



# Beyond the Classroom: Sociolinguistic Factors as Predictors of Vocabulary Breadth in Tanzania's Diverse Linguistic Landscape

**Godson Robert Mtallo**

*College of Business Education, Dodoma, Tanzania*

## Article History

*Received: 29.08.2025*

*Revised: 20.02.2026*

*Accepted: 24.02.2026*

*Published: 26.02.2026*

## Keywords

Education

Language learning

Sociolinguistics

Vocabulary

## How to cite:

Mtallo, G. R. (2026). Beyond the Classroom: Sociolinguistic Factors as Predictors of Vocabulary Breadth in Tanzania's Diverse Linguistic Landscape. *Eastern African Journal of Humanities and Social Sciences*, 5(1), 120-133.

## Abstract

This paper examines the impact of sociolinguistic factors on vocabulary breadth among secondary school learners in Tanzania. While research on sociolinguistics and language acquisition exists, the interaction between sociolinguistic factors and English vocabulary breadth in Tanzania's educational context remains underexplored. This is significant given that English, a foreign language spoken by a minority, is the medium of instruction from secondary school onwards in a multilingual country with over 150 ethnic community languages (ECLs), and Kiswahili is spoken by the majority. The study addresses: How do sociolinguistic factors influence English vocabulary breadth among learners in Tanzania's diverse linguistic landscape? A convergent parallel mixed-methods design was adopted to triangulate quantitative and qualitative data. Using stratified random sampling, the New Vocabulary Levels Test was administered to 320 Form 3 and 4 students from eight secondary schools in Dodoma, and interviews were conducted with the eight English teachers to provide deeper insight. Multiple regression and thematic analyses revealed that sociolinguistic factors significantly influenced vocabulary breadth. Learners who used English at home or in class, interacted with English-speaking peers, started learning English in kindergarten, and had parents with at least Form 4 education had broader vocabularies. Findings suggest that sociolinguistic factors play a crucial role in vocabulary development. Policymakers should enhance structured English exposure beyond the classroom and adopt additive bilingual approaches that consider learners' language ecology. This study contributes to understanding the interplay between sociolinguistic factors and vocabulary mastery in a diverse linguistic setting and recommends further research on this interaction.

Copyright © 2026



## Introduction

Vocabulary knowledge is a key element for learners' mastery of comprehension and production skills (Afzal, 2019; Brock-Utne, Desai, & Qorro, 2003; Dagnaw, 2023). In a diverse linguistic landscape, learners bring valuable sociolinguistic factors that should be identified and utilised to promote vocabulary mastery and other elements of the target language (Mapunda et al., 2024). However, the role of sociolinguistic backgrounds in developing mastery of English vocabulary breadth in Tanzania remains unclear and warrants further research. Sociolinguistic factors include social variables such



as ethnicity, age, social class, gender, and education that influence language use and acquisition (Mekacha, 1993; Mtallo & Kimambo, 2024; Rubanza, 1979). This paper, therefore, analyses how Tanzanian learners' sociolinguistic backgrounds influence vocabulary breadth scores.

With over 120 Ethnic Community Languages (ECLs), Tanzania's linguistic landscape is characterised by English as a foreign language and Kiswahili as a lingua franca (LOT, 2009; Mtallo, 2015). Kiswahili is the medium of instruction in primary schools, while English is taught as a subject; in secondary and higher education, English becomes the medium of instruction (Mtallo, 2015; Qorro, 2013). This shift poses challenges for acquiring English vocabulary breadth, a hallmark of language proficiency (Mtallo & Kimambo, 2024). Given Tanzania's linguistic and sociocultural dynamics (Brock-Utne, Desai & Qorro, 2004; Mapunda et al., 2024), it is crucial to examine sociolinguistic factors underlying variations in vocabulary development.

Research highlights the influence of home language, cultural development, exposure to linguistic input, and parental education on second-language acquisition (Ghalebi et al., 2021; Mtallo & Kimambo, 2024). Sociolinguistic diversity also shapes pedagogical practices and classroom processes (Baker & Jones, 2021; Mekacha, 1993). In Tanzania's multilingual context, the coexistence of many languages presents challenges and opportunities for learners' vocabulary development (Afzal, 2019). English vocabulary breadth is particularly affected, as learners from different linguistic settings show varying rates of mastery (Mtallo & Kimambo, 2022).

Although substantial research exists (Zaretsky & Lange, 2015; Mtallo & Kimambo, 2022; Ghalebi et al., 2021), the interaction between sociolinguistic factors and vocabulary breadth in Tanzania remains underexplored. This study addresses this gap by examining variables such as home language, exposure to English-speaking environments, parental education, interaction with native speakers, and age of first learning English.

Vocabulary breadth is central to communication and proficiency (Nation, 2013). A stronger vocabulary repertoire enhances comprehension, communication, and expression (Beck et al., 2013). Despite its foundational role (Afzal, 2019), vocabulary breadth has received limited attention in Tanzania's diverse linguistic landscape, especially given the transition from Kiswahili to English as medium of instruction (Mtallo, 2015). This study therefore asks: How do sociolinguistic factors influence English vocabulary breadth among learners in Tanzania's diverse linguistic landscape? By addressing this question, it illuminates pathways for improving vocabulary teaching and learning in Tanzania and similar contexts worldwide.

### **Literature Review**

This section reviews literature on vocabulary development among learners. It is organised thematically, focusing on key concepts and theoretical framing, the influence of sociolinguistic factors on vocabulary acquisition, multilingual exposure and vocabulary breadth, methodological approaches, and the research gap this paper addresses.

Vocabulary breadth refers to the number of words a learner knows at a surface level (Nation, 2001). It is central to language proficiency, particularly in multilingual contexts such as Tanzania. Learners' sociolinguistic backgrounds—including home language use, socioeconomic status, and cultural norms—influence vocabulary acquisition (Cummins, 1979; Gass & Selinker, 2008). Vygotsky's (1978) Sociocultural Theory explains that social interaction and cultural tools mediate language learning. Learners' interactions with family, peers, teachers, and the community shape their linguistic repertoire and mastery of English vocabulary. Understanding sociolinguistic influences on vocabulary breadth is therefore essential in multilingual settings.



Several sociolinguistic factors affect vocabulary breadth, including first-language proficiency (Cummins, 1979), learning context (Peets et al., 2022), access to learning materials (Biseko, 2025; Mtallo & Kimambo, 2024), and learning strategies (Kilangi, 2020). First-language proficiency may facilitate literacy and vocabulary development in the target language. Socioeconomic status and cultural resources also shape access to richer linguistic input, often giving learners from more literate families broader vocabularies (Mtallo & Kimambo, 2024). These dynamics reflect Vygotsky's (1978) view that access to cultural tools varies across social contexts.

Multilingual exposure may enhance cognitive flexibility and metalinguistic awareness, yet it can also slow English vocabulary growth when instructional support is limited (Mtallo & Kimambo, 2022). Language development depends on the quality and quantity of input within learners' communities (Biseko, 2025). While Cummins (1979) highlights cross-linguistic transfer, Vygotsky (1978) emphasises sociocultural mediation; both underscore the role of linguistic environments in shaping vocabulary breadth.

Methodologically, studies use quantitative vocabulary tests and qualitative interviews or observations (Oladini et al., 2024). Given the complexity of vocabulary development, mixed-methods approaches provide more comprehensive insights.

Despite growing research, most studies focus on Western or urban contexts, overlooking Tanzanian learners whose language use diverges from classroom instruction. This study addresses this gap by examining how sociolinguistic factors influence vocabulary breadth to inform contextually relevant English language instruction in Tanzania.

### **Theoretical framework**

This study is theoretically informed by Vygotsky's (1978) Sociocultural Theory (SCT) of human development. SCT posits that cognitive development is mediated by social interaction and that language and other cultural tools play a central role in this process. Learning is therefore not an isolated individual activity but a social construct emerging through participation in socially organised practices.

A key construct in SCT is mediation, whereby language serves social and mental functions that enable individuals to acquire shared knowledge. Another central concept is the Zone of Proximal Development (ZPD), defined as the gap between what a learner can accomplish independently and what can be achieved with guidance from more knowledgeable others (Vygotsky, 1978). Through scaffolding within the ZPD, learners gradually develop linguistic skills, including vocabulary.

In Tanzania's diverse linguistic landscape, learners' sociolinguistic experiences—such as home language use and engagement with English in and out of school—constitute social conditions that mediate vocabulary learning. As English serves as the medium of instruction in post-primary education, opportunities for meaningful interaction in English are crucial for vocabulary development. From an SCT perspective, vocabulary breadth is not merely an internal mental lexicon but a product of socially mediated language engagement.

It is therefore posited that differences in learners' sociolinguistic experiences predict differences in English vocabulary breadth. By examining vocabulary breadth in relation to diverse linguistic experiences, this study applies SCT to conceptualise vocabulary knowledge as socially constructed, culturally situated, and distributed across learners' interactional contexts (Vygotsky, 1978).



## Method

This section presents the methods employed in conducting this study. It presents the research approach and design, the participants, the data-collection instruments, the data-analysis procedures, and, lastly, the ethical issues observed in this study.

### *Research approach and design*

This paper adopted a mixed-methods approach, simultaneously collecting and analysing qualitative and quantitative data, drawing on Tashakkori and Teddlie's (2003) view of their complementarity. With a dominant quantitative strand, the study examined the relationship between sociolinguistic factors (independent variables) and English vocabulary breadth scores (dependent variable). A convergent parallel mixed-methods design was employed, consistent with Creswell's (2014) recommendation that this design yields richer descriptions when identical variables are measured across strands. Quantitative and qualitative data were analysed separately and then compared to confirm or challenge findings.

This approach enabled the integration of statistical results with qualitative themes, offering a comprehensive understanding of how sociolinguistic backgrounds shape vocabulary breadth. Insights were drawn from teachers' classroom perspectives and students' performance outcomes. The mixed-methods design was therefore essential in providing a holistic account of vocabulary acquisition, with each strand complementing the other.

### *Participants*

The study was conducted in selected secondary schools in Dodoma, Tanzania, from both urban and rural settings, to accommodate diverse sociolinguistic backgrounds, given the region's central position and rapidly growing population. It involved 320 Tanzanian-born secondary school learners (aged 15 to 17 years old) from 8 schools, where 4 schools are public, and the other 4 are privately owned, and 8 English language teachers from the same schools, targeting Forms 3 and 4. The justification for the sample size in this study draws on Mason (2002: 13), with a sample size slightly below 384 chosen for practical reasons, specifically the rule of thumb that a minimum threshold of 30 is acceptable. It used respondent-driven sampling to select experienced English teachers and stratified random sampling for learners, ensuring representation across sociolinguistic groupings and school types, including both teachers and students.

### *Instruments*

The study employed both quantitative and qualitative data collection instruments to effectively capture the experiences and voices of learners and teachers, respectively. In particular, the study gathered quantitative data using the New Vocabulary Levels Test (NVLTL) by McLean and Kramer (2015) and qualitative data using a semi-structured interview guide. The NVLTL measured the breadth (size) of the English vocabulary knowledge by estimating the total number of words a learner knows (McLean & Kramer, 2015; Webb & Sasao, 2017). This estimation is useful when comparing groups of learners or measuring long-term vocabulary growth (Nation, 2013). The test required the learners to match the target word (English vocabulary) with the corresponding English definitions or translations. A good example is seen in an excerpt presented by McLean and Kramer (2015: 10), as follows:

- (1) **See:** They saw it.
- A. Cut
  - B. Waited for
  - C. Looked at
  - D. Started
- C the correct answer is C.



In this test, the student chooses A, B, C, or D based on the meaning closest to the vocabulary in **boldface**. It had 150 matching items, grouped into six different Levels (NVL T Part 1–6). It required the test taker to match the given definitions/explanations with the target word in bold. Each correct answer was worth 1 point, and the maximum score per level was 30. In total, the test taker could score up to 150 points (Bardakci, 2016). The scores were then converted into percentages for analysis. Qualitative interviews with eight teachers followed the quantitative data collection to explain the quantitative data and provide descriptions that numbers could not. Utilising both tools corroborated findings from both quantitative and qualitative strands.

### *Data analysis procedures*

Quantitative data from the New Vocabulary Levels Test were analysed using SPSS 25. Correlations assessed relationships between sociolinguistic factors and English vocabulary breadth, while multiple linear regression identified the strongest predictors. Each variable was regressed against vocabulary scores to determine its contribution.

Semi-structured interview data were analysed following Braun and Clarke's (2006) six-stage process: familiarisation, coding, theme development, review, definition, and reporting. Emerging themes were compared with existing theories and studies to understand how sociolinguistic factors influence vocabulary breadth.

Integrating quantitative and qualitative data provided richer insights into the relationship between sociolinguistic variables and English vocabulary mastery among secondary school students.

### *Ethical considerations*

Before the actual data collection, research clearance was sought from the University of Dar es Salaam, which is mandated to issue research clearances on behalf of Tanzania's Commission for Science and Technology (COSTECH) for researchers conducting research in Tanzania. Teachers' and students' consent to participate in the study was then obtained. The participants were assured of confidentiality and given the freedom to withdraw from the study without negative consequences. The data were anonymised and securely stored on a password-protected computer to ensure confidentiality and privacy during data analysis and reporting. The following section presents the results.

### **Results**

This section presents mean differences in vocabulary breadth scores across learners' sociolinguistic backgrounds, considering languages used at home, in class, and on the playground; parental education levels; interaction with English-speaking peers; and age of first exposure to English. Multiple linear regression was employed to test the main effects of sociolinguistic background on breadth scores. Selected qualitative excerpts complement the quantitative results, while detailed interpretation is reserved for the discussion section.

The study addressed the question: How do sociolinguistic factors shape learners' English vocabulary breadth in Tanzania's diverse linguistic landscape? Multiple regression and thematic analyses revealed that sociolinguistic factors significantly influenced vocabulary breadth. Learners who used English at home or in class, interacted with English-speaking peers, began learning English in kindergarten, and had parents with at least Form Four education demonstrated broader vocabulary knowledge. Overall, the findings confirm the significant role of sociolinguistic factors in shaping learners' English vocabulary breadth.

### *Mean breadth score differences according to home language*

Table 1 displays the difference in vocabulary breadth scores between students who used English at home and those who used Kiswahili or Ethnic Community Languages (ECLs). Students who used



English at home (n = 155) had a higher average score (M = 60.16) compared to those who used Kiswahili/ECLs (n = 165, M = 56.24). The difference was statistically significant (t = 2.307, p = .012).

*Table 1: Mean breadth scores by home language*

Home language	N	Mean	t	Sig.
Kiswahili/ECLs	165	56.24		.012
English	155	60.16	2.307	

In the qualitative examination, one teacher noted a corresponding observation: “Students who use English at home exhibit significantly higher proficiency compared to those who use either Swahili or Ethnic Community Languages (ECLs)” (T6, August 2023). (Note that the participants are anonymised throughout by combining letters and numbers. For example, T6 refers to a teacher (T) participant, with the number (6) indicating their sequence in the interviews.) Likewise, in a semi-structured interview, another participant echoed this sentiment, arguing: “Most of the learners whose households communicate in English demonstrated higher proficiency compared to their counterparts” (T4, July 2023).

While many participants could not explain their experiences in this regard, they shared a common observation: learners from households where English is frequently used tend to exhibit greater vocabulary proficiency than those who mostly use either Kiswahili or Ethnic Community Languages (ECLs) at home. However, one participant diverged from this consensus, expressing the following viewpoint: “I am unable to discern a clear distinction, although I observe that they encounter similar challenges.” (T3, July 2023)

Therefore, the qualitative findings are consistent with the quantitative results, affirming that the use of English within the household plays a significant sociolinguistic role in enhancing English vocabulary breadth.

#### *Mean breadth score differences based on the language used in the classroom*

As shown in Table 2, learners using English in the classroom (n = 166) scored slightly higher (M = 56.86) than those using Kiswahili/ECLs (n = 154, M = 54.22). However, the difference was not statistically significant (t = 1.048, p = .295). Thus, the data fail to support the notion that speaking English or other languages in the classroom impacts performance on the breadth test.

*Table 2: Mean breadth scores by the language used in the classroom*

Classroom language	N	Mean	t	Sig.
Kiswahili/ECLs	154	54.22	1.048	.295
English	166	56.86		

The findings contradict qualitative data. For instance, one of the interviewees stated:

*There is no way you can find these students at an equal level of language proficiency. Those who [...] use English during class, whether willingly or otherwise, are not inherently superior to those who opt for either ethnic community languages or Kiswahili, Tanzania's predominant language. It is like the saying, 'Practice makes perfect.' Hence, the more a learner engages with a specific language, the more proficient they become, and vice versa. (T8, August 2023)*



From the observed disparities, using English in the classroom could indeed play a sociolinguistic role in enhancing the development of English vocabulary breadth. However, for this factor to have a significant impact, other factors, such as learners' motivation for the language, the availability of appropriate teaching resources, teachers' proficiency in English, and intelligence, may be necessary. This perspective aligns with the sociocultural theory, which emphasises the importance of both social and cognitive factors in learning. Therefore, considering these factors is essential as they collectively contribute to the learning experience.

#### ***Mean breadth score differences according to parents' educational levels***

Table 4 reveals that students whose parents had at least a Form 4 level of education ( $n = 163$ ,  $M = 61.55$ ) scored significantly higher than those whose parents had less than a Form 4 level of education ( $n = 157$ ,  $M = 51.36$ ). The difference was statistically significant ( $t = 6.759$ ,  $p = .000$ ).

*Table 4: Mean breadth scores by parents' educational level*

Parents' educational level	N	Mean	t	sig
Below Form 4	157	51.36	6.759	.000
Form 4 or above	163	61.55		

The majority of interviewees recounted instances of facing such scenarios within their classrooms. One teacher noted that learners whose parents had at least a Form 4 education tended to have stronger English vocabulary knowledge than those whose parents had lower levels of education. According to the teacher, more educated parents were generally more supportive of their children's learning, often providing books and exercise materials and staying involved in school matters. In contrast, dropout and truancy rates appeared higher among students whose parents had less formal education. (T1, July 2023)

Another teacher added:

*Students whose parents have attained a minimum of Form 4 education exhibit a comparatively higher knowledge of vocabulary (T7, August 2023)*

Nevertheless, not all teachers fully supported this perspective. For example, one teacher partly agreed with the perspective. He remarked:

*The difference in English proficiency between students with more and less educated parents isn't always significant. It depends on the language spoken at home. If students regularly speak English at home and their parents are proficient in English, it can be an advantage. (T4, July 2023)*

The extracts show that the parental level of education is indeed a significant sociolinguistic factor influencing the development of English vocabulary size by Tanzanian learners of English. This is also in close agreement with the More Knowledgeable Other (MKO) principle of sociocultural theory, which suggests that learners learn better when scaffolded by individuals with more knowledge than they have, in this case, their parents and guardians.

#### ***Mean breadth score differences based on opportunities to interact with English-speaking children***

The data show that students who had opportunities to interact with English-speaking children during the first ten years of life ( $n = 150$ ) scored significantly higher ( $M = 61.13$ ) than those who did not ( $n = 170$ ,  $M = 53.80$ ), as shown in Table 5 ( $t = 4.543$ ,  $p = .000$ ).



*Table 5: Mean breadth scores by English-speaking peer interaction*

Interacting with English-speaking children	N	Mean	t	Sig
No	170	53.80	4.543	.000
Yes	150	61.13		

This finding suggests that early interactions with English-speaking peers are a crucial sociolinguistic factor in acquiring breadth of English vocabulary in a diverse linguistic landscape. When asked about observing differences in vocabulary knowledge among students in their classes, T1 said that the social environment was extremely important. Anyone who had the opportunity to be in such a setting would inevitably adopt the language, behaviour, and culture of those around them. Although they had not observed this in their class, learners who had the chance to interact with English-speaking children developed an interest in the language, which facilitated vocabulary development. (July 2023)

Another teacher added:

*Perhaps those who have been exposed to an English-speaking environment may find it easier to acquire a larger English vocabulary than those who have never had such exposure. However, I have not observed this situation in my classes. (T3, July 2023)*

Both quantitative and qualitative data align strongly, indicating that opportunities to interact with English-speaking children significantly shape the breadth of English vocabulary mastery. Students with such exposure tend to perform better than those without. This is also supported by teachers' observations, which highlight that exposing L2 learners to an L2-speaking environment is crucial to their success in learning the language.

***Mean breadth score differences based on the age of commencing learning English***

An analysis of variance (ANOVA) was used to test differences in scores based on when English learning began. The following table presents the grouped means of breadth scores by the level at which learners began learning English in Tanzania.

*Table 6: Mean breadth scores by the level of starting to learn English*

English learning starting level	N	Subset for alpha = 0.05		F value	Sig.
		1	2		
Primary school	127	52.04		23.536	.000
Secondary school	103	56.53			
Kindergarten	90		63.40		

As shown in the table, students who began learning English at kindergarten (n = 90, M = 63.40) performed significantly better than those who began learning it at primary (n = 127, M = 52.04) or secondary school (n = 103, M = 56.53), with a significant difference (F = 23.536, p = .000). The difference indicates that commencing second language learning at an early age positively impacts vocabulary acquisition level. The age at which English is first learned is a crucial sociolinguistic factor that influences vocabulary breadth in Tanzania's diverse linguistic landscape. This observation was echoed in the qualitative data. One teacher said:



*I agree with the saying “the earlier, the better” based on my experience. Individuals who began learning English at a younger age tend to exhibit greater overall language proficiency than those who began later. This is evident in my classes. (T1, July 2023)*

While most of the interviewed teachers affirmed that learners who had started English learning at the kindergarten level exhibited considerable proficiency in English vocabulary, a few expressed differing views, indicating uncertainty. T4, for instance, found no notable difference in English vocabulary size between learners who began at kindergarten and those who began at primary or secondary school levels. Some of the teachers explained that differences might initially manifest when learners transition to secondary school, but that they gradually diminish over time until both groups reach a comparable level of language proficiency.

Generally, teachers' experiences aligned with vocabulary breadth test scores, indicating that the age at which one starts learning English determines how broad one's vocabulary becomes. This means that a person exposed to English at an early age has a greater chance of acquiring a broader vocabulary than those who start later. This finding shows the interplay between the starting age of learning English as a sociolinguistic factor and the learner's age as another factor that influences second-language development. While the individual analyses above show significant sociolinguistic differences, the regression model below assesses the combined predictive power of these variables.

#### ***Regression analysis across sociolinguistic factors***

To examine the combined influence of sociolinguistic factors on vocabulary breadth, a multiple regression model was developed. The independent variables were selected based on their theoretical relevance and empirical significance in earlier subsections (4.1–4.5). Specifically, they capture learners' exposure to English across different domains, including early learning experiences (e.g., nursery school attendance), medium of instruction, and participation in English classes. The regression model thus provides a consolidated analysis of how these sociolinguistic experiences predict vocabulary breadth scores.

#### ***Model summary***

The regression model assessed the effect of sociolinguistic background on vocabulary breadth. The model showed a strong positive correlation ( $R = 0.847$ ) and explained 71.7% of the variance in breadth scores (Adjusted  $R^2 = 0.701$ ), with  $F(4, 315) = 24.454$ ,  $p < .05$ .

*Table 7: Model summary*

Model	R	R <sup>2</sup>	Adjusted R <sup>2</sup>	Std. Error of the Estimate	F	Sig.
1	.847 <sup>a</sup>	.717	.701	12.661	24.454	.000 <sup>b</sup>

The coefficient of determination ( $R^2$ ) measures the strength of the association between breadth scores and learners' sociolinguistic backgrounds by indicating the proportion of variance explained by the model. As shown in Table 7,  $R^2$  was 0.717, meaning the model explains 71.7% of the variation in breadth scores. The Adjusted  $R^2$ , which accounts for the number of predictors, was 0.701, indicating that sociolinguistic background explains 70.1% of the variance, while 29.9% is attributable to other factors. This finding supports Sociocultural Theory, which posits that language development emerges from the interplay of social and cognitive influences. The regression model was statistically significant and appropriate for explaining the relationship between breadth scores and sociolinguistic background,  $F(4, 315) = 24.454$ ,  $p \leq 0.05$ .



*Model coefficients*

The regression model equation employed to examine the correlation between breadth scores and the learner’s sociolinguistic background is as follows:

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4$$

The definitions of the variables utilised in the aforementioned equation are provided in the variable summary outlined in Table 8.

*Table 8: Summary of variables used in the breadth model*

<i>Variable</i>	<i>Description</i>
$X_1$	Were all the subjects in primary school taught in English?
$X_2$	Did you attend a nursery school during the first three years of your life?
$X_3$	Did you happen to attend a daycare centre?
$X_4$	Have you attended/ participated in an English class course?
$Y$	Breadth scores

In the regression model,  $\beta$  values represent the change in breadth scores associated with a one-unit change in a learner’s sociolinguistic background, holding other variables constant. These coefficients ( $\beta_0$ - $\beta_4$ ) are drawn from the unstandardised coefficients in Table 9. Substituting these values into the regression equation produces a model describing the relationship between breadth scores and sociolinguistic variables. Each coefficient indicates the expected increase in breadth scores for a one-unit increase in the corresponding independent variable, assuming other factors remain constant. The significance column shows the statistical contribution of each predictor. Variables  $X_1$ ,  $X_3$ , and  $X_4$  were significant at  $\alpha < 0.05$ . The final column in Table 9 presents the 95% confidence intervals for each coefficient, indicating their lower and upper bounds.

*Table 9: Breadth model coefficients*

Model	Unstandardised Coefficients		Standardised Coefficients	t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
(Constant)	53.994	1.069		50.529	.000	51.892	56.097
$X_1$	14.725	1.764	.456	8.347	.000	11.254	18.196
$X_2$	1.604	1.701	.053	.943	.346	1.743	4.951
$X_3$	3.817	2.675	.079	1.427	.045	1.446	9.079
$X_4$	3.035	1.517	.104	2.001	.046	.050	6.021

Therefore, the results reveal that learners’ sociolinguistic backgrounds significantly influence their breadth of English vocabulary. Specifically, using English at home, having parents with higher levels of education, interacting with English-speaking peers, and beginning to learn English at an early age were all associated with higher vocabulary scores. While not all individual factors reached statistical significance, the overall regression model showed that these sociolinguistic variables collectively accounted for over 70% of the variance in vocabulary breadth. These findings underscore the complex



and cumulative impact of learners' linguistic environments on second-language development in a diverse linguistic landscape such as Tanzania. The following section presents a discussion of the findings, drawing on both quantitative and qualitative data from the preceding sub-sections.

### **Discussion**

The findings indicate that several sociolinguistic factors contribute to learners' English vocabulary breadth in Tanzania's diverse linguistic landscape. Each factor is discussed in relation to the data and existing literature.

Regarding home language use, students who used English at home demonstrated significantly greater vocabulary breadth than peers using Kiswahili or ECLs. Teachers attributed this to frequent exposure and practice in domestic settings. This aligns with Cheung et al. (2018), who found that home language strongly shapes vocabulary development in bilingual learners.

Although classroom language did not yield statistically significant differences, qualitative insights suggest English use can enhance proficiency when supported by adequate resources, teacher competence, and learner motivation. This reflects Vygotsky's (1978) sociocultural emphasis on environmental mediation and concurs with Khan (2016) and Tahir et al. (2020), who stress the importance of meaningful language use for vocabulary growth.

No significant effect was found for playground language; interviews revealed limited informal English use. This suggests that English in Tanzania remains largely confined to formal domains, unlike contexts described by Zoubi (2018). Vocabulary breadth appears closely tied to the frequency of meaningful engagement.

Parental education level showed a strong relationship with vocabulary breadth, with more educated parents providing better linguistic and academic support, consistent with Dubow et al. (2009).

Interaction with English-speaking peers and early exposure to English (kindergarten level) also significantly enhanced vocabulary mastery, supporting Adelman (2014) and Wallin and Cheevikumjorn (2020).

Regression analysis confirmed that these sociolinguistic variables collectively explain substantial variance in vocabulary scores, reinforcing Vygotsky's (1978) concept of the More Knowledgeable Other and the centrality of social context in second-language development.

### **Conclusion**

This paper has analysed how sociolinguistic factors influence English vocabulary breadth among learners in Tanzania's diverse linguistic landscape, where English is a foreign language. Home language use, parental education, early exposure to English, and interaction with English-speaking peers strongly predicted vocabulary breadth. Quantitative results showed differences in breadth scores across these variables, while qualitative findings provided contextual support. Although classroom and playground language use were not statistically significant, interviews suggest these domains can contribute to vocabulary development when supported by appropriate pedagogy and resources. Regression analysis confirmed that sociolinguistic variables collectively explain substantial differences in learners' vocabulary scores.

These findings support Vygotsky's (1978) concept of the More Knowledgeable Other, emphasising learning through interaction in supportive social environments. They further demonstrate that second language development is both cognitive and social; meaningful exposure to English enables learners to expand vocabulary through engagement with cultural artefacts and contextual distinctions.



The results highlight the importance of early, consistent, and meaningful English exposure from home to school. Policymakers and educators should therefore integrate sociolinguistic realities into curriculum design and teacher education, strengthen structured English exposure beyond the classroom, and adopt additive bilingual approaches that reflect learners' language ecology in Tanzania's multilingual context.

However, the paper did not consider the role of affective variables in developing vocabulary breadth, and it covered only the Dodoma region in Tanzania. Another study can thus focus on the interaction between these sociolinguistic factors and affective variables in determining vocabulary breadth mastery and whether these patterns are evident in similar linguistic landscapes across the Global South, particularly where English is a foreign language. Moreover, the study's reliance on interviews with only teachers limited the ability to gather student data on vocabulary development. Adding students' perspectives will offer a more comprehensive, real-time picture of vocabulary knowledge.

### References

- Adelman, A. S. (2014). *Conversations with native speakers: Acquiring Japanese as a second language* (Doctoral dissertation). University of California, Oakland.
- Afzal, N. (2019). A study on vocabulary-learning problems encountered by BA English majors at the university level of education. *Arab World English Journal*, 10(3), 81–98.  
<https://dx.doi.org/10.24093/awej/vol10no3.6>
- Baker, C., & Jones, S. P. (2021). *Multilingual education in practice: A guide to effective policies and practices*. Multilingual Matters.
- Bardakçi, M. (2016). Breadth and depth of vocabulary knowledge and their effects on L2 vocabulary profiles. *English Language Teaching*, 9(4), 239–250. <http://dx.doi.org/10.5539/elt.v9n4p239>
- Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction*. Guilford Press.
- Biseko, J. M. (2025). Vocabulary learning in the EFL context: Do primary school English subject textbooks provide structured support? *Cogent Education*, 12(1), 2455047.  
<https://doi.org/10.1080/2331186X.2025.2455047>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brock-Utne, B., Desai, Z., & Qorro, M. (Eds.). (2003). *Language of instruction in Tanzania and South Africa (LOITASA)*. E&D Limited.
- Brock-Utne, B., Desai, Z., & Qorro, M. (Eds.). (2004). *Focus on fresh data on the language of instruction debate in Tanzania and South Africa*. African Minds.
- Cheung, S., Kan, P. F., Winicour, E., & Yang, J. (2018). Effects of home language input on the vocabulary knowledge of sequential bilingual children. *Bilingualism: Language and Cognition*.
- Creswell, J. W. (2014). *Research design*. Sage.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2), 222–251.
- Dagnaw, A. T. (2023). Revisiting the role of breadth and depth of vocabulary knowledge in reading comprehension. *Cogent Education*, 10(1), 2217345.
- Dubow, E. F., Boxer, P., & Huesmann, L. R. (2009). Long-term effects of parents' education on children's educational and occupational success: Mediation by family interactions, child aggression, and teenage aspirations. *Merrill-Palmer Quarterly*, 55(3), 224–249.  
<https://dx.doi.org/10.1353/mpq.0.0030>
- Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). Routledge.



- Ghalebi, R., Sadighi, F., & Bagheri, M. S. (2021). A study of vocabulary learning strategies among high and low Iranian English vocabulary learners. *Cogent Education*, 8(1), Article 1834933. <https://doi.org/10.1080/2331186X.2020.1834933>
- Khan, M. S. (2016). The impact of native language use on second language vocabulary learning by Saudi EFL students. *English Language Teaching*, 9(5), 134-140.
- Kilangi, C. T. (2020). *Language learning strategies use and vocabulary knowledge among Tanzanian learners of Chinese as a foreign language* (Doctoral thesis). University of Dar es Salaam.
- Laufer, B. (2018). Vocabulary acquisition in a second language: Do learners acquire most vocabulary by reading? *The Canadian Modern Language Review*, 74(2), 132-147.
- LOT. (2009). *Atlasi ya lugha za Tanzania*. Mradi wa Lugha za Tanzania, Chuo Kikuu cha Dar es Salaam.
- Mapunda, G., Costley, T., Gibson, H., Kiula, N., & Reilly, C. (2024). Rethinking language: The need for language-supportive pedagogies in teacher training in Tanzania. *Nordic Journal of African Studies*, 33(3), 184-202. <https://doi.org/10.53228/njas.v33i3.1025>
- Mason, J. (2002). *Qualitative researching*. Sage.
- Mekacha, R. D. (1993). *The sociolinguistic impact of Kiswahili on ethnic community language in Tanzania*. African Studies.
- McLean, S., & Kramer, B. (2015). The creation of a new vocabulary levels test. *Shiken*, 19(2), 1-11.
- Mtallo, G.R. (2015). Teaching and Learning English in Tanzania: Blessing or Curse? A Practical Review of Phan Le Ha's Teaching English as an International Language. *Journal of Education and Practice*, 6(17), 118-123.
- Mtallo, G. R., & Kimambo, G. (2022). Acquisition of breadth and depth of vocabulary knowledge by secondary school learners of English in the multilingual context: A case of Dodoma, Tanzania. *East African Journal of Education and Social Sciences*, 3(1), 38-48.
- Mtallo, G.R., & Kimambo, G. (2024). Sociolinguistic factors influencing the acquisition of English vocabulary in multilingual contexts: Insights from selected secondary school teachers in Dodoma, Tanzania. *Journal of Modern Research in English Language Studies*, 11(4), 51-78. <https://doi.org/10.30479/jmrels.2024.19603.2292>
- Nation, I. S. P. (2013). *Learning vocabulary in another language*. Cambridge University Press.
- Oladini, R., Mazlum, F., & Dasta, M. (2024). The relationship between Iranian EFL learners' vocabulary learning strategies and their breadth and depth of vocabulary knowledge. *Journal of Modern Research in English Language Studies*, 11(2), 76-99.
- Peets, K. F., Yim, O., & Bialystok, E. (2022). Language proficiency, reading comprehension, and home literacy in bilingual children: The impact of context. *International Journal of Bilingual Education and Bilingualism*, 25(1), 226-240. <https://doi.org/10.1080/13670050.2019.1677551>
- Qorro, M. (2013). Language of instruction in Tanzania: Why are research findings not heeded? *International Review of Education*, 59(1), 29-45. <https://doi.org/10.1007/s11159-013-9329-5>
- Rubanza, Y. I. (1979). *The relationship between Kiswahili and other African languages: The case of Kihaya* (Unpublished master's dissertation). The University of Dar es Salaam.
- Tahir, M. H. M., Albakri, I. S. M. A., Adnan, A. H. M., Shaq, M. S. Y., & Shah, D. S. M. (2020). The application of visual vocabulary for ESL students' vocabulary learning. *Arab World English Journal*, 11(2), 323-338. <https://dx.doi.org/10.24093/awej/vol11no2.23>
- Tashakkori, A., & Teddlie, C. (2003). *Handbook of mixed methods in social and behavioural research*. Sage.
- Vygotsky, L. S. (1978). *Mind in society: Development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press. <https://doi.org/10.2307/j.ctvjf9vz4>
- Wallin, J., & Cheevakumjorn, B. (2020). Learning English as a second language: Earlier is better. *Journal of English Education Society*, 5(1), 1-8.



- 
- Webb, S., & Sasao, Y. (2017). The updated vocabulary levels test. *International Journal of Applied Linguistics*, 168(1), 33-39.
- Zaretsky, E., & Lange, B. P. (2015). Sociolinguistic factors influencing language acquisition in German preschoolers in three follow-up studies. In *The European Conference on Language Learning* (pp. 1-13). Official Conference Proceedings.
- Zoubi, M. A. (2018). The impact of exposure to the English language on language acquisition. *Journal of Applied Linguistics and Language Research*, 5(4), 151-162.