, 1992). An alternative measure called D was also introduced which is "less vulnerable to the influence of sample size effect" compared to NDW and TTR (Owen & Leonard, 2002 p. 929) and mathematically measures the alteration of repeated type-token ratios over a series of words measured in tokens, e.g. 35-50 words (Malvern, Richards, Chipere, & Duran, 2004). According to Owen & Leonard (2002), D can substitute for NDW because NDW "does not provide a measure of lexical diversity, since it confuses volubility with lexical skills" (Owen & Leonard, 2002 p. 928), but not when NDW is based on a fixed number of word tokens (see, e.g. Klee 1992). The results of the current study are consistent with previous findings for NDW (Klee, 1992; Leonard, Miller, & Gerber, 1999; Watkins et al., 1995). NDW in complete and intelligible utterances showed a promising outcome in diagnosing children with PLI from their TDL peers, as children with PLI had a significantly lower number of different words in their lexicon compared to their TDL counterparts. Children with PLI show problems in morpho-syntactic structure, which is assumed to be interacting with the size of vocabulary. The mutual relationship between semantics and syntax is logical due to the fact that words are the units governed by the rules of syntax (Perez-Leroux, Castilla-Earls, & Brunner, 2012). The diagnostic measures of NDW, however, are not informative because all fall outside the acceptable range in terms of sensitivity, specificity and LRs. Other Semantic LSMs are not discussed individually here due to the fact that they do not hold informative diagnostic accuracies and also they were merged to create composites that were discussed previously. So, a discussion on the results of the survey on Iranian SLTs integrated with the edifying outcomes of the diagnostic accuracy study will form the conclusion chapter as follows.

CHAPTER 5

Conclusion and future directions

5.1 Introduction

This study examined the diagnostic potential of language sample measures (LSMs) derived from Persian-speaking children's conversational language production which had previously been shown to differentiate children with and without Primary Language Impairment (PLI) in other languages, such as English. The other measures included either those Persian-specific measures which had been previously mentioned by Iranian SLTs as important in clinical diagnosis of primary language impairment in children, or those which demonstrated some empirical evidence in differentiating Persian-speaking children with and without PLI (e.g. direct object marker ra in Foroodi-Nejad's study, 2011). While the history of education for speech therapy in Iran dates back over 38 years, standardised assessments for language impairment still do not exist. Consequently, Iranian SLTs mostly rely on diverse, personally-developed procedures of language evaluation which make reliable comparison of research and intervention results impossible. The current study was informed by previous studies on child language assessment globally and seeks to contribute to Iranian SLTs' access to reliable ways of evaluating language through robust novel clinical research frameworks such as Evidence-Based Practice (EBP) and the International Classification of Functioning, Disability and Health (ICF). The study, therefore, followed the fundamentals of these approaches to achieve the main aims as follows.

5.2 Contribution to the new era of clinical research in Iran

An EBP framework, recently introduced in the field of speech-language pathology, was adopted to examine the accuracy of language sample measures in identifying Iranian Persian-speaking pre-school children with language impairment. Supported by a meta-analysis on detecting promising LSMs in identifying children with PLI in non-Persian studies, a diagnostic accuracy study was conducted. It was designed based on the principles which require that participants are recruited by applying a previously-studied assessment procedure with adequate diagnostic values and high agreement among researchers and clinicians, called the reference standard. A commonly agreed definition of what constitutes primary language impairment is also essential to incorporate a reliable clinical diagnosis. These two components, reference standard and a reliable definition of PLI, have also been problematic in non-Persian studies insofar as the researchers have frequently raised the issue of necessity to access a consensus among child language experts to eliminate different sources of bias (e.g. subjective bias and

validity reduction). They emphasize that availability of such reliable and standard devices would indirectly improve the process of clinical decision making (for instance, the following research conducted in different countries to study the issue: Lyons et al., 2008 in Ireland; Thordardottir et al., 2011 in Canada; Roulstone, 2001 and Roulstone, Peters, Glogowska, & Enderby, 2008 in the UK). With regard to the situation in Iran, neither a commonly accepted reference standard nor a definition of PLI currently exists in clinical practice. So, one aim of the research reported here was to survey SLTs with the aim of understanding their personal methods of child language assessment. Then, these personally-developed assessment methods were organised within a clinically practical framework like ICF to be implemented in the later phases of the diagnostic accuracy study as the best reference standard available. Two methods of data collection were adopted: questionnaire survey and focus groups. The results were analysed according to theme coding procedures. They were organised using the framework of the WHO ICF due to its well-known properties of being comprehensive in including body functions and structures as well as contextual factors in determining human health conditions. Iranian SLTs' qualitative data (collected through open-ended questions and focus groups) were analysed by the method of 'content analysis' (Hsieh & Shannon, 2005) which categorised SLTs' assessment items and procedures within more than 30 global themes. The resulting themes were organised within the framework of ICF as the first evidence-based methods of pre-school child language assessment in Iran. The resulting framework was then employed as the reference standard in defining the sample characteristics in the first two phases of the diagnostic accuracy (DA) study.

Naturalistic language sampling during children's free-play with their mothers was selected as the data collection procedure in the DA study. Twenty four children with PLI were recruited to the study through referrals from Iranian SLTs with the above-mentioned reference standard used to identify children with PLI. Twenty seven children with typically developing language (TDL) were randomly recruited from ordinary nurseries. These children met the criteria of TDL as specified in the reference standard. The children's language samples were transcribed following a newly-developed set of conventions for transcribing Persian language samples and the SALT computer program was adapted to accommodate these conventions (Miller & Iglesias, 2012). Out of 76 language sample measures (LSMs), 36 showed promise in differentiating children with and without PLI, including those investigated in previous studies as well as those that were mentioned as being useful by Iranian SLTs. Two other criteria in selecting suitable LSMs were 'being age-correlated' and 'being a feature exclusive to Persian'. The effect

sizes of all the measures were calculated using Cohen's *d* and ranged from 0.00 to 2.26. The diagnostic potential of the LSMs was examined through a phase I DA study (Sackett & Haynes, 2002) in which 36 measures were screened as possibly capable of making a distinction between children with and without PLI. The measures were compared across two groups of children with and without PLI, resulting in between-group differences for 29 measures. With the purpose of determining the clinical diagnostic accuracy of these measures, they were scrutinized using ROC curve analysis and Youden's Index, which resulted in the optimal cut-offs and estimations of sensitivity, specificity, likelihood ratios (LRs), and diagnostic odds ratios. These values are indicative of which measures are clinically capable of identifying children with PLI (sensitivity) and children with TDL (specificity) along with a prediction on which score might belong to which group of children (LRs).

The new frameworks open new windows to the phenomenon of child language assessment for Iranian SLTs, through which they can examine the unstudied field of Persian language by employing different clinically-adapted methodologies.

5.3 Clinical contributions

An interpretation of the diagnostic accuracy results of 29 measures with capability in distinguishing between two groups of children with and without PLI shows that 14 measures hold an area under curve (AUC) of .800 and higher which is interpreted as a powerful property in appraising the clinical utility of a test (Haynes et al., 2006 p. 351). Out of 29 measures, 12 showed fair to good sensitivity (minimum=.74, maximum=.98) but a wider range of specificity (minimum=.62, maximum=.98), indicating that they are most effective at identifying children with PLI than those with TDL (see table 4-29). Although for clinicians, measures that have high sensitivity are more desirable, they also seek measures with low false positive rate to prevent "time and economic burden as well as a potential psychological burden for parents" (Eisenberg & Guo, 2012 p.26).These 14 measures included General LSMs and General Errors whilst no Persian-specific measure held an AUC better than the .769 recorded for Total Use of Noun Plural Marker /*ha*.

With the criteria of area under the curve as "an overall measure of a test's accuracy" suggested by Haynes et al. (2006 p.286), the most clinically powerful measures are Grammaticality, Ungrammaticality, Ungrammatical Utterances, Mean Length of Utterances in words for language samples excluding one word utterances (MLUw-exc.), Total Semantic Errors, Mean Length of Utterance in words for all complete and

intelligible utterances (MLUw), Total Errors, Mean Length of Utterances in morphemes for language samples excluding one word utterances (MLUm-exc.), and Mean Length of Utterance in morphemes for all complete and intelligible utterances (MLUm). These measures hold AUCs extremely close to 1 and higher than .900 with desirable confidence intervals between .828 and 1. Because the CIs are narrow and again fall within the acceptable defined range, the clinician can easily decide which one of these competitors is the best in identifying children with PLI. By looking at table 4-29, the answer is apparent: Grammaticality and Ungrammaticality. These two measures have been shown to be diagnostically capable in similar studies by Eisenberg and Guo (2012) and Simon-Cereijido and Gutierrez-Clellen (2007), respectively (see also tables 2-5 and 2-6). In these studies, however, the accuracy measure was not AUC but sensitivity and specificity and both were within the fair to good range (80% and over) for both LSMs, as interpreted by Plante and Vance's (1994) criteria. The resultant diagnostic accuracies of Grammaticality/Ungrammaticality (AUC, sensitivity and specificity) document them as two "comprehensive measures" (term suggested by Eisenberg and Guo, 2012 p.6, and signifies a measure that encompasses several aspects of grammar) that would be able to correctly identify children with PLI. In search of a measure which is cross-linguistically reliable (Klee et al., 2007) and comprehensive enough (as defined above) in screening (Eisenberg & Guo, 2012), measures of Grammaticality might be considered appropriate due to their promising results in at least two languages so far, English (Eisenberg & Guo, 2012) and Persian (current study). Moreover, it seems that these measures are independent of linguistic-specific features of a given language so they might comply with the expectations of accessing measures insensitive to language types. A disadvantage with regard to other diagnostic metrics of Grammaticality/Ungrammaticality, however, is the unacceptably wide CIs, particularly in specificity. Although the CIs of AUC and sensitivity also appear wide, they do not exceed the acceptable range (all are greater than .800). On the other hand, the CIs of specificity are both wide and unacceptable due to falling below .800. This provides Grammaticality/Ungrammaticality as perfect measures in identifying children with PLI rather than children with TDL because they also possess very low negative LRs which provide high confidence for the clinician in allocating normal scores to those children who are unimpaired.

Other LSMs are also corroborated by previous diagnostic accuracy studies or discriminant analysis. Both MLUm and MLUw had documented fair to good diagnostic metrics (sensitivity and specificity) in combination with other factors or measures but

not when examined independently (see table 2-5 for a detailed review). All variants of MLUw and MLUm, MLUw-exc., and MLUm-exc. are within the first nine powerful LSMs with an interesting pattern of the precedence of 'excluding one-word utterances' type before the original one. This pattern might intensify the notion that pragmatic constraints impact on the child's language output (Klee, 1991). Recall that single-word utterances in Persian might include more than one morpheme and they may be of interest for being investigated individually; however, the current results are congruent with other studies when one-word utterances are excluded from the sample and result in higher MLUs (for a review, see Klee, 1992). In the present study, the measures of MLU with one-word utterances excluded also documented higher MLUs within each group (see results of phase I DA in chapter 4) as well as significantly higher MLUs in children with TDL compared to children with PLI (see results of phase II DA in chapter 4). However, the only measure among them which also demonstrated fair to good sensitivity and specificity was MLUw-exc, but its unacceptably wide CIs suggest that clinicians should be cautious in interpreting its results when identifying children with PLI (due to fair sensitivity without acceptable range of 95% CI). On the other hand, children with TDL can be identified with confidence due to good specificity with acceptable range of 95% CI (see table 4-29). This is also supported by high positive LR, which indicates high certainty in assigning atypical scores to those children who are actually impaired. From the remaining nine good LSMs, Total Semantic Errors records powerful AUC and good sensitivity and specificity with acceptable range of CIs calculated for all but sensitivity. As was described in the discussion section of chapter 4, both of its LRs also hold promising values while showing broad and unacceptable CIs. The specifications of its diagnostic accuracies are very similar to Ungrammatical Utterances except for its better negative LR (Table 4-29).

Under Dollaghan's criteria for choosing a better index test, LRs are preferable metrics to prioritise a diagnostic test and if the clinicians would also rather select the best LSM based on both LRs, Total Semantic Errors appears to be the best option. MLUw-exc., however, exhibited the best positive LR and Grammaticality/Ungrammaticality possess the best negative LR. The results suggest that the most clinically functional measure considering all types of diagnostic accuracy metrics is Total Semantic Errors which could provisionally be applicable by clinicians.

Four important points should not be overlooked by clinicians, as the target users of these measures. Firstly, the specific goal of the policy-makers for the clinician's workplace might help determine the right choice; for instance if the policy of their

workplace dictates that children with PLI are identified with the highest certainty possible, measures with high sensitivity and positive LR from those nine LSMs would be the best options only if their CIs fall within the acceptable range. Secondly, clinicians should seek the most appropriate LSM which is compatible with their specific goal in language assessment; for example assessing MLUw-exc. might not be the best choice for evaluating a child with high frequency of one-word utterances insofar as it requires ruling out the majority of utterances, leaving an unrepresentative language sample to be analysed. Thirdly, all CIs should be carefully checked with regard to encompassing the acceptable range defined for each diagnostic accuracy measure, i.e. higher than .800 for AUC, higher than 80% for sensitivity and specificity, higher than 10 for positive LR, and lower than .1 for negative LR. Fourthly, both Iranian and non-Iranian clinicians should consider that the location of the children and small sample size impose limits on the generalisability of the results to Isfahan city in light of the broad and unacceptable CIs. The researcher did not encounter in transcription any major variation in accents used in mother-child communicative language that had significant effect on the resultant LSMs. However, more research is needed to fortify and verify the results in terms of being generalised to the entire Iranian pre-school population in terms of detecting PLI. However, given the concerns and difficulties that inconsistent language assessment procedures have caused for Iranian SLTs and researchers, these measures can provide all Iranian SLTs with a substantial collection of natural language sample measures that they can rely on to differentiate children with and without PLI with some degree of certainty, especially as the outcomes support the claim that naturalistic language sampling (e.g. free-play in this study) rather than structured contexts can be effectively used for identifying impairment within young children (Heilmann, 2010). If used in conjunction with clinical observation, such naturally-derived language samples could provide a more coherent picture of a child's expressive language. A particular group of SLTs, nonetheless, can directly benefit from the results of the current study: SLTs who work locally in Isfahan city. It should be noted here, however, that no diagnostic accuracy study is complete unless the confirmed tests or measures involved have been tested in a phase III study (replication study) containing an independent sample of the population and similar or improved results for diagnostic accuracy have been achieved (Sackett & Haynes, 2002). The phase III DA attempts to confirm or reject the previously-obtained results within the context of a new sample of children so that this phase would resemble the clinical setting (i.e. each individual child needs to be assessed by the index test approved in phase II in search of re-approving its diagnostic potential

in finding new cases of impairment). This is further discussed in the section 'future directions'.

None of the Persian-specific measures, individually, proved to be acceptable as diagnostic measures. Even those measures that Iranian SLTs had mentioned in the survey as important indexes for diagnosis of PLI (e.g. noun plural marker /*ha*, wrong agreement, problems with verbal markers in terms of number, tense, and mode, direct object marker and so forth) did not prove to be diagnostically effective. Moreover, even such measures as direct object marker *ra*, progressive verb marker *mi*\, and objective clitics which had been shown to be able to differentiate between children with and without PLI (Foroodi-Nejad, 2011) did not provide appropriate diagnostic potential to clinically identify children with PLI. Iranian SLTs would not be advised to continue using these measures as quantitative indicators of primary language impairment unless more research with similar design and a larger sample size provides some empirical evidence for including them in the Persian-speaking child language assessment repertoire. Nevertheless, these measures can still be used qualitatively to show the child's weaknesses in expressive grammar and semantics, or they can be quantitatively added together, with the same pattern as in the current study, to produce composites with better diagnostic accuracies like Total Errors or Clitic Errors. Alternatively, a better way of quantitatively including them in the child's language assessment based on the results of the current study would be counting all the child's utterances that contain errors in these morphemes to produce the Grammaticality/Ungrammaticality measures which have shown the best diagnostic potential (AUC-based) compared to all other measures.

5.4 Theoretical implications

The outcomes of this study have provided support for several accounts of morpho-syntactic and semantic difficulties in PLI. Significantly higher numbers of Missing Grammatical and Semantic Errors were detected in the conversational language of children with PLI compared to their TDL peers. Children with PLI also scored significantly lower than their TDL peers in general measurements of the size of syntactic output (e.g. MLU and its variants) and vocabulary (e.g. NTW, NDW, and their variants). The error analysis of morpho-syntactic and lexico-semantic differences between children with and without PLI in this study reveals that errors of omission are more common than wrong usage of grammatical morphemes or lexical items (Tables 4-12 to 4-14). In addition, children with PLI generally use short utterances when

conveying their communicative intentions and a small vocabulary size directly affects their lexical richness. Existence of more omission errors compared to substitution errors on both morpho-syntactic and semantic measures can be attributed to optional use of related morphological and lexical units as well as their high level of reliance on processing strings of units (either morphological or lexical) as whole units. Two claims which are compatible with these data are the Extended Optional Infinitive (EOI) theory (Rice, Wexler, & Cleave, 1995) and the Direct Activation (Rispoli, Hadley, & Holt 2012) (as detailed in chapter 3). A later modified version of the EOI account by Wexler (2003) suggests extending the *Optionality* of the account to provide an explanation for observed omission errors in such grammatical morphemes as clitics in children with PLI (as done by Simon-Cereijido & Gutierrez-Clellen, 2007). Optionality is considered to be the expected result of limited processing capacity and is motivated by the child with PLI's difficulty in decoding morphological and lexical units of the input language individually. With a different point of view, Rice (1993) attributed affixation problems to a potential underlying morpho-syntactic deficit. This account assumes syntactic fundamentals are represented through the morphological system and children's difficulty is located in "clusters of syntactically related morphemes" (Rice, 1994 p.72); for instance the *–s* morpheme in English that syntactically shares the marked plurality (of the noun) with the unmarked agreement (of the verb); the morphemes of plural and agreement, therefore, are considered syntactically related morphemes. Due to the fact that in Persian, the morphemes have not been investigated by employing this notion, attributing the underlying morpho-syntactic deficit to Persian-specific errors observed in children with PLI is inappropriate. Further research would provide more evidence for or against this account from a newly-studied language like Persian. The current results, however, support the Surface Hypothesis of Leonard and colleagues (Leonard et al., 1992), because it is likely that the non-salient (unstressed) nature of some Persian-specific bound morphemes makes them difficult to perceive among other phonological strings which are more phonologically salient in the input language. Examples of such morphemes which were investigated in phase I DA are direct object marker *ra* (Karimi, 2003, Megerdoomian, 2012), addition or genitive sign */e-Ezafeh* (Parsafar, 2010), progressive marker *mi* (Megerdoomian, 2012), and objective clitics (Amini, 1997, Sadat-Tehrani, 2008). All these morphemes, except one (*ra*), are bound morphemes or affixes; *ra* is a closed-class grammatical morpheme (see chapter 1 for detailed description). In the Surface Account, "unstressed bound morphemes and closed-class morphemes are systematically filtered from the input" (Rice, 1994 p. 72) by children

with PLI. The stress pattern in colloquial Persian needs more investigation, particularly to find out its clinical applicability; however, two examples of stress location in a sentence carrying *ra* and the other without *ra* are provided in example 5-1 to show the non-salient nature of stress on such morphemes (Amini, 1997; Megerdoomian, 2012; Parsafar, 2010; Sadat-Tehrani, 2008). Both sentences are identical with respect to subject and verb but the object in the first sentence is definite whilst in the second it is indefinite. Neither subject nor DO marker *ra* accepts stress. The stress pattern of the colloquial Persian morphological system should be clinically investigated, which would help to support or refute this notion.

Example 5-1 No stress on *ra* but on the verb:
Man-havij-ra-xo´rd/am. (example modified from Megerdoomian, 2012)
subject (pronoun)-definite direct object-direct object marker *ra*-verb
I carrot DO marker *ra* ate/1st person single verbal marker
I ate the carrot.

Stress is on the noun:
Man-havi´j-xord/am.
subject (pronoun)-indefinite direct object-verb
I carrot ate/1st person single verbal marker
I ate carrot(s).

5.5 Future directions

Possible future directions may start with two suggestions derived from the immediate results of the present study. Firstly, although the results show an overall promising diagnostic accuracy for five measures (Grammaticality, Ungrammaticality, Total Number of Ungrammatical Utterances, MLUw-excluding one-word utterances, and Total Semantic Errors), their corresponding CIs do not fall within acceptable ranges (Dollaghan, 2007). Replication of the second phase of this study with a larger sample of the population would help to improve the accuracy values and CIs. The current results, however, are clinically applicable for current child language evaluation in Iran as there is currently no access at all to a reliable assessment measure. Nonetheless, the CIs of the accuracy metrics must be kept in mind. Secondly, according to Sackett and Haynes (2002), the next phase to complete the quartet of the diagnostic accuracy of language sample measures would include an independent sample of 42 to 54 month old Iranian Persian-speaking children. The reference standard and the index tests (29 LSMs) would be administered to all children to check consistency between the results of that provisional study and phase II of the current study (chapter 4). The outcomes of phase III of DA would be more clinically applicable because phase III is more similar to a real

clinical situation. Time constraints and limits on financial and human resources prevented the researcher of the current study from carrying out the further stages of DA for the current LSMs; however, this is provisionally the next research plan of the author. Moreover, the association between age and language sample measures was examined through the pre-accuracy study in order to find the measures that change by age development. The age-associated language measures had been shown to be able to differentiate between children with and without PLI (for instance, see Gavin et al., 1993). Although there was no meaningful correlation between age and language sample measures in the current study, in some cases we observed that accuracy measures showed lower or higher values than those seen in languages other than Persian and with different age ranges. Replication of the pre-accuracy phase with a wider age range is recommended to trace any age-related developmental association of the measures in Persian. Then the DA phases should be replicated according to the results of that pre-accuracy study to explore accuracy measures across different age groups and compare them to other studies, cross-linguistically. Those results would help in finding out whether the differences observed were age-related or language-related (Klee et al, 2007).

Another point of research is that in this study, the children with PLI were not classified in terms of severity, or the contingent subcategories of PLI. The results might have appeared differently if the PLI population studied was more homogenous in terms of severity or type. The heterogeneity of PLI is has been shown to have an impact on the results of studies; hence, it is suggested that new approaches to the severity and typology of PLI are taken into account in future studies with similar objectives. For instance, in a new perspective on the categorisation of SLI, taxometric studies have evidenced that children with SLI at three, four and six years of age are quantitatively different from their non-impaired peers rather than qualitatively (Dollaghan, 2004). Supported by the mean above minus below a cut (MAMBAC) procedure introduced by Meehl & Yonce (1994), Dollaghan suggests a dimensional rather than a categorical feature of SLI (Dollaghan, 2004, 2011). Examining Dollaghan's notion in Persian would provide a different set of data for error analysis, as this study showed that children with PLI showed more omission errors rather than substitution errors; Persian might further provide a linguistically-different context for the researchers to examine this opinion among others.

The next suggestion is to focus on the nature of language measures. The mutual relationship between syntactic and semantic measures could be explored; NDW, for

instance, could be examined using syntactic categories such as nouns, verbs or conjunctions. The results would support the notion of an inter-correlation of morpho-syntax and semantics. The different behaviour of morphemes in various grammatical contexts, e.g. the English plural noun marker *–s* in determiner or quantifier contexts, has been shown to be a result of underlying grammatical structure (Rice & Oetting, 1993). If the inconsistency of morphemes' behaviour in Persian is determined by changes in semantic categories, the underlying linguistic structure might be responsible. A further area for future research on language measures could be to use combinations of different language sample measures to create novel syntactic and semantic models for justifying the difficulties experienced by children with PLI. Composite measures created out of those promising measures with strong effect sizes as well as high accuracy measures could have improved diagnostic values. Another measure-specific area of research appears when paying attention to the point that some measures (i.e. Missing Objective Clitic + Missing Possessive Clitic + Missing /*e-Ezafeh*) were combined to make the category of Clitic Errors. The frequency of a child's exposure to these morphemes would be affected by the frequency of production of each morpheme individually in everyday language input. This will consequently determine the strengths or weaknesses in the child's language performance. The more frequent morphemes could obscure the child's probable weaknesses in low frequent morphemes. Since there is no information about the frequency of morphemes in colloquial Persian, further investigation of this issue is recommended, with a complementary study on the developmental emergent trajectory of each morpheme. It is also recommended that error types, i.e. omission or substitution, be studied in detail to find out how congruent errors appear compared to other languages. This would also be helpful in mapping error types cross-linguistically. Focusing on the language sampling approach, it is suggested that this study be replicated with smaller language samples with the aim of introducing a faster assessment procedure and developing a screening tool embedded in reliable language sample contexts. One of the reasons that SLTs are reluctant to employ LSA in their routine assessment procedure is the length of time it takes to administer it. They even mentioned that they would collect language samples from their clients, but not analyse them due to the length of time taken to perform the analysis. Nevertheless, in a literature review by Casby (2011), smaller samples showed MLU-related results similar to those of longer samples in children with TDL; the samples ranged from ten utterances to 150 with reliability coefficients more than .80 for a measure of mean length of responses in words as well as for MLU. Although longer samples would enhance the reliability of

language measures, shorter samples are more realistic in some clinical settings because their reliability would be affected by the "type of sample collected, measures of interest, child's age, and child's diagnosis" (Heilmann et al., 2010 p.402). The suggestion is that Persian MLU values (and other LSMs) should be investigated for the purpose of screening within small language samples and comparing them to longer samples. In the case of high correlations between two samples, the resultant LSMs of smaller samples should be examined in a properly-designed diagnostic accuracy study.

Thinking more widely about the elicitation procedures used while sampling language in children, the natural context of sampling in this study can be combined with some less-frequently occurring communication opportunities e.g. narratives, to examine the language of children with PLI in terms of different grammatical forms. Some grammatical problems in children with PLI are only revealed when elicited sampling is used and specific syntactic contexts are included in language sampling. This has been observed in languages such as English and Cantonese (Simon-Cereijido & Gutierrez-Clellen, 2007). Elicitation tasks would also be helpful in eliciting less common grammatical structures (Eisenberg, 2005).

Following the previous point, it is suggested that processing tasks are used in conjunction with linguistic tasks (Conti-Ramsden, 2003; Klee et al., 2007; Thordardottir et al., 2011). A combination of highly-structured tasks and time pressure would be a challenge for children with PLI whose deficit is not apparent in ordinary language settings such as free-play and conversation. In this case, several tasks such as "describing pictures, retelling complex stories, or producing expository discourse with time pressures imposed" (Moyle et al., 2011 p. 557) are helpful in disclosing child's problems.

A final suggestion is about the nature of errors. Coding error types would help in specifying the linguistic problems qualitatively and provide enough detail for planning intervention goals (Eisenberg & Guo, 2012). As a result, it is strongly suggested that the error types used in the current study are defined and organised with the aim of designing a practical clinical assessment based on a descriptive developmental model of language impairment.

5.6 Limitations

In addition to the topic-specific concerns which were discussed previously in this chapter, procedural and technical constraints will now be discussed.

5.6.1 Procedural constraints

Extremely low questionnaire response rate in the SLTs was a serious limitation and was completely out of the researcher's control (see chapter 3 for a full explanation). However, a more convenient sampling procedure, snowball sampling, was chosen to compensate for the small number of SLTs recruited. The response rate might be improved if a pre-planned awareness program was established to increase Iranian SLTs' knowledge of the importance of getting involved in this type of project and inform them of how their career would be improved through active engagement in similar research.

The first procedural limitation in the diagnostic accuracy study was related to insufficient human resources to be employed in the process of diagnosing children with and without PLI. Transcription and coding of the language samples was also restricted to the main researcher due to the fact that no trained Persian transcriber was available. This has led to some degree of subjective bias that the researcher attempted to overcome by frequently checking the child's diagnostic condition against the reference standard. Also reliability constraints in transcribing and coding the samples were improved by listening to the sample three times if there was any doubt about the intelligibility. Human resources can be made more available if Iranian SLTs are taught to the new ICF-based reference standard and are trained in the newly-adapted conventions for transcribing language samples in SALT.

Another limitation was related to the small number of children recruited for the diagnostic study. As shown in chapter 4, only 24 out of 110 children with PLI were recruited to the study. There were 27 children with TDL recruited in spite of the researcher initially voice-recording 55. The attrition of children with PLI was probably because SLTs were not informed enough about the initial inclusion criteria and they may have thought that more referrals would benefit the research more. This might be improved through the same pre-planned programs for educating SLTs about their important role in enhancing the quality of clinical research. The attrition of language samples included from children with TDL was due to the time constraints in transcribing the samples. The samples are available and will be used in future research.

5.6.2 Technical limitations

The necessity of developing a new set of transcription conventions was apparent because no Persian transcribing protocol had been comprehensively prepared for any software. Consequently, SALT was selected due to its more straightforward nature in transcribing and analysing language samples, although a new set of rules was needed to

accommodate to the affix-rich nature of Persian, particularly prefixes. From the onset of raising the issue to the SALT technical team to getting access to the version adapted for such a prefix-rich language as Persian, the time constraints played a significant role. Another technical constraint is related to the procedure of decreasing subjective bias in diagnostic accuracy studies: blinding (Haynes et al., 2006). Blinding is the main method for achieving the highest objectivity in diagnosing cases with and without the target condition. In studying communication problems, however, blinding is very difficult, if not impossible. According to Dollaghan (2007), most of the communication disorders, including PLI, require some subjectivity in rating and judgement. Subjective bias, nonetheless, can potentially be controlled "first by ensuring that the index test and reference measure are not administered by the same examiner, and second by ensuring that examiners have no information about the participants that could systematically influence the way in which they administer, score, or interpret the results" (Dollaghan, 2007 p. 86). For the interest of the current study, the majority of children with PLI (n=22) were recruited by referrals from SLTs other than the researcher and only two children with PLI were diagnosed by the researcher through the procedure of recruiting children with TDL, so the administrations of the index test and reference standard were highly controlled for subjective bias. On the other hand, the second criterion in controlling this bias was not fully met due to the fact that the researcher/examiner was aware of all demographics and conditions of the children. To prevent knowledge about the children interfering with the process of transcribing, coding and diagnosis based on index tests (LSMs), however, the children were dual-coded at the beginning of the study. One set of codes included the information about children's condition in terms of being PLI or not and the second code set was randomly assigned to the first code set. The data about children's LSMs was recorded in SPSS-19 by looking at the second codes. In this way, an attempt to control the subjective bias was made at the second level of Dollaghan's definition.

5.7 Conclusion

An observed significant difference between children with PLI and their typically developing peers on any language sample measure should not be interpreted as that measure being appropriate for use in clinical diagnosis. The clinical diagnostic value of measures must be weighed against the defined best measure of diagnosis including AUC, sensitivity and specificity, LRs and DOR as well as their corresponding CIs cross-linguistically. This evaluation should follow a scientifically-designed framework

such as EBP, integrating frameworks such as the ICF, in order to gain a more united approach towards improving assessment procedures in child language assessment and intervention.

Implementing the aforementioned advantages of EBP regulations, this study could contribute to motivating Iranian SLTs to replace English assessment criteria within their clinical decision making with the solely available set of national-driven, evidence-based measures for evaluating Iranian Persian-speaking pre-school children. The measures have been reliably assayed through this first diagnostic accuracy study in Persian, not only in the field of child language but also in the field of Iranian speech therapy in general. This study has provided well-grounded evidence on the clinical applicability of two robust frameworks of EBP and ICF for child language assessment. The suggested empirical ICF-based framework for the assessment of primary language impairment in pre-school years has created a direct connection between research and practice by providing practice-based evidence of what is currently used as an assessment of primary language impairment in Iran. It may also encourage Iranian researchers to implement more scientific frameworks, in the forms of EBP and ICF, in investigating new approaches or improving the current research designs to more standardised scientific designs within child language assessment and intervention in Iran. Contrary to current SLTs' belief in dissociation between what the researchers pursue and find and what the clinicians intuit and experience in a real clinical setting, this study has documented the fact that carefully examining language sample measures through evidence-based practice approaches like diagnostic accuracy can provide more certainty for clinicians that the gap between theory and practice is removable. This congruence between research outcomes and clinical goals is mutually assisted by implementing strong diagnostic methodologies, which will finally result in a better consensus among researchers and clinicians.

APPENDICES

Appendix A-1

Persian Questionnaire

به نام خدا

همکار محترم

این پرسشنامه ، قسمتی از یک طرح پژوهشی درباره ارزیابی رشد و تکامل زبان در کودکان پیش دبستانی ایرانی است که همکاری شما در به نتیجه رسیدن آن نقش به سزایی خواهد داشت. ما علاقمندیم که نظر شما را در این مورد بدانیم و از همکاری شما در تکمیل این پرسشنامه پیشاپیش سپاسگزاریم.

لطفاً با تجربه کار بالینی خودتان به هر سوال پاسخ دهید و سپس آن را با استفاده از پاکتی که هزینه پستی آن از قبل پرداخت شده و به ضمیمه این فرم برایتان فرستاده شده است، برای مجری طرح بفرستید.

پاسخهای شما به سوالات زیر کاملاً محرمانه و بدون نام بوده و بیش از 15 دقیقه از وقت شما را نخواهد گرفت.

لازم به ذکر است که نتایج این بررسی به شکل چاپ مقاله یا سخنرانی ارائه خواهد شد و برای جامعه گفتار درمانی ایران بسیار ارزشمند خواهد بود. اگر شخصاً مایل به دریافت نسخه ای از نتایج ان هستید، لطفاً با تلفن یا پست الکترونیکی مجری طرح که در زیر صفحات پرسشنامه چاپ شده است، تماس بگیرید.

1. سن: ----------

2. جنس:

☐ مرد ☐ زن

3. سطح تحصیلات:

☐ کارشناسی ☐ کارشناسی ارشد ☐ دکترا

4. سابقه کار گفتار درمانی شما چند سال است؟

5. در کدام استان و شهرستان مشغول به کار هستید؟

6. وضعیت اقتصادی-اجتماعی منطقه ای که در ان کار می کنید، چگونه است؟

☐ سطح پایین

☐ متوسط پایین

☐ متوسط

☐ متوسط بالا

☐ سطح بالا

7. تعداد بیماران شما طی سه ماه گذشته در هر کدام از گروههای سنی زیر چقدر بوده است؟

☐ کودکان زیر 3 سال: ---------- نفر

☐ کودکان پیش دبستانی (3:1 تا 6 سال): ---------- نفر

☐ کودکان دبستانی (6:1 تا 12 سال): ---------- نفر

8. لطفاً علت اولیه ارجاع کودکان پیش دبستانی به کلینیک خود را با رتبه بندی از «1=کمترین ارجاع» تا «5=بیشترین ارجاع» مشخص کنید. ضمناً در هر مورد تعداد تقریبی مراجعان را به درصد ذکر فرمائید.

☐ مشکلات تلفظی

☐ تاخیر در گفتار

☐ مشکلات ارتباطی

☐ مشکلات زبانی

☐ مشکلات حافظه

☐ لکنت

☐ مشکلات صوت

9. لطفاً در سوال قبل، در هر مورد درصد تقریبی مراجعان را در جلوی هر کدام ذکر کنید.

10. در بین کودکان ارجاع شده به شما تحت عنوان تاخیر در گفتار، شما بعد از ارزیابی ، چند درصد از آنها را در طبقه بندی های زیر قرار داده اید؟ لطفاً سنی را نیز که فکر می کنید در آن سن می توان روی کودک تشخیص مورد نظر را گذاشت، ذکر فرمائید.

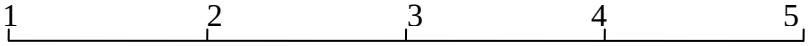
- ☐ **اختلال در تولید: ----------٪ ؛ در سن ---------- سالگی**
- ☐ **تاخیر در گفتار یا دیر شکوفائی زبان: ----------٪ ؛ در سن --------- سالگی**
- ☐ **اختلال زبانی به دلیل MR : ----------٪ ؛ در سن ---------- سالگی**
- ☐ **اختلال زبانی به دلیل HI : ----------٪ ؛ در سن ---------- سالگی**
- ☐ **اختلال زبانی به دلیل طیف اوتیسم: ----------٪ ؛ در سن ---------- سالگی**
- ☐ **اختلال زبانی بدون دلیل مشخص یا آسیب زبانی ویژه(SLI) : ------٪ ؛ در سن ------- سالگی**
- ☐ **سایر موارد (لطفاً مشخص کنید) : ----------٪ ؛ در سن ---------- سالگی**

11. در جدول زیر، عناوین حوزه های ارزیابی احتمالی برای تشخیص آسیب های زبانی در کودکان آمده اند. لطفاً مشخص کنید که معمولاً ارزیابی کدام حوزه ها به شما در امر تشخیص آسیب زبانی کمک می کند و شما آن را با چه روشی ارزیابی می کنید؟

	گزارش یا مصاحبه با والدین	مشاهده کودک	استفاده از چک لیست	ارزیابی غیر رسمی	تستهای استاندارد (لطفاً نام ببرید)	نظر خواهی از همکاران درباره مراجع	روشهای دیگر (لطفاً نام ببرید)
تاریخچه کودک							
جنسیت							
سابقه وجود احتلال مشابه در خانواده							
تاریخچه پزشکی کودک							
سابقه عفونت گوش میانی							
دوزبانگی							
رشد زبان (به ویژه مرحله babbling و کیفیت و کمیت آن)							
بازی تخیلی							
میزان محرکهای زبانی در محیط کودک							
تعامل اجتماعی با والدین و همسالان							

سابقه حضور در مهد کودک و مدت زمان آن							
سطح تحصیلات والدین							
وضعیت شنوائی							
غربالگری شنوائی							
تست نجوا							
تست PTA							
تست SRT							
شناخت							
تست IQ							
ارزیابی بازی							
ارزیابی نقاشی							
ارزیابی حافظه							
وضعیت عصبی							
نامه ارجاع از سوی نورولوژیست							
ارزیابی مهارتهای حرکتی (ظریف و زمخت)							
رشد دهانی – حرکتی							
پردازش زبانی							
ارزیابی تکرار ناکلمه ها							
زبان درکی							
آگاهی واجی							
خزانه واژگانی							
روابط معنائی							
نحو							
صرف							
قوانین مکالمه							
زبان بیانی							
توجه مشترک							
استفاده از ژست							
فهرست واجهای کودک							
خزانه واژگانی							
روابط معنائی							
نحو							
صرف							
قوانین مکالمه							
مهارتهای داستانگوئی و استدلال							

12. کدام یک از موارد ذکر شده درر جدول بالا به شما در تشخیص کودک مبتلا به اختلال زبانی بدون علت مشخص یا SLI بیشتر کمک می کنند؟ لطفاً میزان کمک را از«1 = خیلی کم» تا «5 = خیلی زیاد» روی محورهای زیر علامت بزنید.



- سابقه وجود احتلال مشابه در خانواده
- سابقه عفونت گوش میانی
- دوزبانگی
- میزان محرکهای زبانی در محیط کودک
- تعامل اجتماعی با والدین و همسالان
- سطح تحصیلات والدین وضعیت شنوائی
- شناخت
- ارزیابی بازی
- ارزیابی حافظه وضعیت عصبی
- رشد دهانی – حرکتی
- پردازش زبانی
- ارزیابی تکرار ناکلمه ها
- اندازه خزانه واژگان درکی
- اندازه خزانه واژگان بیانی
- درک و بیان روابط معنائی
- درک و بیان صرف
- درک و بیان نحو
- مهارتهای کاربرد شناختی

13. آیا شما از تستهای استاندارد استفاده می کنید؟

☐ بله

☐ خیر – در این صورت به سوال 15 پاسخ دهید.

14. از تستهای استاندارد به چه منظوری استفاده می کنید؟

☐ درمان ☐ غربالگری

☐ تشخیص افتراقی ☐ تشخیص

15. دلیل یا دلایل شما برای عدم استفاده از تستهای استاندارد چیست؟

☐ کمبود آنها در ایران

☐ درباره آنها چیزی نمی دانم.

☐ دلایل دیگر (لطفاً نام ببرید)

16. **آیا در ارزیابی زبان بیانی، از تحلیل نمونه زبانی استفاده می کنید؟**

- ☐ **بله**
- ☐ **خیر – در این صورت به سوال 23 پاسخ دهید.**

17. **در نمونه گیری زبانی خود معمولاً از چه روشی استفاده می کنید؟ کدام روش به شما در تصمیم گیری درباره مشکل کودک ببیشتر کمک می کند؟ به ترتیب شماره مشخص کنید.**

- ☐ **تقلید کردن از یک بزرگسال**
- ☐ **برانگیختن زبان بیانی**
- ☐ **نمونه گیری طبیعی از زبان**

18. **طول نمونه زبانی که ضبط می کنید معمولاً چقدر است؟**

- **_____ دقیقه**
- **_____ گفته**

19. **معمولاً نمونه زبانی را چگونه ضبط می کنید؟**

- ☐ **ضبط صوتی**
- ☐ **ضبط ویدئویی**
- ☐ **یادداشت نمونه، همزمان با صحبت کردن کودک**
- ☐ **یادداشتهای روزانه والدین از صحبتهای کودک**

20. **معمولاً از کدام مقیاس برای تحلیل نمونه زبانی استفاده می کنید؟**

- **LARSP**
- **میانگین طول گفته یا MLU**
- **نسبت نوع به تعداد واژگان یا Type-Token Ratio (TTR)**
- **تنوع واژگانی یا Lexical Diversity**
- **مقیاس خاصی را استفاده نمی کنم.**
- **مقیاسهای دیگر (لطفاً نام ببرید):**

21. **از تحلیل نمونه زبانی به چه منظوری استفاده می کنید؟**

- ☐ **غربالگری**
- ☐ **تشخیص**
- ☐ **درمان**
- ☐ **تشخیص افتراقی**

22. **تقریباً چند نمونه زبانی را در سال تحلیل می کنید؟**

23. **لطفاً بنویسید که چرا از تحلیل نمونه زبانی استفاده نمی کنید؟**

- ☐ **عدم آموزش**
- ☐ **کمبود مهارت**
- ☐ **فشارهای مالی**
- ☐ **کمبود زمان**
- ☐ **عدم دسترسی به سخت افزارهای کامپیوتری**
- ☐ **عدم دسترسی به نرم افزارهای کامپیوتری به منظور ثبت و تحلیل داده ها**

24. **لطفاً هرگونه مشخصه زبان شناختی در زبان کودک، به ویژه دستور زبان، را که به شما در تشخیص کودک به عنوان دچار آسیب زبانی یا SLI کمک می کند، در جدول زیر وارد کنید.**

25. **آیا به نظر شما حوزه دیگری از ارزیابی زبان در کودکان پیش دبستانی باقی مانده که در قسمتهای قبل به آنها اشاره نشده باشد؟ اگر بله؛ آنها را در جدولهای زیر بنویسید.**

26. **لطفاً ذکر کنید که خود شما شخصاً چگونه تشخیص می دهید که کودکی دچار اختلال زبانی است؟ در جدول زیر روند ارزیابی اختلال زبانی در کودکان پیش دبستانی را از نظر خودتان ذکر فرمائید.**

27. اگر مایل به دریافت خلاصه ای از نتایج این بررسی هستید، نشانی و شماره تماس خود را در اینجا یادداشت فرمائید.

با سپاس فراوان از وقتی که برای همکاری در اختیار این پژوهش قرار دادید.

Appendix A-2

English-translated Questionnaire

This questionnaire is a part of a research study about clinical assessment of preschool language development in Iran. We are very interested in your views and would appreciate it if you would complete this short questionnaire.

Please answer each question as it relates to your own clinical work. It should not take more than 15 minutes. You are not forced to reply all the questions.

Your responses to the questions below will be completely anonymous since we are not asking you to put your name on the questionnaire form.

The results of this survey will be summarized and presented in the form of a research paper and/or presentation. If you are interested in the findings, you may email me and I will send you a summary in due course.

Please return the questionnaire after completing it using the enclosed stamped envelope.

1. Age: ----------

2. Gender:
   - □ Female
   - □ Male

3. What is your level of education in speech therapy?
   - □ BSc
   - □ MSc
   - □ PhD

4. How long have you worked as a speech and language therapist?

5. Which province and city do you currently work in?

6. What is the socioeconomic condition of the area you work in?
   - □ Low
   - □ Below average
   - □ Average
   - □ Above average
   - □ High

7. What is the size of your current caseload, during the past 3 months, in any of the following groups of age?
   - Infants and toddlers (birth to 3 year-olds) ……
   - Pre-schoolers (3;1 to 6 year-olds) …….
   - School-aged children (6;1 to 12 year-olds) …….

8. Please rate the reasons of referrals to you in preschool range of age from "1=the least referrals" to "5=the most referrals".
   - □ Pronunciation problems
   - □ Late-talking
   - □ Communication problems
   - □ Language difficulty
   - □ Memory problems
   - □ Stuttering
   - □ Voice problems

9. In previous question, please specify the percentage of each reason in front of them.

10. If children are referred to you as late-talkers, what percentage do you actually diagnose with the following conditions and at what age?
    - Articulation disorder ……%; .....years old
    - Late-blooming or late-talking…….%; .....years old
    - Language disorder due to MR ……%; ..... years old
    - Language disorder due to HI …….%; ..... years old
    - Language disorder due to ASD ……..…%; .....years old
    - Language disorder without specific reason or Specific Language Impairment (SLI) ……..%; ..... years old
    - Others (please specify)……%; .....years old

11. Below is a list of some assessment areas that may be evaluated in children referred for assessment of language impairment. Please indicate which area you usually evaluate to help you diagnose a child as language impaired and how do you evaluate it?

	Parents' report or interview	**Observation**	**Checklists**	**Informal assessment**	**Standardized tests (please specify)**	**Colleagues' independent judgement**	**Other procedures (please specify)**
Child history							
Gender							
History of the difficulty in the family							
Child's medical history							
History of Otitis Media							
Bilingualism							
Language development (specifically age of babbling and its quantity and quality)							
Pretend play							
Quality and quantity of Language stimulation							
Social interaction with parents and peers							
History of attending at the nursery and its duration							
Parents' educational level							
Hearing status							
Pure tone Screening							
Whispering test							
PTA (threshold) test							
SRT test							
Cognition							
IQ tests							
Play assessment							
Painting assessment							
Auditory and visual perception							
Memory assessment (short- or long-term)							

	Parents' report or interview	Observation	Checklists	Informal assessment	Standardized tests (please specify)	Colleagues' independent judgement	Other procedures (please specify)
Neurological status							
Neurologist's referral letter							
Test of motor skills (fine and gross movements)							
Oro-motor development							
Language Processing							
Non-word repetition tasks							
Pre-verbal skills (joint attention, imitation etc.)							
Receptive language							
Phonological awareness							
Vocabulary							
Semantic relations							
Syntax							
Morphology							
Conversation rules							
Expressive language							
Joint attention							
Use of gestures							
Phonetic inventory							
vocabulary							
Semantic relations							
Syntax							
Morphology							
Conversation rules							
Narrative and reasoning skills							

12. Which of the above-mentioned areas would help you to identify the child as language impaired without specific reason or SLI? Please rate them from "1=the least" to "5=the most" on the lines below: *(Note that the Likert scale has been repeated for each item in the main questionnaire sent to the SLTs.)*

1 2 3 4 5

- History of difficulty in the family
- History of Otitis Media
- Bilingualism

- Quality of Language stimulation
- Social interaction
- Parents' educational level
- Hearing condition
- Cognitive
- Play assessment
- Memory assessment
- Neurological status
- Oro-motor development
- Language Processing
- Non-word repetition tasks
- Phonological awareness
- Receptive and Expressive vocabulary
- Receptive and Expressive Semantic relations
- Receptive and expressive syntax
- Receptive and expressive morphology
- Pragmatic skills

13. Do you use standardized language tests?
    - ☐ Yes
    - ☐ No- if no, please skip to Q. 15

14. For which of the following purposes do you use standardized language tests?
    - ☐ Screening
    - ☐ Treatment
    - ☐ Diagnosis
    - ☐ Differential diagnosis

15. What is/are the reason/s you do not use standardized language tests?
    - ☐ Lack of them in Iran
    - ☐ I don't know much about them
    - ☐ Other reasons:

16. Do you use language sample analysis in the assessment of expressive language?
    - ☐ Yes
    - ☐ No – if no please skip to Q. 23

17. How do you usually elicit language production in language sampling? And which one helps you the most to make a decision about the child's problem?
    - ☐ Imitation tasks
    - ☐ Elicitation tasks
    - ☐ Natural language sampling

18. What is the typical length of the language sample you record?
    - Minutes ----------
    - Utterances ----------

19. How do you usually record the language sample?
    - ☐ Audio recording
    - ☐ Video recording
    - ☐ Real-time transcription
    - ☐ Parents' diary

20. Which language measures do you usually analyse from the language sample?
    - ☐ LARSP
    - ☐ Mean Length of Utterance (MLU)
    - ☐ Type-Token Ratio (TTR)
    - ☐ Lexical Diversity
    - Nothing special
    - ☐ Other – please specify

21. For which of the following purposes do you use language sample analysis?
    - ☐ Screening
    - ☐ Diagnosis
    - ☐ Treatment
    - ☐ Differential diagnosis

22. Approximately how many language samples do you analyse in a year?

23. If you don't use language sample analysis, indicate why.
    - ☐ Lack of training
    - ☐ Lack of expertise
    - ☐ Financial constraints
    - ☐ Time constraints
    - ☐ Lack of computer hardware
    - ☐ Lack of computer software for recording and analysing

24. Please use the box below to outline any linguistic features in child language, particularly grammar, which leads you to judge about the child as language impaired or SLI.

25. Is there any other area of language assessment in preschool children which you think has not been addressed in this questionnaire? Please specify in the boxes below.

26. Please describe how you decide that a preschool child is language impaired by own? You can use the following box to explain your own assessment process.

27. Please leave your contact details if you would like to receive a summary of the results of this study.

Thank you very much in advance for your collaboration.

Appendix B-1

Persian Information Sheet

به نام خدا

به: والدین (نام کودک) تاریخ

با سلام و احترام

فرزند شما به منظورانجام یک کار تحقیقاتی ، از بین همکلاسیهایش به صورت تصادفی انتخاب شده است. از شما درخواست می شود که در یک جلسه بازی با فرزند خود که در محل دفتر اینجانب در دانشکده علوم توانبخشی برگزار می شود، شرکت فرمائید. در این جلسه، صدای شما و فرزندتان طی بازی با یکدیگر ضبط شده و سپس صدای کودک صرفاً به منظور بررسی مشخصات زبانی ، تحلیل خواهد شد. نمرات همه کودکان شرکت کننده، به صورت انفرادی گزارش نخواهد شد اما شما می توانیدادر صورت تمایل یک گزارش کلی از ان را درخواست کنید. تلفن تماس xxxxxxx می باشد.

نام هر کودک محرمانه باقی می ماند و هر کودک فقط با یک کد مشخص می شود. اطلاعات بدست آمده در یک مکان امن نگهداری شده و فقط بنده، دو استاد راهنما، و دو ممتحن اینجانب می توانند به آنها دسترسی داشته باشند.

شرکت شما در این مطالعه، کاملاً اختیاری است و شما اجازه دارید در هر زمان که مایل باشید، از ادامه همکاری در این مطالعه، صرف نظر کنید اما مشارکت شما در این امر جهت ارتقای دانش گفتار درمانی در ایران و به منظور کمک به بهداشت رشد و تکامل زبان و گفتار کودکان ایرانی بسیار ارزشمند خواهد بود.

از شما به خاطر وقتی که برای خواندن این نامه گذاشتید بسیار متشکرم. اگر مایل به همکاری با این طرح هستید، لطفاً قرارداد پیوست را امضا کرده و به مدیر مهد کودک تحویل دهید. نسخه دیگر قرارداد را نزد خودتان نگه دارید.

این پژوهش دارای مجوز از کمیته اخلاق در پژوهش دو دانشگاه علوم پزشکی اصفهان و دانشگاه نیوکاسل می باشد.

با سپاس

یلدا کاظمی

Appendix B-2

English-translated Information sheet for parents

Date: ----------

To (child's name) parents,

My name is Yalda Kazemi and I am a PhD student of speech and language therapy at Newcastle University, where I am doing a study on child language assessment.

You child has been invited to join the study because I have selected her/him randomly among her classmates in this nursery.

You will be asked to play with your child in a free-play setting at my office and I will record your voices during the session to analyse it in terms of your child's language components.

Joining this study is entirely voluntary. You have the right to withdraw your consent without affecting your right to future service.

Data scores of each person will not be disclosed but a summary of scores will be made available on request. The phone number to contact is xxxxxxx.

Each child will remain anonymous and will be identified with a number. The data will be stored securely in my computer. It will only be available to my supervisors, my examiners and myself and will not be made available to anyone else. It will be used for research purposes only.

Your contribution to this study is very important to help enhancing children's speech and language health in Iran.

Thank you very much for reading this information. If you agree to join the study, please sign one of the consent forms enclosed and return it to the head of nursery. Please keep the other consent form for your reference.

This research project was approved by Isfahan University of Medical Sciences as well as the School of Education, Communication and Language Sciences Research Ethics Committees.

Thank you

Yalda Kazemi

Appendix C-1

Persian consent form

به نام خدا

نام کودک:

من ، مادر/پدر (نام کودک)، با شرکت در مطالعه «کفایت تشخیصی مقیاسهای نمونه گیری زبانی در کودکان فارسی زبان ایرانی» که توسط خانم یلدا کاظمی و با استاد راهنمایی دکتر توماس کلی و دکتر هلن استرینگر انجام می شود، موافقم.

من برگه راهنمای شرکت کنندگان را کاملاً خوانده ام و فهمیده ام وخانم کاظمی به همه سوالاتی که داشته ام پاسخ داده اند. می دانم که اطلاعات مربوط به این مطالعه، در مکان امنی در دانشگاه اصفهان نگهداری شده و فقط برای مقاصد تحقیقاتی مورد استفاده قرار خواهد گرفت. همچنین می دانم که کودک من در هرگونه ارائه گزارش، فقط با یک شماره یا با یک نام مستعار شناسایی می شود. من همچنین می دانم که هر موقعی که بخواهم می توانم حتی بدون توضیح دادن برای کسی یا جریمه شدن، از ادامه مشارکت در بررسی انصراف دهم.

نام و امضاء

_____

تاریخ

_____

Appendix C-2

English-translated consent form

Child's name:

I, (Child's Name) mother/father, agree to participate in the study: **The diagnostic accuracy of language sample measures in Iranian Persian-speaking children** being conducted by Mrs. Yalda Kazemi under the supervision of Dr. Thomas Klee and Dr. Helen Stringer.

I have read and understood the information sheet for participants. Mrs. Kazemi has answered any questions that I have had. I understand that the data collected for this study will be stored in a secure location in the Speech and Language Sciences section at Newcastle University and that the data will be used only for research purposes. I understand that all participants will be given an identification number or a given name and that no participant will be referred to by name in any presentation of the study findings. I also understand that I can withdraw from the study at any time without explanation or penalty.

Name and Signature

Date

Appendix D-1

English-translated Basic information form

به نام خدا کد پرسشنامه -----

فرم اطلاعات پایه

نام و نام خانوادگی کودک: ----------

تاریخ تولد کودک: ---------- تاریخ امروز:----------

آدرس : ----------تلفن: ----------

کودک چندمین فرزند است؟ ---------- تعداد فرزندان خانواده : ----------

وزن هنگام تولد کودک : ----------

بررسی دو زبانگی

آیا کودک شما به طور مرتب در معرض زبان دیگری به غیر از زبان فارسی قرار دارد؟ بله☐ خیر☐

اگر پاسخ بله است چه زبانی ؟---------- توسط چه کسی؟ ----------

چند روز در هفته ؟------ چند ساعت در روز؟---------- از چه سنی؟---------

سلامت

آیا مادر در هنگام زایمان یا در دوران بارداری مشکلی داشته است؟ بله☐ خیر☐ اگر پاسخ بله است، لطفاً توضیح دهید:

آیا زایمان شما زودرس بوده است؟ (قبل از تاریخ معین شده) بله☐ خیر☐ اگر پاسخ بله است، چند هفته ---------

آیا کودک شما عفونت مزمن گوش داشته است؟ (5 بار یا بیشتر) بله☐ خیر☐

آیا نسبت به شنوایی کودکتان شک دارید؟ بله☐ خیر☐

آیا کودک شما سابقه بیماری، بستری شدن در بیمارستان یا ناتواناییهای مشخصی داشته است؟

بله ☐خیر☐

اگر پاسخ بله است، لطفاً توضیح دهید:

آیا شما یا عضوی از خانواده شما (برادر یا خواهر کودک، مادربزرگ یا پدر و .....) دارای اختلال رفتاری، نورولوژیکی، ناتواناییهای زبانی یا اختلال شنوایی هستند؟ بله☐ خیر☐

اگر پاسخ بله است، لطفاً نام ببرید.

اطلاعاتی راجع به مراقبین کودک

کودک شما با چه کسی زندگی می کند؟ ----------

چه کسی در نگهداری روزانه از کودک شما مشارکت دارد؟

مادر☐ آمادگی ☐ تعداد ساعات در هفته:

پدر ☐ اشخاصی غیر از والدین (مثلاً مادربزرگ، پرستار بچه)☐ تعداد ساعات در هفته:

موارد دیگر ☐ لطفاً توضیح دهید. (تعداد ساعات در هفته .....)

میزان تحصیلات والدین

پدر :---------- مادر : ----------

شغل والدین

پدر :---------- مادر : ----------

نام و امضاء--------- تاریخ ----------

Appendix D-2

English-translated Basic information form

Child's name ----------

DOB ---------- Today's date ----------

Address ---------- Phone ----------

Child's birth rank ---------- Number of children in family----------

Birth weight ----------

Bilingualism

Is the child being represented to other languages than Persian? Yes □ No □

If yes, what language (s)? ---------- With whom? ----------

How many days per week? ---------- How many hours per day? ----------What age? ----

Health

Does mother has had any disease during pregnancy or complications during delivery?

Yes □ No □ If yes, please explain:

Was it a premature delivery? Yes □ No □ If yes, how many days? ----------

Does the child have any history of Otitis Media more than 5 times? Yes □ No □

Do you worry about your child's hearing? Yes □ No □

Does your child have any history of hospitalization or specific disabilities?

Yes □ No □ If yes, please specify:

Is anybody at your family suffering from behaviour problems, neurological disorder, language impairment or hearing impairment? Yes □ No □

If yes, please specify:

Child's carers

Whom does the child live with? ----------

Who helps you to take care of your child?

Mother □ Father □

Nursery □ Relatives other than parents □ Others □

How many hours per week? ----------

Parents' educational level

Father: ---------- Mother: ----------

Parents' occupation

Father: ---------- Mother: ----------

Name and sign ---------- Date ----------

Appendix E

Transcription Example with SALT Analysis Results

$ Child, Examiner
+ Language: Persian
+ SubjectId: 56
+ Name: Hani ----
+ Gender: F
+ DOB: 12/20/1383
+ DOE: 4/30/1388
+ CA: 4;4
+ Ethnicity: Iranian
+ Context: Free play with mother
+ Subgroup: TDL
+ Examiner: Mother
+ Transcriber: Yalda
- 0:12
e Azin inha che ast?
e inha ra dorost kon.
C {IA}.
C in che ast?
e in che ast maman?
C tv.
e tv ra koja be\zar/im?
C in ham xx be\zar/im.
e na, in boxari/esh ast.
C boxari/esh.
e koja bayad be\zar/im?
C bayad be\zar/im pahlu/esh.
e be\zar/im bala?
C be\zar/im pahlu/esh?
e bale.
C in {IA}.
e in fil ast.
e heyvun/ha ra payin be\zar/im.
C heyvun/ha ham x?
e bale, esm/ha/eshun ra boland be\gu va be\zar.
C in fil ast.
e xob~
C in>
C in gorg.
e in che ast?
e in che ast?
C gorg.
e gorg ku?
C hamin ast ke raft/im mosaferat bara/am1 xarid/i.
e ahan, asb ast.
C asb.
C asb.
C asb/e3 chera ne\mi\vays/d?
C in arusak/ha ra.
C {xob} harf ne\mi\zan/an.

e xob xod/et ba/eshun harf be\zan.
e masalan in ra be\bin.
e che xoshgel ast.
e che lebas/e2 qashang/i1 dar/d.
e mu/ha/esh ra chikar kard/e1, azin?
C mu/ha/esh ra?
e han?
C ne\mi\dun/am.
e mu/ha/esh ra baft/e1.
C baft/e1.
e maman/esh bara/esh baft/e1.
e esm/esh ra che be\zar/im?
e dust na\dar/i?
e xob gerye mi\kon/d mi\gu/d man esm mi\xa/am.
e man esm mi\xa/am, (esm mi\xa/am).
e esm/am1 ra mi\xa/am.
C {eeee}.
e che be\zar/im esm/esh ra?
e esm/esh ra che be\zar/im?
C maman.
e bale.
C mi\xa/am in ra dorost/esh kon/am.
e bi\ya dorost/esh kon.
e {xob} dige chikar kon/im?
e hala yani sobh shod/e1.
C bale.
e xanum bidar mi\sho/d ba bache/ha/esh~
e chikar mi\kon/d?
e mi\ro/d chikar mi\kon/d?
C mi\ro/d ~
e che bo\xor/d?
C x.
e sobhune bo\xor/d.
C sobhune bo\xor/d.
e sobhune che dasht/e1 bash/im?
C sobhune?
C toxme.
e han?
e tu sobhune/emun che bash/d?
C {cough}.
e Azin, tu sobhune/emun che bash/d?
C toxmemorq, xiyar va goje.
C be\band in ra maman.
e xiyar va dige?
C goje.
e dige?
e nun va ~
C nun va panir.
e (asal va) kare va asal ham bo\xor/im.
C bo\xor/im.
e ke bozorg be\sho/im.
C be\sho/im.

e {xob} chayi ham ba/esh~
C bo\xor/im.
e bo\xor/im.
e bad chikar kon/im?
C bad?
C bad in otobus/emun (in) in mosafer ra savar mi\kon/im [WRONG-AGREEMENT]~
e xob~
C bad ba in mashin/emun mi\ya/im.
e mashin/emun mi\ya/im?
C mashin/emun ham inja ast dobare.
e {han} dadash ba otobus mi\ro/d kelas.
e chetor/i2 mi\ro/d?
e be\gu.
C intor/i1.
e chikar mi\kon/d?
e az xune birun mi\ya/d~
C az xune birun mi\ya/d~
e xob~
C mi\ro/d (in) savar/e2 in mi\sho/d~
e savar/e2 che mi\sho/d?
C in.
e esm/esh ra be\gu.
C [MISSING-WORD] otobus mi\sho/d~
e xob va koja mi\ro/d?
e be\gu ba/esh koja mi\ro/d masalan?
C maman unha mi\gu/an madrese x.
e masalan madrese ra ye ja/i2 be\zar.
C inja.
C maman dorost/eshun mi\kon/i?
C bezar inja.
e afarin, saat/et ham oftad/ø.
e dorost/esh kon.
C in shekl/i1 bud/ø ya1 in shekl/i1?
e azin in che ast?
C in ashpazxune.
e {xob} in ashpazxune ra payin be\chin.
C xob.
C mi\yoft/d maman.
e xob inja be\zar/im.
e ashpazxune ra koja mi\zar/i maman?
C yani inja ashpazxune/emun ra <be\zar/im>.
e <jelo> tv mi\sho/d mi\sho/d maman.
C na.
e inja be\zar.
e payin be\zar/i?
C are.
e bad, in che ast?
C in mal/e2 qaza.
e {xob} koja be\zar/im?
C inja.
e <afarin>.
C [MISSING-PREPOSITION] <ashpazxune> be\zar/am.

transcription Example
Analysis Set: C&I Verbal Utts

WORD AND MORPHEME SUMMARY

	Child			Examiner	
	Analysis Set	Total Utterances		Analysis Set	Total Utterances
MLU in Words	2.33	2.26		3.18	3.18
MLU in Morphemes	3.65	3.48		4.83	4.83
Brown's Stage	L IV/E V	Early IV		Post V	Post V
Expected Age Range (months)	31 - 50	28 - 45		41 - --	41 - --
Number Different Words	68	74		93	93
Number Total Words	126	140		245	245
Type Token Ratio	0.54	0.53		0.38	0.38
50 Utterances					
Number Different Words	65	63		73	73
Number Total Words	119	111		163	163
Type Token Ratio	0.55	0.57		0.45	0.45
Number of Bound Morphemes	71	76		127	127
Number of Maze Words	2	2		4	4
Number of Omitted Words	0	0		0	0
Number Omitted Bound Morphemes	0	0		0	0

NUMBER OF UTTERANCES BY UTTERANCE LENGTH
C&I Verbal Utts

Utterance Length in Words

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15+	Total
Child	0	25	10	8	4	4	1	1	1	0	0	0	0	0	0	0	54
Examiner	0	11	14	20	19	10	1	2	0	0	0	0	0	0	0	0	77

Utterance Length in Morphemes

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15+	Total
Child	0	17	5	8	9	3	4	1	2	3	1	1	0	0	0	0	54
Examiner	0	9	3	9	16	11	11	10	3	3	0	1	0	1	0	0	77

NUMBER OF UTTERANCES BY UTTERANCE LENGTH
Total Utterances

Utterance Length in Words

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15+	Total
Child	3	27	10	9	5	5	1	1	1	0	0	0	0	0	0	0	62
Examiner	0	11	14	20	19	10	1	2	0	0	0	0	0	0	0	0	77

Utterance Length in Morphemes

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15+	Total
Child	3	19	5	8	10	3	5	2	2	3	1	1	0	0	0	0	62
Examiner	0	9	3	9	16	11	11	10	3	3	0	1	0	1	0	0	77

transcription Example
Analysis Set: C&I Verbal Utts

UTTERANCE CODE TABLE
Table Expanded by Utterances
Total Utterances
Main body
1st Speaker

	Child	Examiner	
[MISSING-PREPOSITION]	1	0	
C [MISSING-PREPOSITION] <ashpazxune> be\zar/am.			
[MISSING-WORD]	1	0	
C [MISSING-WORD] otobus mi\sho/d~			
[WRONG-AGREEMENT]	1	0	
C bad in otobus/emun (in) in mosafer ra savar mi\kon/im [WRONG-AGREEMENT]~			

transcription Example
Analysis Set: C&I Verbal Utts

WORD ROOT TABLE
Table Expanded by Words
C&I Verbal Utts
Main body
1st Speaker

	Child		Examiner
	Total	Expanded	Total
ARE	1		0
ARUSAK	1		0
ARUSAK/HA		1	
ASB	3		1
ASB		2	
ASB/E3		1	
ASHPAZXUNE	3		2
ASHPAZXUNE		2	
ASHPAZXUNE/EMUN		1	
AST	4		10
AZ	1		1
BA	1		5
BAD	3		2
BAFT	1		2
BAFT/E1		1	
BALE	1		3
BAND	1		0
BE\BAND		1	
BARA	1		1
BARA/AM1		1	
BAYAD	1		1
BIRUN	1		1
BOXARI	1		1
BOXARI/ESH		1	
BUD	1		0
BUD/ø		1	
CHE	1		16
CHERA	1		0
DOBARE	1		0
DOROST	2		3
DOROST/ESH		1	
DOROST/ESHUN		1	
DUN	1		0
NE\MI\DUN/AM		1	
FIL	1		1
GOJE	2		0
GORG	2		1
HAM	1		3
HAMIN	1		0
HARF	1		1

IN	16			9
INJA	5			2
INTOR	1			0
INTOR/I1		1		
KE	1			1
KON	3			9
KON/AM		1		
MI\KON/I		1		
MI\KON/IM		1		
MAL	1			0
MAL/E2		1		
MAMAN	4			4
MASHIN	2			1
MASHIN/EMUN		2		
MOSAFER	1			0
MOSAFERAT	1			0
MU	1			2
MU/HA/ESH		1		
NA	1			1
NUN	1			1
OTOBUS	2			1
OTOBUS		1		
OTOBUS/EMUN		1		
PAHLU	2			0
PAHLU/ESH		2		
PANIR	1			0
QAZA	1			0
RA	6			15
RAFT	1			0
RAFT/IM		1		
RO	2			5
MI\RO/D		2		
SAVAR	2			1
SAVAR		1		
SAVAR/E2		1		
SHEKL	2			0
SHEKL/I1		2		
SHO	3			5
BE\SHO/IM		1		
MI\SHO/D		2		
SOBHUNE	2			4
TØXME	1			0
TØXMEMORQ	1			0
TV	1			2
VA	2			5
VAYS	1			0
NE\MI\VAYS/D		1		
XA	1			3

MI\XA/AM		1		
XARID	1			0
XARID/I		1		
XIYAR	1			1
XOB	1			7
XOR	3			4
BO\XOR/D		1		
BO\XOR/IM		2		
XUNE	1			1
YA	2			3
MI\YA/D		1		
MI\YA/IM		1		
YA1	1			0
YANI	1			1
YOFT	1			0
MI\YOFT/D		1		
ZAN	1			1
NE\MI\ZAN/AN		1		
ZAR	5			14
BE\ZAR		1		
BE\ZAR/AM		1		
BE\ZAR/IM		3		

transcription Example
Analysis Set: C&I Verbal Utts

BOUND MORPHEME TABLE
Table Expanded by Words
C&I Verbal Utts
Main body
1st Speaker

	Child		Examiner
	Total	Expanded	Total
/AM	4		3
BE\ZAR/AM		1	
KON/AM		1	
MI\XA/AM		1	
NE\MI\DUN/AM		1	
/AM1	1		1
BARA/AM1		1	
/AN	1		0
NE\MI\ZAN/AN		1	
/D	8		20
BO\XOR/D		1	
MI\RO/D		2	
MI\SHO/D		2	
MI\YA/D		1	
MI\YOFT/D		1	
NE\MI\VAYS/D		1	
/E1	1		5
BAFT/E1		1	
/E2	2		2
MAL/E2		1	
SAVAR/E2		1	
/E3	1		0
ASB/E3		1	
/EMUN	4		3
ASHPAZXUNE/EMUN		1	
MASHIN/EMUN		2	
OTOBUS/EMUN		1	
/ESH	5		14
BOXARI/ESH		1	
DOROST/ESH		1	
MU/HA/ESH		1	
PAHLU/ESH		2	
/ESHUN	1		2
DOROST/ESHUN		1	
/HA	2		5
ARUSAK/HA		1	
MU/HA/ESH		1	
/HA/ESH	1		3
MU/HA/ESH		1	
/I	2		3

MI\KON/I		1		
XARID/I		1		
/I1	3			1
INTOR/I1		1		
SHEKL/I1		2		
/IM	9			16
BE\SHO/IM		1		
BE\ZAR/IM		3		
BO\XOR/IM		2		
MI\KON/IM		1		
MI\YA/IM		1		
RAFT/IM		1		
/ø	1			1
BUD/ø		1		
BE\	7			21
BE\BAND		1		
BE\SHO/IM		1		
BE\ZAR		1		
BE\ZAR/AM		1		
BE\ZAR/IM		3		
BO\	3			4
BO\XOR/D		1		
BO\XOR/IM		2		
MI\	13			20
MI\KON/I		1		
MI\KON/IM		1		
MI\RO/D		2		
MI\SHO/D		2		
MI\XA/AM		1		
MI\YA/D		1		
MI\YA/IM		1		
MI\YOFT/D		1		
NE\MI\DUN/AM		1		
NE\MI\VAYS/D		1		
NE\MI\ZAN/AN		1		
NE\	3			0
NE\MI\DUN/AM		1		
NE\MI\VAYS/D		1		
NE\MI\ZAN/AN		1		
NE\MI\	3			0
NE\MI\DUN/AM		1		
NE\MI\VAYS/D		1		
NE\MI\ZAN/AN		1		

Appendix F

Operational definitions of General Language Sample Measures (LSMs) and Errors

General LSMs

Number of Total Complete and Intelligible (C&I) Utterances: total number of utterances in a fixed length sample (in 20 minutes) that are not partially transcribed, i.e. included missing parts (coded by X character), or abundant utterances (coded by >).

Example of included utterances: *be\ro/d tu koja?*

Example of excluded utterances: *in x/esh ast.* (Unintelligible)

man xers ra> (abundant)

Mean Length of Utterances C&I in morphemes or words: the average count of the morphemes or words in a sample of fixed length (in 20 minutes) based on C&I utterances.

Number of Total Words C&I: Total number of words either inflected or bared roots in a fixed length sample (in 20 minutes) based on C&I utterances.

Number of Different Words C&I: Total number of different uninflected words in a fixed length sample (in 20 minutes) based on C&I utterances.

Total Number of One-Word Utterances: Total number of one- word utterances including either single-morpheme utterances (e.g. nouns) or single-word utterances (e.g. an inflected verb).

Examples: *mi\xor/d* (inflected verb, single-word utterance).

biskuit (noun, single-morpheme utterance).

Mean Length of Utterances – excluding one-word utterances, either in morphemes or words: same as MLU above but excluding one-word utterances.

Number of Total Words - excluding one-word utterances: same as NTW above but excluding one-word utterances.

Number of Different Words - excluding one-word utterances: same as NDW above but excluding one-word utterances.

Total Number of Verbal Morphemes: number of total verb affixes including prefixes (*mi*, *na*, *be*) and suffixes (verbal markers of tense and number).

Intelligibility: Percentage of intelligible utterances in a fixed length sample (in 20 minutes).

General Errors

Clitics Errors: Total number of errors in all Persian clitics listed in chapter 1.

Verb Inflectional Errors (Finite Verb Morphology): Total number of all errors, either missing or wrong usage, on verbal morphemes.

Example: *na\dar/*am* (missing verb number suffix)

be\xa/am[wrong prefix: *be* instead of *mi*]

Semantic Errors: Total number of all semantic errors as listed in below.

Total Errors: Total number of all errors, either missing or wrong usage.

Total Number of Grammatical Utterances: Number of total C&I utterances without error codes.

Total Number of Ungrammatical Utterances: Number of total C&I utterances with error codes.

Percent of Grammaticality: Percentage of Total Number of Grammatical Utterances.

Percent of Ungrammaticality: Percentage of Total Number of Grammatical Utterances.

Missing verb markers: Total number of missing verb suffixes, i.e. tense and number.

Missing prepositions: Total number of missing prepositions, e.g. *az, ba, tu*.

Missing conjunctions: Total number of missing conjunctions, e.g. *va, ama, ta*.

Missing verbs: Total number of omitted verbs.

Wrong agreement: disagreement between subject and verb in terms of number either in the same utterance or in relation to a reference in the previous utterance.

Example: *man xast*[wrong agreement].

Wrong word order: syntactically wrong alignment of the words in the sentence which affects meaning.

Example: *taxt in ast*[wrong word order].

Correct form is: *in taxt ast*.

Nonsense string of words: A syntactical word order with no meaningful content.

Example: inja mi\xa/d tu xab ast.

Meaning: here wants in bed is.

Wrong responses to the questions: meaningfully wrong answer to the antecedent question.

REFERENCES

Abbot-Smith, K., Lieven, E., & Tomasello, M. (2001). What preschool children do and do not do with ungrammatical word orders. *Cognitive Development, 16*(2), 679-692.

Akhlaghi, F. (2008). Three modal Verbs in Contemporary Persian. *Name-ye Farhangestan, 3*(3).

Akhtar, N. (1999). Acquiring basic word order: evidence for data-driven learning of syntactic structure. *Journal of Child Language, 26*(02), 339-356.

Aljenaie, K., & Farghal, M. (2009). Comprehension of three word orders in Kuwaiti Arabic child language. *Language Sciences, 31*(4), 494-506. doi: 10.1016/j.langsci.2007.10.003

Allen, D. V., Bliss, L. S., & Timmons, J. (1981). Language evaluation: science or art? *Journal of Speech and Hearing Disorders, 46*(1), 66-68.

American Speech-Language-Hearing Association. (2005). *Principles for speech-language pathologists serving persons with mental retardation/developmental disabilities* [Technical Report]. Retrieved from http://www.asha.org/policy.

Amini, A. (1997). On stress in Persian. *Toronto Working Papers in Linguistics (Proceedings of the Montréal-Ottawa-Toronto Phonology Workshop), 16*(1), 1-20.

Aram, D. M., Morris, R., & Hall, N. E. (1993). Clinical and research congruence in identifying children with specific language impairment. *Journal of Speech and Hearing Research, 36*(3), 580-591.

Assessing scientific admissibility and merit of published articles: Critical appraisal form. (2012). Retrieved 22/03/2012, from http://peds.stanford.edu/Tools/handouts_forms_guides.html

Babarczy, A. (2006). Negation and word order in Hungarian child language. *Lingua, 116*(3), 377-392. doi: 10.1016/j.lingua.2004.08.010

Bahar, M. T. (1996). *Sabk Shenāsi (The Development of Persian Language)* (7th ed.). Tehran: Amir Kabir.

Bates, E., Friederici, A., & Wulfeck, B. (1987). Grammatical morphology in aphasia: evidence from three languages. *Cortex, 23*(4), 545-574.

Bedore, L. M., & Leonard, L. B. (1998). Specific language impairment and grammatical morphology: a discriminant function analysis. *Journal of Speech, Language, and Hearing Research, 41*, 1185–1192.

Bishop, D. V. M. (1992). The underlying nature of specific language impairment. *Journal of Child Psychology and Psychiatry, 33*(1), 3-66. doi: 10.1111/j.1469-7610.1992.tb00858.x

Bishop, D. V. M., & McDonald, D. (2009). Identifying language impairment in children: combining language test scores with parental report. *International Journal of Communication Disorders, 44*(5), 1-16.

Blake, B. J. (2001). *Case.* Cambridge: Cambridge University Press.

Bland, J. M., & Altman, D. G. (2011). Correlation in restricted ranges of data. *British Medical Journal, 342*(d556). doi: 10.1136/bmj.d556

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Fixed-effect versus random-effects models. In M. Borenstein, L. V. Hedges, J. P. T. Higgins & H. R. Rothstein (Eds.), *Introduction to Meta-Analysis*. West Sussex: John Wiley & Sons Ltd.

Bornstein, M. H., Painter, K. M., & Park, J. (2002). Naturalistic language sampling in typically developing children. *Journal of Child Language, 29*(03), 687-699. doi: doi:10.1017/S030500090200524X

Brown, R. (1973). *A First Language: The Early Stages*. Massachusetts: Harvard University Press.

Carstairs-McCarthy, A. (1992). *Current morphology*, London: Routledge

Clahsen, H., Bartke, S., & Göllner, S. (1997). Formal features in impaired grammars: A comparison of English and German SLI children. *Journal of Neurolinguistics, 10*(2–3), 151-171. doi: 10.1016/s0911-6044(97)00006-7

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Comrie, B. (1989). *Language universals and linguistic typology: syntax and morphology.* University of Chicago Press.

Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impairment (SLI). *Journal of Speech Language and Hearing Research, 46*(5), 1029-1037. doi: 10.1044/1092-4388(2003/082)

Jones, M., & Conti-Ramsden, G. (1997). A comparison of verb use in children with SLI and their younger siblings. *First Language, 17*(51), 165-184. doi: 10.1177/014272379701705108

Conti-Ramsden, G., & Windfuhr, K. (2002). Productivity with word order and morphology: a comparative look at children with SLI and children with normal

language abilities. *International Journal of Language & Communication Disorders, 37*(1), 17-30. doi: doi:10.1080/13682820110089380

Costanza-Smith, A. (2010). The clinical utility of language samples. *Perspectives on Language Learning and Education, 17*(1), 9-15. doi: 10.1044/lle17.1.9

Creswell, J. W., Shope, R., Clark, V. L. P., & Green, D. O. (2006). How interpretive qualitative research extends mixed methods research. *Research in the Schools, 13(1),* 1-11.

Damico, J. S., & Simmons-Mackie, N. N. (2003). Qualitative research and speech-language pathology: a tutorial for the clinical realm. *American Journal of Speech-Language Pathology, 12*, 131-143. doi:10.1044/1058-0360(2003/060)

Dempsey, L., & Skarakis-Doyle, E. (2010). Developmental language impairment through the lens of the ICF: An integrated account of children's functioning. *Journal of communication disorders, 43*(5), 424-437. doi: 10.1016/j.jcomdis.2010.05.004

Deville, W., Buntinx, F., Bouter, L., Montori, V., de Vet, H., van der Windt, D., & Bezemer, P. (2002). Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Medical Research Methodology, 2*(1), 9.

Diagnostic Study Appraisal Worksheet. (2010). University of Oxford: Centre for Evidence-Based Medicine. Retrived from http://www.cebm.net/index.aspx?o=1157

Dollaghan, C. A. (2004). Taxometric analyses of specific language impairment in 3- and 4-year-old children. *Journal of Speech Language and Hearing Research, 47*(2), 464-475. doi: 10.1044/1092-4388(2004/037)

Dollaghan, C. A. (2007). *The handbook for evidence-based practice in communication disorders*. Maryland: Paul H. Brooks Publishing Company.

Dollaghan, C. A., & Horner, E. A. (2011). Bilingual language assessment: a meta-analysis of diagnostic accuracy. *Journal of Speech Language and Hearing Research, 54*(4), 1077-1088. doi: 10.1044/1092-4388(2010/10-0093)

Dressler, W. U. (2010). A Typological approach to first language acquisition. In M. Kail & M. Hickmann (Eds.), *Language Acquisition across Linguistic and Cognitive Systems* (pp. 109-124). Amsterdam; Philadelphia: John Benjamins Publishing Company [NetLibrary version]. Retrieved from http://NCL.eblib.com/patron/FullRecord.aspx?p=623308.

Dressler, W. U. (2012). On the acquisition of inflectional morphology: introduction. *Morphology*, 1-8. doi 10.1007/s11525-011-9198-1

Dromi, E., Leonard, L. B., Adam, G., & Zadunaisky-Ehrlich, S. (1999). Verb agreement morphology in hebrew-speaking children with specific language impairment. *Journal of Speech Language and Hearing Research, 42*(6), 1414-1431.

Dunn, M., Flax, J., Sliwinski, M., & Aram, D. (1996). The use of spontaneous language measures as criteria for identifying children with specific language impairment: an attempt to reconcile clinical and research incongruence. *Journal of Speech and Hearing Research 39*, 643-654.

Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Paediatric Psychology, 34*(9), 917-928. doi: 10.1093/jpepsy/jsp004

Eisenberg, S. L., Fersko, T. M., & Lundgren, C. (2001). The use of mlu for identifying language impairment in preschool children: a review. *American Journal of Speech-Language Pathology, 10*, 323-342. doi:10.1044/1058-0360(2001/028)

Eisenberg, S. L., & Guo, L. Y. (2012). Differentiating children with and without language impairment based on grammaticality. *Language, Speech and Hearing Services in Schools*, 0161-1461_2012_0111-0089. doi: 10.1044/0161-1461(2012/11-0089)

Eisenberg, S. L., Guo, L. Y., & Germezia, M. (2012). How grammatical are 3-year-olds? *Language, Speech, and Hearing Services in Schools, 43*(1), 36-52. doi: 10.1044/0161-1461(2011/10-0093)

Eslami, M., & Alizadeh-Lemjiri, S. (2009). Inflectional structure of the word in Persian. *Persian Language and Literature, Journal of the Tabriz Faculty of Literature and Human Sciences, 52* (211), 1-18

Estigarribia, B., Martin, G. E., & Roberts, J. E. (2012). Cognitive, environmental, and linguistic predictors of syntax in fragile x syndrome and down syndrome. *Journal of Speech Language and Hearing Research*, 1092-4388_2012_1010-0153. doi: 10.1044/1092-4388(2012/10-0153)

Evans, J. L., & Craig, H. K. (1992). Language sample collection and analysis: interview compared to free play assessment contexts. *Journal of Speech and Hearing Research, 35*(2), 343-353.

Fahim, M. (n. d.) *First language learning.* Paper presented at the Third Iranian Conference on Linguistics.

Fey, M. E. (1986). *Language intervention with young children.* San Diego, California: College-Hill Press.

Fletcher, P., & Garman, M. (1988). LARSPing by numbers. *British Journal of Disorders of Communication, 23(3),* 309-321.

Foroodi-Nejad, F. (2011). *Towards the identification of linguistic characteristics of specific language impairment in Persian.* (Doctoral dissertation), Retrived from University of Alberta Library. http://hdl.handle.net/10048/2391

Gass, S. M. (1989). *Linguistic perspectives on second language acquisition* (2nd ed.). Cambridge: Cambridge University Press.

Gavin, W. J., Klee, T., & Membrino, I. (1993). Differentiating specific language impairment from normal language development using grammatical analysis. *Clinical Linguistics & Phonetics, 7*(3), 191-206.

Girolametto, L., & Weitzman, E. (2002). Responsiveness of child care providers in interactions with toddlers and preschoolers. *Language, Speech, and Hearing Services in Schools, 33*(4), 268-281. doi: 10.1044/0161-1461(2002/022)

Gladfelter, A., & Leonard, L. B. (2012). Alternative tense and agreement morpheme measures for assessing grammatical deficits during the preschool period. *Journal of Speech Language and Hearing Research*, 1092-4388_2012_1012-0100. doi: 10.1044/1092-4388(2012/12-0100)

Glas, A. S., Lijmer, J. G., Prins, M. H., Bonsel, G. J., & Bossuyt, P. M. (2003). The diagnostic odds ratio: a single indicator of test performance. *Journal of Clinical Epidemiology, 56*(11), 1129-1135.

Glascoe, F. P. (1991). Can clinical judgment detect children with speech-language problems?. *Pediatrics, 87*(3), 317-322.

Golpour, L., Nilipour, R., & Roshan, B. (2007). A comparison between morphological and syntactic features of 4 to 5 years old in education severe to profound hearing impaired and normal children. *Audiology, 15*(2), 23-29.

Greenhalgh, T. (2010). *How to Read a Paper: The Basics of Evidence-Based Medicine*, London: BMJ Books.

Greenslade, K. J., Plante, E., & Vance, R. (2009). The diagnostic accuracy and construct validity of the structured photographic expressive language test--preschool: second edition. *Language, Speech, and Hearing Services in Schools, 40*(2), 150-160. doi: 10.1044/0161-1461(2008/07-0049)

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*(1), 29-36.

Hasanati, F., Agha Rasouli, Z., Mahmoudi Bakhtiyari, B., & Kamali, M. (2011). Sentence repetition test for measurement of grammatical development in Farsi speaking children. *Audiology, 20*(1), 73-81.

Hasanzadeh, S., & Minaei, A. (2000). *Norming Test Of Language Development-Primary (TOLD-P) in Persian*. Tehran: Research Centre for Exceptional Children.

Haynes, R. B., Sackett, D. L., Guyatt, G., & Tugwell, P. (2006). *Clinical epidemiology: how to do clinical practice research* (3rd ed.). Philadelphia: Lippincott Williams and Wilkins.

Haynes, W. O., Purcell, E., & Haynes, M. D. (1979). A pragmatic aspect of language sampling. *Language, Speech, and Hearing Services in Schools, 10*(2), 104-110.

Heilmann, J. J., Miller, J. F., & Nockerts, A. (2010). Using language sample databases. *Language, Speech, and Hearing Services in Schools, 41*(1), 84-95. doi: 10.1044/0161-1461(2009/08-0075)

Hewitt, L. E., Hammer, C. S., Yont, K. M., & Tomblin, J. B. (2005). Language sampling for kindergarten children with and without SLI: mean length of utterance, IPSYN, and NDW. *Journal of Communication Disorders, 38* (3), 197-213.

Heydari, S., Torabi Nezhad, F., Agha Rasouli, Z., & Hoseyni, F. (2011). Development of speech intelligibility measurement test for 3 to 5 years old normal children. *Audiology, 20*(1), 47-53.

Hoff, E. (2011). *Research Methods in Child Language: A Practical Guide*. West Sussex: John Wiley & Sons.

Howe, T. J. (2008). The ICF Contextual Factors related to speech-language pathology. *International Journal of Speech-Language Pathology, 10*(1-2), 27-37. doi: doi:10.1080/14417040701774824

Howick, J., Chalmers, I., Glasziou, P., Greenhalgh, T., Heneghan, C., Liberati, A., . . . Hodgkinson, M. (2011). The Oxford 2011 Levels of Evidence. Retrieved 22/03/2012, from http://www.cebm.net/index.aspx?o=5653

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277-1288. doi: 10.1177/1049732305276687

Huang, C.-t. J., & Li, Y.-h. A. (1996). *New horizons in Chinese linguistics*. Netherland: Kluwer Academic Publishers.

Hutchins, T. L., Brannick, M., Bryant, J. B., & Silliman, E. R. (2005). Methods for controlling amount of talk: Difficulties, considerations and recommendations. *First Language, 25*(3), 347-363. doi: 10.1177/0142723705056376

Hux, K., Morris-Friehe, M., & Sanger, D. D. (1993). Language sampling practices: a survey of nine states. *Language, Speech, and Hearing Services in Schools, 24*(2), 84-91.

Jabbar Baghcheban, Iran's sign language pioneer, remembered (2007). Retrived from *www.payvand.com/news/07/nov/1248.html* 17/02/2012.

Jalilevand, N., Ebrahimipur, M., & Purqarib, J. (2012). Mean length of utterance and grammatical morphemes in the speech of two Farsi-speaking children. *Audiology*, 21(2), 96-108.

Joffe, V., & Pring, T. (2008). Children with phonological problems: a survey of clinical practice. *International Journal of Language & Communication Disorders, 43*(2), 154-164. doi: 10.1080/13682820701660259

Kalbasi, I. (2008). *The derivational structure of word in Modern Persian*. Tehran: Institute for humanities and cultural studies.

Karimi, S. (1996). Case and specificity: Persian ra revisited. *Linguistic Analysis, 26*(3/4), 173-194.

Karimi, S. (2003). *Word order and scrambling*. Cornwall: Wiley-Blackwell

Kazemi, Y., Afsharianzadeh, E., Mirzaei, B., Baghbani, M., Sademirinejad, M., Gheleyempour, L., . . . Yabandeh, N. (2007). Children's communication checklist : the study of persian children. *Iranian Journal of Research in Rehabilitation Sciences, 2*(3), 1-5.

Kazemi, Y., & Derakhshandeh, F. (2007). Exploring the normal range of scores of Persian-speaking children in Robbins-Klee oral/speech motor control protocol. *Quarterly Journal of Rehabilitation, 8*(3), 50-56.

Kazemi, Y., Ghasisin, L., Rezaei, P., Samadi, A., Sharifi, A. (2007). *A primary validity and reliability study on Test of Language Development-Intermediate (TOLD-I):3 in Persian children.* Paper presented at the 27 World Congress of the International Association of Logopedics and Phoniatrics, Copenhagen: Technical University of Denmark.

Kazemi, Y., Nematzadeh, S., Hajian, T., Heidari, M., Daneshpajouh, T., & Mirmoeini, M. (2008). Finding the validity and reliability coefficient of MacArthur-Bates Communicative Development Inventory (CDI-Infant form) in Persian. *Journal of Research in Rehabilitation Sciences, 3*(4), 45-51.

Kazemi, Y., Nockerts, A., Klee, T., Stringer, H., & Miller, J. F. (2012). *Challenges in transcribing and analysing Persian child language samples using SALT*. Paper

presented at the Symposium on Research in Child Language Disorders, Wisconsin, Madison.

Kazemi, Y., Taheri, A., Kianfar, F., Shafiei, M., Eslamifard, R., Pirmoradian, M., … Nourian, F. (2012). Mean Length of Utterance (MLU) in typically-developing 2:6-5:6 year-old Farsi-speaking children in Iran. *Journal of Research in Rehabilitation Sciences, 8(5)*.

Kemp, K., & Klee, T. (1997). Clinical language sampling practices: Results of a survey of speech-language pathologists in the United States. *Child Language Teaching and Therapy, 13*(2), 161-176.

Klee, T. (1991). Measuring children's conversational language. In S. F. Warren & J. Reichle (Eds.), *Causes and effects in communication and language intervention* (pp. 315-330). Baltimore: P.H. Brookes.

Klee, T. (1992). Developmental and diagnostic characteristics of quantitative measures of children's language production. *Topics in Language disorders, 12*(2), 28-41.

Klee, T., & Fitzgerald, M. D. (1985). The relation between grammatical development and mean length of utterance in morphemes. *Journal of Child Language, 12*(2), 251-269.

Klee, T., Gavin, W. J., & Stokes, S. F. (2007). Utterance length and lexical diversity in american- and british-english speaking children: what is the evidence for a clinical marker of sli? In R. Paul (Ed.), L*anguage disorders from a developmental perspective: essays in honour of Robin S. Chapman* (pp. 103-140). Mahwah, NJ: Lawrence Erlbaum

Klee, T., Stokes, S. F., Wong, A. M. Y., Fletcher, P., & Gavin, W. J. (2004). Utterance length and lexical diversity in cantonese-speaking children with and without specific language impairment. *Journal of Speech, Language and Hearing Research, 47*(6), 1396-1410. doi: 10.1044/1092-4388(2004/104)

Klee, T., Wong, A. M.-Y., Stokes, S. F., Fletcher, P., & Leonard, L. B. (2009). Assessing Cantonese-speaking children with language difficulties from the perspective of Evidence-based Practice: current practice and future directions. In S.-p. Law, B. Weekes & A. M.-Y. Wong (Eds.), *Language disorders in speakers of chinese* (pp. 89-111). Bristol: Multilingual Matters.

Klein, H. B., Moses, N., & Jean-Baptiste, R. (2010). Influence of context on the production of complex sentences by typically developing children. *Language, Speech and Hearing Services in Schools, 41*(3), 289-302. doi: 10.1044/0161-1461(2009/08-0080)

Kunnari, S., Savinainen-Makkonen, T., Leonard, L. B., MÄKinen, L., Tolonen, A.-K., Luotonen, M., & Leinonen, E. (2011). Children with specific language impairment in finnish: the use of tense and agreement inflections. *Journal of Child Language, 38*(05), 999-1027. doi: doi:10.1017/S0305000910000528

Lahey, M. (1988). *Language disorders and language development*. New York: Macmillan.

Laiacona, M., & Caramazza, A. (2004). The noun/verb dissociation in language production: varieties of causes. *Cognitive Neuropsychology, 21*(2-4), 103-123. doi: 10.1080/02643290342000311

Law, J. (2003). The close association between classification and intervention for children with primary language impairments. In L. Verhoeven & H. van Balkom (Eds.), *Classification of developmental language disorders: theoretical issues and clinical implications*: Mahwah, N.J. : Lawrence Erlbaum Associates.

Law, J., Garrett, Z., & Nye, C. (2003). *Speech and language therapy interventions for children with primary speech and language delay or disorder*. (Publication no. 10.1002/14651858.CD004110). The Cochrane Collaboration: Published by John Wiley & Sons, Ltd. doi: 10.1002/14651858.cd004110

Leonard, L. B. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Leonard, L. B. (2009). Is expressive language disorder an accurate diagnostic category? *American Journal of Speech and Language Pathology, 18*(2), 115-123. doi: 10.1044/1058-0360(2008/08-0064)

Leonard, L. B., Caselli, M. C., Bortolini, U., McGregor, K. K., & Sabbadini, L. (1992). Morphological deficits in children with specific language impairment: the status of features in the underlying grammar. *Language Acquisition, 2*(2), 151-179. doi: 10.1207/s15327817la0202_2

Leonard, L. B., Miller, C., & Gerber, E. (1999). Grammatical morphology and the lexicon in children with specific language impairment. *Journal of Speech, Language and Hearing Research, 42*(3), 678-689.

Likelihood ratios. (2012). Retrieved 10/10/2012, from http://www.cebm.net/index.aspx?o=1043

Lof, G. L. (2011). Science-based practice and the speech-language pathologist. *International Journal of Speech-Language Pathology, 13*(3), 189-196. doi:10.3109/17549507.2011.528801

Lyons, R., Byrne, M., Corry, T., Lalor, L., Ruane, H., Shanahan, R., & McGinty, C. (2008). An examination of how speech and language therapists assess and diagnose children with specific language impairment in Ireland. *International Journal of Speech-Language Pathology, 10,* 425-437. doi: doi:10.1080/17549500802422569

MacWhinney, B. (1994). Computational tools for analysing language. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of child language research*. London: Blackwells.

Mahmoudi Bakhtiyari, B., Soraya, M., Badiee, Z., Kazemi, Y., & Soleimani, B. (2012). The size of expressive lexicon in 18-to-36-month-old children raised in farsi-speaking families: a comparative study. *Journal of Neuroscience and Behavioural Health, 4*(4), 33-36.

Mainela-Arnold, E., Evans, J. L., & Coady, J. A. (2010). Explaining lexical-semantic deficits in specific language impairment: the role of phonological similarity, phonological working memory, and lexical competition. *Journal of Speech, Language and Hearing Research, 53*(6), 1742-1756. doi: 10.1044/1092-4388(2010/08-0198)

Maleki Shahmahmood, T., Soleymani, Z., & Faghihzade, S. (2011). The study of language performances of Persian children with specific language impariment. *Audiology, 20*(2), 11-21.

Maleki Shahmahmood, T., Soleymani, Z., & Jalaei, S. H. (2009). A comparison study in Test of Language Development (TOLD) and speech samples between children with specific language impairment and their MLU matched group. *Modern Rehabilitation, 2*(3,4), 5-5.

Malvern, D., Richards, B., Chipere, N., & Duran, P. (2004). *Lexical diversity and language development : quantification and assessment* [NetLibrary version]. Retrieved from http://NCL.eblib.com/patron/FullRecord.aspx?p=358640

Marinis, T. (2011). On the nature and cause of Specific Language Impairment: A view from sentence processing and infant research. *Lingua, 121*(3), 463-475. doi: 10.1016/j.lingua.2010.10.010

Marshall, J., Goldbart, J., & Phillips, J. (2007). Parents' and speech and language therapists' explanatory models of language development, language delay and intervention. *International Journal of Language & Communication Disorders, 42*(5), 533-555. doi: 10.1080/13682820601053753

McCauley, R. J., & Swisher, L. (1984). Psychometric review of language and articulation tests for preschool children. *Journal of Speech and Hearing Disorders, 49*(1), 34-42.

Megerdoomian, K. (2012). The status of the nominal in Persian complex predicates. *Natural Language & Linguistic Theory, 30*(1), 179-216. doi: 10.1007/s11049-011-9146-0

Megerdoomian, K., & MansouriRad, H. (2000). Acquisition of Persian resources: Corpora and dictionary development in the Shiraz project *NMSU, CRL, Memoranda in Computer and Cognitive Science (MCCS-00-323).*

Mehdipour, N., Shirazi, T. S., & Nematzadeh, S. *Development of expressive vocabulary and sentences in 18 - 24 months old Persian-speaking children.* Paper presented at the First Conference of Iranain Linguistics Society, Tehran.

Meshkato-Dini, M. (1996). *Development of grammatical components until five years old*. Mashad: Ferdowsi University.

Meshkato-Dini, M. (2001). Early speech and grammar considering Persian-speaking children. *Mashad School of Literature and Human Sciences Journal, 34*(3-4), 447-478.

Meshkato-Dini, M. (2004). Inflectional morphology and structure links in child developing language. *Journal of Linguistics, 19*(2); 10-26.

Meshkato-Dini, M. (2008). *Persian Grammar, the lexical categories and merge* (3rd ed.). Tehran: SAMT.

Miller, J. F. (1981). *Assessing language production in children*. Maryland: University Park Press.

Miller, J. F., & Fletcher, P. (2005). Developmental theory and language disorders, background issues. In P. Fletcher & J. F. Miller (Eds.), *Developmental theory and language disorders*. Philadelphia: John Benjamin Publishing Company.

Miller, J. F., & Iglesias, A. (2008). *Systematic Analysis of Language Transcripts (SALT), Version 2008*. Muscoda, WI SALT Software LLC.

Miller, J. F., & Iglesias, A. (2012). Systematic Analysis of Language Transcripts (SALT), Research Version 2012.0.1: SALT Software, LLC.

Modarres Zadeh, A. (2010). A descriptive analysis of responses given to persian picture naming test by 2.5 year-old children. *Journal of Research in Rehabilitation Sciences, 6*(2), 1-9.

Moghaddam, M. D. (2001). Word order typology of Iranian languages. *The International Journal of Humanities (The Journal of Humanities), 8*(2), 17-24.

Mohammad Ebrahimi Jahromi, Z., & Haghshenas, A. M. (2004). The structure of VP core in standard Persian. *Journal of Literature and Language of Faculty of Letters and Humanities-Shahid Bahonar University of Kerman, 12*(15), 153-177.

Moyle, M. J., Ellis Weismer, S., Evans, J. L., & Lindstrom, M. J. (2007). Longitudinal relationships between lexical and grammatical development in typical and late-talking children. *Journal of Speech, Language and Hearing Research, 50*(2), 508-528. doi: 10.1044/1092-4388(2007/035)

Naderi, E., & Seifenaraghi, M. (1993). *Persian language development from birth to two years old.* Paper presented at the First Iranian Conference in Theoretical and Applied Linguistics, Tehran.

Narasimhan, B., & Dimroth, C. (2008). Word order and information status in child language. *Cognition, 107*(1), 317-329. doi: 10.1016/j.cognition.2007.07.010

Natel-Khanlari, P. (1994). *Persian grammar*. Tehran: Tous Publishing. pp. 162, 168.

Neeleman, A., & Szendroi, K. (2005). *Pro drop and pronouns.* Paper presented at the The 24th West Coast Conference on Formal Linguistics, Somerville, MA.

Nelson, H. D., Nygren, P., Walker, M., & Panoscha, R. (2006). Screening for speech and language delay in preschool children: systematic evidence review for the US preventive services task force. *Pediatrics, 117*(2), e298-e319. doi: 10.1542/peds.2005-1467

Nilipour, R. (2002). Emerging issues in speech therapy in Iran. *Folia Phoniatrica et Logopedica, 54* :65-68 (doi: 10.1159/000057916)

Nilipour, R., & Paradis, M. (1995). Breakdown of functional categories in three Farsi-English bilingual aphasic patients. In M. Paradis (Ed.), *Aspects of bilingual aphasia* (pp. 123-138). Oxford: Pergamon Press.

Nilipour, R., & Raghibdoust, S. (2001). Manifestations of aphasia in Persian. *Journal of Neurolinguistics, 14*(2-4), 209-230.

O'Shannessy, C. (2011). Competition between word order and case-marking in interpreting grammatical relations: a case study in multilingual acquisition. *Journal of Child Language, 38*(04), 763-792. doi: doi:10.1017/S0305000910000358

Oryadi-Zanjani, M. M., & Ghorbani, R. (2005). Study of speech quality indexes in 4 to 5 years old Farsi-speaking children in Semnan, Birjand and Tonekabon. *Journal of the University of Mazandaran Medical Sciences, 15*(50), 90-96.

Oryadi-Zanjani, M. M., Ghorbani, R., & Keikha, F. (2006). Standardization of total numbers of word, mean length of utterance and mean length of 5 long sentences

in normal Persian language children between 2 to 5 years old in Semnan city. K*oomesh, 7*(3), 177-182.

Owen, A. J., & Leonard, L. B. (2002). Lexical diversity in the spontaneous speech of children with specific language impairment: application of d. *Journal of Speech, Language and Hearing Research, 45*(5), 927-937. doi: 10.1044/1092-4388(2002/075)

Pan, B. A. (2011). Assessing vocabulary skills. In E. Hoff (Ed.), *Research methods in child language* (pp. 100-112). West Sussex: Wiley-Blackwell.

Pantcheva, M. (2007). Language diversity and typology: LIN-1050 Word Order. Retrieved 05/03/2012, from ansatte.uit.no/mpa001/teach/WO-pantcheva.pdf

Parsafar, P. (2010). Syntax, morphology, and semantics of ezafe. *Iranian Studies, 43*(5), 637-666. doi: 10.1080/00210862.2010.518029

Paul, R., & Norbury, C. (2012). *Language disorders from infancy through adolescence: listening, speaking, reading, writing, and communicating* (4th ed.). St. Louis, Missouri: Mosby Elsevier Health Sciences.

Paul, R. (2007). *Language Disorders from Infancy through Adolescence* (3rd ed.). St. Louis: Mosby Elsevier.

Peña, E. D., Reséndiz, M., & Gillam, R. B. (2007). The role of clinical judgements of modifiability in the diagnosis of language impairment. *International Journal of Speech-Language Pathology*, *9*(4), 332-345. doi: doi:10.1080/14417040701413738

Penke, M. (2008). Morphology and language disorder. In M. J. Ball, Michael R. Perkins, Nicole Müller & S. Howard (Eds.), *The handbook of clinical linguistics*: Blackwell Reference Online. 30 January 2012 Retreived from http://www.blackwellreference.com/subscriber/tocnode?id=g9781405135221_chunk_g978140513522115.

Perez-Leroux, A. T., Castilla-Earls, A. P., & Brunner, J. (2012). General and specific effects of lexicon in grammar: determiner and object pronoun omissions in child Spanish. *Journal of Speech, Language and Hearing Research, 55*(2), 313-327. doi: 10.1044/1092-4388(2011/10-0004)

Plante, E., & Vance, R. (1994). Selection of preschool language tests: a data-based approach. *Language Speech and Hearing Services in Schools, 25*(1), 15-24.

Pouladi, S., & Khoddam, A. (2002). *Mean Length of Utterance in four to five years old children.* Paper presented at the 7th Iranian Congress on Speech Therapy, Tehran.

QUADAS-2: A quality assessment tool for diagnostic accuracy studies. (2010). Retrieved 11/08/2012, from http://www.bris.ac.uk/quadas/resources/quadas2.pdf

Rahmany, R., Marefat, H., & Kidd, E. (2010). Persian speaking children's acquisition of relative clauses. *European Journal of Developmental Psychology, 8*(3), 367-388. doi: 10.1080/17405629.2010.509056

Rasekh-Mahand, M. (2010). Persian clitics in verb structure. *Journal of Research in Linguistics, 2* (2). 75-85

Rasooli, M. S., Moloodi, A., Kouhestani, M., & Minaei-Bidgoli, B. (2011). *A syntactic valency lexicon for Persian verbs: the first steps towards Persian dependency Treebank.* Paper presented at the 5th Language and Technology Conference (LTC): Human language technologies as a challenge for computer science and linguistics, Poznan, Poland.

Redmond, S. M., Thompson, H. L., Goldstein, S. (2011). Psycholinguistic profiling differentiates specific language impairment from typical development and from attention-deficit/hyperactivity disorder. *Journal of Speech, Language, and Hearing Research, 54,* 99-117. doi: 10.1044/1092-4388(2010/10-0010)

Reitsma, H. (2009). *Introduction to meta-analysis of accuracy data*. Retrieved from http://srdta.cochrane.org/sites/srdta.cochrane.org/files/uploads/Copenhagen_5_StatisticalMethodsIntro.pdf

Rezaei, E., Shavaki, Y. A., Arshi, A., & Keyhani, M. R. (2011). The perception and expression of nouns in 2.5 to 4 year-old normal Persian-speaking children in Arak, Central Iran. *Audiology, 20*(2), 54-62.

Rice, M. L. (1994). Grammatical categories of children with specific language impairments. In R. V. Watkins & M. Rice (Eds.), *Specific language impairments in children*. Baltimore: MD Brookes Publishing Co.

Rice, M. L., & Oetting, J. B. (1993). Morphological deficits of children with SLI: Evaluation of number marking and agreement. *Journal of Speech and Hearing Research, 36*(6), 1249-1257.

Rice, M. L., Redmond, S. M., & Hoffman, L. (2006). Mean Length of Utterance in children with Specific Language Impairment and in younger control children shows concurrent validity and stable and parallel growth trajectories. *Journal of Speech, Language, and Hearing Research 49. 793*-808.

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific Language Impairment as a Period of Extended Optional Infinitive. *Journal of Speech and Hearing Research, 38*(4), 850-863.

Rispoli, M., Hadley, P. A., & Holt, J. K. (2012). Sequence and system in the acquisition of tense and agreement. *Journal of Speech, Language and Hearing Research, 55*(4), 1007-1021. doi: 10.1044/1092-4388(2011/10-0272)

ROC Curve. (2012). Retrieved 27/04/2012, from http://www.medicalbiostatistics.com/ROCCurve.pdf

Roulstone, S. (1997). What's driving you? A template which underpins the assessment of preschool children by speech and language therapists. *International Journal of Language & Communication Disorders, 32*(3), 299-315. doi: 10.3109/13682829709017897

Roulstone, S. (2001). Consensus and variation between speech and language therapists in the assessment and selection of preschool children for intervention: a body of knowledge or idiosyncratic decisions?. *International Journal of Language & Communication Disorders, 36*(3), 329-348. doi: 10.1080/13682820010019928

Roulstone, S., Peters, T. J., Glogowska, M., & Enderby, P. (2008). Predictors and outcomes of speech and language therapists' treatment decisions. *International Journal of Speech-Language Pathology, 10*(3), 146-155. doi: 10.1080/17549500801894362

Rowe, M. L. (2011). Recording, transcribing, and coding interaction. In E. Hoff (Ed.), *Research methods in child language* (pp. 193-207). West Sussex: Wiley-Blackwell.

Sackett, D. L., & Haynes, R. B. (2002). Evidence base of clinical diagnosis; the architecture of diagnostic research. *British Medical Journal, 324* (7336), 539-541.

Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. *British Medical Journal, 312*(7023), 71-72.

Sadat-Tehrani, N. (2008). *The structure of Persian intonation.* Paper presented at the Proceedings of Speech Prosody.

Sahakian, S., & Snyder, B. (2012). *Automatically learning measures of child language development*. Paper presented at the 50th Annual Meeting of the Association for Computational Linguistics, Jeju Island, Korea. Short Paper Retrieved from http://www.aclweb.org/anthology-new/P/P12/P12-2019.pdf

Samadi, H., & Perkins, M. R. (1998). P-LARSP: A developmental language profile for Persian. *Clinical Linguistics & Phonetics, 12*(2), 83-103.

Sayyahi, F., Soleymani, Z., Mahmoudi Bakhtiyari, B., & Jalaie, S. (2011). Providing a non word repetition test in 4-year-old Persian children and determining its validity and reliability. *Audiology, 20*(2), 47-53.

Schlosser, R. W., Wendt, O., & Sigafoos, J. (2007). Not all systematic reviews are created equal: Considerations for appraisal. *Evidence-Based Communication Assessment and Intervention, 1*(3), 138-150. doi: 10.1080/17489530701560831

Schneider, P., Dubé, R. V., & Hayward, D. (2005). The Edmonton narrative norms instrument. Retrieved 13/02/2012, from http://www.rehabmed.ualberta.ca/spa/enni

Sheng, L., & McGregor, K. K. (2010). Lexical-semantic organization in children with Specific Language Impairment. *Journal of Speech Language and Hearing Research, 53*(1), 146-159. doi: 10.1044/1092-4388(2009/08-0160)

Shokouhia, H., & Kipkab, P. (2003). A discourse study of Persian raˆ. *Lingua, 113*, 953-966. doi: doi:10.1016/S0024-3841(02)00145-6

Simon-Cereijido, G., & Gutierrez-Clellen, V. F. (2007). Spontaneous language markers of Spanish language impairment. *Applied Psycholinguistics, 28*(02), 317-339. doi: doi:10.1017/S0142716407070166

Skarakis-Doyle, E., Campbell, W., & Dempsey, L. (2009). Identification of children with language impairment: investigating the classification accuracy of the MacArthur-Bates Communicative Development Inventories, Level III. *American Journal of Speech-Language Pathology*, *18*(3), 277-288. doi: 10.1044/1058-0360(2009/08-0035)

Soleimani, Z., & Dastjerdi Kazemi, M. (2005). Study of validity and reliabity of phonological awareness test. *Journal of Psychology*, *33*(1), 82-100.

Solomon, S. I. (2009). *Ibn Sina: a treatise on arabic phonetics.* München: LINCOM publishers.

Soraya, M., Bakhtiyari, B. M., Badiee, Z., Kazemi, Y., & Soleimani, B. (2012). A comparative study of size of expressive lexicon in prematurely born children with full-term 18-36 month's children. *Audiology, 21*(1), 76-82.

Southwood, F., & Russell, A. F. (2004). Comparison of conversation, freeplay, and story generation as methods of language sample elicitation. *Journal of Speech, Language and Hearing Research, 47*(2), 366-376. doi: 10.1044/1092-4388(2004/030)

Statistical Centre of Iran, *Selected finding of the 2011 national population and housing census*. (2011). Tehran: Author.

Study Designs. (2012, 22 January 2009). Retrieved 22/03/2012, from http://www.cebm.net/index.aspx?o=1039

Taylor, R. (1990). Interpretation of the correlation coefficient: a basic review. *Journal of Diagnostic Medical Sonography, 6*(1), 35-39. doi: 10.1177/875647939000600106

Thordardottir, E., Kehayia, E., Mazer, B., Lessard, N., Majnemer, A., Sutton, A., . . . Chilingaryan, G. (2011). Sensitivity and specificity of French language and processing measures for the identification of Primary Language Impairment at Age 5. *Journal of Speech, Language and Hearing Research, 54*(2), 580-597. doi: 10.1044/1092-4388(2010/09-0196)

Tomasello, M., & Stahl, D. (2004). Sampling children's spontaneous speech: how much is enough? *Journal of Child Language, 31*(01), 101-121. doi: doi:10.1017/S0305000903005944

Tomblin, J. B., Zhang, X., Buckwalter, P., & O'Brien, M. (2003). The stability of Primary Language Disorder: Four years after kindergarten diagnosis. *Journal of Speech, Language and Hearing Research, 46*(6), 1283-1296. doi: 10.1044/1092-4388(2003/100)

Tschanz, D. (2003). Ibn-Sina: the prince of physicians. *Journal of the International Society for the History of Islamic Medicine (JISHIM), 3*.

Valency Lexicon for Persian Verbs. (2012). Retrieved 27/10/2012, from http://dadegan.ir/en/pervallex

Wagner, C. R., Nettelbladt, U., Sahlen, B., & Nilholm, C. (2000). Conversation versus narration in pre-school children with language impairment. *International Journal of Language & Communication Disorders, 35*(1), 83-93. doi: 10.1080/136828200247269

Wanska, S. K., Bedrosian, J. L., & Pohlman, J. C. (1986). Effects of play materials on the topic performance of preschool children. *Language Speech and Hearing Services in Schools, 17*(3), 152-159.

Washington, K. N. (2007). Using the ICF within speech-language pathology: Application to developmental language impairment. *International Journal of Speech-Language Pathology, 9*(3), 242-255. doi: doi:10.1080/14417040701261525

Watkins, R. V., Kelly, D. J., Harbers, H. M., & Hollis, W. (1995). Measuring children's lexical diversity: differentiating typical and impaired language learners. *Journal of Speech and Hearing Research, 38*(6), 1349-1355.

Weismer, S., & Robertson, S. (2006). Focused stimulation approach to language intervention. In R. J. McCauley & M. E. Fey (Eds.), *Treatment of language disorders in children*. Baltimore: Paul H. Brookes Pub.

Wexler, K. (2003). Lenneberg's dream: Learning, normal language development, and specific language impairment. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of specific language impairment*. Mahwah, NJ: Erlbaum.

Wong, A. M. Y., Klee, T., Stokes, S. F., Fletcher, P., & Leonard, L. B. (2010). Differentiating Cantonese-speaking preschool children with and without SLI using MLU and lexical diversity (D). *Journal of Speech, Language and Hearing Research, 53*(3), 794-799. doi: 10.1044/1092-4388(2009/08-0195)

World Health Organization. (2002). *Towards a common language for functioning, disability and health ICF.* Geneva: Author.

Yarshater, E. (1989). Persia or Iran. *Iranian Studies, XXII*(1).

Zamora, J., Abraira, V., Muriel, A., Khan, K., & Coomarasamy, A. (2006). Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Medical Research Methodology, 6*(1), 31.

Ziatabar Ahmadi, S. Z., Arani Kashani, Z., Mahmudi Bakhtiari, B., & Keyhani, M. R. (2010a). Development and comparison of tasks for assessment of first phoneme of word in normal 5 to 6 year-old Persian-speaking children. *Koomesh, 11*(2), 121-128.

Ziatabar Ahmadi, S. Z., Arani Kashani, Z., Mahmudi Bakhtiari, B., & Keyhani, M. R. (2010b). The development and evaluation of Persian rhyme awareness tasks for normal 5-6 year-old Persian-speaking children. *Audiology, 19*(2), 47-56.

Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry, 39*(4), 561-577.