Development of the Urdu Speech Assessment Test

Saira Ambreen, Carol K. S. To

The University of Hong Kong, HKSAR

For correspondence:
Saira Ambreen, 7/F Meng Wah Complex, The University of Hong Kong, Pok Fu Lam Road, Hong Kong
saira.ambreen90@gmail.com

[This paper appeared in the Journal of Clinical Practice in Speech-Language Pathology - JCPSLP, Vol. 24, No. 3]

How to cite: Ambreen, S., & To, C. K. S. (2022). Development of the Urdu speech assessment test. *Journal of Clinical Practice in Speech-Language Pathology*, *24*(3). 138–145.

Development of the Urdu Speech Assessment Test

ABSTRACT

This paper outlines the development of the Urdu speech assessment test (USAT). There is a pressing need for standardized tests focusing on the articulation and phonological skills of the Urdu-speaking population. The USAT was developed for Urdu-speaking children aged up to eight years. More specifically, it focuses on Urdu spoken in Pakistan. Throughout its development, the recommended guidelines and procedures for test development were carefully followed (Eisenberg & Hitchcock, 2010; McLeod, 2012[a]). The USAT elicits all Urdu consonants and vowels at all word positions, and for each position, two production opportunities are provided. During word selection, language-specific stress patterns and other constraints were carefully considered. A scoresheet is designed to assist with the sample transcription, and the analyses of speech sound acquisition, production accuracy, and occurrence of phonological processes. This paper discusses the conceptualization of USAT and its content validity. Further operationalization is currently ongoing and once the USAT is standardized, it will be available for clinical and research use.

(*word count:* 160)

Keywords. Urdu, test development, speech sound acquisition, phonological development

INTRODUCTION

Urdu is an Indo-Aryan language. With more than 230 million speakers, it is the 10th most spoken language in the world (Eberhard et al., 2022). It is spoken in many countries including Australia, Afghanistan, Bahrain, Bangladesh, India, Pakistan, the UK, and the USA. There are different dialects of Urdu including Awadhi, Bhojpuri, Dakani/Deccani, Hyderabadi, Rekhta, and Khari Boli. Most of these dialects are spoken in India. Since the independence of Pakistan in 1947, Urdu has been its national language, where it is also used as the lingua franca. Over time, Urdu spoken in Pakistan has been influenced by the provincial and regional languages, and "has become different from Urdu spoken in India" (Khanum & Hussain, 2017, p.18). USAT is mainly developed based on Urdu spoken in Pakistan. No specific name for this dialect could be found in the literature, however, it can be referred to as "Pakistani Urdu" (Khanum & Hussain, 2017, p.18).

In Pakistan, the field of speech-language pathology only began to flourish after 2005 when higher education institutes began to offer degree programs in this domain. Before that, clinicians were mostly trained in post-graduate courses; for example, the National Institute of Rehabilitation Medicine (NIRM), Islamabad has been running a training program since 1991 (c.f. Ambreen & To, 2021; Shafqat, 2020). The field awaits more standardized speech and language assessments with rigorous design.

In the field of speech sound disorders, single-word articulation tests have been developed and widely used for the assessment of speech sound production ability. The assessment process is time-efficient such that clinicians can get the most information about the child's speech sound production ability from the test within a reasonable amount of time. These tests generally include speech sounds in the possible (initial, medial, and final) word positions which provide a more comprehensive picture of a child's production abilities. For a

long time, there was limited research available to support clinical practice in Pakistan. Clinicians mainly rely on informal tools and instruments devised from other languages. The first tool for speech sound assessment was developed by Shabbir and Zahra (2016), which however is unpublished. Later, two more tests were developed by Noveen et al. (2017) and Muzaffar et al. (2019).

To avoid reinventing the wheel (McLeod, 2012a), we consulted an extensive list of speech assessments available for different languages (McLeod, 2012b). This list identified Bilingual Speech Sound Screen (BiSSS; Stow & Pert, 2020). It was developed for young children from the Pakistani heritage community in the UK and focuses on Urdu, Mirpuri, and Punjabi. Twenty-one target words are used to elicit consonants in response to their corresponding picture illustrations.

Next, we consulted Ambreen & To (2021), which includes a review of available speech assessments for Urdu. Two of the before-mentioned tests (Urdu wordlist and TAAPU) were found to be focusing on consonants. The Urdu wordlist (Shabbir & Zahra, 2016) included 95 words to elicit Urdu consonants in the following positions: syllable-initial word-initial (SIWI), syllable-initial within word (SIWW), syllable-final within word (SFWW), and syllable-final word-final (SFWF). TAAPU (Noveen et al., 2017) elicits consonants at three word positions: initial, medial, and final.

During our search, we found one more test, namely the Urdu Articulation Screening Tool (UAST; Muzaffar et al., 2019). UAST (Muzaffar et al., 2019) was developed for children aged from four to 11 years. It includes a total of 96 words that elicit 36 consonants on initial, medial, and final word positions. Similar to TAAPU (Noveen et al., 2017) and Shabbir and Zahra's wordlist (2016), UAST (Muzaffar et al., 2019) also focuses on Urdu

consonants only and provides a single production opportunity for each consonant at each word position.

Eisenberg and Hitchcock (2010) reviewed and compared 11 standardized tests that aimed to inventory speech sounds of a language. They concluded that speech assessment tests should provide "sufficient coverage of consonants or vowels for establishing a phonetic inventory and making conclusions about the segments that a child can and cannot produce," and suggested a set of criteria for the phonetic makeup of the words to be included in such tests, calling these words "phonetically-controlled" (p. 494). These criteria help guard against the confounding aspects that may contribute to scaffolding or the inhibition of correct sound productions. The assessment tools discussed above, TAAPU (Noveen et al., 2017), UAST (2019), and the wordlist (Shabbir & Zahra, 2016), have not incorporated these important criteria in their development. We aimed to develop an assessment test that would carefully consider these important aspects. This motivated the development of USAT, which includes Urdu consonants as well as vowels. It uses phonetically controlled words, in varied phonetic contexts. Moreover, it provides two production opportunities for all consonants, at each word position. This paper outlines the process that was followed during the development of USAT and reports the considerations that were made during this process.

Guidelines for Test Development

McLeod (2012a) discussed the guidelines for creating speech sampling tools based on Frytak (2000). These consist of two main stages: 1. Conceptualization, and 2. Operationalization.

Phase 1. Conceptualization

McLeod (2012a) listed the following dimensions as part of the test conceptualization process.

1. Purpose
2. Intended population
3. Target skills
4. Scope

(a) *Phonetic inventory and phonotactic constraints.* Preparation of the phonetic inventory of speech sounds that exist in a language is a crucial step in test development (McLeod, 2012a). This step also includes the review of phonotactic constraints specific to that language.

(b) *Production opportunities.* Eisenberg and Hitchcock (2010) suggested that confounding factors should be controlled to obtain a representative picture of children's phonological skills. These confounding factors include the lack of sufficient production opportunities for the target sounds, difficult phonetic contexts that may inhibit children's production, or the phonetic contexts that may support production that is otherwise not present in the child's inventory. In terms of production opportunities, McLeod (2012a) suggested providing two to five production opportunities. Eisenberg and Hitchcock (2010) suggested that at least two opportunities should be provided for the production of each sound, which should be in different phonetic contexts. For example, a consonant should be produced in different vowel contexts (e.g., front, and back). Similarly, production opportunities for vowels should be provided in different post-vocalic contexts (e.g., open, and closed syllables). As mentioned earlier, Eisenberg and Hitchcock (2010) provided guidelines that can be used to select phonetically controlled words for eliciting speech sounds, which are as follows:

Consonants.

i. should be elicited in a singleton context,

ii. should be elicited in either mono- or bisyllabic words,

iii. for bisyllabic words, the target consonant should be in the stressed syllable,

iv. should be tested in non-harmonic contexts, which means the target consonant should be different from the other consonants in the word,

v. if in the word-final position, should be tested in a monomorphemic context, which means the targeted consonant is not part of a separate phoneme,

vi. should be elicited in different word positions,

vii. for each position (initial, medial, and final), at least two opportunities should be provided,

viii. and these opportunities should be in different vowel contexts, for example, front and back.

Vowels.

i. should be elicited in mono- or bisyllabic words,

ii. for bisyllabic words, the target vowel should be in the stressed syllable,

iii. should be elicited in different postvocalic contexts and should not be followed by a liquid consonant,

iv. at least two production opportunities should be provided for each targeted vowel.

(c) *Word selection* – words should be frequently used by native speakers, familiar to young children, culturally appropriate, and picturable (McLeod, 2012a).

(d) *Picture illustrations* – should be culturally appropriate and engaging.

(e) *Test administration* – the aspects related to the test administration should be clearly discussed, for instance, whether the test should be administered individually or in a group, etc.

(f) *Sample elicitation* – how the sample will be elicited, for instance, will the elicitation be spontaneous or imitated? can the elicitation be facilitated with prompts? etc.

(g). *Aspects related to recording, transcription, scoring, and analyses* – what considerations should be made during these steps? should the sample be transcribed using narrow or broad transcription? What kind of analyses can be done with the elicited data and what material will be available to support this process?

Phase 2. Operationalization

Operationalization is a key step in the standardization of a test. It includes the establishment of test norms and the examination of the test's validity (content validity, criterion validity, construct validity), reliability (internal consistency, test-retest reliability, inter-rater reliability, and intra-rater reliability), item analyses, and analyses of sensitivity and specificity (Fabiano-Smith, 2019; Friberg, 2010; McLeod, 2012a).

RESULTS

This section presents detailed information about the process that was followed during the phases of conceptualization and operationalization.

Phase 1. Conceptualization

1. Purpose

The USAT is developed to assess the production of Urdu consonants and vowels. It is also designed to test Urdu speech sounds in phonetically controlled contexts so that it can assist in presenting a more reliable picture of the phonological skills of Urdu-speaking children.

2. Intended population

It is developed for Urdu-speaking children aged between 2:07 and 7:12 years.

3. Target skill

The USAT is a single-word test that uses picture-naming to elicit Urdu speech sounds. It assesses speech sound acquisition, production accuracy, and the occurrence of phonological error patterns.

4. Scope

The aspects of language that were considered during the test development are discussed below:

(a) *Urdu phonetic inventory and phonotactic constraints.* Ambreen and To (2022) reviewed the Urdu phonological system and identified a total of 38 Urdu consonants, of which 36 are initial, 37 are middle, and 31 are final consonants. Urdu allows consonant clusters at the syllable-final position only, with a maximum of two consonants in a cluster. There are 15 commonly used consonant clusters (Ranjha, 2014). Most consonants can exist at the word-medial position as geminates. Urdu syllabification rules govern that if there are two consonants at the WM syllable boundary, then the first consonant goes to the coda of the preceding syllable and the second one goes to the onset of the next syllable (Akram, 2002; Ranjha, 2012). That is why during syllabification, these geminate consonants are broken, for example, for the word /ɖəbbəh/ ‘ڈبہ’ (box), there is a geminated /b/ at the WM position, and its syllabification is done as /ɖəb.bəh/. Because of frequent occurrences of geminates at the WM position, it was difficult to completely avoid such words, hence USAT includes some words where the target consonant is in a geminated context in the WM position. The overall percentage of such words is 5.61%. Similarly, some of the target consonants at the WF position are in a clustered context (7.14%), because no other suitable words could be found.

There are 23 monophthongs (Centre for Language Engineering – CLE, 2015) and 15 diphthongs (Bhatti et al., 2016). No triphthongs have been found in Urdu (Khurshid et al.,

2003). Urdu has many bisyllabic and polysyllabic words. Nazar (2002) and Ghazali (2002) reported 11 syllable structures in Urdu, in which vowel is the essential factor and consonants can exist at the syllable initial or final positions. These syllable structures were distinguished based on vowel length, which means they considered CVC and CVVC as two separate structures. Hence, if the possible syllable structures in Urdu are counted regardless of the vowel length, it can be said that there are a total of six syllable structures in Urdu. Urdu does not have any tones, and generally follows the syllable structure $C_{(0-1)}V_{(1-2)}C_{(0-2)}$. Syllables can be light, heavy, or superheavy based on their stress. However, stress does not influence the meaning (Nyyar, 2003).

The USAT includes all Urdu consonants at all possible word positions and uses all Urdu syllable structures in the included words. Similarly, all monophthongs and 13 out of 15 diphthongs are included. The remaining two diphthongs /ɪ͡ũː, æ͡ɑː/, are not included yet because no illustrations could be found for the words that were deemed suitable to elicit these sounds, for example, for /ɪ͡ũː/, the word was /kɪ͡ũː/ 'کیوں' (why).

The speech sounds included are as follows:

- 36 initial consonants /p, b, pʰ, bʰ, m, n, f, v, ʈ, ɖ, ʈʰ, ɖʰ, ʧ, ʤ, ʧʰ, ʤʰ, r, k, g, kʰ, gʰ, q, ʃ, ʒ, j, ɣ, x, ʔ, h/
- 37 middle consonants /p, b, pʰ, bʰ, m, n, ŋ, f, v, ʈ, ɖ, ʈʰ, ɖʰ, ʧ, ɖ, ʧʰ, s, z, l, ʧ, ʧʰ, ʤ, ʤʰ, r, ɽ, k, g, kʰ, gʰ, q, ʃ, ʒ, j, ɣ, x, ʔ, h/
- 31 final consonants /p, b, bʰ, m, n, f, ʈ, ɖ, ʈʰ, ɖʰ, ʧ, ɖ, ʧʰ, s, z, l, ʧ, ʧʰ, ʤ, ʤʰ, r, k, g, kʰ, gʰ, q, ʃ, ɣ, x, ʔ, h/
- 23 monophthongs /iː, ĩː, uː, ũː, oː, õː, eː, ẽː, ɔː, ɔ̃ː, æː, æ̃ː, ɑː, ɑ̃ː, ɪ, ɪ̃, ʊ, ʊ̃, o, e, ə, ə̃, æ/
- 13 diphthongs /ɑ͡eː, ɑ͡e, ɑ͡ẽː, ɑ͡ɪ, ɑ͡iː, ɑ͡oː, ə͡eː, ə͡iː, ə͡ĩː, e͡ɑː, ɪ͡õː, o͡iː, ɑ͡ĩː/

(b) *Production opportunities*. We carefully followed the criteria suggested by Eisenberg and Hitchcock (2010) and McLeod (2012a), for example, we included two production opportunities for each phoneme at each word position. These opportunities are in different vowel contexts. The only exceptions are the following: /ɣ/ (WI and WF), /q/ (WF), and /ʤʰ/, where despite multiple attempts (web searches and searches of other available resources), suitable words could not be found to address varied CV interaction. For example, /ɣ/ at the word-initial position is elicited in /ɣʊbɑːrɑː/ 'غبارہ' (balloon) and /ɣɑːr/ 'غار' (cave). In both words, /ɣ/ is followed by back vowels; however, both still include different vowels: /ʊ/ and /ɑː/. The USAT elicits all Urdu consonants at three word positions (initial, medial, and final) while keeping in view the language-specific constraints. Each consonant is elicited twice at each word position except for sounds for which no more than one suitable word could be found. For example, /ʤʰ/ at the word-final position is only elicited by /boːʤʰ/ 'بوجھ' (load). Similarly, all of the monophthongs are elicited twice, except for the ones where no more than one suitable word could be found, for example, for /ẽː/, only one suitable word could be found: /gẽːn̪ˌ/ 'گیند' (ball). As mentioned earlier, USAT elicits 13 out of 15 Urdu diphthongs. Following points should be noted: Firstly, from these 13 diphthongs, six /ɑːe͡ː, ɑːẽ͡ː, ɑːi͡ː, ɑːo͡ː, ɪõ͡ː, ɪõ͡ː/ are elicited twice. However, it should be noted that all of their target words are open-syllable, and these target sounds are at the syllable-final position, which means that although two production opportunities are being provided, these are in similar phonetic contexts. Secondly, the remaining seven /ɑːe͡, ɑːɪ͡, əe͡ː, əi͡ː, əĩ͡ː, əɑ͡ː, ɑːẽ͡ː/ are elicited once only. The reason for both the above points is that no other suitable words could be found to elicit these sounds.

(c) *Word selection*. To find suitable test words, we reviewed the available Urdu wordlists from Noveen et al. (2017) and Shabbir and Zahra (2016) and consulted the Urdu

textbooks used in Pakistani primary schools. We also consulted *Feroz-ul-Lughat* (an Urdu dictionary) and several online dictionaries to find suitable words. As a result, USAT consists of 175 words. Although most of the words (93.7%) are either mono- or bisyllabic, some trisyllabic and tetrasyllabic words (6.3%) are included when no suitable mono- or bisyllabic words could be found (see Figure 1), such as /mũːŋg.pʰəliː/ 'مونگ پھلی' (peanut) for /pʰ/ at the word-middle position. The USAT includes both nouns (87%) and verbs (13%).

(d) *Picture illustrations*. After the wordlist was prepared, pictures were selected for each word. The pictures had to be clear and attractive to help engage the children in the task. All the pictures were selected from online royalty-free/stock images. The USAT thus includes 175 colorful and culturally appropriate pictures. The pictures shown to the children are like the one presented in Figure 2. The backside of each page contains a picture of the target word and its orthography, as well as hierarchical prompts, item number, target sound(s), English meaning, International Phonetic Alphabet (IPA) transcription, and syllable structure (see Figure 3).

(e) *Test administration*. The USAT is designed for standardized administration. It requires each child/subject to be tested individually. It can be administered by speech-language pathologists and other personnel associated with this field who are familiar with speech transcription. The testing material will include supplementary information to facilitate non-Urdu-speaking professionals. This may include the syllabification of the items, audio recordings of native productions of the test words, and also detailed guidelines so that the same standardized method can be followed during the test administration.

(f) *Sample elicitation/Cueing hierarchy*. The USAT requires children to spontaneously produce the test words in response to the pictures shown to them. To facilitate the testing process, it uses hierarchical cueing consisting of three-level prompts. Upon the stimulus

presentation, for example, for the item shown in Figure 2, “pencil,” the examiner will ask a question to elicit a spontaneous response: “what is this (that we use to write)?” In the case of a verb, the question would be, “what is he/she doing?” Children will be encouraged to respond spontaneously in Urdu. The second-level prompt is the binary cue. It will only be provided if the child is unable to respond to the previous prompt. For instance, if the child is unable to respond spontaneously to the picture of the pencil, the examiner will say, “is it a pencil or a notebook?” If the child remains unable to respond correctly, the examiner will elicit the response by imitation. To keep the testing process standardized, we have prepared hierarchical prompts for all test items, and these prompts will be available on the back page of each item illustration along with other relevant information (see Figure 3).

For most of the words that were selected for eliciting Urdu diphthongs, we could not find any suitable pictures. A cue question task was therefore designed for their elicitation. Stimuli for such diphthongs consist of two pictures, where the examiner will use the first picture to elicit a response regarding the second. The examiner will instruct the child as follows: “You will be shown two pictures. I will tell you what is happening in the first picture, and you will complete the sentence by telling me about the other one.” For example, to elicit the word /roːiː/ ‘روئی’ (cried) in Figure 4, the examiner will say, “بچہ رویا اور بچی ــــــــ” (the boy cried and the girl _____), and the child’s response should be /roːiː/ ‘روئی’ (cried).

(g). <u>Recording, transcription, scoring, and analysis</u>. During the assessment, the speech samples should be audio and video recorded (with prior permission from the child’s parent/guardian) to aid transcription. These recordings should be carefully made so that the factors related to the testing environment and the recording equipment do not compromise the sample quality. The test should be administered in a quiet and well-lit place for better-quality recordings. A stand-held or collar microphone can be used for audio recording,

whereas for video recording, a camera can be set up to clearly record the child's face and articulatory movements. These recordings can also be used to establish inter-rater reliability and test-retest reliability.

The speech samples should be transcribed using narrow transcription with Urdu IPA symbols and diacritics. For scoring, we have designed a sheet that records information about the child (e.g., age, and gender). It includes Urdu orthography and the IPA transcription of each test item. The child's response can be transcribed for each item along with the response type: whether the response was spontaneous, or was elicited by using a binary cue or imitation. The target positions of each speech sound are listed on the sheet, and the examiner can mark which sound (at what position) was not correctly produced. The last column can be used to record the presence of the phonological process(es). The sheet is color-coded for ease of use and for making the scoring process time efficient. On the last page, the incorrect productions can be calculated for each speech sound. In addition to the independent analyses, the scoring sheet facilitates production accuracy (relational) analyses, as it calculates PPC, PCC, PICC, PFCC, and PVC[1]. Finally, it includes a table to calculate the occurrence of phonological processes.

Researchers and clinicians have been using relational analyses comprising production accuracy and phonological patterns/processes to see how children's speech varies from adults. PCC analysis was first introduced by Shriberg and Kwiatkowski (1982) as a component of phonological assessment. Since then, several variations of this measure have been employed, such as PPC, PVC, etc. These measures have the potential to provide a snapshot of the overall accuracy of a child's production. Research has found that production accuracy can serve as a predictor of a child's phonological skills (Li & To, 2017; McLeod

[1] PPC = Percentage of Phonemes Correct; PCC = Percentage of Consonants Correct; PICC = Percentage of Initial Consonants Correct; PFCC = Percentage of Final Consonants Correct; PVC = Percentage of Vowels Correct.

& Crowe, 2018; McLeod & Verdon, 2014; Pham & McLeod, 2019; Shriberg & Kwiatkowski, 1982).

Following phonological processes were found in the adult native speakers of Urdu (Nawaz, 2002), and can be expected in the speech of typically-developing children: /ʔ/ is deleted and generally substituted with a long vowel; /h/ is deleted at the syllable boundary if it follows a long vowel; /h/ is deleted at syllable boundary if follows a short vowel and that vowel is produced as long; /q/ is mapped as /k/; aspiration may be deleted from consonants at the word-final position; and /ʒ/ can be mapped as /j/ or /z/ (Khan, 1997).

Phase 2. Operationalization

The conceptualization of the USAT has been completed, and its operationalization is in process.

Content Validity

Content validity evaluates the relevance and representativeness of the test items to determine whether the test actually measures what it aims to measure (McLeod, 2012a). As mentioned above, the item selection for the USAT included a detailed literature search, so content coverage is discussed earlier in detail (see headings “production opportunities” and “word selection”). The USAT includes 175 different words to provide sufficient coverage for all Urdu consonants and vowels, and it thus meets the recommendation made by Weston, Shriberg, and Miller (1989), that to evaluate a child’s speech productions, a speech test should include at least 90 different or 250 total words.

McLeod (2012a) identified expert consultation as an important step in determining content validity. For this, we designed a survey questionnaire and sent it to nine SLPs who are native Urdu speakers and have experience working with children. Five of them returned the completed surveys. The survey questions and the details of the responses are presented in

Table 1. The participants were asked to select “Yes” or “No” to indicate their judgment about the test words. For the survey, we selected the two most suitable picture illustrations for each test word, and the survey participants were asked to judge their appropriateness. They were given the following options: 1) Picture-a better depicts the test item; 2) Picture-b better depicts the test item; 3) Both pictures are equally appropriate to depict the test item, and 4) Neither of the pictures is a suitable illustration of the test item. All of the illustrations except the ones listed below were judged as appropriate by at least four of the five respondents. Therefore, all of the other test words and their illustrations were accepted as appropriate. The exceptions were the illustrations of: (1) laborer /məzd̪uːr/ ‘مزدور’; (2) (girls) came /ɑːi̯ː/ ‘آئیں’; (3) charpai /ʧɑːrpɑːi̯ː/ ‘چارپائی’

The illustrations for the first and second words were modified, and the third word, چارپائی, was replaced by “(girl) came” /ɑːi̯ː/ آئی, which was equally suitable for eliciting the target sound and easier to illustrate.

The evaluation of content validity also involved piloting with 4 native Urdu-speaking children from Lahore, Pakistan. The purpose was to evaluate the efficiency of the testing instructions and other testing aspects. The speech samples were transcribed and scored using the pre-designed scoring sheet. The modifications made to the test items based on the pilot study are presented in Table 2. The only new words included after the pilot study were /ʃɑːd̪iː/ ‘شادی’ (wedding) and /ʧiːx/ ‘چیخ’ (scream). For the other words, only the word forms were modified, for example, in the case of ‘باندھنا’, only ‘نا’ was removed because the children did not produce it this way during elicitation, and instead said /bɑ̃ːnd̪ʰ/ 'باندھ رہی ہے۔' (tying up). Moreover, the children could not identify the illustration of /məzd̪uːr/ ‘مزدور’ (laborer), thus, new illustrations were found for this word and the other two new words. Two pictures for

each were selected and sent for expert judgment, and those judged appropriate were included in the USAT.

Future Direction & Conclusion

This paper provides detailed information on the development of the USAT. The complete process followed during the conceptualization phase and for content validity (from the operationalization phase) is discussed. USAT is currently being used to conduct a speech sound acquisition study. Based on the available data, the average time to complete the USAT is 15–20 minutes. In the later stages, the norms, test validity, and reliability will be established.

As discussed throughout this paper, the assessment tools not only should be carefully developed based on a language's specific features and phonotactic constraints, but these should also provide enough production opportunities for each phoneme. It is also necessary that these production opportunities are provided in varied phonetic contexts to get a true representation of a child's phonological skills. The absence of carefully designed assessment tools may contribute to an over- or under-diagnosis, which may, in turn, lead to late intervention. SLPs must have thorough and accurate information about the assessment tools to make reliable clinical judgments. USAT is an attempt to develop an instrument that carefully follows essential guidelines of speech sound assessment tests, and this paper is an attempt to share a detailed review of the decisions and considerations made throughout the test development. With this paper, we intended to transparently communicate what aspects contributed to the final version of USAT. It is hoped that USAT will be a valuable addition to the assessment tools focusing on the Urdu language which has been understudied despite having such a large number of speakers worldwide.

References

Akram, B. (2002). *Analysis of Urdu syllabification using maximal onset principle and sonority sequence.* Center for Research in Urdu Language Processing, National University of Computers and Emerging Sciences.

Ambreen, S. & To, C. K. S. (2021). Phonological development in Urdu-speaking children: A systematic review. *Journal of Speech, Language, and Hearing Research, 64*(11), 4213-4234. http://dx.doi.org/10.1044/2021_JSLHR-21-00148.

Ambreen, S. & To, C. K. S. (2022, May 22 – 25). A review of the Urdu phonological system [Poster presentation]. *National Conference of Speech Pathology Australia (SPA 2022)*. Australia.

Bhatti, R., Mumtaz, B., & Hussain, S. (2016). Identification of Diphthongs in Urdu and their Acoustic Properties [Poster presentation]. *Conference on Language and Technology (CLT 16)*, Pakistan.

Centre for Language Engineering (2015). *Urdu phonetic inventory*. Retrieved from www.cle.org.pk

Eberhard, D. M., Gary, F. S., & Charles, D. F. (Eds.). (2022). *Ethnologue: Languages of the world* (25th ed.). SIL International. Retrieved July 18, 2022, from https://www.ethnologue.com/.

Eisenberg, S. & Hitchcock, E. R. (2010). Using standardized tests to inventory consonant and vowel production: A comparison of 11 tests of articulation and phonology. *Language Speech and Hearing Sciences in Schools, 41*, 488-503. http://dx.doi.org/10.1044/0161-1461(2009/08-0125).

Fabiano-Smith, L. (2019). Standardized tests and the diagnosis of speech sounds disorders. *Perspectives of the ASHA Special Interest Groups, 4*(1), 58-66. http://dx.doi.org/10.1044/2018_PERS-SIG1-2018-0018

Friberg, J. C. (2010). Considerations for test selection: How do validity and reliability impact diagnostic decisions? *Child Language Teaching and Therapy, 26*(1), 77-92.

Frytak, J. C. (2000). Measurement. *Journal of Rehabilitation Outcomes, 4*(1), 15-31.

Ghazali, M. A. (2002). Urdu syllable templates. *CRULP Annual Student Report 2001-2002.* Center for Language Engineering, University of Engineering and Technology.

Khan, M. A. (1997). *Urdu ka sauti nizam.* National Language Promotion Department, Pakistan

Khanum, K., & Hussain, S. (2017). Development of noun inflectional morphology in children acquiring Urdu and Punjabi. *Journal of Research (Urdu)*, 32.

Khurshid, K., Usman, S. A., & Butt, N. J. (2003). Possibility of existence and identification of diphthongs and triphthongs in Urdu language. *CRULP Annual Student Report 2003–2004*. Center for Language Engineering, University of Engineering and Technology.

Li, X. X. & To, C. K. S. (2017). A review of phonological development of Mandarin-speaking children. *American Journal of Speech-Language Pathology, 26*(4), 1262–1278. http://dx.doi.org/10.1044/2017_AJSLP-16-0061.

McLeod, S. (2012a). Translation to practice: Creating sampling tools to assess multilingual children's speech. In S. McLeod and B. A. Goldstein (Eds.). *Multilingual Aspects of Speech Sound Disorders in Children.* Multilingual Matters.

McLeod, S. (2012b). *Multilingual speech assessments.* Bathurst, NSW, Australia: Charles Sturt University. Retrieved from http://www.csu.edu.au/research/multilingual-speech/speech-assessments.

McLeod, S. & Crowe, K. (2018). Children's consonant acquisition in 27 languages: A cross-linguistic review. *American Journal of Speech-Language Pathology, 27*, 1546–1571. http://dx.doi.org/89.187.163.213 on 03/29/2021.

McLeod, S. & Verdon, S. (2014). A review of 30 speech assessments in 19 languages other than English. *American Journal of Speech-Language Pathology, 23*. 706–723. http://dx.doi.org/10.1044/2014_AJSLP-13-0066.

Muzaffar, H., Iftikhar, N., Khan, M. S. G., & Malik, S. N. (2019). Development of articulation screening tool to assess the articulation errors in Urdu-speaking school-going children. *UMT Education Review, 2*(1), 23–37.

Nawaz, S. (2002). Deletion rules in Urdu language. *CRULP Annual Student Report 2001–2002*. Centre for Language Engineering, University of Engineering and Technology.

Nazar, M. N. (2002). Syllable templates in Urdu language. *CRULP Annual Student Report 2001-2002.* Centre for Language Engineering, University of Engineering and Technology.

Noveen, S., Butt, A. K., & Alam, M. B. (2017). Development of a test for articulation and phonological disorders in Urdu-speaking children. *The Journal of Riphah College of Rehabilitation Sciences (JRCRS), 5*(2), 89-93.

Nyyar, S. (2003). Syllable stress in Urdu. *CRULP Annual Student Report 2001-2002.* Centre for Language Engineering, University of Engineering and Technology.

Phạm, B. & McLeod, S. (2019). Vietnamese-speaking children's acquisition of consonants, semivowels, vowels, and tones in northern Vietnam. *Journal of Speech Language and Hearing Research, 62*(8), 2645–2670. http://dx.doi.org/10.1044/2019_JSLHR-S-17-0405.

Ranjha, M. I. (2012). Urdu syllable: Templates and constraints. *Language & Technology.* Retrieved June 28, 2021, from https://www.tehqeeqat.com/downloadpdfc/22.

Ranjha, M. I. (2014). Stability of consonant clusters in Urdu. *ELF Annual Research Journal, 16*, 137–156.

Saleem, A. M., Kabir, H., Riaz, M. K., Rafique, M. M., Khalid, N., & Shahid, S. R. (2003). Urdu consonantal and vocalic sounds. *CRULP Annual Student Report 2001-2002.* Centre for Language Engineering, University of Engineering and Technology.

Shabbir, H., & Zahra, N. (2016). *Wordlist and manual for articulation and phonological assessment* (Unpublished material). Centre for Clinical Psychology, University of the Punjab.

Shafqat, F. (2020). Struggling field of speech-language pathology in Pakistan (Editorial). *The Rehabilitation Journal, 4*(2), 165-166.

Shriberg, L. D. & Kwiatkowski, J. (1982). Phonological disorders I: A diagnostic classification system. *Journal of Speech and Hearing Disorders, 47,* 226–241. http://dx.doi.org/0022-4677/82/4703-0226501.

Stow, C. & Pert, S. (2020). *Bilingual Speech Sound Screen (BiSSS) Revised and expanded edition: Pakistani Heritage Languages: Mirpuri, Punjabi, and Urdu. The* University of Manchester.

Weston, A. D., Shriberg, L. D., & Miller, J. J. (1989). Analysis of speech samples with SALT and PEPPER. *Journal of Speech and Hearing Research, 32*, 755-766. http://dx.doi.org/0022-4685/89/3204-0755501.00/0

Table 1.

Results of the expert consultation survey

Q1.	Is this word appropriate to elicit the target sound(s) at the mentioned position?
	- 97.5% of test words were judged to be appropriate by all respondents. - 1.9% of test words were judged to be appropriate by 4/5 respondents. - The remaining test word was judged to be appropriate by 3/5 respondents.
Q2.	Does the word exist in the vocabulary of Urdu-speaking children aged three years and above? If this word is not in the vocabulary of young children, can they produce it in imitation?
Q3.	
	- 93.3% of words were judged to be present in the vocabulary of young children. - For the remaining 6.7% of words, 3/5 respondents judged that even though they may not occur in all children's vocabulary, children can produce them in imitation.
Q4.	Is this word culturally appropriate for young children?
	- All of the test words were rated as culturally appropriate by all respondents.

Table 2.

Modifications made to the test words based on the pilot study

Old item			New item		
Urdu Word	Transcription	Old position	Urdu Word	Transcription	New position
جوتا	ʤuːʈaː	ʤ = WI	جوتے	ʤuːʈeː	Same
جھولے	ʤʰuːleː	ʤʰ = WI	جھولا	ʤʰuːlaː	Same
بجھنا	bʊʤʰnaː	ʤʰ = WM	بجھا	bʊʤʰɑː	Same
بچہ	bəʧʧaː	ʧ = WM	بچے	bəʧʧeː	Same
چبھنا	ʧʊbʰnaː	bʰ = WM	چبھ	ʧʊbʰ	bʰ = WF
باندھنا	bãːndʰnaː	ɖʰ = WM	باندھ	bãːndʰ	ɖʰ = WF
سونگھنا	sũːŋgʰnaː	gʰ = WM	سونگھ	sũːŋgʰ	gʰ = WF
نلکا	nəlkaː	n = WI	ناک	naːk	Same
		l = WM	بلی	bɪlliː	Same
شاخ	ʃaːx	ʃ = WI	شادی	ʃaːɖiː	Same
		x = WF	چیخ	ʧiːx	Same
فراک	fərɔːk	r = WM	بارش	baːrɪʃ	Same

Note. The gray boxes show the new words added to the USAT.

Figure 1. Distribution of words based on the number of syllables.

Figure 2. Front page for the test item “pencil” /pæ̃ːnsɪl/ (پینسل)

Figure 3. Back page for the test item “pencil” /pæ̃ːnsɪl/ (پینسل)

Figure 4. Illustration for “cried” /roːi̯ː/ (روئی)