JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Effects of Multimodal Instruction on Students' Performance in English Speaking Skills and Self-efficacy in Colleges of Education, North West Zone, Nigeria

[1]Jamilu Sani, [2]Sule Sale, [3]Mustapha Shettima
[1]Lecturer 1, [2]Lecturer 2, [3]Lecturer 3
[1]Department of English
[1]Jigawa State College of Education, Gumel, Jigawa State, Nigeria

Abstract: This study investigated the effects of multimodal instruction on students' performance in English speaking skills and self-efficacy in Colleges of Education, North West Zone, Nigeria. Quasi-experimental research was adopted involving the 2024/2025 registered 488 NCE 1 English language students of nine Colleges of Education in the study area. Purposive sampling was used to select 4 out of the 9 Colleges of Education with a total of 204 NCE 1 students as the sample size. Three sets of data were collected: pre-test, post-test and post-treatment self-efficacy data. These were analysed using One-Way ANOVA and descriptive statistics respectively. Results indicated significant improvements in English speaking skills. Mean scores rose from 2.543, 9.442, 9.915, 0.000 and 7.413 in pre-test to 1404.914, 885.219, 873.223, 1370.581 and 791.770 in post-test in terms of ESPS, ESSPS, EVPS, EGS and EFS, respectively. Further, subjects demonstrated higher degree of self-efficacy in English speaking skills after the treatment. It was concluded the improvements followed the series of multimodal trainings to which the learners were exposed involving linguistic and non-linguistic modes. Although limited in terms of scope and rating criteria, the study suggested adoption of wider scope and technology-informed rating criteria by future research.

Index Terms - multimodal instruction, English speaking skills, self-efficacy, Colleges of Education, North-West Zone, Nigeria.

I. INTRODUCTION

The capacity to communicate effectively in English has become a crucial predictor of academic success. In Nigeria, specifically, English functions as a second language: a measure of academic achievement in schools. More particularly, in Colleges of Education in Nigeria, the curriculums for teaching and learning of English speaking skills are designed to equip learners with the necessary articulatory skills. In return, the essence is for the students to be able to teach English effectively as a subject and language of instruction at the basic education levels. The curriculums are geared towards equipping pre-service English teachers deliver effective communication skills for primary and lower secondary schools in Nigeria (National Commission for Colleges of Education, 2020). In spite of this, however, the spoken English skills of Nigerian graduates are not in commensurate with the communicational demands of the 21st century society and expectations.

More worrisome is when such problems manifest also persistently in the performance of English language learners in Colleges of Education. This is because such learners would soon, after graduation, become teachers of the English language in primary and lower secondary schools (Nigeria, Federal Republic, 2020). Thus, from the personal observation of the researchers, in Colleges of Education, teachers are still complacent to the fact that pre-service English teachers are already multimodal savvy. Thus, the prevalent use of the mono-modal – traditional textual lecture method would not enable learners perform up to the expectations in the English speaking skills (kummin, Surat, Mydin, Othman & Muslim, 2020). This would also further inform the basis of the persistent low English speaking self-efficacy of college-level learners in the study area. Therefore, this calls for a scientific investigation to establish empirically whether exposure to multimodal instruction can significantly improve college-level learners' performance in English speaking skills and self-efficacy in the North West Zone of Nigeria.

Multimodal classroom approach is a contemporary response to the changes in the 21st century mode of communication. The approach finds reflection and expression in teaching English in non-native contexts (Belabbas, 2024). To elaborate this position, Maghsoudi, Golshan and Naemi (2022) examined multimodal strategies, self-efficacy and English writing skills among Iranian English as a foreign language students. The emphasis was on multimodality as a predictor of high self-efficacy in writing skills. Results showed that the multimodal group scored significantly higher than the monomodal group in terms of content, communication, organization and achievement. Despite this significant revelation, English language teachers in Colleges of Education in North West Zone, Nigeria, are still under-informed in terms of the use of multimodal instruction as a predictor of high self-efficacy in English speaking skills. Both teachers and students in the study area would require the multimodal instructional framework since it has the potentials of integrating both the linguistic and the non-linguistic learning modes during English speaking tasks in the actual classroom.

From the foregone, therefore, the aim of the present study is to establish empirically the effectiveness of multimodal instruction on students' performance in English speaking skills and self-efficacy in Colleges of Education in the North West Zone of Nigeria. Thus, the research objectives of this study are: (1) to investigate the effects of multimodal instruction on students' performance in English speaking skills in Colleges of Education, North West Zone, Nigeria; and (2) to evaluate weather exposure to multimodal instruction increases learners' levels of self-efficacy in interpersonal English speaking contexts.

This study is expected to contribute to theory, practice and curriculum in English Language Education in the following major ways: (1) in terms of theory, the study contributes to the emerging field of research on multimodality in education; (2) practically, it provides teacher-educators in North West Zone, Nigeria, with empirical evidence and guidance on the use of multimodal framework in the actual English speaking classroom for the enhancement of participation, engagement and interest; (3) in terms of curriculum, the study would enlighten the regulatory body of Colleges of Education in Nigeria (NCCE) to re-assess the areas of English pedagogy in the curriculums of Colleges of Education in order to integrate multimodal framework.

2.1 RESEARCH METHODOLOGY

This section contains a description of the methodology used for the study. This includes the research design, the participants, research data and data sources, data collection instruments, validity and reliability strategies of instruments, and data analysis procedures. The essence of the study is to investigate whether exposure to multimodal instructional strategies could enhance learners' English speaking skills and English-speaking self-efficacy. The theoretical orientations of the study are two: Halliday's (2004) Social Semiotic Theory and Bandura's (1997) Self-efficacy Theory. These are essential for the study since the Social Semiotic theory argues in favour of simultaneous integration of multiple modes of communication other than only the mere linguistic mode (Halliday, 2004). Besides, the Self-efficacy theory underscores leaners' intrinsic, self-conceived notion of themselves in terms of abilities to perform tasks (Bandura, 1997).

2.2 Research Design

The research design employed for the present study is quasi-experimental research involving pre-test/post-test experimental-control-group design. Again, a one-week-later follow-up quantitative measure was used to assess the effects of the multimodal treatment on the participants' self-efficacy in English speaking skills.

2.3 Research Population and Sample

This research took place in the North-West Zone of Nigeria comprising 7 states of the Federation. This Zone has 9 accredited state-owned Colleges of Education. The entire population of the study comprised all the 488 registered National Certificate in Education One (NCE 1) English students of the Departments of English in the nine (9) state-owned Colleges of Education in North West Zone, Nigeria. In the 2025/2026 academic session, these Departments registered 488 NCE 1 male and female English students: 279 and 209 males and females respectively. The choice of NCE 1 was informed by the 2020 revised edition of the NCCE English curriculum for Colleges of Education in Nigeria whose English-speaking related courses are exclusively found in NCE 1 (NCCE, 2020).

Purposive sampling was used to select two states namely: Jigawa State and Kano State. This could be justified because the other states were more insecurity-laden: a situation that frequently disrupted academic calendars. Further, each of the selected two states possesses two State Colleges of Education namely: Jigawa State College of Education, Gumel (JSCOE, Gumel); Jigawa State College of Education and Legal Studies, Ringim (JSCOELS, Ringim); Sa'adatu Rimi College of Education, Kano (SRCOE, Kano) and Kano State College of Education and preliminary Studies (KASCEPS). The four selected Colleges of Education registered a total of 204 NCE 1 English students which constituted the sample size of the study. Again, adopting the intact-class approach, the simple random sampling procedure was used to assign the groups to experimental and control groups: JSCOE, Gumel, KASCEPS, JSCOELS, Ringim and SRCOE, Kano, respectively.

2.4 Data Collection Instruments

Two instruments were used for data collection in the present study namely: Multimodal English Speaking Test (MEST) and English Speaking Self- Efficacy Scale (ESSES). MEST was designed to elicit oral productive English performance of the learners in the form of paired interpersonal communication. ESSES was adapted from Maghsoudi *et al* (2022). It was a 21-item Likert type of scale modified to elicit responses from the experimental subjects one week after being exposed to multimodal instruction. The 21 items were divided into three components of English speaking self-efficacy, 7 for each, namely: Linguistic Dimension of Self-Efficacy, Interpersonal Dimension of Self-efficacy and Performance Dimension of Self-efficacy.

2.4.1 Validity and Reliability of Research Instruments

Numerous techniques were used to ensure the two instruments were valid and reliable for the present study. For both MEST and ESSES, content validity and construct validity were ensured via expert validation by specialists in English phonetics and phonology, and Educational Psychology respectively. The validation processes established that both instruments could measure key aspects of English speaking skills such as pronunciation, vocabulary, grammar and fluency, and the essential constructs of self-efficacy in relation to the specifics of the English speaking skills.

The reliability of MEST was established following a pilot test involving 10 NCE 1 students outside the actual sample size of the study. Using test re-test reliability technique, the reliability index of MEST was 0.91. The other instrument was not subjected to reliability check because it was designed as a follow-up quantitative measure.

2.5 Data and Data Sources

In addition to the researcher himself from JSCOE Gumel, three more research assistants were employed: one from each of the three other Colleges of Education which constituted the sample of the present study. Thus, the 54, 60, 47 and 43 subjects in intact classes of JSCOE Gumel, SRCOE Kano, JSCOELS Ringim and KASCEPS, respectively, were paired in twos. This pairing resulted in 27, 30, 23 and 21 pairs, respectively. For ease of management of the participants' English speaking performances and to ensure anonymity, each pair was given a code. The coding ranged from 1a & 1b to 27a & 27b in JSCOE Gumel; 1a & 1b to 30a & 30b in SRCOE Kano; 1a & 1b to 23a & 23b in JSCOELS Ringim; and finally 1a & 1b to 21a & 21b in KASCEPS Kano.

Furthermore, having paired and coded the subjects in each group, and in order to collect data in the form of English speaking performances of the subjects, video recordings were used. Interpersonal English conversations of each pair were recorded both before and after the treatment. The recorded conversations of all the pairs in the form of pre-test and post-test

constituted the main data for the study. Specific components of the English speaking skills examined, transcribed, analysed and scored were: 12 pure vowels, word-level stress, context-relevant vocabulary, verb-tenses, and speed of speech and speech coherence After the post-test, the two experimental classes in JSCOE, Gumel and KASCEPS, Kano were provided with ESSES - the second instrument – which they had filled. The essence was to find out if the intervention was effective on the subjects' English speaking self-efficacy.

3.1 Theoretical Framework

Variables of the study were divided into dependent and independent ones. The essence of the study is to investigate whether exposure to multimodal instructional strategies could enhance learners' English speaking skills and English-speaking self-efficacy. The theoretical orientations of the study are two: Halliday's (2004) Social Semiotic Theory and Bandura's (1997) Self-efficacy Theory. These are essential for the study since the Social Semiotic theory argues in favour of simultaneous integration of multiple modes of communication other than only the mere linguistic mode (Halliday, 2004). From the Social Semiotic perspective, meaning-making is realizable in various modes of communication. All modes (in Cope and Kalantzis, 2022: audio, visual, gestural, tactile, spatial, spoken and written; in Bezemer and Kress, 2016 also layout) potentially can convey full meaning. Thus, in a multimodal perspective, the meaning potential of social semiotics is applicable in all modes. For this, according to Corballis (2017), the multimodal approach is traced to the theoretical underpinnings of social semiotics. Aligning to practice, by implication, in the multimodal sense, language learning is a collaborative; an active process of interaction between learners and multimodal tools

Besides, the Self-efficacy theory underscores leaners' intrinsic, self-conceived notion of themselves in terms of abilities to perform tasks (Bandura, 1997). According to Lopez-Garrido (2023), one of the most effective ways to learn a new skill or to improve one's performance in a given activity is by practicing. In most cases, part of the reason practice becomes effective in the learning process is that it gives people a sense to understand that they are capable of acquiring new skills (Lopez-Garrido, 2023). This is a positive way of thinking about oneself because believing that one is capable of achieving tasks one sets for oneself can offer numerous opportunities for success at new learning tasks. In a multimodal English speaking classroom, opportunities are provided for learners to engage and interact with multiple learning modes to practice and use the English language such as through exposure to visual modes, aural modes, kinaesthetic modes, charts, textual modes and technological gadgets.

4.1 Statistical Analytical Tools

This section describes the appropriate statistical tools used in analyzing the research data from raw scores to research inferences. The details of the statistical methods are given as follows: Quantitative data from the pre-test and post-test were analysed using One-Way ANOVA in order to determine whether there was a significant effect of the multimodal treatment on the English speaking performance of the subjects. On the other hand, descriptive statistics was used to analyse the responses of the experimental subjects to the 21-item Likert type of questionnaire in order to discover how effective multimodal instruction had been on learners' self-efficacy in English speaking skills.

5.1 RESULTS AND DISCUSSION

This section presents the quantitative results of the pre-test and post-test scores obtained from the recorded videos of learners' paired conversations. The data were analysed to examine students' improvement in English speaking skills after the treatment.

5.1 One-Way ANOVA Statistics of Pre-test and Post-test Scores of Students' English speaking skills Before and After the Intervention

Table 1: Summary of One-Way ANOVA of Pre-Test and Post-Test Mean English Segmental Pronunciation Skills Performance Scores of Students Taught Using the Multimodal Instruction and Those Taught Using the Conventional Approach

Source of Variance		Sum of Squares	Df	Mean Square	F	Sig.
Pre-test	Between Groups	1.396	1	1.396	2.543	.114
	Within Groups	54.366	99	.549		
	Total	**55.762**	**100**			
Post-test	Between Groups	5145.566	1	5145.566	1404.914	.000
	Within Groups	362.592	99	3.663		
	Total	**5508.158**	**100**			

(Source: Research Data, 2026)

Table 1 revealed the pre-test scores as $F(1,99) = 2.543, p = 0.114$. This showed that there is no significant difference in the pre-test mean ESPS performance scores of students taught using the multimodal instruction and those taught using conventional approach. The mean post-test performance scores showed $F(1,99) = 1404.914, p = 0.001$. This indicated that there is a significant difference in the post-test mean ESPS performance scores of students taught using the multimodal instruction and those taught using the conventional approach in favour of the former.

Table 2: Summary of One-Way ANOVA of Pre-Test and Post-Test Mean English Supra-Segmental Pronunciation Skills Performance Scores of Students Taught Using the Multimodal Instruction and Those Taught Using the Conventional Approach

Source of Variance		Sum of Squares	Df	Mean Square	F	Sig.
Pre-test	Between Groups	2.290	1	2.290	9.442	.003
	Within Groups	24.007	99	.242		
	Total	**26.297**	**100**			
Post-test	Between Groups	1856.210	1	1856.210	885.219	.000

	Within Groups	207.592	99	2.097
	Total	**2063.802**	**100**	

(Source: Research Data, 2026)

Table 2 presented the mean pre-test scores as $F(1,99) = 9.442, p = 0.003$. This indicated that there is a significant difference in the pre-test mean ESSPS performance scores of students taught using the multimodal instruction and those taught using the conventional approach. However, the mean post-test scores showed that $F(1,99) = 885.219, p = 0.001$. This revealed that there is a significant difference in the post-test mean ESSPS performance scores of students taught using the multimodal instruction and those taught using the conventional approach in favour of the multimodal group.

Table 3 Summary of One-Way ANOVA of Pre-Test and Post-Test Mean English Vocabulary Pronunciation Skills Performance Scores of Students Taught Using the Multimodal Instruction and Those Taught Using the Conventional Approach

Source of variance		Sum of Squares	Df	Mean Square	F	Sig.
	Between Groups	6.799	1	6.799	9.915	.002
Pre-test	Within Groups	68.574	100	.686		
	Total	**75.373**	**101**			
	Between Groups	609.461	1	609.461	873.223	.000
Post-test	Within Groups	69.794	100	.698		
	Total	**679.255**	**101**			

(Source: Research Data, 2026)

Table 3 presented the mean pre-test scores as$F(1,100) = 9.915, p = 0.002$. This implied there is a significant difference in the pre-test mean English vocabulary pronunciation skills achievement scores of students taught using the multimodal instruction and those taught using the conventional approach. The mean post-test scores showed that $F(1,100) = 873.223, p = 0.001$. This showed that there is a significant difference in the post-test mean English vocabulary pronunciation skills performance scores of students taught using the multimodal instruction and those taught using the conventional approach in favour of the experimental learners.

Table 4: Summary of One-Way ANOVA of Pre-Test and Post-Test Mean English Grammar Skills Performance Scores of Students Taught Using the Multimodal Instruction and Those Taught Using the Conventional Approach

Source of variance		Sum of Squares	Df	Mean Square	F	Sig.
	Between Groups	.000	1	.000	.000	.998
Pre-test	Within Groups	50.158	99	.507		
	Total	**50.158**	**100**			
	Between Groups	7031.727	1	7031.727	1370.581	.000
Post-test	Within Groups	507.917	99	5.130		
	Total	**7539.644**	**100**			

(Source: Research Data, 2026)

Table 4 presented the mean pre-test scores as follows:$F(1,99) = 0.000, p = 0.998$. This showed that there is no significant difference in the pre-test mean English grammar pronunciation skills performance scores of students taught using the multimodal instruction and those taught using the conventional approach. However, the mean post-test scores showed that $F(1,99) = 1370.581, p = 0.001$. This confirmed that there is a significant difference in the post-test mean English grammar pronunciation skills achievement scores of students taught using the multimodal instruction and those taught using conventional approach in favour of the multimodal instruction subjects.

Table 5: Summary of One-Way ANOVA of Pre-Test and Post-Test Mean English Fluency Skills Performance Scores of Students Taught Using the Multimodal Instruction and Those Taught Using the Conventional Approach

Source of variance		Sum of Squares	Df	Mean Square	F	Sig.
	Between Groups	3.104	1	3.104	7.413	.008
Pre-test	Within Groups	41.451	99	.419		
	Total	**44.554**	**100**			
	Between Groups	1715.183	1	1715.183	791.770	.000
Post-test	Within Groups	214.460	99	2.166		
	Total	**1929.644**	**100**			

(Source: Research Data, 2026)

Table 5 presented the mean pre-test scores as follows: $F(1,99) = 7.413, p = 0.008$. This implied that there is a significant difference in the pre-test mean EFS performance scores of students taught using the multimodal instruction and those taught using the conventional approach. The mean post-test scores showed that $F(1,99) = 791.770, p = 0.001$; this confirmed that there is a significant difference in the post-test mean English fluency pronunciation skills performance scores of students taught using the multimodal instruction and those taught using the conventional approach.

5.2 Descriptive Statistics Involving Frequencies and Percentages of Students' Responses on the Effects of Multimodal Instruction on Students' Self-efficacy in English speaking Skills after the Treatment

Table 6: Frequencies and Percentages of Experimental Learners' Responses on the Effects of Multimodal Instruction on Self-Efficacy in English Speaking Skills After the Treatment

Scales	Respondents	Frequencies	Percentages
LSE			
Always true of me	97	62	9.1
Usually true of me	97	305	44.9
Sometimes true of me	97	283	41.7
Rarely true of me	97	24	3.5
Never true of me	97	5	0.7
IPSE			
Always true of me	97	63	9.3
Usually true of me	97	298	43.9
Sometimes true of me	97	276	40.6
Rarely true of me	97	42	6.2
Never true of me	97	0	0
PSE			
Always true of me	97	65	9.6
Usually true of me	97	289	42.6
Sometimes true of me	97	298	43.9
Rarely true of me	97	27	4
Never true of me	97	0	0

(Source: Research Data, 2026)

From Table 6 above, the analyses on LDSE showed that most of the learners (53.1%) believed that multimodal instruction was highly effective on the linguistic dimension of English speaking self-efficacy. Others (41.7%) were neutral (sometimes true). However, only few of them (4.2%) showed the approach was rarely effective.

Further, in terms of IPDSE, the results showed that the largest percentage of the responses (53.9) represented higher confidence of the learners in IPDSE after exposure to the multimodal English speaking treatment. Contrastively, 40.6% showed neutrality while only 6.2% showed little confidence in the approach. From the above, it can be concluded that multimodal instruction was effective on students' IPSE dimension of English speaking skills.

Finally, it can be concluded that most of the responses (52.2%) showed that participants had a higher level of confidence in their abilities to accomplish PDSE of the English speaking skills after exposure to multimodal instruction. However, a good number of the responses (43.9%) revealed the neutrality of the participants in whether multimodal instruction could increase PSE dimension of the English speaking skills. The least of all fell on very low level of confidence of the participants in PSE dimension of the English speaking skills after exposure to multimodal instruction. This was represented by 4% of the total responses.

5.3 Discussions

The finding relating to the performance of students in terms of the ESPS component of the English speaking skills showed that the experimental group performed significantly higher th an those not exposed to the treatment. This finding supports the discovery of Belabbas (2024) who also found out that multimodal instruction was effective in enhancing learners' English pronunciation skills.

Again, in terms of ESSPS component of the English speaking skills, the experimental group's performances were observed to be more significantly improved at the post-test. However, such significant improvements were missing in the performance of the control group at the post-test. This finding proved that learning the English word stress and accentual intonation tunes through multimodal instruction gave the students a greater chance of internalizing the patterns of word stress and intonation of the language. Hence, the treatment allowed students to explore, rehearse, practice and produce the language in interpersonal communication. This finding echoes the conclusion of Kummin, Surat, Mydin, Othman and Muslim (2020) that multimodal texts enhanced the teaching of English language oral skills.

The finding relating to the effectiveness of multimodal instruction in enhancing learners' English vocabulary skills is in agreement with the finding of Anonymous (2016) who also investigated the impact of multimodal instruction on the acquisition of vocabulary. However, the difference between the two studies could be traced to the emphasis of the present study on teaching the vocabulary skills in context. This practice supported the effective utilization of the multimodal tools in the classroom with which learners progressed from exploration of English speaking models, rehearsals, and practices to interpersonal oral English vocabulary production.

The finding of the present study on English grammar skills is consistent with the findings of Aggrainy (2016) and Ngagu-Dominic (2021) who both accounted for the effectiveness of multimodal approaches in enhancing language skills. However, while Ngagu-Dominic's study was on German Grammar and, while Aggrainy's work centered on reading and writing skills of students, the present study is particularly different. Using multimodal instruction, the present study emphasized the relevance of teaching grammar skills in the contexts of the other English speaking skills. This is essential in that language is primarily an oral exercise that involves the simultaneous conglomeration of all the five major components which constituted the dependent variables of the present investigation.

In terms of fluency, after the treatment, the experimental learners demonstrated a greater deal of efforts at speed of speech and speech coherence than the control group. This followed the multimodal treatment to which the experimental students were exposed. During the treatment phase, activities involving a series of exploration of multimodal resources, rehearsals between partners in every pair, practicing and automatic production of the fluency-related skills were carried out. Because the control group did not receive similar treatment, their mean performance score in English fluency skills was significantly lower than that of the experimental group at the post-test. Earlier, Guo and Li (2023) also echoed that it was crucial for teachers to engage students using multimodal instructional procedures.

From the descriptive analyses on the self-efficacy component of the investigation, LDSE scored the highest. This was followed by PDSE and, finally, the IDPSE. Thus, the new ground broken by the finding of the present study is that multimodal instruction increases learners' performance across the three levels of the English speaking self-efficacy after the treatment. This finding is evident also in the works of Darrington and Dousay (2015) and Maghsoudi, Golshan and Naeimi (2022). The difference lies in that the former found out that multimodal writing procedures possessed motivational tendencies. In the same vein, the latter discovered that the digital multimodal composition increased learners' writing self-efficacy. Unlike the above, thus, the present study found out the relevance of multimodal instruction in enhancing learners' English speaking self-efficacy.

6.1 CONCLUSION

This aimed to investigate the effects of multimodal instruction on students' performance in English speaking skills and self-efficacy in Colleges of Education in North West Zone, Nigeria. The findings demonstrated that multimodal instruction improved learners' English speaking performance and English speaking self-efficacy. One-Way ANOVA statistics revealed a significant increase in experimental students' post-test performance in ESPS, ESSPS, EVS, EGS and EFS after the treatment. Still, descriptive statistics had shown improvements in leaners' self-efficacy in English speaking skills across linguistic, interpersonal and performance dimensions of self-efficacy. These improvements followed the series of multimodal trainings to which the learners were exposed involving linguistic and non-linguistic modes of communication such as textual, pictorial, visual, audio-visual, kinaesthetic and technology integration.

In addition, descriptive statistical findings from students' responses to the Likert type of questionnaire revealed positive increases in leaners self-efficacy in English speaking skills after the treatment. Students reported that the engagement, participation and negotiation during multimodal activities had helped them focus on context-relevant pronunciation, vocabulary, grammar and speed variables during interpersonal communication.

Despite these positive outcomes, this study has some limitations. First, the study was limited to the North West Zone of Nigeria. NCE 1 from the other regions of Nigeria might demonstrate different dispositions and abilities. Thus, the findings might not be generalizable to all Colleges of Education in Nigeria. Second, transcription and scoring of video-recorded pairwise conversations of participants were done by inter-rater technique: the researcher himself along with research assistants. Technology-driven speech analysers could have been faster and more automatic. Finally, the purposive selection of Kano State and Jigawa State as a representative sample of the seven states of the North West Zone of Nigeria is another limitation of the present investigation.

Based on these limitations, several suggestions are proposed for future research and practice. Subsequent empirical studies are recommended to incorporate all regions of Nigeria in the application of multimodal instruction in teaching and learning of English speaking skills and self-efficacy in Colleges of Education. Again, in addition to the use of inter-rater technique only, subsequent research should integrate technology-based speech analysers in the transcription and scoring of participants' paired oral communication.

REFERENCES

Aggrainy, S. (2016). Multimodal Media and Its Application to Improve Reading and Writing of Students. *English Education: Journal Tadris Bahasa Inggris*, 9 (2), 2016, 444-460.

Anonymous, (2016). The Impact of Multimodal Instruction on the Acquisition of Vocabulary. *ScieLO-Brazil*. https://doi.org/10.1590/010318134842170942. Artigos. Trab. Linguist. Apl. 55(1).Jan-Apr 2016.

Balebbas, B. (2024). Investigating the Effectiveness of the Multimodal Approach in Developing Speaking Skills among EFL Students. The Case of Second-Year Students of English Department of Mohamed Kheider Biskra University, People's Democratic Republic of Algeria. Unpublished M A Dissertation.

Bandura, A. (1997). Self-Efficacy: The Exercise of Control. *Journal of Cognitive Psychotherapy,* Volume 13, Issue 2; DOI: 10.1891/0889-8391.13.2.158.

Bezemer, J., and Kress, G., (2016). *Multimodality, learning and communication: A social semiotic frame*. Oxon Routledge.

Cope, B., and Kalantis, M., (2022). "*Towards educational justice: A pedagogy of multi literacies, Revised."* Cambridge University Press.

Corballis, M. (2017). *The Truth about Language: What It Is and Where It Came From.* Chicago: The University of Chicago Press.

Darrington, B. and Dousay, T. (2015). Using Multimodal Writing to Motivate Struggling Students to Write. *TechTrends, Vol 59 (6)*

Guo, M., and Li, J. (2023). An Empirical Study on the Application of Multimodal Teaching Model on English Writing Teaching in China Senior High School. Asian Social Science Science. Vol. 19 (5) 2023 https://doi:10.5539/ass.v19n5pl

Halliday, M. A. and Matthlessa, (2004). *An introduction to functional grammar,* (3rd ed.). London: Hodder Arnold.

Kummin, S., Surat, S., Mydin, F., Othman, Z. and Muslim, N. (2020). The Use of Multimodal Texts in Teaching English Oral Skills. Universal Journal of Education Research. Vol. 8 (12): 7015-7021, 2020.

Laadem, M. and Makena, H., (2022). Multimodal pedagogies in teaching English for specific purposes in higher education; Perceptions, challenges and strategies. *International journal on studies in education,* 1, (1), 488-496, doi10.46328/ijonse.

Lopez- Garrido, G. (2023). Self-Determination Theory: How It Explains Motivation. *Simply Psychology*, 2023 www.simplycomplicated.me/self-determination.

Maghsoudi, N., Golshan, M. and Naemi, A. (2022). The Impact of Digital Multimodal Composition on Writing Ability of Iranian EFL Learners with Respect to Their Self-Efficacy. *Journal of Modern Research in English Language Studies, 9 (1),* https//:doi.10.30479/jmrels.2020.

National Commission for Colleges of Education. (2020). *Nigeria certificate in education minimum standards for languages*. Abuja, Nigeria: National Commission for Colleges of Education.

Nigeria, F.R (2020) *National Language Policy*. Lagos: Nigerian Educational Research and Development Council.

Ugagu-Dominic, N. (2021). Multimodal Approaches in Teaching and Learning of German Grammar: The Implications. *Journal of African Studies, Vol. 9 (1)* Nov. 2021.