

The Impact of Code Switching on Multicultural Iranian English as a Foreign Language Learners' Oral Production

Keivan Seyyedi, Fereydoon Rasouli

Department of Translation, Cihan University-Erbil,
Kurdistan Region, Iraq

Abstract—This quasi-experimental research aims to fill a gap in the literature on classrooms with multilingual populations. It investigates the effect of systematic code-switching (CS) on Iranian English as a Foreign Language (EFL) learners' oral production. Using the Preliminary English Test (PET), the researcher selected 60 subjects, who were then assigned evenly to the experimental and control groups. Strategy sessions for the experimental group included six sessions on the CS of metalinguistic awareness and on the CS of task sensitivity. This research was assessed using a test designed by the researcher to measure the subjects' oral proficiency. The data are analyzed using an independent t-test. The experimental group showed a significant improvement in oral production, whereas the control group showed no significant improvement. The current study examines the requirement for CS in teacher education programs. In Iran's multicultural classrooms, it is imperative to achieve a balance between language equality and instruction effectiveness. Thus, the current study proposes revisiting curriculum-level changes to enable CS that connects its multiculturalism, without focusing on Persian-language dominance, by stressing language diversity. Therefore, the current study enlightens multilingual EFL contexts concerning the controversies over the pedagogical efficacy of CS.

Keywords—Code-switching, Cognitive scaffolding, Iranian EFL learners, Multicultural classrooms, Oral production proficiency.

I. INTRODUCTION

Iran is a multilingual country with diverse ethnic groups, including mainly Persians, Azeris, Kurds, Arabs, and Balochs. This multilingual phenomenon, in turn, has a unique relationship to the country's language education, in which its ethnicities and sub-ethnicities form the basis of this multilingualism. Code-switching (CS), a pattern that shifts between the first language and the target language, has emerged as a pedagogical practice and a sociolinguistic phenomenon in the context of English as a Foreign Language (EFL) learning, where some see it as a means of enhancing understanding and connecting the two languages, while others see it as negatively impacting SLA because of overdependence on the first language (Noorbar and Mamaghani, 2016; Massoudi, 2025; Rezvani and Eslami Rasekh, 2011).

Fluency, accuracy, and complexity of oral production are also important aspects of the process of forming communicative ability within the framework of Second Language Acquisition (Seyyedi, 2025). In the multicultural educational context of

Iranian culture, in which learners are accustomed to different language structures, the potential function of CS may serve as a psychological factor that enables learners to deliver their ideas more effectively (Namaziandost et al., 2019). However, over richness and the lack of regulatory functions, in the form of excessive CS, may endanger the autonomy of second language (L2) acquisition (Lyser and Saito, 2010; Seyyedi and Amin, 2020). Although there has been extensive research on the application of CS in monolingual EFL contexts, more research is needed on its application in the Iranian educational context, as well as on learners' identity in those settings.

Moreover, this study aims to explore the effects of CS on oral production measurements among Iranian EFL learners of different ethnic groups. In consideration of both sociolinguistic and cognitive perspectives, this integration of ideas will seek to reconcile the disparities within the current body of research and offer recommendations for the optimal use of CS in a multilingual classroom.

Based on the purpose of the study, the following research question is proposed: How does the use of CS

affect the oral production performance of Iranian EFL learners in a multilingual classroom? Accordingly, the following hypothesis is formulated: The use of CS does not significantly affect the oral production performance of Iranian EFL learners in multilingual classroom settings.

II. RESEARCH BACKGROUND

The involvement of CS in facilitating and inhibiting speaking ability in EFL education has been disputed by researchers. CS, according to Noorbar and Mamaghani (2016), reduces the cognitive load of linguistic production in the target language and therefore provides cognitive bridging for non-advanced learners. In this research project with 60 participants in an EFL class consisting of Iranian learners of English, they supported the notion that learners would find improved syntactical precision and fewer traces of oral hesitation, comparing organized CS activities to translation activities in which learners used Persian to structure intricate sentences for translation to English. Conversely, excessive use of CS negatively affects vocabulary development among advanced learners of the target language. In this longitudinal research project conducted on learners in CS-reliant educational environments, vocabulary development was slower among CS-reliant learners than among immersion learners because learners who rely on L1 have limited autonomy in the target language.

Various aspects of fluency, accuracy, and complexity are strongly mediated by proficiency level when examining CS results. Mugableh and Alruwaili, in a 2023 study, concluded that unlimited CS negatively influenced the vocabulary of upper-intermediate Saudi Arabian students for 6 months, thereby affecting L2 interactions. The study by Namaziandost et al., published in 2019, demonstrated that unlimited CS improved affective involvement among marginalized ethnic groups, including Kurds and Balochis, who were actively participating in oral interactions, facilitated by the teacher's awareness of the significance of FL and L1 in these groups. The study by Rezvani and Eslami Rasekh, published in 2011, showed that the subtype involving abstract reasoning could have benefited from CS, and L1 helped the connectivity between the narratives. In contrast, the subtype involving abstract reasoning was hindered by discourse unity arising from excessive use of CS.

CS in Iranian multilingual EFL settings generates a paradox between the advantages and fairness of language use. Rezvani and Eslami Rasekh's (2011) Iranian study illustrated how teachers use CS with Persian and ignore minority dialects of Persian spoken by Azeri and Gilaki Iranians. They used their observations of CS in class, focusing on when CS could help learners benefit from what was being taught. For example, "teaching Persian idioms in learners' languages could enhance verbal fluency by filling gaps between L1 and L2 knowledge." The use of CS is consistent with Vygotsky (1980) on how learners should adopt a sociocultural approach when acquiring a new language, using their native language as a psychological tool.

III. METHODOLOGY

A. Participants

The research employed a quasi-experimental design with a population of 100 subjects. The subjects enrolled in the research were intermediate English Foreign Language learners. There were 50 males and 50 females aged 13–18. They underwent rigorous selection from the "Zaban Sara Language Institute" in Mahabad, Iran. The subjects had similar characteristics, which requires a filter to pick suitable subjects. They had to take "The Cambridge English: Preliminary test." The test was prepared by Cambridge Assessment English in 2020. They selected 60 subjects whose results were closest to the average and within 1 standard deviation of it. They made sure to pick equal subjects from each group. They obtained permission to conduct research.

B. Instruments

The sample was standardized by using the PET test. The oral test, an instrument designed by the researcher, assessed fluency, lexical complexity, and grammatical accuracy. The test consisted of 20 items, divided into four categories: job-related items (Items 1–5), internet-related items (Items 6–10), health-related items (Items 11–15), and friends-related items (Items 16–20). To establish the content validity of this test, each subsection was repeatedly analyzed by experts in applied linguistics. The internal consistency of the test parameters was determined using the Cronbach alpha formula through SPSS v.27. As illustrated in Table I, each domain met the good consistency threshold of $\alpha > 0.70$.

C. Procedures

Both groups underwent an initial oral test. Both sets of performances were marked using an analytical marking rubric aligned with the IELTS speaking band descriptors. Interrater reliability for the first set of performances was calculated using intraclass correlation coefficient (ICC). The results showed a high degree of agreement (ICC = 0.92, $P < 0.001$). Independent samples t-tests on the results did not provide evidence of statistical significance in pre-test score differences between groups ($t[58] = 0.47$, $P = 0.64$), thereby supporting equivalence. The experimental group received 30 days of instruction comprising six 60-min sessions focusing on strategies for CS. The instruction targeted metalinguistic awareness, strategic lexical replacement, and syntactic interlanguage blending in contextually supported conditions, including negotiation of meaning and handling of discourse gaps. The sessions included controlled practice activities, such as acting out dialogues with a partner, and semi-controlled activities, such

TABLE I
CRONBACH'S ALPHA COEFFICIENTS OF RESEARCH VARIABLES

Variable	Components	Number of questions	Cronbach's alpha coefficients
Code switching	Job	1–5	0.78
	Internet	6–10	0.77
	Health	11–15	0.88
	Friends	16–20	0.86

as peer-debating sociocultural issues. It also included whole-class explicit instruction on high-frequency collocations. Those manifesting anxiety or having problems with lexical retrieval were provided with scaffolded feedback through think-aloud protocols. The control group received the standard curriculum provided by the educational institution. This was monolingual instruction.

After this intervention, an oral post-test was given similarly for both groups. To enhance inter-rater reliability, discussions emphasizing inter-rater agreement were held for the post-test assessment. However, differences between groups were tested using independent t-tests, and Cohen's *d* was calculated to assess the magnitude of the observed differences. Normality assumptions in t-tests were tested using the Shapiro-Wilk test, but Levene's test tested the equality of group variances before any analysis.

IV. RESULTS

A. Descriptive Statistics

Table II shows the descriptive statistics for the control and experimental groups on their pre-test and post-test scores. Both the control and experimental groups performed equally well on the pre-test, with mean scores of 71.40 (standard deviation [SD] = 6.85) and 71.60 (SD = 7.77), respectively. Meanwhile, a significant improvement was observed in the mean scores of the experimental groups ($M = 90.23$, $SD = 4.6$) on oral production, while a slight improvement was observed in the control groups ($M = 73.40$, $SD = 8.27$).

B. Normality of Research Variables

Table III presents the results of the Kolmogorov-Smirnov tests for normality of all variables. According to the findings, because all *p*-values for the pre-control, pre-experimental,

post-control, and post-experimental variables exceed the $\alpha = 0.05$ level of significance (*P*-values: 0.139, 0.200, 0.227, and 0.200, respectively), this study rejects the hypothesis that the data are normally distributed.

C. Research Hypothesis

According to the independent t-tests as shown in Table IV, there is a statistically significant difference in post-test scores between the experimental and control groups. The value of *t* and the level of significance are $t(29) = -10.51$ and $P = 0.00$. However, for the same group, the post-pre difference in *t* was insignificant. *T* (29) is -0.15 , with the *P*-value being 0.881.

Therefore, null hypothesis (H_0): The use of CS does not significantly affect the oral production performance of Iranian EFL learners in multilingual classroom settings was rejected ($P < 0.001$).

V. DISCUSSION

The application of the systematic CS in this research served the same purpose as the cognitive scaffolding theory presented by Noorbar and Mamaghani (2016). This implies that through systematic CS, the experimental group made it easier for learners to understand complex L2 grammar and instructions, thereby reducing their cognitive load and enabling them to concentrate on language production. Therefore, from an empirical standpoint, the results prove the validity of the cognitive scaffolding theory in explaining how systematic CS serves as an important structure for helping learners develop their speaking skills. The experimental group experienced positive gains in fluency, complexity, and grammar, which appears to support the sociocultural models proposed by Vygotsky (1980) concerning L1 support for L2 acquisition. The application of multilingual capacities is beneficial for learners engaged in the complicated act of speaking, allowing time to judge errors, thus supporting Cook's (2001) proposal on multilingual acquisition.

Socially, the CS component widened students' multicultural identities, especially among the Kurds and Baloch, thereby enhancing their affiliation and engagement, as Namaziandost et al. (2019) verified. Nevertheless, the issue arising from the Persian-centric component in the educational context of the CS, according to Rezvani and Eslami Rasekh in 2011, is linguistic equality. The lecturers applied the component to explain idioms in Persian for heritage languages, but the educational context diminished the minority dialects.

It might be said that, in view of the homogeneous data set of this investigation, which included participants from a language institute, this could affect the results' applicability to Iranian society at large. Furthermore, this investigation revealed that perhaps a 30-day period may not be sufficient to observe a pronounced effect of CS on autonomous L2 proficiency. Future investigations might help clarify concepts of lexical extension and syntactic development, enabling a more detailed analysis of the effects of CS on language acquisition.

TABLE II
DESCRIPTIVE STATISTICS

Variable	N	Minimum	Maximum	Mean	Std. deviation
Pre-control	30	61.00	82.00	71.400	6.84609
Pre-experimental	30	55.00	83.00	71.600	7.76642
Post-control	30	60.00	88.00	73.400	8.27376
Post-experimental	30	80.00	97.00	90.233	4.59898
Valid N (listwise)	30				

TABLE III
ONE-SAMPLE KOLMOGOROV-SMIRNOV TEST

Normality results	Pre-con	Pre-exp	Post-con	Post-exp
N	30	30	30	30
Normal parameters ^{a,b}				
Mean	71.4000	71.6000	73.4000	90.2333
Std. deviation	6.84609	7.76642	8.27376	4.59898
Most extreme differences				
Absolute	0.140	0.105	0.190	0.114
Positive	0.140	0.099	0.116	0.072
Negative	-0.116	-0.105	-0.190	-0.114
Test Statistic	0.140	0.105	0.110	0.114
Asymp. Sig. (2-tailed)	0.139 ^c	0.200 ^{c,d}	0.227 ^c	0.200 ^{c,d}

a: Test distribution is Normal; b: Calculated from data; c: Lilliefors significance Correction; d: This is a lower bound of the true significance

TABLE IV
INDEPENDENT T-TEST

Variable	Paired differences					t	Df	Sig. (2-tailed)
	Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
				Lower	Upper			
Pre-con and pre-exp	-0.200	7.24640	1.32301	-2.90585	2.50585	-0.151	29	0.881
Post-con and post-exp	-16.83	8.76743	1.60071	-20.1071	-13.559	-10.51	29	0.000

VI. CONCLUSION

The findings of this study suggest several specific conclusions based on the empirical results. To begin with, participants exposed to systematic CS instruction showed a statistically significant and notable increase in their oral production scores from 71.60 to 90.23. In contrast, the subjects in the control group that did not receive such instruction showed insignificant progress from 71.40 to 73.40. Therefore, the null hypothesis of no significant effect of CS on oral production is rejected based on the independent t-test results ($t[29] = -10.51, P < 0.001$). As shown by t-tests, pre-tests of the two groups were statistically similar ($t(58) = 0.47, P = 0.64$); however, the data obtained in connection with all four variables – the pre-control, pre-experimental, post-control, and post-experimental – failed to meet normal distribution (Kolmogorov–Smirnov test; $P > 0.05$).

Furthermore, the significance of the 30-day six-session CS approach's impact on participants' performance is highlighted by a mean difference between the experimental and control groups of 16.83, accompanied by a narrow confidence interval (between -20.11 and -13.56). In addition, the participants who received the CS treatment showed lower cognitive processing load for their L2 performance, along with a higher level of affective involvement in L2, L1, and culturally sensitive tasks; in other words, the lack of change in performance observed in the control group suggests that the monolingual treatment did not help students improve their oral production during this period. Moreover, the reliability of the results obtained from the oral test was confirmed by satisfactory Cronbach's alpha coefficients, ranging from 0.77 to 0.88 for each of the four assessed areas (job, internet, health, friends). Finally, it should be noted that the described findings apply only to the context of intermediate Iranian

EFL learners (aged 13–18) enrolled at a single language institute in Mahabad, Iran; thus, they cannot be generalized to conclusions about autonomous L2 proficiency.

REFERENCES

- Cook, V. (2001). Using the first language in the classroom. *Canadian Modern Language Review*, 57(3), 402-423.
- Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition*, 32(2), 265-302.
- Massoudi, A.H. (2025). Predicting the power of entrepreneurial orientation in improving the level of product innovation. Business education as a mediator. *Journal of Management and Business Education*, 8(1), 58-73.
- Mugableh, A.I., & Alruwaili, T.F. (2023). A longitudinal investigation of the impact of code-switching versus target language only on Saudi EFL students' written receptive vocabulary knowledge development in terms of breadth. *Forum for Linguistic Studies*, 5(3), 1719.
- Namaziandost, E., Neisi, L., & Banari, R. (2019). The impact of code-switching on vocabulary learning among Iranian upper-intermediate EFL learners. *International Journal of Linguistics, Literature and Translation*, 2(5), 309-318.
- Noorbar, S., & Mamaghani, H.J. (2016). The effect of code-switching on Iranian elementary EFL learners' oral fluency, accuracy, and willingness to communicate. *Journal of Modern Research in English Language Studies*, 3(4), 75-103.
- Rezvani, E., & Eslami Rasekh, A. (2011). Code-switching in Iranian elementary EFL classrooms: An exploratory investigation. *English Language Teaching*, 4(1), 18-25.
- Seyyedi, K. (2025). The effect of mock conference on EFL learners' speaking lexical complexity. *Zanco Journal of Human Sciences*, 29(5), 243-253.
- Seyyedi, K., & Amin, N.M.H. (2020). The effect of immediate and delayed error correction on accuracy development of intermediate EFL learners' writing. *International Journal of Linguistics, Literature and Translation*, 3(2), 100-108.
- Vygotsky, L.S. (1980). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge: Harvard University Press.