

# **The Use of Non-Word Key Words in the Speech Assessment of English Second Language Learners**

Lauren Ross<sup>1,\*</sup>, Salomé Geertsema<sup>1</sup>, Mia le Roux<sup>1</sup>, and Marien Alet Graham<sup>2</sup>

<sup>1</sup>Department of Speech-Language Pathology and Audiology, University of Pretoria;

<sup>2</sup> Department of Science, Mathematics and Technology Education, University of Pretoria.

\*Correspondence concerning this article should be addressed to:

Lauren Ross, Cluster Box 6615, Le Domaine 3626, South Africa.  
Email: u14012554@tuks.ac.za; laurenross0441@gmail.com

## **Abstract**

Use of non-words are a potentially more appropriate method of assessment for EL2 learners. A mixed comparative design was used to compare the effects when using non-words instead of picture-based stimuli to assess articulation of EL2 learners. Sub-aims were to compare results between two tests and age groups. 16 Setswana L1 children assigned to two age cohorts were assessed using the GFTA-2, and non-word list created via a registered Speech Motor Learning website. Results of the two assessments differed significantly, indicating lack of semantic information may yield different outcomes for articulation assessments of EL2 learners. Speech sound differences on the GFTA-2 were sounds not found in L1. This agrees with previous research indicating incorrect diagnosis due to speech and language differences. There was no significant difference between the two age cohorts. This research forms the basis for investigations into non-words as a more accurate method for assessment of EL2 learners.

**Keywords:** assessment, cultural/linguistic, diversity, as a second language (ESL) /bilingualism/dialects, English/languages, articulation, speech/sound, phonemes/phonetics/phonology, speech/sound, speech-language pathologists (SLPs)

## Introduction

When Speech Sound Disorders (SSD) occur in conjunction with second language (L2) learning, long-term negative effects can result. These negative results are especially evident when intervention is not received before 5-6 years (McLeod et al., 2013; Miltiņa & Augstkalne, 2015). The impact of SSD may be greater when the Language of Learning and Teaching (LoLT) differs from the first language, as is the case with English second language (EL2) learners in South Africa (Miltiņa & Augstkalne, 2015). These differences create challenges for a speech-language pathologist (SLP) when trying to identify children with disorders in this population (Preston & Seki, 2011). The main difficulties identified by SLPs in the assessment of EL2 learners with SSD fall within three broad categories: a) a lack of norms for multilingual speech acquisition, b) a lack of confidence in differential diagnosis between SSD and speech difference, c) and a lack of culturally appropriate tools for assessment (McLeod et al., 2013).

The multilingual nature of South Africa is evident from the 11 official languages and additional unofficial languages (South African Statistics, 2011; Van Biljon et al., 2015). Only 9.8 % of the South African population are English first language (EL1) learners (South African Statistics, 2011). Despite this fact, in most South African classrooms English is the LoLT (ELoLT) regardless of the heterogeneous language context representing more than 20 different L1s (Málek, 2013; Van Biljon et al., 2015). In this multilingual environment, it is to be expected that the sound systems of the different languages influence one another. The 11 official languages themselves differ from one another in a number of ways (Niesler et al., 2005).

Differences in accents and dialect are mainly attributed to the different phoneme inventories of the languages (Niesler et al., 2005). English would differ from African languages such as Setswana, which is a Bantu language spoken in South Africa (Van der

Merwe & Le Roux, 2014). Some single consonants that occur in English words do not occur in the African languages. Examples of the sounds which appear in English but not Setswana are [z], [v], [θ] and [ð] (Niesler et al., 2005; Snyman, 1989). Setswana and English do not only differ in sound inventory, but also manner of articulation.

Aspiration occurs frequently in Setswana, while in English aspiration happens but is not phonemic. Ejection does not occur in English, but it is part of the production process of consonants in Setswana (Niesler et al., 2005). The dorso-velar fricative [x] occurs in Setswana but not in English, while the lamino-alveolar fricative [z] occurs in English but not in Setswana. Setswana uses the trill [r], instead of the [ɹ] found in English (Cole, 1955). The number of affricates in African languages is significantly greater than the number used in English. Furthermore, English employs none of the clicks used in many African languages (Niesler et al., 2005). English contains many consonant clusters, such as [bl], [bɹ], [spɹ] and [dɹ] which do not occur in Setswana. In addition, Setswana frequently uses clusters (for example [tʰ]), which do not appear in English (Snyman, 1989; Ziervogel, 1967).

In addition to the differences present in the consonant inventories of English and the African languages, vowels differ across these languages. Research established that EL2 speakers find it difficult to distinguish all the vowels of English (Le Roux, 2016; Seeff-Gabriel, 2003). Nineteen vowel sounds have been distinguished in English while many African languages have a reduced vowel system (Bekker, 2009). An example of this limited vowel inventory is that of Setswana which consists of only seven basic vowels, four raised vowels and no diphthongs (Cole, 1955; Snyman, 1989). The four tense vowels [a], [i], [u] and [o] are used in Setswana, but are infrequent in English (Niesler et al., 2005). In contrast, English frequently contains lax vowels which do not occur in Setswana (Niesler et al., 2005). Additionally, languages also differs concerning syllabic structure, the CVC structure (closed

syllable) is found in English but not Setswana (Van der Merwe & Le Roux, 2014). On the other hand, Setswana can have a syllable made up of only one consonant (C) which does not occur in English (Van der Merwe & Le Roux, 2014). Multilingualism is affected not only by the differences between languages, but also by the interaction between the two languages.

Speech differences and variation in pronunciation can occur due to interaction between two speech systems (Zajac, 2016). Specifically, bilingual children keep the two languages separate to a large extent, but the two systems can still influence each other (Zajac, 2016). Bekker and Eley (2007) describe South African English spoken by EL1 speakers as being distinct from the English spoken by other ethnic groups in South Africa. This distinction is shown by the different accents linked to the different first languages spoken by EL2 speakers (Moonsamy & Kathard, 2015). Multilingualism is also affected by the speakers' language proficiency.

The complications in assessment of English SSDs presented by multilingualism, differing phoneme inventories, and differing syllabic structures are exacerbated by the different levels of English competency, as seen in the most recent assessment of the literacy skills of international learners, the Progress in International Reading Skills Study (PrePIRLS) (Howie et al., 2017). South Africa scored the lowest out of all participating countries (Howie et al., 2017). The 49% of participants (Grade 4 and 5) who wrote the test in a LoLT which was different from their first language scored significantly lower than those who wrote in their first language (i.e., their LoLT and first language are the same) (Howie et al., 2017). The difference in scores between EL1 and EL2 learners was equivalent to approximately two years of schooling (Howie et al., 2017). These results show that when children are taught in a second language, it can impact their literacy skills. Furthermore, difficulties with precise pronunciation of the sounds of a language also affect literacy skills (Miltiņa & Augstkalne, 2015). An in-depth SLP assessment is important to prevent or minimize these effects of L2

learning. Unfortunately these assessments come with unique challenges in the multilingual and multicultural context.

Kadyamusuma (2016) reports that even after an articulation assessment an EL2 learner could be misdiagnosed with the presence (or absence) of an SSD due to the occurrence of speech differences, dialects and variation in first and subsequent languages. Preston and Seki (2011) therefore stress the importance of differentiating between SSD and speech differences associated with speaking a second language.

Assessment of this population is problematic, not only because of the interaction between the different languages, but also due to assessments being administered by speakers of the mainstream dialect, and assessments being designed exclusively for this dialect (Southwood, 2013). A better understanding of an EL2 child's speech system could be obtained through a bilingual assessment. However, this is an unrealistic goal in the South African context (Pascoe & Norman, 2011). According to Caesar and Kohler (2007) the homogeneity of a predominantly monolingual SLP workforce is also frequently reported globally, especially in English-speaking countries. Preferably, an SLP should have a similar language background to the client and speak the same languages, and even have the same dialect (Goldstein & Gildersleeve-Neumann, 2015). Unfortunately the situation in South Africa is not ideal, as the majority of SLPs speak English or Afrikaans. However, they are aware of their lack of knowledge of African languages, and its effect on assessment (Wilsenach, 2016).

The multilingual and -cultural society, coupled with the lack of culturally, linguistically, and contextually appropriate assessment tools creates assessment challenges for the homogenous SLPs in South Africa (Barratt et al., 2012; Pascoe et.al., 2013; Southwood & Van Dulm, 2015). Pascoe and Norman (2011) suggest the consideration of informal dynamic assessment methods to address the challenges posed by inappropriate

assessment tools. Unfortunately, the lack of norms against which the results can be compared, and limited knowledge about typical development of many African languages, result in dynamic assessment being only a partial solution (Pascoe & Norman, 2011). Southwood and Van Dulm (2015) acknowledge the need for valid and reliable normative assessment tools as these are the cornerstones of effective intervention. The proposed partial solution of the use of informal dynamic assessment methods in the South African context therefore leads to less valid and reliable assessment outcomes. We suggest that present methods of formal and normative assessment should be revisited to investigate possible adaptations to these reliable tools, rather than augmentation with informal methods.

The Goldman-Fristoe Test of Articulation (Goldman & Fristoe, 2000) is a picture-based articulation test which is the most commonly used test in the urban Gauteng area (Van Biljon et al., 2015). Picture naming is affected by abilities other than articulation, including memory, visuo-spatial skills, vocabulary, visual perception, and lexical-semantic abilities (Yochim et al., 2009). Wen-hui (2016) found that a picture-naming task slows the EL2 learner down and demonstrates a high rate of errors in this population. During a picture-naming task, EL2 learners produce words at a slower rate, and therefore have reduced word output. This word output disadvantage can be due to competition between the two languages (Gollan et al., 2008). If the sounds of the target words in the second language are similar to those of the first language, the ability to name pictures in the second language will be reduced (Hermans et al., 1998; Lee & Williams, 2001). This phenomenon is referred to as cross-linguistic interference, which purports that the activation of the first language affects the processing of the second language (Wen-hui, 2016). The disadvantage of word output can be attributed to this competition between the two languages (Gollan et al., 2008). Wen-hui (2016) stated that the phoneme inventory/sound repertoire of the first language will have a greater impact on the second language than vice versa.

In addition to the influence of the phoneme inventory of the L1, South African authors Van Biljon et al. (2015) found that when assessing EL2 learners using the picture-based GFTA-2, the target words were not always elicited as expected from the stimulus. An example of this is the picture of the wagon. The wagon is an object most children in the urban South African context have little experience with, they may instead call it a wheelbarrow, as this is more familiar to them (Van Biljon et al., 2015). All these challenges result in the need for innovative research to bridge the gap between SSD and speech sound difference.

Non-word keywords may eliminate misinterpretation of results associated with different dialects, tools that rely on culturally inappropriate world knowledge, and limited vocabulary (Shriberg et al., 2009). As such, these words have been used in the treatment of motor speech disorders such as apraxia of speech (AOS) (Van der Merwe, 2011). The rationale for the use of non-words in the treatment of AOS consists of three parts. Firstly, when using non-words in treatment there is freedom to vary the speech sounds included. This freedom means all consonants and vowels can be targeted in the CVCV syllable structure (Van der Merwe, 2011). Secondly, non-words still allow for interaction between adjacent syllables (Van der Merwe, 2011) therefore sounds are still produced in varied phonetic contexts. These non-words will contain familiar syllable structures, as it must adhere to the phonotactic rules of a language (Van der Merwe, 2011). Thirdly, non-words have no semantic meaning and are therefore more likely to activate the neural motor areas, without additional activation of language areas (Van der Merwe, 2011). Finally, the second language learner will not struggle to find the necessary vocabulary when assessed. It is therefore reasonable to assume that non-words may assist with the challenges presented by dialect, culturally inappropriate tools and vocabulary.

Non-word repetition is a task used not only in the treatment of motor speech disorders, but also the assessment of children with language and literacy impairments, such as

Specific Language Impairment (SLI) (Bishop et al., 1996; Shriberg et al., 2009). Non-words have been used to diagnose children with possible SLI, and other impairments where semantic content may interfere with performance (Reuterskiöld & Grigos, 2015). The rationale for their use in this regard is that the participant cannot access lexical information from long-term memory, which would support their performance (Reuterskiöld & Grigos, 2015). The lexical units comprise the vocabulary of a specific language. The lexical semantic value dictates the way in which the lexical units correlate with the syntax of that language. The value of non-words, therefore, lies not only in the lack of semantic information, but also in the lack of specific syntax or structure.

Non-word keywords are mostly made up of two or more syllables. This syllabic composition is due to their use in the treatment of people with AOS, as these speakers use shorter words than normal speakers (Edmonds & Marquardt, 2004). This structure generally represents the CVCV structure at the lowest difficulty level (Van der Merwe, 2011). In almost all world languages, infant babbling and early words include the CV syllable. The basic syllable structure has therefore become the preferred basic unit of speech articulation (Van der Merwe, 2011). Reduplicated babbling, CVCV, is also common in the first 50 words of children from different language groups (Edwards & Shriberg, 1983). Therefore, children from different linguistic backgrounds could be assessed using non-words to identify speech sound differences and articulation errors not associated with learning a second language.

Results of the Shriberg et al. (2009) study indicate that non-word tasks could possibly discriminate between SSD and speech differences. Shriberg and his colleagues (2009) stated that non-word repetition difficulties indicate a deficit in memorial processes. Such deficits have been identified as contributing to impairments such as language delays and other verbal traits such as SSD (Shriberg et al., 2009). Children with SSD present with poor performance on non-word repetition tasks (Shriberg et al., 2009). Shriberg et al. (2009) concluded that



some of these errors are also due to habitual misarticulations, another indication of SSD. Shriberg et al. (2009) results agree with previous findings by Kovas et al. (2005) which stated that the results of a non-word repetition task are sensitive to articulatory accuracy in addition to phonological short term memory. These research studies suggest there is a possibility to use non-words for assessment of SSD, as speech sound differences and articulation errors would still show up in results, but those productions made due to language difference would not be seen.

Van der Merwe (2011) suggested that the use of non-words of a CVCV structure will not activate the language areas of the brain. It should therefore be possible to assess the speech production of an EL2 learner without semantic or linguistic interference. English differs from Setswana in both speech sound inventory and grammatical structure (Bekker, 2009; Niesler et al., 2005; Snyman, 1989; Van der Merwe & Le Roux, 2014). A non-word list can be varied to include speech sounds and phonological rules specific to a language. As such, it is possible that the use of non-words could be used to identify only misarticulations due to SSD, as opposed to those resulting from interaction between the first language of the learner and English, as indicated by Shriberg et al. (2009). Therefore, the aim of the present research was to identify the difference in results when non-words are used, as SSDs could be identified without semantic and linguistic interference resulting from EL2 learning.

The null hypothesis ( $H_0$ ) stated that despite the lack of semantic information, the use of novel, non-word keywords instead of a picture-based SSD test such as the GFTA-2 would not yield significantly different results in the assessment of articulation skills of EL2 learners. The alternative hypothesis ( $H_a$ ) stated that the use of novel non-word stimuli instead of a picture-based SSD test such as the GFTA-2 would yield significantly different results in the assessment of articulation skills of EL2 learners due to the lack of semantic information.

## **Method**

### **Research Design**

This study employed a mixed comparative group design. This design allowed the researcher to compare each participant's results when assessed using the non-word test as opposed to the GFTA-2 (Each participant is assessed using both tests to allow comparison). In addition comparison could be made between the different groups to identify differences based on age.

### **Ethical considerations**

Ethical clearance was obtained from the Research Committee of the Department of Speech-Language Pathology and Audiology, University of Pretoria, South Africa. The researchers received written permission from the private school, allowing the research to be conducted on their premises. Written consent was granted by the parents or legal guardians for their children to participate in the study. The participants granted written consent by writing their names on the consent form, as well as assent by selecting the correct illustration (tick vs. cross/ happy vs. sad face) on a form of assent. Participant anonymity was ensured through the use of numeric codes.

### **Participants**

#### **Selection criteria for participants.**

The study included 16 participants assigned into two groups using purposive sampling as the groupings were based on age. Group 1 consisted of seven 6-7 year old EL2 learners. Group 2 consisted of nine 8-10 year old EL2 learners. All participants were required to have English as their Language of Learning and Teaching (LoLT) and Setswana as their first language. Male and female participants were selected. Participants came from similar socio-economic backgrounds. This socio-economic status is assumed due to all participants being

selected from the same private school. Participants were aged between 6;0 and 9;11 [years; months] at the time of the study

Potential participants were excluded if they had structural deficits (e.g., cleft palate), hearing loss, autism spectrum disorder, a language disorder in their first language, or a previously diagnosed SSD (phonological disorder, articulation disorder or dysarthria). Learners who display these challenges were excluded as the present study only aimed to determine the assessment value for typically developing EL2 speakers. These excluded learners were referred for further assessment and treatment by the qualified SLP working at the school.

### **Sample size and sampling method.**

A convenience sampling process was used. Participants were selected from a single private primary school (Gr 1-Gr 3). This private school is located in Hammanskraal, a rural area in Northern Gauteng, South Africa. The intended sample size was approximately 24 Setswana speaking children. Unfortunately, due to limited willingness of the parents to allow their children to participate, and varying first languages in the school, only 19 children were assessed. Of the 19 participants, two were excluded from the research results as they presented with speech delays or disorders. One participant withdrew from the research due to illness. Therefore, only 16 participants were included in the results.

### **Data collection**

#### **Data collection material.**

During the pre-experimental assessment quantitative data were obtained from a parental questionnaire, the Test of Auditory Processing Skills - Third edition (TAPS-3) (Martin & Brownell, 2005), the Phonological Awareness Test- Second edition (PAT-2) (Robertson & Salter, 2007), the Test of Auditory Comprehension of Language - Fourth

edition (TACL-4) (Carrow-Woolfolk, 2014), the Tswana Expressive Receptive Language Assessment (TSWERLA) (Bortz, 1997), the Oral Facial Examination (OFE) (Shipley & McAfee, 2004), and the Gilliam Autism Rating Scale - Second edition (GARS-2) (Gilliam, 2005). The experimental assessment materials consisted of the Goldman-Fristoe Test of Articulation - Second Edition (GFTA-2) (Goldman & Fristoe, 2000), and an articulation assessment using a non-word list compiled on the Van der Merwe (2011) SML website (Appendix A).

### **Data collection procedures.**

#### ***Pre-experimental assessment***

The researcher sent out letters of consent to parents before data collection commenced. When consent forms were received, potential participants were assessed in a venue at the school. The researcher, an EL1 SLP, assessed all potential participants using the material as described. For the TSWERLA assessment (Bortz, 1997), a Setswana first language speaker pre-recorded the stimuli sentences and phrases. These recordings were used by the EL1 SLP to administer and record responses to the Setswana language test. Potential participants identified as having delays or disorders were excluded from the study and referred for further testing.

#### ***Experimental assessment***

During the experimental assessment, participants were assessed in the same venue by the EL1 SLP, using the material as described. The SLP transcribed each response from the auditory recording using the International Phonetic Alphabet (IPA).

Participants were assessed using the GFTA-2 Sounds in word section. If a participant presented with an SSD in their first language their results were excluded from the study, this was determined through a comparison of the GFTA-2 results and their speech in Setswana

(i.e. if sounds found in Setswana were incorrectly produced on the GFTA-2, the participant was not stimuable, and the same sound was produced with a difference during the TSWERLA test, the participant was excluded due to an SSD).

The testing procedure for the non-word test was different for the two age groups:

*6-7 year olds:* With access to the face of the examiner, the words were read out loud once.

The participant was requested to repeat the word. Request for clarity or repetition was allowed once.

*8-10 year olds:* The participants were presented with cards displaying single non-words, and were asked to read the word. This method follows that in the decoding section of the PAT-2.

In cases where the participants were unable to accurately read the word, the examiner assisted the participant by “reading” the word to the participant once, and then asked for oral repetition. Allowing for oral repetition follows the administration guidelines used in the GFTA-2 (Goldman & Fristoe, 2000).

### **Data analysis**

The results of the different tests performed were compared in regards to ease of first time elicitation of the target words. In a study on dynamic assessment, it was found that children with a possible language disorder required more effort on the part of the examiner (Peña, Quinn & Iglesias, 1992). The amount of effort needed from the examiner can be an important measure to determine the amount and intensity of mediation needed to facilitate future learning (Gutierrez-Clellen, Conboy, Brown & Robinson-Zañartu, 1998). In cases where the participant struggled to read the word (older participants) or asked for repetition (younger participants), the assistance with reading as well as the once-off allowed repetition, were accepted as the first time elicitation of the target words. In other words, first time elicitation can be defined by how accurately the child produced the target word when

presented with the stimulus. The percentage correct production for each target word was calculated and statistically analyzed. This analysis made use of non-parametric methods due to the small sample size. The mean and SD were calculated and the Mann-Whitney U test and Wilcoxon Signed-rank test were run to compare the assessments. The same procedure was followed for the novel non-word list. Within- and between-participant comparisons were made and statistical differences calculated if present. Between-group comparisons were also made. The Wilcoxon signed-rank test is a nonparametric statistical test used to compare two related samples. Since the same participants completed Test 1 and Test 2, this test is used to compare the difference in results. The Mann-Whitney U is a nonparametric test used to assess for significant differences in a scale or ordinal dependent variable by a single dichotomous independent variable. The Mann-Whitney U test was used to compare the differences in results between the two age groups. The independent variable here is dichotomous, since the groups take on the values of either 'younger group' or 'older group'.

## **Results and Discussion**

### **Test Comparison**

Each sound with speech sound differences in either test was individually compared across the two tests. The comparison identifies the consistency of differences across the two tests, by identifying sounds with significant variation in the number of speech sound differences. This comparison was made to identify if sounds are produced incorrectly in one test but not the other and vice versa. In this way the effect of using non-words can be looked at for each sound. Test 1 (T1 to the GFTA-2 and Test 2 (T2) refers to the novel non-word list.

If the p-value is less than 0.05, there is a statistically significant difference between the tests for that individual sound. This would mean there is a significant variation in the number of differences for that sound on the different tests. On the other hand, if the p-value is

greater than 0.05, the differences between the tests are not statistically significant. The Wilcoxon signed-rank test statistics as well as the corresponding p-values are shown in Table 1.

**Table 1.** Comparison of individual sounds assessed in the GFTA-2 and non-word list

	Wilcoxon signed-rank test	p-value
i_initial_T2 - i_initial_T1	0.00	1.000
i_medial_T2 - i_medial_T1	-0.45	1.000
i_Final_T2 - i_Final_T1	-2.24	0.063
bɪ_T2 - bɪ_T1	-2.50	0.031*
dɪ_T2 - dɪ_T1	-1.34	0.375
tɪ_T2 - tɪ_T1	-1.00	0.625
ʃ_initial_T2 - ʃ_initial_T1	-1.41	0.500
ʃ_medial_T2 - ʃ_medial_T1	-1.00	1.000
ʃ_final_T2 - ʃ_final_T1	-1.00	1.000
ð_initial_T2 - ð_initial_T1	-1.41	0.500
ð_medial_T2 - ð_medial_T1	-1.41	0.500
θ_medial_T2 - θ_medial_T1	-1.00	1.000
θ_final_T2 - θ_final_T1	-1.00	1.000
dʒ_medial_T2 - dʒ_medial_T1	-1.00	1.000

\* Significant difference (p-value < 0.05)

Note: T1 = GFTA-2; T2 = Non-word list; GFTA=Goldman-Fristoe Test of Articulation

As the p-value for bɪ\_T1 and bɪ\_T2 is less than 0.05, the results for this sound differ significantly between the two tests. Of all the speech sound differences compared, only /bɪ/ showed a significant variation in number of differences between the two tests. This means the other speech sound differences had a similar number of variations in both tests, and the results between the two tests were statistically consistent for these sounds.

To identify if there is a significant difference between the speech sound differences on each test, an overall comparison was run. To compare the overall consistency of results across the two tests, the overall statistical differences between T1 and T2 are shown in Table 2.

**Table 2.** Comparison between the GFTA-2 and non-word List

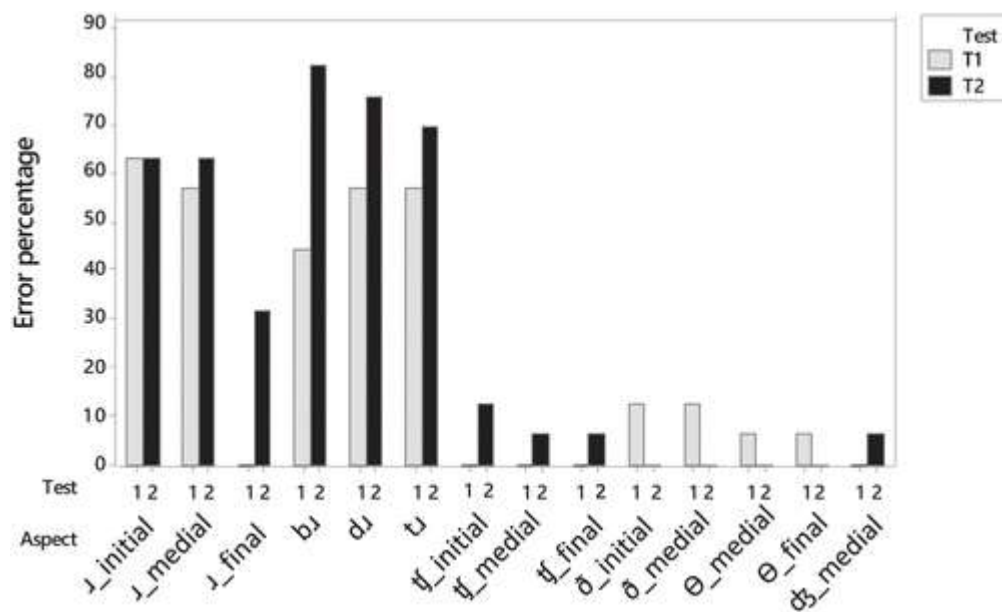
	Wilcoxon signed-rank test	p-value
T1 – T2	-2.469	0.020*

\* Significant difference (p-value < 0.05)

As the p-value equals 0.020, there is a statistically significant difference between T1 and T2. Although only /bɹ/ had a statistically significant difference across the tests, the overall comparison shows the two tests statistically have significantly different results. It is necessary to look at the individual sound differences regardless of the significance at an individual level. If an individual sound does not have a statistically significant difference between the two tests, it does not mean that minor differences on each sound will not add up to result in an overall significant difference across the two tests.

In order to more closely examine the variation between individual sounds, the percentage of differences for each sound in each test is represented in Figure 1.





**Figure 1.** Chart of Error Percentage

Figure 1 indicates that, although the variation is not significant for all sound differences, there are differences in results for each sound which, when combined, result in a significant variation between the two tests. This means that, overall, there is a significant difference in the number of sound differences on each sound across the two tests.

The null hypothesis (Ho) stated that the use of novel non-word keywords instead of a picture-based SSD test such as the GFTA-2 would yield significantly different results in the assessment of articulation skills of EL2 learners due to the lack of semantic information. This hypothesis is accepted as the results indicate that there is a significant difference in results between these two tests.

Figure 1 shows all the speech sounds which had differences and the percentage of participants who presented with these differences. These speech sound differences include /ɹ/ in all positions, /tʃ/ in all positions, /ð/ in initial and medial positions, /θ/ in medial and final positions, /dʒ/ in the medial position, and consonant clusters /bɹ/, /dɹ/, /tɹ/. Speech sounds /ɹ/, /ð/, /θ/ and consonant clusters /bɹ/, /dɹ/, /tɹ/ are English sounds, and do not occur in

Setswana, they were substituted for sounds from this sound inventory. For example alveolar approximant /ɹ/ was substituted with the alveolar trill /r/ which occurs in Setswana. It is important to note that the mentioned sounds do not occur in Setswana, as these speech sound differences would not be present if the non-word list followed the Setswana sound system. These findings agree with the report of Shriberg et al. (2009); removing the linguistic content from a speech assessment may allow SLP's who do not speak an African language to accurately assess the participants' articulation in their first language as they revert back to using that sound system. The speech sounds are /dʒ/ and /tʃ/ were also produced differently. These speech sounds do occur in Setswana. The /dʒ/ sound was changed to /j/ by one participant in the older group. The /tʃ/ sound was produced by one older participant as /k/, and by two older participants as /f/. For those participants who produced the non-words differently, the administrator gave the participants an auditory model of the word and asked for repetition. The participants were able to produce these sounds accurately once a correct model was produced. This accurate repetition supported the need for a single verbally modeled production in addition to the initially read attempt. We suggest that these speech sound differences were made not because of an articulation disorder or language difference, but rather due to limited literacy skills in their L2. Likewise, the participants produced the /dʒ/ and /tʃ/ sounds correctly in the GFTA-2, indicating production success and supporting possible reading failure.

The results of the GFTA-2 and non-word test were compared. It was found that /bɹ/ was the only sound with a statistically significant difference. However, the other speech sounds produced differently also present with variation between the two tests. This leads to an overall significant difference between the two tests. In summary, the use of non-words does have an effect on results, not on an individual sound level, but overall. Van der Merwe

(2011) states that the lack of semantic content is an important aspect regarding non-word application in the clinical setting. The present study suggests that the removal of semantic information induces significant differences when compared with a traditional picture-based test. The two tests contained exactly the same speech sounds therefore the significant difference in results can only be caused by the lack of semantic information.

The significant difference between the two tests implies that when the semantic information is removed, the participants reverted back to the Setswana sound system. As a result sounds such as /ð/ and /θ/ which the participants produced correctly in English words, was pronounced as /t/ in the non-words. The potential purpose of using non-words as an articulation test is to be able to assess a participant in their L1. The Setswana speakers used Setswana sounds when pronouncing the non-words; this supports the potential use of non-words as, without semantic information, the sounds of a specific language can still be targeted.

### **Age comparisons**

The participants of this study were assigned to two age cohorts to determine the functional effect for both groups. Results in the use of non-words may differ according to age groups due to literacy levels.

The two age groups are compared in regards to each individual speech sound difference in Table 3. The Mann-Whitney U test statistics as well as the corresponding p-values are shown in Table 3.

**Table 3.** Comparison between the different age groups for individual error sounds

	Mann-Whitney U	p-value
ɹ_initial_T1	26.5	0.633
ɹ_medial_T1	23.0	0.358
ɹ_Final_T1	31.5	1.000
bɹ_T1	31.0	1.000
dɹ_T1	31.0	1.000
tɹ_T1	31.0	1.000
ʃ_initial_T1	31.5	1.000
ʃ_medial_T1	31.5	1.000
ʃ_final_T1	31.5	1.000
ð_initial_T1	30.5	1.000
ð_medial_T1	30.5	1.000
θ_medial_T1	27.0	0.437
θ_final_T1	27.0	0.437
dʒ_medial_T1	31.5	1.000
ɹ_initial_T2	26.5	0.633
ɹ_medial_T2	26.5	0.633
ɹ_Final_T2	14.0	0.034*
Bɹ_T2	26.0	0.550
Dɹ_T2	21.5	0.262
Tɹ_T2	17.0	0.106
ʃ_initial_T2	24.5	0.475
ʃ_medial_T2	28.0	1.000
ʃ_final_T2	28.0	1.000
ð_initial_T2	31.5	1.000
ð_medial_T2	31.5	1.000
θ_medial_T2	31.5	1.000
θ_final_T2	31.5	1.000
dʒ_medial_T2	28.0	1.000

\* Significant difference (p-value < 0.05)

The only p-value that is less than 0.05 and therefore statistically different is for ‘I\_Final\_T2’. Thus, there is a statistically significant difference between the younger and older groups, only for ‘I\_Final’ in Test 2.

When assessed using the GFTA-2, none of the participants made speech sound differences on /ɹ/ final position. During the non-word assessment, only participants in the older group made changes to this sound by varying the approximant /ɹ/ to the trill /r/. The younger group was provided with an auditory model which used the approximant /ɹ/ and as a result did not switch to the trill sound. This implies that when the older participants were able to freely read the sounds (not provided with a verbal model), many of them reverted back to the Setswana speech sound inventory and produced the trill /r/. These sound transfers from one language to another indicates a cross-linguistic effect (Goldstein & Gildersleeve-Neumann, 2015). A cross-linguistic effect occurs when a bilingual child uses a specific sound in their second language because it is part of their first language speech sound inventory (Goldstein & Gildersleeve-Neumann, 2015). In the case of the Setswana speaking participants cross-linguistic transfer occurred as the /r/ sounds from Setswana is used in the English speech production. The results of this study confirm Goldstein and Gildersleeve-Neumann’s (2015) statement that, although children keep the two languages separate, there is interaction between the two systems. For the overall differences between the age groups for T1 and T2 the statistics are shown in Table 4.

**Table 4.** Comparison between the age groups when assessed using the GFTA-2 and non-word list

	Mann-Whitney U	p-value
T1	5824.0	0.335
T2	5404.0	0.054

Since both p-values in Table 4 are greater than 0.05, there are no statistically significant differences between Test 1 for the younger and older groups as well as no statistically significant difference between Test 2 for the younger and older groups. The comparison between the two age groups provides information about the effect of removing semantic information at different age and literacy levels.

Despite the different methods used for the different age groups it is important to note that there is no significant difference in results. Non-words cannot be represented pictographically, therefore they need to be read. Since the younger group (grade 1) does not have the necessary literacy level, the examiner is required to provide an auditory model for these participants. In assessments such as the GFTA-2 the use of an auditory model can only be used to assess stimulability of the sound, but the error is still recorded. The lack of a significant difference between the age groups implies that the use of an auditory model for the younger group did not significantly affect the results obtained. As can be seen in Figure 1, it is evident that initial incorrect productions of the /dʒ/ and /tʃ/ may relate to the incorrect reading of these sounds. It is therefore important to consider not asking older children to read the non-words and instead have all children repeat after an auditory model. In other words, due to inadequate literacy levels of the children assessed in this study, it may be necessary to remove literacy skills from the assessment. Further research is needed to address this finding as, although the comparison between the age groups indicated no significant difference, use of an auditory model may impact the validity of the assessment.

### **Limitations or weaknesses**

Due to the explicit inclusion criteria for this study specifically (language background and matched socio-economic status determined by being in the same school) there was a very limited pool of potential participants. This limited group was further diminished by a lack of

willingness for parents to allow their children to participate. As a result, this study has a very small sample size, and the results may not be representative of a larger portion of the target population. Additionally there are other factors, which may have limited the strength of the results. There was variation in the elicitation methods from picture naming to reading to repetition and in the canonical structure of the test stimuli between the two test conditions. Furthermore, there was a small difference in the number of stimulus items between the two test conditions. Finally, the sound differences resulting from second language influence were not calculated as errors.

### **Recommendations for future studies**

If the non-word list is created using only Setswana speech sounds, and aligned with the linguistic rules of this language, these speech sound differences would not be present. The results indicate that in the presence of SSD the speech sound differences found in the GFTA-2 would also be present when assessed using a non-word list. If the participants' results on the non-word list were interpreted using the Setswana sound system, no errors would be identified. This would reduce the risk of over-diagnosing children as having SSD, in the presence of a speech difference.

In addition, when the non-word list is used to assess older children, they were required to read the non-words; therefore some speech sound differences made were the result of literacy errors as opposed to articulation difficulties. More research is needed to identify the effect when EL2 learners read non-words, as well as the assessment validity when an auditory model of the non-words is used. Furthermore, there are languages internationally that have not been included in traditionally used standardized norming samples. Therefore, research into the use of non-words with other language groups is recommended.

Finally, the limitations mentioned previously should be considered in future studies. The challenges of using non-words as stimuli with African languages having different structures than English call for rigorous research designs.

### **Conclusion**

This investigation shows that speech assessment results of EL2 learners differ when non-words are used instead of the GFTA-2 as assessment items. We suggest that these differences may be directly related to the lack of semantic information because the GFTA-2 and novel non-word list contain the same speech sounds in the same structure - the only difference is the lack of semantic information. In addition, when assessed using non-words EL2 learners reverted back to the sound system of their L1. We propose that the use of a novel non-word list containing only speech sounds found in the participants' L1, may offer improved articulation assessment opportunities in regards to diagnosis of SSD. Further research is needed into this potential use of non-words.

### **References**

- Barratt, J., Khoza-Shangase, K., & Msimang, K. (2012). Speech-language assessment in a linguistically diverse setting: preliminary exploration of the possible impact of informal “solutions” within the South African context. *The South African Journal of Communication Disorders. Die Suid-Afrikaanse Tydskrif Vir Kommunikasieafwykings*, 59(1), 34–44. <https://doi.org/10.7196/SAJCD.51>
- Bekker, I. (2009). *The vowels of South African English*. Potchefstroom University for Christian Higher Education.
- Bekker, I., & Eley, G. (2007). An acoustic analysis of White South African English (WSAfE) monophthongs. *Southern African Linguistics and Applied Language Studies*, 25(1),



107–114. <https://doi.org/10.2989/16073610709486449>

Bishop, D., North, T., & Donlan, C. (1996). Nonword Repetition as a Behavioural Marker for Inherited Language Impairment: Evidence From a Twin Study. *The Journal of Child Psychology and Psychiatry*, 37(4), 391–403.

Bortz, M. A. (1997). *Tswana Expressive Receptive Language Assessment*.

Caesar, L., & Kohler, P. (2007). The State of School-based Bilingual Assessment: Actual Practice Versus Recommended Guidelines. *Language, Speech, and Hearing Services in Schools*, 38(3), 190–200.

Carrow-Woolfolk, E. (2014). *Test of Auditory Comprehension of Language* (4th ed). Pro-ed.

Cole, D. T. (1955). *An introduction to Tswana grammar*. Capetown: Longman Penguin S.A.

Edmonds, L. A., & Marquardt, T. P. (2004). Syllable use in apraxia of speech: Preliminary findings. *Aphasiology*, 18(12), 1121–1134. <https://doi.org/10.1080/02687030444000561>

Edwards, M., & Shriberg, L. (1983). *Phonology: Applications in communicative disorders*. San-Diego: College-Hill Press.

Gilliam, J. E. (2005). *Gilliam Autism Rating Scale* (2nd ed). Pearson.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation* (2nd ed). Pearson.

Goldstein, B. A., & Gildersleeve-Neumann, C. E. (2015). Bilingualism and Speech Sound Disorders. *Current Developmental Disorders Reports*, 2(3), 237–244.

<https://doi.org/10.1007/s40474-015-0049-3>

Gollan, T., Montoya, R., Cera, C., & Sandoval, T. (2008). More use almost always a means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal*

*of Memory and Language*, 58(3), 787–814. <https://doi.org/10.1021/nl061786n>. Core-Shell

Gutierrez-Clellen, V., Conboy, B., Brown, S., & Robinson-Zañartu, C. (1998). Modifiability  
A Dynamic Approach to Assessing Immediate Language Change. *Communication Disorders Quarterly*, 19, 31-42. 10.1177/152574019801900204.

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a  
foreign language: Can speakers prevent interference from their first language?  
*Bilingualism: Language and Cognition*, 1(3), S1366728998000364.  
<https://doi.org/10.1017/S1366728998000364>

Howie, S., Combrinck, C., Tshele, M., Roux, K., Mcleod Palane, N., & Mokoena, G. (2017).  
*PIRLS Literacy 2016 : South African Highlights Report*. Pretoria.

Kadyamusuma, M. R. (2016). Limitations to Practicing as a Bilingual Speech-Language-  
Pathologist in the United States. *Journal of Speech Pathology & Therapy*, 1(2), 1–2.  
<https://doi.org/10.4172/2472-5005.1000e103>

Kovas, Y., Hayiou-Thomas, M., Oliver, B., Dale, P., Bishop, D., & Plomin, R. (2005).  
Genetic Influences in Different Aspects of Language Development: The Etiology of  
Language Skills in 4.5-year-old Twins. *Child Development*, 76(3), 632–651.

Le Roux, M. (2016). *An acoustic investigation of English vowels as produced by English L1  
and Setswana L1 foundation phase learners*. University of Pretoria.

Lee, M., & Williams, J. (2001). Lexical Access in Spoken Word Production by Bilinguals:  
Evidence From the Semantic Competitor Priming Paradigm. *Bilingualism: Language  
and Cognition*, 4(3), 233–248.

- Málek, H. L. (2013). *The acquisition of English prepositions in first language speakers of Northern Sotho and Afrikaans : A cognitive linguistic study*. University of Johannesburg.
- Martin, N. A., & Brownell, R. (2005). *Test of Auditory Processing Skills* (3rd ed). Academic Therapy Publications.
- McLeod, S., Verdon, S., & Bowen, C. (2013). International Aspirations for Speech-Language Pathologists' Practice with Multilingual Children with Speech Sound Disorders: Development of a Position Paper. *Journal of Communication Disorders.*, 46, 375–387.
- Miltiņa, I., & Augstkalne, D. (2015). Factors, Influencing The Emergence Of Child Speech Disorder. *Society. Integration. Education.*, 3, 153–164.
- Moonsamy, S., & Kathard, H. (2015). *Speech-Language Therapy in a School Context: Principles and Practices*. Van Schaik Publishers.
- Niesler, T., Louw, P., & Roux, J. (2005). Phonetic Analysis of Afrikaans, English, Xhosa and Zulu Using South African Speech Databases. *Southern African Linguistics and Applied Language Studies*, 23(4), 459–474. <https://doi.org/10.2989/16073610509486401>
- Pascoe, M., & Norman, V. (2011). Contextually relevant resources in speech-language therapy and audiology in South Africa--are there any? *South African Journal of Communication Disorders*, 58, 2–5.
- Pascoe, M., Rogers, C., & Norman, V. (2013). Are we there yet ? On a journey towards more contextually relevant resources in speech-language therapy and audiology. *South African Journal of Communication Disorders*, 60, 2–9. <https://doi.org/10.7196/SAJCD.256>
- Peña, E., Quinn, R., & Iglesias, A. (1992). The application of dynamic methods to language assessment: A nonbiased procedure. *The Journal of Special Education*,

26(3), 269-280.

Preston, J. L., & Seki, A. (2011). Identifying Residual Speech Sound Disorders in Bilingual Children: A Japanese-English Case study. *American Journal of Speech-Language Pathology, 20*, 73–86. [https://doi.org/10.1044/1058-0360\(2011/10-0057\)](https://doi.org/10.1044/1058-0360(2011/10-0057))

Reuterskiöld, C., & Grigos, M. I. (2015). *Non-word repetition and speech motor control in children*. New York University.

Robertson, C., & Salter, W. (2007). *Phonological Awareness Test* (2nd edition). Pro-ed.

Seeff-Gabriel, B. (2003). Phonological processing: A platform for assisting second-language learners with English spelling. *Child Language Teaching and Therapy, 19*(3), 291–310. <https://doi.org/10.1191/0265659003ct256oa>

Shipley, K. G., & McAfee, J. G. (2004). *Oral Facial Examination*. Delmar: Cengage Learning.

Shriberg, L. D., Lohmeier, H. L., Campbell, T. F., Dollaghan, C. A., Green, J. R., & Moore, C. A. (2009). A nonword repetition task for speakers with misarticulations: The Syllable Repetition Task (SRT). *Journal of Speech, Language, and Hearing Research, 52*(5), 1189–1212. [https://doi.org/10.1044/1092-4388\(2009/08-0047\)](https://doi.org/10.1044/1092-4388(2009/08-0047))

Snyman, J. W. (1989). *An introduction to Tswana phonetics*. Houtbaai: Marius Lubbe Publishers.

South African Statistics. (2011). *Statistics South Africa*. Pretoria: Statistics South Africa.

Southwood, F. (2013). Towards a dialect-neutral assessment instrument for the language skills of Afrikaans-speaking children: The role of socioeconomic status. *Journal of Child Language, 40*(2), 415–437. <https://doi.org/10.1017/S0305000912000037>

- Southwood, F., & Van Dulm, O. (2015). The Challenge of Linguistic and Cultural Diversity : Does Length of Experience Affect South African Speech-Language Therapists' Management of Children With Language Impairment. *South African Journal of Communication Disorders*, 62(1), 1–14. <https://doi.org/10.4102/sajcd.v62i1.71>
- Van Biljon, S., Nolte, B., Van der Linde, J., Zsilavec, U., & Naude, E. (2015). Performance of EAL learners on a non-South African Articulation Test (GFTA-2). *Southern African Linguistics and Applied Language Studies*, 33(2), 129–139. <https://doi.org/10.2989/16073614.2015.1047388>
- Van der Merwe, A. (2011). A speech motor learning approach to treating apraxia of speech: Rationale and effects of intervention with an adult with acquired apraxia of speech. *Aphasiology*, 25(10), 1174–1206. <https://doi.org/10.1080/02687038.2011.582246>
- Van der Merwe, A., & Le Roux, M. (2014). Idiosyncratic sound systems of the South African Bantu languages: Research and clinical implications for speech-language pathologists and audiologists. *South African Journal of Communication Disorders*, 61(1), 1–8.
- Wen-hui, Q. (2016). A Review of Studies on Bilingual Children's Word Awareness Development. *Sino-US English Teaching*, 13(5), 324–331.
- Wilsenach, C. (2016). Identifying phonological processing deficits in Northern Sotho-speaking children: The use of non-word repetition as a language assessment tool in the South African context. *South African Journal of Communication Disorders*, 63(2), 1–11. <https://doi.org/10.4102/sajcd.v63i2.145>
- Yochim, B. P., Kane, K. D., & Mueller, A. E. (2009). Naming Test of the Neuropsychological Assessment Battery : Convergent and Discriminant Validity. *Archives of Clinical Neuropsychology*, 24, 575–583.

<https://doi.org/10.1093/arclin/acp053>

Zajac, M. (2016). *Variability in L2 English pronunciation examined through the prism of phonetic imitation*. (A. C. E. Waniek-Klimczak, Ed.), *Variability in English Across Time and Space*, 5.

Ziervogel, D. (1967). *Handbook of the Speech Sounds and Sound Changes of the Bantu Languages of South Africa*. Pretoria: University of South Africa.

Appendix A: Non-word list

**Non-word Articulation Test:**

Participant number: \_\_\_\_\_

Age: \_\_\_\_\_

Date: \_\_\_\_\_

Key words	Phonetic Transcription of Key words	Response
Wibe	/w/ /b/	
Hoti	/h/ /t/	
pami	/p/ /m/	
dapem	/d/ /p/ /m/	
Nefa	/n/ /f/	
Sona	/s/ /n/	
Geson	/g/ /s/ /n/	
Tob	/t/ /b/	
Blagu	/b/ /g/	
Feg	/f/ /g/	
Brek	/b/ /k/	
Dref	/d/ /f/	
Bod	/b/ /d/	
Fludich	/f/ /d/ /tʃ/	
Ret	/r/ /t/	
Yeela	/j/ /l/	
Shil	/ʃ/ /l/	
plangu	/p/ /l/ /ŋ/	
kreng	/k/ /r/ /ŋ/	
lishas	/l/ /ʃ/ /s/	
Kep	/k/ /p/	
Mir	/m/ /r/	
slekesh	/s/ /k/ /ʃ/	

Compiled by Lauren Ross  
Using Van Der Merwe (2010) SML  
website: [www.apraxia-anitavandermerwe.co.za](http://www.apraxia-anitavandermerwe.co.za)

chom	/tʃ/ /m/	
Jecha	/dʒ/ /tʃ/	
thave	/θ/ /v/	
vozan	/v/ /z/ /n/	
zaje	/z/ /dʒ/	
thaja	/ð/ /dʒ/	
Klithi	/kl/ /θ/	
sqivath	/skw/ /v/ /θ/	
stath	/st/ /ð/	
trizz	/t/ /z/	
whothin	/hw/ /ð/ /n/	

Sound:	Initial	Medial	Final
p			
m			
n			
w			
h			
b			
g			
k			
f			
d			
ŋ			
j			
t			
ʃ			
tʃ			
l			
r			
dʒ			
θ			
v			
s			
z			
ð			

bl		pl	
bɹ		skw	
dɹ		sl	
fl		st	
kl		tɹ	
kɹ		hw	