Effects of language experience on selective auditory attention and speech-in-noise perception among English second language learners: Preliminary findings

Lianca Strydom [a,*], Lidia Pottas [a], Maggi Soer [a], Marien Alet Graham[b]

[a]Department of Speech-Language Pathology and Audiology, Faculty of Humanities, University of Pretoria, Pretoria, Gauteng, South Africa
[b]Department of Science, Mathematics and Technology Education, Faculty of Education, University of Pretoria, Pretoria, Gauteng, South Africa

* Corresponding author. Email: liancastrydom62@gmail.com

Abstract

Objective: The purpose of the study was to examine the effects of language experience on selective auditory attention and speech-in-noise perception in English Second Language (ESL) learners aged seven to eight years.

Method: A quantitative, descriptive, comparative cross-sectional research design was used to determine the effect of age of exposure to English on the selective auditory attention abilities and speech-in-noise perception skills of 40 children with normal hearing in first or second grade (aged seven to eight years). The control group comprised of 20 English first language (EFL) learners (mean age = 7.35 years ±0.49) and the research group included 20 s language learners (mean age = 7.70 years ±0.47). In order to compare the control and research groups with respect to the age of exposure to English through various sources, the Mann Whitney test was used. Information regarding the age of exposure was gathered by a case history questionnaire, completed by the parents/guardians of the participants. The Selective Auditory Attention Test (SAAT) and Digits-in-Noise (DIN) test were performed in one sitting.

Results: No statistically significant differences between the EFL and ESL groups were found for the SAAT and DIN. However, a statistically significant difference was obtained between the SAAT lists 1 and 3 & the DIN: diotic listening condition for the ESL group only ($r_s = -0.623$; $p = 0.003$). The difference between the EFL and ESL groups in the mean age of exposure to English was statistically significant ($p = 0{,}019$), with mean age of exposure to English in the ESL group (mean age = 2.82 ± 0.53) being higher than the mean age of exposure in the EFL group (mean age = 1.81 ± 1.53). However, this difference did not influence the results of the SAAT and DIN significantly.

Conclusion: The main finding was that selective auditory attention and speech-in-noise perception were not significantly affected in the ESL learners who participated in the study – learners who were recruited from private schools located in an urban area and thus from higher socio-economic status (SES) households. There is a need for additional research with a larger sample size to determine the selective auditory attention abilities and speech-in-noise perception skills of ESL learners in government-funded schools located in rural areas and from various socio-economic backgrounds.

Keywords: Selective auditory attention; Speech-in-noise perception; Language experience; Age of exposure; English second language; English first language

1. Introduction

Multilingualism is a defining characteristic of the African continent, where roughly 2100 languages - which amounts to 30% of the world's languages - are spoken [1]. Generally considered to be among the most multilingual countries in the world, South Africa has 11 official languages that are recognized in its democratic constitution [2]. The terms ‘multilingualism’ and ‘bilingualism’ are used interchangeably in the South African context, and the abbreviations EFL (English first language) and ESL (English second language) refer to English as the primary home language and additional/second language respectively. Most of the schools in South Africa follow one of two different monolingual educational programmes as suggested by Siegel (2003) [3,4]. In the first type, the learner's first language is used as the language of learning and teaching (LoLT) (i.e., English – EFL learner), and additional languages are taught as subjects. English- and Afrikaans-speaking learners in South Africa are enrolled in this type of educational programme, where English (or Afrikaans) is used as the LoLT and they learn other South African languages (e.g., isiZulu) as subjects [4]. In the second type monolingual programme, the learner's second (or third) language is used as the LoLT (i.e., English – ESL learner), and other languages are taught as subjects, as in the case of the majority of learners in South Africa [4]. During the 2007 Annual School Survey, it was determined that 65.3% of South African learners are enrolled in English medium schools [5], yet less than 10% of these learners are EFL speakers [6]. This implies that more than half of the learners in South Africa are ESL speakers who are not receiving education in their first language [[7], [8], [9]]. These learners are labelled as educationally at-risk or disadvantaged as a result of the linguistic transition they have to make, and often do not succeed in the academic domain [10].

It is widely accepted that language competence and proficiency play a fundamental role in literacy development and are basic to achieving academic success [[11], [12], [13]]. Internationally, and especially in developing countries such as South Africa, numerous school-aged learners have inadequate English language proficiency to succeed within formal academic settings [[14], [15], [16]]. Kotzé and Hibbert (2010) identified the use of English as the LoLT in schools, especially primary schools, as a major contributing factor to the underdevelopment of academic skills in South Africa [17]. There is adequate proof to support the reality that language development is influenced by auditory processing skills [18,19] – a significant challenge exists in attempting to separate the influence of auditory processing on language processing and academic performance [20]. Therefore, since learning takes place in English, the second (or third) language of many South African learners, these learners need to develop their auditory processing skills as rapidly and as comprehensively as possible in order to process the additional language as the LoLT [21].

Research on second language learning suggests that age and age-related factors such as age of exposure to the additional language are major variables in the acquisition of a second language for learning in school [22]. As early as in the 1980s, studies determined that age of exposure to a second language has a significant effect on a child's academic achievement. The majority of urban listening environments are clamorous, and a substantial proportion of individuals must function in these environments every day, experiencing the pressure of having to perceive speech in their second language [23]. While the typical classroom environment overflows with abundant auditory and visual distractions [[24], [25], [26]], the influence of noise and reverberation on speech perception cannot be underestimated [27]. In addition to noise and reverberation, a child's familiarity with a specific language also affects his or her ability to make sense of incoming speech stimuli [27,28]. Therefore, children learning in their additional/second language appear to be at a distinct disadvantage when listening in classrooms where background noise and reverberation are inevitable [27,29].

In order for learners to perceive and understand a verbal message in the classroom environment, it is necessary for them to attend to a signal whilst simultaneously suppressing the competing noise. ESL learners experience even further challenges – in addition to the fact that they experience problems with the language of instruction, they have to process it in a non-optimal listening environment [30]. Auditory processing abilities, such as selective auditory attention and speech-in-noise perception, are important skills for school-age children to master as these competencies support learning in noisy classroom environments. Selective auditory attention means that a specific input is extracted and focused on for further processing, whilst irrelevant or distracting information such as noise is simultaneously suppressed [25,31]. Selective auditory attention plays an important role in an individual's orientation to environmental stimuli and in maintaining an alert state in order to detect signals for subsequent detailed processing. Competent listeners have the ability to segregate different stimuli into different streams and subsequently decide which streams are most pertinent to them. This skill is crucial for speech perception in noise.

Noise masks and interferes with selective attention to a primary stimulus [21]. A study by Koopsman et al. (2018) showed that in order for children to understand speech in noisy settings so that they can follow and participate in classroom discussions, children need to be able to separate speech from noise [32]. The impact of selective auditory attention and optimal speech-in-noise perception on academic achievement of ESL learners cannot be disregarded. If educators are made aware of the influence of noise on speech perception abilities of ESL learners, they may realize the importance of reducing classroom noise and increasing speech intensity levels in order to improve access to spoken language for these ESL learners [27].

Despite evidence from research both in Africa and in other countries across the world that education in a child's first language is critical for true learning to occur at a deeper and profound cognitive level, English remains the chosen medium of education in many countries [33]. To date, few studies have been conducted to investigate the impact of language proficiency on selective auditory attention abilities and speech-in-noise perception in young ESL learners in a multilingual country. Given the range of auditory demands with which learners are faced in the typical classroom environment, along with the importance of auditory skills for achieving academic success, additional research is warranted that may support the development of effective strategies for addressing the challenges these learners face in the South African context. Therefore, the purpose of the study was to examine the effect of language experience on selective auditory attention and speech-in-noise perception of ESL learners in a multilingual country, to assist with teaching and learning in the classroom.

2. Method

Ethical clearance was obtained from the Research Ethics Committee of the Faculty of Humanities at the University of Pretoria (reference number: HUM006/0320), prior to data collection. The schools and parents/guardians were informed of the study aims and provided their consent for the learners to participate. Furthermore, the learners provided assent before data collection procedures commenced.

2.1. Research design and participants

Using a quantitative, descriptive, comparative cross-sectional research design, 40 participants (7–8 years old) with the same socio-economic status (SES) were purposively selected from three English private schools in the City of Tshwane, South Africa. Participants were assigned either to the research group (ESL learners; n = 20; mean age 7.70 years ±0.47) or the control group (EFL learners; n = 20; mean age 7.35 years ±0.49) based on their first language (their mother tongue). A first language is identified as the language to which a person had been exposed since birth and which was learned within the critical period of language development [34]. If the participant's first language was English (i.e., English was spoken at home), he/she was classified as an EFL participant. If the participant's first language was a language other than English, the participant was categorized in the ESL participant group. There was no statistically significant difference between the mean age of participants from the EFL group and the mean age of participants from the ESL group. None of the participants had any known neurological or cognitive disorder, as determined by a question included in the case history questionnaire that was completed by the parents/guardians of all participants. All participants met the inclusion criteria of normal hearing (bilateral PTA ≤20 dB HL across 1–4 kHz) [35], and normal outer and middle ear function (type A tympanograms and at least one present acoustic reflex at 1 kHz) [36,37]. Two auditory processing tests were conducted in order to determine the selective auditory attention abilities and speech-in-noise perception of each participant.

2.2. Instrumentation and procedures

The Selective Auditory Attention Test (SAAT) and the Digits-In-Noise (DIN) test were performed during one individual sitting. The order in which these tests were administered was alternated in order to avoid order effects. A break was taken between testing if necessary. The research group (ESL learners) and the control group (EFL learners) were assessed in the same manner, time frame, and setting.

2.2.1. Selective auditory attention test (SAAT)

The SAAT is a monaural low-redundancy speech test that was developed as a speech-in-competing-message test for the early identification of learners who may have a problem attending to auditory stimuli, especially in the presence of background noise [38]. Cherry and Rubinstein (2006) established that binaural presentation when using the SAAT resulted in better performance due to the binaural advantage when listening to stimuli binaurally. Therefore, during data collection in the current study, the SAAT was also presented binaurally. The SAAT is a closed-set picture-pointing task, consisting of two parts: two lists of Word Intelligibility by Picture Identification (WIPI) words in quiet (lists one and three), and two lists of WIPI words presented in a competing noise (lists two and four). The competing noise comprised a speaker telling a story that was identified by children to be interesting, thus causing a semantic distraction. The WIPI words and competing story were recorded at a signal-to-competition ratio of 0 dB, which increased the test's difficulty [38]. Each of the four lists consists of 25 monosyllabic words [31]. The SAAT was conducted in a quiet setting through headphones at a comfortable listening level. The comfortable listening level was held constant at approximately 50 dB HL [39,40]. Participants were requested to indicate the corresponding picture on the page by pointing to it as the word was heard over the headphones. During testing the researcher counted the words that were correctly identified. Four percent was given for each word correctly identified, and a percentage of correct scores was computed for each of the four lists. A score of less than 88% on the wordlists read in quiet invalidated the entire test and therefore precluded administration of the word lists imbedded in a semantic distractor.

2.2.2. Digits-In-Noise (DIN) test

Speech recognition in noise is assessed through the DIN test [[41], [42], [43]]. The DIN was presented binaurally using diotic (digits and masking noise presented interaurally in phase) and antiphasic (digit stimuli presented 180-degree phase inverted to the masking noise) stimulus presentations on a smartphone with headphones in a quiet setting [41,42]. Recent studies concluded that antiphasic stimulus presentations improves the ability to detect speech in the presence of a diotic masker [41,44]. The difference in signal-to-noise ratio (SNR) between diotic and antiphasic stimulus presentations is known as the binaural intelligibility level difference (BILD). As early as 1948, this phenomenon was described as the ability to spatially segregate speech from noise and understand speech in the presence of background noise [45,46].

During data collection for the DIN test, participants were requested to complete the test in diotic and antiphasic listening conditions. The first digit-triplet was presented at an intensity that was based on the participant's selected comfortable listening level. After the response was entered on the keypad, the following digit-triplet was presented automatically at a 2 dB higher SNR for an incorrect response, or at a 2 dB lower SNR for a correct response. A digit-triplet was only judged as correct when all three digits were entered correctly [42]. The Speech Reception Threshold (SRT) was calculated as the average SNR of the last 19 of 23 triplets presented in total [41,42]. Since the SNR is a measure of signal strength relative to background noise, the ratio is usually measured in decibels using a SNR formula which is expressed in a logarithmic scale. A 0 dB SNR indicates that the power of the signal (speech) is equal to the noise power. When the signal power is poorer than the power of the noise, it will result in a negative SNR in dB. Thus, more negative dB SNRs where a listener can identify 50% of the digits correctly, indicate better SRTs, as well as better performance [41,42]. Results for both diotic and antiphasic listening conditions were then recorded and the BILD was determined.

2.3. Data analysis

Raw data were edited, coded and categorized in a Microsoft Excel spreadsheet. Data analysis was done by means of the software G*Power version 3.1.9.4 [47]. for the power analysis, and Statistical Package for Social Sciences (SPSS) version 26.0 [48] for all other statistical analyses. Descriptive statistics included the mean, standard deviation (SD), median and inter-quartile range (IQR) for the SAAT and DIN for the two groups. The Mann-Whitney test was used to determine the overall outcome of the continuous data of the SAAT and DIN between the two groups.

3. Results

Firstly, the comparison between the EFL and ESL groups regarding the age of exposure to English is provided. Where the median age of exposure is indicated as 0,00, the learners were for the most part exposed to English through this source from birth. The results of the SAAT are followed by the DIN test's results. Nominal results of both tests are provided subsequently.

3.1. Comparison between EFL and ESL learners regarding the age of exposure to English

The mean age of exposure did not differ statistically significantly between the EFL and the ESL groups for exposure via television (p = 0,383), radio (p = 0,068), nursery/day care (p = 0,190) and Grade R (p = 0,812). However, statistically significant differences were found between the EFL and the ESL group for exposure via caregivers (p < 0.001), family and friends (p = 0,001) and books (p = 0,014). In addition, the mean age of exposure was computed (by averaging all the measures of exposure) and the difference between the EFL and ESL groups was also found to be statistically significant (p = 0,019). In all cases where statistically significant differences were found, the mean age of exposure to English for ESL was significantly higher than the mean age of exposure to English for EFL (see Table 1).

Table 1. Age of exposure to English.

Age of exposure to English *(in years)* through various sources	**EFL**		**ESL**		**Mann-Whitney**	**p-value**
	Mean (SD)	**Median (IQR)**	**Mean (SD)**	**Median (IQR)**	Empty Cell	Empty Cell
Via caregivers	0,00 (0,00)	0,00 (0,00)	1,30 (1,04)	1,00 (1,63)	6000	<0,001[a]
Via family and friends	0,50 (1,40)	0,00 (0,00)	1,93 (1,00)	2,00 (1,25)	27,500	0,001[a]
Via books	1,41 (1,92)	0,50 (2,50)	3,30 (2,20)	3,50 (3,75)	64,000	0,014[a]
Via television	1,39 (1,26)	1,25 (3,00)	1,70 (0,97)	2,00 (1,00)	135,000	0,383
Via radio	1,42 (2,31)	0,00 (2,50)	2,30 (1,54)	2,00 (2,50)	29,000	0,068
Via nursery/day care	2,01 (1,71)	2,00 (3,13)	2,56 (1,04)	2,50 (1,25)	114,000	0,190
Via Grade R	5,78 (0,44)	6,00 (0,50)	5,73 (0,46)	6,00 (1,00)	64,500	0,812
Mean age exposed to English through all sources	1,81 (1,53)	1,86 (2,94)	2,82 (0,89)	2,67 (1,30)	94,000	0,019[a]

[a]$p \leq 0.05$.

3.2. Selective auditory attention test (SAAT)

The Mann-Whitney test was used to determine whether any significant differences were present in the SAAT results between the lists presented in quiet (lists 1 and 3) and the lists presented in the presence of a competing stimulus (lists 2 and 4) (Table 2). The descriptive statistics for the lists used in the SAAT are presented with the results for lists 1 and 3 averaged, and for lists 2 and 4 averaged.

Table 2. Averaged results for lists 1 and 3 & lists 2 and 4.

	EFL		**ESL**		**Mann-Whitney**	**p-value**
SAAT	**Mean (SD)**	**Median (IQR)**	**Mean (SD)**	**Median (IQR)**		
Average of lists 1 & 3 – no competing stimulus (%)	96,30 (3,26)	97,00 (4,00)	96,60 (3,05)	98,00 (4,00)	188,500	0,758
Average of lists 2 & 4 – competing stimulus (%)	67,80 (8,03)	68,00 (10,00)	63,10 (7,44)	64,00 (12,00)	128,500	0,052

*$p \leq 0.05$.

The mean and standard deviation of the EFL (mean = 96,30; SD = 3,26) and ESL (mean = 96,60; SD = 3,05) groups for lists 1 and 3 were similar. For lists 2 and 4, the mean and standard deviation for the EFL group (mean = 67,80; SD = 8,03) were higher than for the ESL group (mean = 63,10; SD = 7,44). However, no statistically significant differences were found between the EFL and ESL groups. It is important to note, though, that the p-value of 0,052 for the average of lists 2 and 4 is very close to 0,05.

3.3. Digits-In-Noise (DIN) test

The results of Mann-Whitney test as performed for the DIN are presented in Table 3.

Table 3. DIN hearing test results – diotic & antiphasic listening conditions.

	EFL		**ESL**		**Mann-Whitney**	**p-value**
DIN	**Mean (SD)**	**Median (IQR)**	**Mean (SD)**	**Median (IQR)**		
Diotic (SNR)	−8,79 (0,95)	−8,90 (1,35)	−8,94 (0,98)	−8,90 (1,60)	184,000	0,664
Antiphasic (SNR)	−16,01 (1,71)	−15,90 (1,55)	−15,90 (1,42)	−15,90 (2,30)	193,500	0,860
BILD (SNR)	−7,22 (2,16)	−7,30 (3,95)	−6,96 (1,68)	−6,70 (3,00)	180,000	0,597

*p ≤ 0.05.

The SNR obtained in the diotic listening condition by the EFL group (mean = −8,79 dB) was higher than the SNR obtained in the diotic listening condition by the ESL group (mean = −8,94 dB). For the antiphasic listening condition the SNR obtained by the EFL group (mean = −16,01 dB) was lower than the SNR obtained by the ESL group (mean = −15,90 dB). However, no statistically significant difference was found between the EFL group and the ESL group for either the diotic (p = 0,664) or the antiphasic (p = 0,860) listening condition. The BILD calculated for the EFL group (mean = −7,22) was lower than the BILD calculated for the ESL group (mean = −6,96). No statistically significant difference was found between the EFL group and the ESL group for the BILD (p = 0,597).

3.4. Integration and correlation of results

In order to determine the strength of association between the SAAT and DIN, the Spearman rank correlation was used. For the purpose of discussion, the researcher will only focus on the correlations established between the averaged results of lists 1 and 3 and lists 2 and 4 of the SAAT (due to small sample size) and the diotic and antiphasic listening conditions of the DIN (Table 4).

A statistically significant negative correlation ($r_s = -0.623$) was established within the ESL group between the DIN in the diotic listening condition and the average of SAAT lists 1 and 3 in the non-competing listening condition ($p = 0.003$). This means that as the value of the average of lists 1 and 3 of the SAAT increased, the value of the DIN: diotic SNR decreased. As stated earlier regarding the DIN, more negative SNRs indicate better test performance [42]. Therefore, as the value of the SAAT lists 1 and 3 increased, the results for the DIN: diotic listening condition will also be identified as better performance. For the EFL group, no statistically significant correlations were found between the tests.

Table 4. Strength of association between the SAAT and DIN.

Tests		Values	SAAT: List 1 & 3 (non-competing)	SAAT: List 2 & 4 (competing)	DIN: diotic (SNR)	DIN: antiphasic (SNR)
Control (EFL) Group	**SAAT: List 1 & 3 (non-competing)**	r_s	1000	0,115	−0,120	−0,158
		p-value		0,629	0613	0,507
	SAAT: List 2 & 4 (competing)	r_s	0,115	1000	0,224	−0,208
		p-value	0,629		0,342	0378
	DIN: in-phase	r_s	−0.120	0,224	1000	−0,272
		p-value	0,613	0342		0,246
	DIN: out-of-phase	r_s	−0,158	−0,208	−0,272	1000
		p-value	0,507	0378	0,246	
Research (ESL) group	**SAAT: List 1 & 3 (non-competing)**	r_s	1000	−0,167	−0,623[a]	0,349
		p-value		0,482	0003	0,132
	SAAT: List 2 & 4 (competing)	r_s	−0,167	1000	−0,014	0058
		p-value	0,482		0,953	0809
	DIN: in-phase	r_s	−0,623[a]	−0,014	1000	−0,026
		p-value	0,003	0953		0,912
	DIN: out-of-phase	r_s	0,349	0058	−0,026	1000
		p-value	0,132	0809	0,912	

[a]$p \leq 0.05$.

4. Discussion and conclusion

4.1. Summary of results and contributions of the study

In order to achieve academic success when the LoLT is an additional language rather than the first language, learners need to master various auditory processing skills including selective auditory attention and speech-in-noise perception [49,50].

4.1.1. The comparison between EFL and ESL learners regarding the age of exposure to English

The findings seem to indicate that selective auditory attention and speech-in-noise perception were not adversely affected in the ESL learners who participated in this study. These findings are not in agreement with those of previous studies investigating either selective auditory attention [51] or speech-in-noise perception [21,27,52,53] in ESL speakers. These studies found that proficiency (or lack of proficiency) in the LoLT plays an influential role in selective auditory attention and speech-in-noise perception in ESL learners.

Due to the mediating role of language proficiency and selective auditory attention, second language learners are negatively influenced to a greater extent than first language learners by noise in speech perception tasks [54]. Bovo et al. (2018) and Florentine (1985) determined that speech-in-noise perception in ESL speakers correlates significantly with years of exposure to the second language. Speech perception in noise was found to improve as the period of exposure to the second language increased. Therefore, earlier exposure to the second language is associated with improved understanding of speech in noise from an early age [23].

Age of exposure to an additional language is influenced by several external factors, such as choice of schools (privately- or government-funded), geographic location (i.e. urban, suburban, rural) and SES [55,56]. Three private schools were included in this study, and it is possible that the parents of these learners are from a different educational and socio-cultural background than parents of ESL learners located in rural areas, from lower SES households and in government funded schools [11]. The discrepancy between the current and previous studies may be due to the participants' more favourable exposure to English at an early age when compared to ESL learners located in rural areas and from lower SES households, with majority who have minimal English exposure when they reach school age. Regarding the statistically significant differences found between the mean age of first exposure to English through caregivers, as well as family and friends, it is important to keep in mind that the mean age of exposure of the ESL learners (mean age = 1.30 years ± 1.04) is still lower than that of the majority ESL learners located in rural areas and from lower socio-economic backgrounds whose first exposure to English is generally only when they reach school age [56].

Previous research investigated various factors that contribute to the difference of performance in English between learners located in rural and urban areas [55]. Hossain (2016) found that factors such as parental education status, SES, and the availability of adequate books to read contribute to the poor proficiency in English of ESL learners located in rural areas. On average, young children from lower SES households where a language other than English is spoken have language trajectories different from those of children from middle or high SES, bilingual households [57]. Al-Zoubi (2018) recommended that ESL learners should be frequently exposed to English through various sources, including by watching English programmes on television, surfing the internet, listening to the radio, reading English books, and communicating with EFL speakers daily to improve their proficiency in English [58]. In 2019 it was determined that South Africa has very high child poverty rates, where 76% of children between 0 and 17 years of age are living below the poverty line, are from lower socio-economic backgrounds located in rural areas and do not have access to sufficient educational resources [57,59]. In the current study, the ESL learners were exposed to English through reading English books at approximately three years of age (mean age = 3.30 years ±2.20), which might be an earlier age than the majority of South African learners who have limited access to educational resources, such as books [56]. These learners who have access to more facilities (i.e. nursery/daycare, grade R) or resources (i.e. television, radio) are identified as having an advantage when compared to those from lower SES households [55,57]. This can be substantiated by the insignificant differences established within the current study between the two participant groups for the age of exposure to English through television, radio, nursery/daycare and through Grade R. Therefore, it is recommended that this study be replicated in a rural area and include learners from government-funded schools and various socio-economic backgrounds, as the current findings are based on learners in private schools located in an urban area and thus from higher SES households and learners in rural areas might present with lower English proficiency when compared to learners in urban areas. This discrepancy relates to their ideal access to educational resources.

4.1.2. Selective auditory attention test (SAAT)

In the non-competing and ideal conditions (lists 1 and 3), the EFL and ESL groups achieved similar results. There is a definite trend for the EFL group to perform slightly better in the conditions where a semantic distractor was present (lists 2 and 4). Although not statistically significant, the calculated probability (p-value) of 0,052 established for the average of lists 2 and 4, which is extremely close to 0,05, may be indicative that if a larger sample size was included in the study, a statistically significant difference might have been present for the lists presented in the presence of a competing story. Thus, a recommendation for future research is to replicate this study, but with a larger sample size to investigate this probability further and shed more light on this finding.

These results do not correlate, however, with findings from previous research regarding selective auditory attention abilities in ESL learners. Venter, Pottas and Soer (2019), determined that ESL learners have greater difficulty attending to the target stimuli whilst suppressing the competing noise to understand speech [51]. Warzybok, Brand, Wagener and Kollmeier (2015) also found significant differences between first and second language speakers on tasks which require selective attention to speech as the listening conditions became more demanding [60]. It is important to note that should a larger sample size be included in the current study, a statistically significant difference might be present and the results would then correlate with findings from previous research.

4.1.3. Digits-in-noise (DIN) test

The EFL group did not perform significantly better than the ESL group in either the diotic or the antiphasic listening conditions. The results within the diotic listening condition agree with previous research that established no significant effects on ESL speakers' ability to recognise digit-triplets in noise presented interaurally in-phase [49,61]. The results obtained within the antiphasic listening condition are in agreement with the findings of De Sousa et al. (2019) and Wolmarans et al. (2021), who established that when evaluating SRT the use of the antiphasic listening condition was shown to improve the DIN SRTs in normal hearing listeners [41], and also in normal hearing children from seven years of age [44]. The insignificant difference in the BILD established between the two participant groups and derived from the diotic and antiphasic results, substantiates the finding that the ESL learners who participated in the current study do not appear to be at a distinct disadvantage both to spatially segregate speech from noise and to understand speech in the presence of background noise.

Despite the fact that the DIN test is a well-known and often used speech-in-noise test, the results of the current study do not correlate with previous research on speech-in-noise perception in second language speakers. Several studies have found that bilingual speakers perform worse in their second language in terms of perceiving speech in noise [52,[62], [63], [64]]. Some studies have suggested that an early age of exposure to and acquisition of an additional language could mediate speech-in-noise processing in the second language, with earlier exposure and acquisition being associated with performance similar to first language speakers [21,52,65]. Another study also suggested that children's abilities to recognise speech in noise develop well with age into adolescence [44]. Although there is a statistically significant difference between the EFL and ESL groups for the mean age of exposure to English, it is important to keep in mind that the current study has a small sample size, and also that the participants included in the study are from higher socio-economic backgrounds than most of the learners in South Africa and most probably exposed to English from an earlier age than learners from lower socio-economic backgrounds located in rural areas.

4.1.4. Correlations between the SAAT and DIN

The negative correlation established between the ESL learners' test results for the non-competing lists in the SAAT and the diotic listening condition for the DIN might be due to the varying linguistic demands of the stimuli used in the respective tests. The choice of speech material used when testing auditory processing skills can have a significant influence [49].

For the DIN test, it was essential to use easy, familiar words within a closed-set paradigm instead of open-set sentences or words in order to decrease the impact of linguistic skills on the test outcome [43]. Digits have been identified as one category of highly familiar words. They are in the lists of commonly spoken words and therefore are not known to be linguistically demanding, since numbers are some of the first words a child learns when acquiring a second language [43]. The WIPI words imbedded in the lists used when completing the SAAT are simple monosyllabic words which result in decreased linguistic demand of the test. Keep in mind that participants are also provided with pictures to choose from. The use of pictures adds a visual component that can aid the listener, especially given that the test is a four-alternative forced-choice (4AFC) closed set task [31].

Regarding the negative correlation between the ESL learners' results for the SAAT and DIN, it is important to remember that more negative DIN results refer to better test performance. Therefore, as the participants performed better in the SAAT lists 1 and 3, their performances in the DIN: diotic listening condition also improved. These results correlate with research that previously determined a relationship between improved speech-in-noise perception abilities and the development of auditory processing efficiency with age, such as selective attentional control [21,[66], [67], [68]]. However, with regard to attention, the results of the current study do not support the findings of Klatte et al. (2013), who determined that the immature selective auditory attention abilities of children add to their struggle with speech perception in noise [54].

4.2. Conclusion

The finding that the differences between the EFL and ESL groups for the SAAT and DIN are statistically insignificant can most likely be ascribed to the fact that ESL learners in private schools and from higher SES households are exposed to English at an earlier age than ESL learners from lower SES households located in rural areas. Another reason might be the small sample size included in this study. There is a need for additional research with a larger sample size to determine the selective auditory attention abilities and speech-in-noise perception skills of ESL learners in government-funded schools located in rural areas and from various socio-economic backgrounds.

Declaration of competing interest

None declared.

Acknowledgements

We are grateful to the principals and learners of the three private schools included in this study for their willingness to participate in the study. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] M.P. Lewis, Table 1: Distribution of languages by area of origin, in: Ethnol. Lang. World, Sixt. Ed., SIL International, Dallas, Texas, 2009.

[2] Constitution of the Republic of South Africa, Constitution of the Republic of South Africa [South Africa], [South Africa], 1996, https://www.refworld.org/docid/3ae6b5de4.html. (Accessed 14 October 2021).

[3] J. Siegel, Social context, in: C.J. Doughty, M.H. Long (Eds.), Handb. Second Lang. Acquis., Blackwell Publishing, Berlin, Germany, 2003, pp. 193–194.

[4] H. Jordaan, Multilingualism and speech-language therapy in education: theory and research, First, in: S. Moonsamy, H. Kathard (Eds.), Speech-Language Ther. A Sch. Context - Princ. Pract., Van Schaik Publishers, Pretoria, South Africa, 2015, pp. 55–64.

[5] Department of Basic Education, Report on the 2009/2010 Annual Survey for Ordinary Schools, South Africa Dep, Basic Educ., 2010. www.gov.za/sites/default/files/gcis_document/201409/report-2009-2010-annual-surveys-ordinary-schools. pdf. (Accessed 24 August 2018).

[6] Statistics South Africa, Census 2011: Census in Brief, Stat. South Africa, 2012. http://www.statssa.gov.za/census/census_2011/census_products/Census_2011_Census_in_brief.pdf.

[7] B. Brock-Utne, I. Skattum, Languages and Education in Africa: A Comprehensive and Transdisciplinary Analysis, Symposium Books, Oxford, 2009.

[8] K. Heugh, Into the cauldron: an interplay of indigenous and globalised knowledge with strong and weak notions of literacy and language education in Ethiopia and South Africa, Lang, Matters 40 (2009) 166–189, https://doi.org/10.1080/10228190903188559.

[9] N. Spaull, Disentangling the language effect in South African schools: measuring the impact of "language of assessment" in grade 3 literacy and numeracy, South Afr. J. Child. Educ. 6 (2016) 1–20, https://doi.org/10.4102/sajce.v6i1.475.

[10] R. Kapp, Reading on the line: an analysis of literacy practices in ESL classes in a South African township school, J. Lang. Educ. 18 (2004) 246–263.

[11] E. Hoff, How social contexts support and shape language development, Dev. Rev. 26 (2006) 55–88, https://doi.org/10.1016/j.dr.2005.11.002.

[12] R.E. Owens, R.E. Owens, Language Development: an Introduction, eighth ed., Pearson, United States of America, 2012.

[13] I. Hay, R. Fielding-Barnsley, Social learning, language and literacy, Aust. J. Early Child. 37 (2012) 24–29, https://doi.org/10.1177/183693911203700104.

[14] A. Hugo, M. Nieman, Using English as a second language as the language of instruction: concerns and needs of primary school teachers in South Africa, J. Lang. Teach. 44 (2010) 59–69, https://doi.org/10.4314/jlt.v44i1.71781.

[15] S. Taylor, M. von Fintel, Estimating the impact of language of instruction in South African primary schools: a fixed effects approach, Econ. Educ. Rev. 50 (2016) 75–89, https://doi.org/10.1016/j.econedurev.2016.01.003.

[16] Z. Wildschut, T. Moodley, S. Aronstam, The baseline assessment of Grade 1 learners' literacy skills in a socio-economically disadvantaged school setting, South Afr. J. Child. Educ. 6 (2016) 1–9, https://doi.org/10.4102/sajce.v6i1.340.

[17] E. Kotzé, L. Hibbert, Are multilingual education policies pipe dreams? Interdiscip. J. Study Arts Humanit. South Afr. 17 (2010) 4–25.

[18] B. Wible, T. Nicol, N. Kraus, Correlation between brainstem and cortical auditory processes in normal and language-impaired children, Brain 128 (2005) 417–423, https://doi.org/10.1093/brain/awh367.

[19] M. Sharma, S.C. Purdy, A.S. Kelly, Comorbidity of auditory processing, language, and reading disorders, J. Speech Lang. Hear. Res. 52 (2009) 706–722, https://doi.org/10.1044/1092-4388(2008/07-0226.

[20] G.J. Richard, language processing versus auditory processing, in: D. Geffner, D. Ross-Swain (Eds.), Audit. Process. Disord. - Assessment, Manag. Treat., Third, Plural Publishing Inc, San Diego, CA, 2019, pp. 224–226.

[21] R. Bovo, E. Lovo, L. Astolfi, S. Montino, S. Franchella, S. Gallo, N. Prodi, D. Borsetto, P. Trevisi, Speech perception in noise by young sequential bilingual children, ACTA Othorhinolaryngol. Ital. 38 (2018) 536–543, https://doi.org/10.14639/0392-100x-1846.

[22] V.P. Collier, The effect of age on acquisition of a second language for school, New Focus, NCBE Occas. Pap. Biling. Educ. 2 (1987).

[23] S. Kousaie, S. Baum, N.A. Phillips, V. Gracco, D. Titone, J.K. Chen, X.J. Chai, D. Klein, Language learning experience and mastering the challenges of perceiving speech in noise, Brain Lang. 196 (2019) 1–9, https://doi.org/10.1016/j.bandl.2019.104645.

[24] American National Standards Institute, Acoustic performance criteria, design requirements and guidelines for schools, S12, ANSI 60 (2002).

[25] D.L. Strait, N. Kraus, Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise, Front. Psychol. 2 (2011) 1–10, https://doi.org/10.3389/fpsyg.2011.00113.

[26] J.B. Talcott, C. Witton, G.S. Hebb, C.J. Stoodley, E.A. Westwood, S.J. France, P. C. Hansen, J.F. Stein, On the relationship between dynamic visual and auditory processing and literacy skills; results from a large primary-school study, Dyslexia 8 (2002) 204–225, https://doi.org/10.1002/dys.224.

[27] P. Nelson, K. Kohnert, S. Sabur, D. Shaw, Classroom noise and children learning through a second language: double jeopardy? Lang. Speech Hear. Serv. Sch. 36 (2005) 219–229, https://doi.org/10.1044/0161-1461(2005/022.

[28] A.C. Chang, B.W. Wu, J.C. Pang, Second language listening difficulties perceived by low-level learners, Percept. Mot. Skills 116 (2013) 415–434, https://doi.org/10.2466/22.23.PMS.116.2.415-434%0A.

[29] C.L. Rogers, J.J. Lister, D.M. Febo, J.M. Besing, H.B. Abrams, Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing, Appl. Psycholinguist. 27 (2006) 465–485, https://doi.org/10.1017/s014271640606036x.

[30] B. Kollmeier, A. Warzybok, S. Hochmuth, M.A. Zokoll, V. Uslar, T. Brand, K. C. Wagener, The multilingual matrix test: principles, applications and comparisons across languages: a review, Int. J. Audiol. 54 (2015) 3–16, https://doi.org/10.3109/14992027.2015.1020971.

[31] R. Cherry, A. Rubinstein, Comparing monotic and diotic selective auditory attention abilities in children, Lang. Speech Hear. Serv. Sch. 37 (2006) 137–142, https://doi.org/10.1044/0161-1461(2006/015.

[32] W.J.A. Koopsman, S.T. Goverts, C. Smits, Speech recognition abilities in normal-hearing children 4 to 12 years of age in stationary and interrupted noise, Ear Hear. 39 (2018) 1091–1103, https://doi.org/10.1097/AUD.0000000000000569.

[33] N. Ndimande-Hlongwa, R. Wildsmith-Cromarty, Multilingualism for access, language development and language intellectualisation, Interdiscip. J. Study Arts Humanit. South Afr. 17 (2010) 92–105.

[34] A. Saniei, Who is an ideal native speaker?, in: 2011 Int. Conf. Lang. Lit. Linguist. IPEDR, 26 IACSIT Press, Singapore, 2011, pp. 74–78.

[35] D.W. Swanepoel, H.C. Myburgh, D.M. Howe, F. Mahomed, R.H. Eikelboom, Smartphone hearing screening with integrated quality control and data management, Int. J. Audiol. 53 (2014) 841–849, https://doi.org/10.3109/14992027.2014.920965.

[36] S.E. Hind, R. Haines-Bazrafshan, C.L. Benton, W. Brassington, B. Towle, D. R. Moore, Prevalence of clinical referrals having hearing thresholds within normal limits, Int. J. Audiol. 50 (2011) 708–716, https://doi.org/10.3109/14992027.2011.582049.

[37] J. Katz, Handbook of Clinical Audiology, seventh ed., Wolters Kluwer Health, Philadelphia, 2014.

[38] Auditec, Selective Auditory Attention Test (SAAT), Audit. Inc., 2015. https://auditec.com/2015/08/04/selective-auditory-attention-test-saat/. (Accessed 25 September 2018).

[39] S. Aarabi, F. Jarollahi, S. Jalaei, Development and determination of the validity of Persian version of monaural selective auditory attention test in learning disabled children, Audit. Vestib. Res. J. 25 (2016) 49–54.

[40] T.T.A. Rahman, S.T. Mohamed, M.H. Albanouby, H.F. Bekhet, Central auditory processing in elderly with mild cognitive impairment, Geriatr. Gerontol. Int. 11 (2011) 304–308.

[41] K.C. De Sousa, D.W. Swanepoel, D.R. Moore, H.C. Myburgh, C. Smits, Improving sensitivity of the digits-in-noise test using antiphasic stimuli, Ear Hear. (2019) 1–9, https://doi.org/10.1097/AUD.0000000000000775.

[42] J.M. Potgieter, D.W. Swanepoel, H.C. Myburgh, T.C. Hopper, C. Smits, Development and validation of a smartphone-based digits-in-noise hearing test in South African English, Int. J. Audiol. 55 (2016) 405–411, https://doi.org/10.3109/14992027.2016.1172269.

[43] C. Smits, S.T. Goverts, J.M. Festen, The digits-in-noise test: assessing auditory speech recognition abilities in noise, J. Acoust. Soc. Am. 133 (2013) 1693–1706, https://doi.org/10.1121/1.4789933.

[44] J. Wolmarans, K.C. De Sousa, C. Frisby, F. Mahomed-Asmail, C. Smits, D.R. Moore, D.W. Swanepoel, Speech recognition in noise using binaural diotic and antiphasic Digits-In-Noise in children: maturation and self-test validity, J. Am. Acad. Audiol. 32 (2021) 315–323.

[45] I.J. Hirsh, The influence of interaural phase on interaural summation and inhibition, J. Acoust. Soc. Audiol. Am. 20 (1948) 536–544, https://doi.org/10.1121/1.1906407.

[46] J.C.R. Licklider, The influence of interaural phase relations upon the masking of speech by white noise, J. Acoust. Soc. Am. 20 (1948) 150–159, https://doi.org/10.1121/1.1906358.

[47] F. Faul, E. Erdfelder, A.G. Lang, A. Buchner, G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences, Behav. Res. Methods 39 (2007) 175–191, https://doi.org/10.3758/BF03193146.

[48] IBM Corporation, IBM SPSS Statistics for Windows, 2019. https://www.ibm.com/products/spss-statistics.

[49] K.-L. Anderssen, A. Kritzinger, L. Pottas, Auditory skills and listening comprehension in English second language learners in Grade 1, South African, J. Child. Educ. 9 (2019) a600, https://doi.org/10.4102/sajce.v9i1.600.

[50] B.M. Shield, J.E. Dockrell, The effects of environmental and classroom noise on the academic attainments of primary school children, J. Acoust. Soc. Am. 123 (2008) 133–144, https://doi.org/10.1121/1.2812596.

[51] C. Venter, L. Pottas, M. Soer, Selective Auditory Attention and Listening Comprehension in English Additional Language Learners, University of Pretoria, 2019. http://hdl.handle.net/2263/72711.

[52] D. Tabri, K.M.S.A. Chacra, T. Pring, Speech perception in noise by monolingual, bilingual and trilingual listeners, Int. J. Lang. Commun. Disord 46 (2011) 411–422, https://doi.org/10.3109/13682822.2010.519372.

[53] M. Florentine, Speech Perception in Noise by Fluent, Nonnative Listeners, Proc. Acoust. Soc. Japan, 1985. H-85-16.

[54] M. Klatte, K. Bergstr¨om, T. Lachmann, Does noise affect learning? A short review on noise effects on cognitive performance in children, Front. Psychol. 4 (2013) 1–6, https://doi.org/10.3389/fpsyg.2013.00578.

[55] M.M. Hossain, English language teaching in rural areas: a scenario and problems and prospects in context of Bangladesh, Adv. Lang. Lit. Stud. 7 (2016) 1–12.

[56] R. Alejo, A. Piquer-Pìriz, Urban vs. rural CLIL: an analysis of input-related variables, motivation and language attainment, Lang. Cult. Curric. (2016) 1–18, https://doi.org/10.1080/07908318.2016.1154068.

[57] E. Hoff, Interpreting the early language trajectories of children from low SES and language minority homes: implications for closing achievement gaps, Dev. Psychol. 49 (2013) 4–14, https://doi.org/10.1037/a0027238.

[58] S.M. Al-Zoubi, The impact of exposure to English language on language acquisition, J. Appl. Linguist. Lang. Res. 5 (2018) 151–162.

[59] K. Hall, Child Poverty, Stat. Child. South Africa, 2021. http://childrencount.uct.ac.za/indicator.php?domain=2&indicator=98. (Accessed 5 January 2022).

[60] A. Warzybok, T. Brand, K.C. Wagener, B. Kollmeier, How much does language proficiency by non-native listeners influence speech audiometric tests in noise? Int. J. Audiol. 54 (2015) 88–99, https://doi.org/10.3109/14992027.2015.1063715.

[61] M. Kaandorp, A. De Groot, J. Festen, C. Smits, S. Goverts, The influence of lexical-access ability and vocabulary knowledge on measure of speech recognition in noise, Int. J. Audiol. 55 (2015), 157167, https://doi.org/10.3109/14992027.2015.1104735.

[62] L. Lucks Mendel, H. Widner, Speech perception in noise for bilingual listeners with normal hearing, Int. J. Audiol. 55 (2016) 126–134.

[63] L.H. Mayo, M. Florentine, S. Buus, Age of second-language acquisition and perception of speech in noise, J. Speech, Lang, Hear. Res. 40 (1997) 686–693.

[64] L.F. Shi, Perception of acoustically degraded sentences in bilingual listeners who differ in age of English acquisition, J. Speech, Lang, Hear. Res. 53 (2010) 821–835, https://doi.org/10.1044/1092-4388(2010/09-0081.

[65] R. Reetzke, B.P. Lam, Z. Xie, L. Sheng, B. Chandrasekaran, Effect of simultaneous bilingualism on speech intelligibility across different masker types, modalities, and signal-to-noise ratios in school-aged children, PLoS One 11 (12) (2016) 1–19, https://doi.org/10.1371/journal.pone.0168048.

[66] L.L. Elliot, Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability, J. Acoust. Soc. Am. 66 (1979) 651–653.

[67] C.E. Johnson, Children’s phoneme identification in reverberation and noise, J. Speech, Lang, Hear. Res. 43 (2000) 144–157.

[68] D.R. Moore, Maturation and aging, in: J.J. Eggermont (Ed.), Audit. Temporal Process. Its Disord., Oxford University Press, Oxford, 2015.