Content list available at http://ijltr.urmia.ac.ir

Iranian Journal of Language Teaching Research

ORIGINAL ARTICLE

Urmia University

A Comparative Study of the Developmental Paths of Low and High-achieving Iranian EFL Pairs and Individuals in Terms of L2 Writing Features

Sahar Zahed Alavi [a, *]

[a] *University of Bojnord, Iran*

ABSTRACT

Following Dynamic Systems Theory (Thelen & Smith, 1994) and Sociocultural Theory (Vygotsky, 1978), this study compared the developmental paths of low and high-achieving EFL individuals and dyads, considering production features of complexity, accuracy, and fluency (CAF). The participants included six MA students majoring in Industrial Design at Iran University of Science and Technology. They were assigned to individual and pair groups; one learner with minimum performance on CAF features (Low-achieving Individual; LAI) and one learner with maximum performance on CAF features (High-achieving Individual; HAI) were asked to compose 10 tasks over the semester individually. Furthermore, two learners with minimum performance on CAF features (Low-achieving Pair; LAP) and two learners with maximum performance on CAF features (High-achieving Pair; HAP) were asked to write on 10 tasks over the semester in pairs. Their developmental paths in terms of CAF features were tracked. The findings demonstrated unique developmental paths with oscillations and trade-offs between the components of writing proficiency, and the possible influence of collaborative performance and participants' proficiency level in the differences in the developmental paths. They showed that LAI and HAI outperformed in fluency and complexity over time, and HAP outperformed in accuracy. The study concludes with implications for pedagogy, suggesting the uniqueness of the developmental paths of each learner and the need for evaluating writings in terms of a multidimensional lens.

Keywords: dynamic systems theory; collaborative performance; individual performance; developmental paths

ARTICLE HISTORY

Received: 28 Oct. 2024 **Revised version received:** 25 Apr. 2025

Accepted: 12 June 2025 **Available online:** 10 July 2025

* Corresponding author: University of Bojnord, Iran
Email address: s.zahedalavi@ub.ac.ir

doi 10.30466/ijltr.2025.55682.2843

Introduction

The performance of learners in academic writing can be examined through features of complexity (i.e., the ability to use a variety of advanced and elaborate structures in a text), accuracy (i.e., the ability to use error-free structures), and fluency (i.e., the ability to generate a structure with proper speed). These three distinct aspects of language production validly signal language proficiency in written production which is multi-dimensional. They address meaning (fluency) and form (complexity and accuracy; Skehan, 2009). They can also distinguish a proficient language user from a non-proficient one. In other words, it is assumed that a proficient language user can do tasks fluently through complex and accurate structures (Ellis, 2009; Housen & Kuiken, 2009; Housen et al., 2012). More specifically, not only should the three aspects of language proficiency, namely CAF, be taken into account if researchers are making claims about a learner's level of proficiency, but also it is more profitable to attend to different operationalization of CAF to present a more complete picture (Norris & Ortega, 2009).

Although EFL writing has been performed individually, there is an interest in the involvement of two or more individuals in writing tasks (i.e., collaborative writing). In collaborative writing, individual performance is accompanied by social orientation. Collaboration provides learners with the opportunity to brainstorm ideas, co-construct knowledge, negotiate meaning, discuss language use, scaffold, and offer feedback in all stages of a writing task (Storch, 2019; Villarreal & Gil-Sarratea, 2020).

On the other hand, the development in writing features can be tracked through the Dynamic Systems perspective (Thelen & Smith, 1994), which tracks variations within and between learners. Both continuous and discontinuous changes in the writing performance of learners can be detected through this perspective.

Most studies on writing focused on the cognitive components of writing processes (Akbarzadeh et al., 2014; Storch, 2019); they lack a developmental viewpoint. On the other hand, most of the studies which examined the development of specific features were cross-sectional (e.g., Beers & Nagy, 2009; Crossley & McNamara, 2014; Ellis & Yuan, 2004; Lu, 2011). Moreover, the few studies which investigated the development in longitudinal studies mostly considered one or two aspects of writing proficiency. For instance, Baba and Nitta (2014) examined the development of fluency in individuals' performances, Polat and Kim (2014) and Spoelman and Verspoor (2010) investigated the accuracy and complexity of individuals' performances, and Vyatkina (2012) investigated the development of complexity in individuals' performances. In addition, if they had examined all three CAF features, they would have focused on individuals' performances (e.g., Larsen-Freeman, 2006) and not the collaborative productions.

Thus, few studies have focused on how and when writers resolve conflicts in their writing process and develop during writing tasks. Investigating the developmental paths of learners may have implications in real-world contexts; it offers strategies to support students on their journeys toward development. Thus, the objective of the present study was to adopt a more comprehensive perspective and compare the developmental patterns in writing tasks by both low and high-proficiency learners in individual and pair writing groups, focusing on the production features of language (i.e., CAF). The methodological frameworks followed were the Dynamic Systems Theory (Thelen & Smith, 1994) and the Sociocultural Theory (Vygotsky, 1978). Thus, the research questions are posed as follows:

- What are the developmental paths of low and high-achieving EFL pairs writing collaboratively in terms of CAF?

- What are the developmental paths of low and high-achieving EFL individuals writing independently in terms of CAF?

Review of Literature

Due to the theoretical significance of the Dynamic Systems Theory and the importance of time (10 tasks) in this study, this section elaborates on Dynamic Systems Theory. In addition, empirical investigations on the individual differences in terms of proficiency level (high and low proficiency writers) and collaborative writing performances are provided as follows.

Dynamic Systems Perspective

The Dynamic Systems Theory (Thelen & Smith, 1994) claims that a system (e.g., language) includes several sub-systems (e.g., phonology, semantics, syntax) which are interconnected. Changes in one sub-system affect other sub-systems which may result in a supportive or competitive relationship between them. Hence, Dynamic Systems Theory considers language acquisition a complex and complicated endeavor distinguished from the linear, stage-based, and predictable perspective (de Bot et al., 2008; Larsen-Freeman, 2020).

The Dynamic Systems perspective is outstanding for several reasons. First, it detects continuous and discontinuous changes in a system. If the system is rearranged to a higher level after the discontinuous change, a phase transition in which new properties are added would appear (van Dijk et al., 2024; van Dijk & van Geert, 2007). More specifically, phase transitions represent significant changes and development (Larsen-Freeman & Cameron, 2008). Second, nonlinearity is a common developmental feature of dynamic systems (de Bott, 2008). Third, it considers learning a language a dynamic attempt in which the passage of time is a crucial consideration. In other words, how the components of language are used changes over time (Larsen-Freeman, 2020). Fourth, the non-linearity of the complex system results in self-organizing changes and the emergence of new behaviors (Larsen-Freeman & Cameron, 2008).

Since Dynamic Systems Theory is a relatively new framework, a few studies have taken it in applied linguistics mainly to examine intra-individual variability in the language learning process. Larsen-Freeman (2006) examined the writings of five Chinese students during four tasks in terms of CAF features. She noted that the developmental pattern of the mean of the participants' performances was different from that of individual participant's performance. The mean developmental path obscured each participant's unique developmental path. She highlighted that the variation was not "noise, but rather a natural part of dynamically emergent behavior assembled by the individual" (p. 615). In other words, variations in each individual's performance should not be neglected.

Spoelman and Verspoor (2010) studied the development of complexity and accuracy measures in a Finnish learner over time. They found no significant relationship between the accuracy and complexity of the learner's productions. The accuracy of the performance improved in most tasks over time. In addition, although the sentence complexity of the learners' written productions improved over time, noun phrase complexity did not improve. They noticed the need to consider different aspects of complexity simultaneously, rather than considering complexity as a unidimensional construct.

In his study, Vyatkina (2012) compared the cross-sectional and longitudinal data on the development of complexity. The cross-sectional data showed that despite improvements in sub-clausal complexity, the coordinate complexity measure deteriorated. However, the longitudinal

study of two students demonstrated variations in their developmental paths. The general measure of complexity of the participants improved over time. Although one of the participant's performances were more complex than the mean of complexity measure in the cross-sectional part of the study, the other participant's use of complexity was lower than the mean of complexity in the cross-sectional part of the study. Moreover, in terms of subordination and coordination, variability was observed in the data.

Polat and Kim (2014) examined the two features of complexity and accuracy in the speech of an untutored language learner during one year. They showed that the untutored learner who learned English in a naturalistic context (i.e., without instruction) improved a bit in terms of complexity features. However, there were fluctuations in the data with no gain in the performance of the learner in terms of accuracy over time. Thus, they suggested that improvements in terms of accuracy require instruction (i.e., an external control): In a naturalistic learning context, learners receive feedback on communicative effectiveness rather than correct use of form. Thus, it was not possible to develop accuracy through naturalistic learning.

Liu et al. (2025) also tracked the development patterns of learners in terms of lexical and syntactic complexity. They found that the improvement of these two measures over time was accompanied by phases of compensatory growth, which showed how these measures compete for cognitive resources.

Collaborative Writing

The theoretical background of collaborative writing lies in the Interaction Hypothesis (Long, 1983), the Output Hypothesis (Swain, 1993), and the Sociocultural perspective (Vygotsky, 1978). The Interaction Hypothesis (Long, 1983) emphasizes the efficiency of learning through comprehensible input and negative feedback. The Output Hypothesis (Swain, 1993) focuses on the role played by the learner's output in language development. Furthermore, the key constructs in the Sociocultural Theory (Vygotsky, 1978) include the zone of proximal development, scaffolding, and language as a tool for mediation.

Most of the studies found in the literature on the investigation of the writing performance of collaborative writing groups are cross-sectional. In addition, the quality of learners' writing performances in these studies is examined through different scoring scales. For instance, Davison (2024), Peng (2024), Villarreal and Gil-Sarratea (2020), Tavakoli and Rezazadeh (2014), Dobao's (2012), and Storch (2005) measured writing proficiency in terms of CAF measures. Hsu (2025) measured the quality of content and organization. Ajideh et al. (2016) considered accuracy and cohesion/ coherence improvements. Moreover, Ameri-Golestan and Dousti (2015) and Shehadeh (2011) focused on organization, content, grammar, mechanics, and style.

Moreover, some studies investigated the impact of collaboration in just one phase of writing. For instance, Hsu (2025), Ameri-Golestan and Dousti (2015), and Mazdayasna and Zaini (2015) explored the effect of collaborative pre-writing activities (planning) on EFL learners' writing performance. It was suggested that when learners worked in pairs, they could receive feedback on their word choice, tense, article, word order, mechanics, rhetorical patterns, and discourse markers. Furthermore, Tavakoli and Rezazadeh (2014) found that collaboration in planning did not affect the complexity of the participants' productions. A justification put forward in this regard was the type of task (argumentation) used in the study; this task type demands a certain amount of subordination. Concerning the accuracy of the participants' productions, learners in the collaborative planned situation significantly performed more accurately than those in the individually planned situation. This result was consistent with the Output Hypothesis (Swain, 1995), claiming that interaction can improve grammatical performance. Concerning the fluency of the

participants' writing performance, learners in the individual planning situation performed more fluently than those in the collaborative planning situation. The researchers concluded that collaboration helped the learners discuss different points of view, which assisted them in limiting the length of their productions; however, the learners who wrote individually could not limit their ideas and wrote more detailed texts.

The cross-sectional studies which examined the overall quality of learners' collaborative productions through the pretest-posttest design showed that the collaborative writing group produced high-quality texts in terms of accuracy (Dobao, 2012; Davison, 2024; Mujtaba et al., 2021; Peng, 2024; Storch, 2005; Villarreal & Gil-Sarratea, 2020). Villarreal and Gil-Sarratea (2020) explained the reason for producing accurate texts as the opportunity to discuss problems and find solutions to them during collaboration. More specifically, collaborative learners tried to bridge their zone of proximal development through languaging, dialogues, and discussions. Mujtaba et al. (2021) demonstrated that learners writing collaboratively had fewer verb tense and word choice errors. However, Shehadeh (2011) found an insignificant effect of collaborative activity on writing accuracy. He related this insignificant effect to the students' proficiency level. He admitted that the students did not assist each other with grammatical accuracy due to their low proficiency in English.

Some studies emphasized that learners' writing performances in collaborative writing tasks were different in terms of lexical and grammatical accuracy and mechanical accuracy. Dobao (2012) showed that the grammatical and lexical accuracy of the texts written by the learners in pairs was enhanced; however, mechanical accuracy did not. The pair members collaboratively decided on the structure and vocabulary of their production, but the decision on spelling and punctuation was individually made by the learner writing the text. Shehadeh (2011) also argued that there was no improvement in the students' use of mechanics in collaborative tasks since the use of mechanics is more straightforward and obeys a limited range of rules, which can be more easily dealt with and mastered by students individually.

However, some studies showed that the complexity of the productions of the collaborative writing group did not improve (e.g., Davison, 2024; Dobao, 2012; Peng, 2024; Villarreal & Gil-Sarratea, 2020). Villarreal and Gil-Sarratea (2020) explained the lack of complexity by referring to Pallotti's (2009, cited in Villarreal & Gil-Sarratea, 2020) claim that complexity improves only when the task needs it. When two tasks have similar communicative goals, a similar level of complexity is expected. However, Storch (2005) found that collaborative writing groups outperformed in complexity. This discrepancy in findings can be attributed to the short sample size of Storch's (ibid) study and the use of complexity measures, which were different from other studies.

In line with complexity, most studies showed that the fluency of the productions of collaborative writing groups was lower than that of individual writing groups (e.g., Davison, 2024; Dobao, 2012; Peng, 2024; Storch, 2005; Villarreal & Gil-Sarratea, 2020). According to Dobao (2012), having been given the same amount of time to write, students experiencing independent writing produced longer texts than students experiencing dyadic writing because the students writing together required extra time to reach a consensus on both content and form in their texts. As Storch (2005) admitted, collaboration helped pairs in focusing on different dimensions of the writing process, namely, planning, drafting, and revising. Thus, they were in short supply of time to write longer texts. Furthermore, Villarreal and Gil-Sarratea (2020) referred to the word limit set for learners as the reason for the lack of improvement of collaborative writing groups in terms of fluency measures. Tavakoli and Rezazadeh (2014) also concluded that collaboration helped the learners discuss different points of view, which assisted them in limiting the length of their productions; however, the learners who wrote individually could not limit their ideas and wrote more detailed texts.

Other factors can also affect the ultimate quality of the learners' collaborative writings. Ajideh et al. (2016) discussed the findings of collaborative writing groups in terms of task familiarity and the enculturation of the participants; it was argued that repeated performance on similar tasks might improve the learners' skills in selecting appropriate information from the isomorphic tasks and writing them coherently. In addition, learners from specific cultures might act differently in collaborative writing endeavors. For instance, he showed that Asian learners (compared to Malaysian learners) favored teacher-directed classes and neglected their peers' comments on the grammatical accuracy of the texts. Zhang (2018) admitted the positive role of using the first language (rather than the second language) in group discussions while writing collaboratively in producing more complex texts.

Learners' Performances Considering Proficiency Level

Several studies investigated the relationship between learners' proficiency levels and their improvement in terms of different aspects of language performance. For instance, language proficiency is reported to be related to establishing relationships between words (Christiansen & Arnon, 2017), syntactic complexity (Kim et al., 2024), lexical complexity (Liu et al., 2025), accuracy (Panagopoulos, 2024; Sánchez & Sunesson, 2023), and fluency (Panagopoulos, 2024).

Deciding on the way to group learners for the sake of doing collaborative activities in classes with different proficiency levels has been one of the concerns of teachers (Cao, 2021; Niu et al., 2018; Storch & Aldosari, 2012; Watanabe & Swain, 2008; Zabihi & Ghahramanzadeh, 2022). The studies on the relationship between learners' proficiency levels and the quality of their collaborative productions have divergent findings. Some studies showed that as the overall proficiency of the members of a pair enhances, learners use more language-related episodes (Leeser, 2004; Storch & Aldosari, 2012; Watanabe & Swain, 2007). They claimed that when learners collaborate with a more proficient partner (high-low proficiency members), they use more language-related episodes, whereas other studies claimed that pair members with similar proficiency level (i.e., low-low and high-high proficiency members) are superior to different proficiency level pair members (i.e., low-high proficiency members). In addition, high-high proficiency level pairs engage with more language use than low-low proficiency pairs because they engage with more idea generation and offer more solutions to solve problems during pair work (Dao & McDonough, 2018; Namkung & Kim, 2024; Qiu & Lo, 2017; Zabihi & Ghahramanzadeh, 2022).

This discrepancy in findings is due to considering different aspects of performance. The studies which demonstrated the superiority of heterogenous pairs (high-low proficiency members) to homogenous pairs (low-low or high-high proficiency members) investigated the effect of learners' level of proficiency on the frequency of language related episodes, while the studies which claimed the superiority of similar proficiency level pairs to different proficiency pairs examined the impact of learners' proficiency level on their engagement. Accordingly, the engagement level of partners in a pair (not the use of language-related episodes) shows the members' actual learning (Zabihi & Ghahramanzadeh, 2022).

The proficiency level of pair members also impacts the effectiveness of different types of feedback (i.e., implicit and explicit feedback). In other words, high proficiency level learners benefit more from feedback than low proficiency level learners (Cao, 2021). In collaborative endeavors, high proficiency level learners can notice both implicit and explicit feedback types, whereas low proficiency level learners are more affected by implicit feedback (Li, 2014).

As evident, no study has compared the developmental paths in writing tasks of low and high-proficiency learners in individual and pair writing groups.

Methods

Design

This study followed the Time Series quasi-experimental design which examines a variable in different time intervals (Ary et al., 2010). The measures of the learners' writing performances (i.e., CAF) were examined across time (during 10 tasks) and were demonstrated graphically.

Participants

The participants included six MA students majoring in Industrial Design at Iran University of Science and Technology. They were in the age range of 24-28 years old and were studying in their third semester. They had a two-credit course on English academic writing needed for their MA degree. Since the objective of the study was to investigate the writing performances of learners with maximum variation (i.e., high and low-achieving learners) in different contexts of writing (i.e., individual writing and collaborative writing), these six participants were purposively selected from among their twelve classmates. More specifically, they were purposively selected based on the investigation of CAF features in a pretest. From among a class of 18 MA learners, those who gained the maximum scores (i.e., three learners) and those who gained the minimum scores (i.e., three learners) in terms of CAF features in the pretest were selected. These six learners were assigned to individual and pair groups. One learner (a female) with minimum performance on CAF features (Low-achieving Individual; LAI) and one learner (a female) with maximum performance on CAF features (High-achieving Individual; HAI) were asked to write individually. Furthermore, two learners (two males) with minimum performance on CAF features (Low-achieving Pair; LAP) and two learners (two females) with maximum performance on CAF features (High-achieving Pair; HAP) were asked to write in pairs.

Materials

Oxford Placement Test (the first version, 2001) was administered to examine the learners' level of proficiency and to choose the proper syntactic complexity measure suggested by Norris and Ortega (2009). Cronbach's Alpha index was used to examine the reliability of the Oxford Placement Test (2001); it turned out to be .85, which is proper internal consistency reliability (Pallant, 2007). Furthermore, eighteen paragraphs written individually in the pretest phase by the learners in the 18-member class, and 40 paragraphs written by the participants were the materials of the present study. LAI and HAI were invited to write ten paragraphs throughout the semester independently; however, LAP and HAP were invited to write ten paragraphs during the semester in pairs. It should be noted that to explore the points with a sudden change, which is a critical and discontinuous change, in the time series data, the change point analysis was carried out by a computer program called Change-Point Analyzer (Taylor, 2003).

Data Collection Procedures

At the beginning of the course, the Oxford Placement Test (2001) was administered. Then, each of the learners in the 18-member class was tasked with writing a paragraph individually (i.e., the pretest), which served as the basis for the selection and assignment of the participants to the individual and pair writing groups. The learners were assigned to individual and pair writing groups according to the CAF measures' means in their initial individual compositions (i.e., the pretest). One learner (a male) with minimum performance on CAF features (LAI) and one learner (a male) with maximum performance on CAF features (HAI) were selected to write individually. Furthermore, two learners (two males) with minimum performance on CAF features (LAP) and

two learners (two females) with maximum performance on CAF features (HAP) were selected to write in pairs.

Then, to investigate their developmental paths over time, they were asked to write ten tasks during a semester. More specifically, LAI and HAI were asked to write each task individually in 30 minutes. However, LAP and HAP were asked to collaborate in all processes of paragraph writing (i.e., brainstorming, drafting, and revising) in 40 minutes.

Data Analysis Procedures

Since complexity, accuracy, and fluency provide a full picture of writing performance (Lu, 2011), they were used as the unit of analysis of the participants' productions. Concerning complexity, although it can be investigated in terms of subordination, general complexity, and sub-clausal estimates considering different proficiency levels (Norris & Ortega, 2009), in the current study subordination measures were used because the participants were at the intermediate proficiency level based on the results of Oxford Placement Test (2001). Thus, syntactic complexity was examined by estimating the ratio of clauses to T-units (Foster & Skehan, 1998), and the ratio of dependent clauses to all clauses (Wolfe-Quintero et al., 1998). Considering accuracy, following Wigglesworth and Storch (2009) and Dobao (2012), it was examined by estimating the ratio of error-free T-units to all T-units and the ratio of error-free clauses to all clauses. It should be noted that syntactic errors (e.g., fragmentation) and morphological errors (e.g., articles and verb tenses) were considered. However, errors in spelling and punctuation were not taken into account. Finally, fluency was examined in terms of the average number of words, T-units, and clauses in the text (Wigglesworth & Storch, 2009). The number of words in the compositions was counted by Microsoft Word (2010). It should be noted that the writing performances of the study groups over the 10 tasks were compared graphically in terms of each of the CAF features.

Intra-rater reliability and inter-rater reliability were examined through re-analyzing the units of analysis relevant to CAF measures (i.e., number of T-units, number of dependent clauses, overall clauses, error-free clauses, and error-free T-units). The researcher estimated these units again after five weeks. Kappa agreement coefficients turned out to be .96, .94, .91, .90, and .93, respectively. In addition, inter-rater reliability was examined by a colleague who holds a Ph.D. in TEFL. He was trained regarding the units of CAF measures. He was asked to code the above-mentioned units in 15 paragraphs written individually and 15 paragraphs written in pairs. The agreement indices were estimated to be .94, .92, .91, .88, and 92, respectively.

Results

The performances (i.e., the developmental paths) of the participants (i.e., LAI, HAI, LAP, and HAP) in terms of fluency measures are shown in Figure 1. As is evident, there are fluctuations in the performances of all participants during the ten tasks.

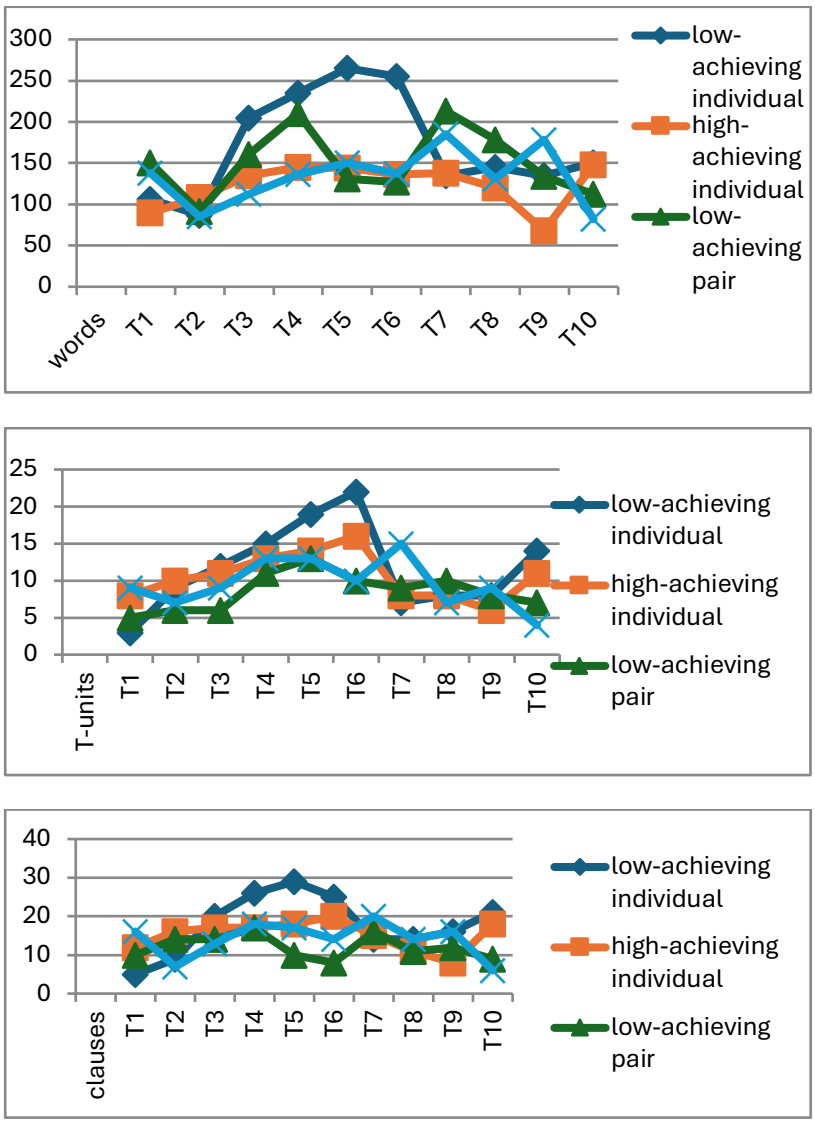
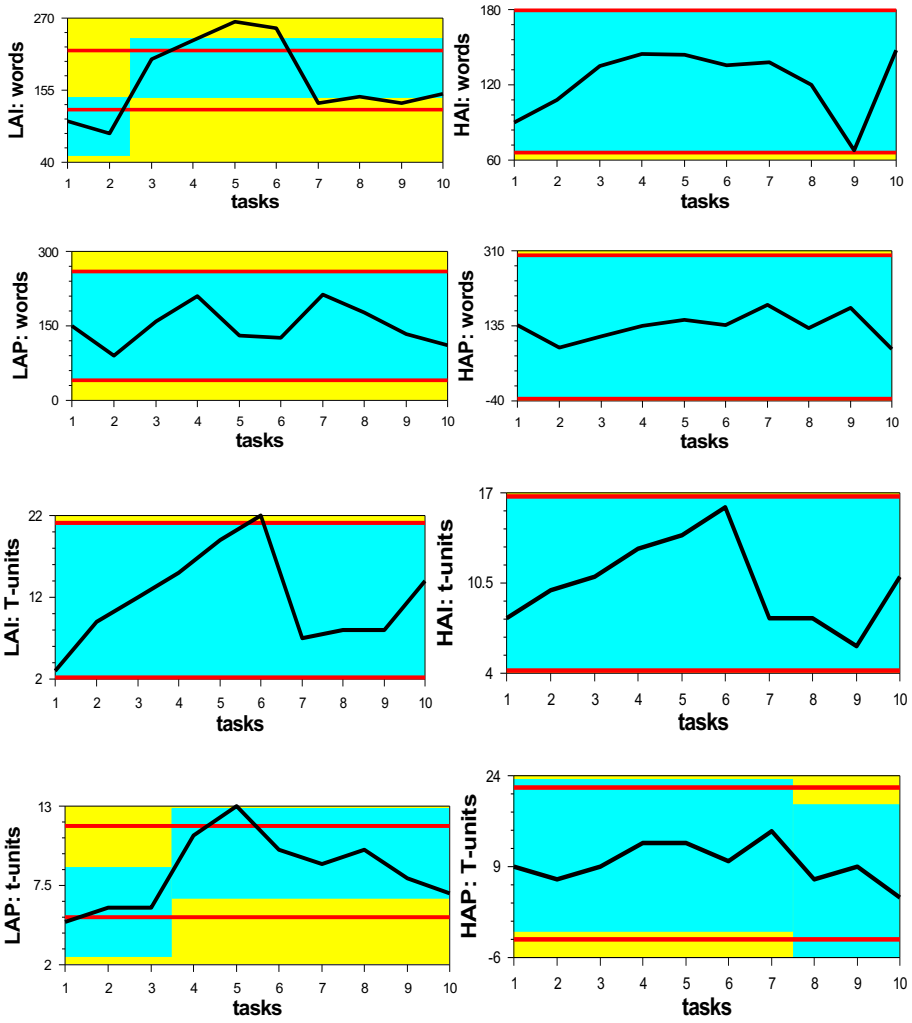


Figure 1. The participants' developmental paths in terms of fluency measures

Concerning the number of words, LAI produced more words than HAI during the ten tasks. Moreover, LAP produced more words than HAP in seven tasks. In addition, in the second, third, and fourth tasks, HAI produced longer texts than HAP; however, in the seventh, eighth, ninth, and tenth tasks, HAP produced longer texts than HAI. In addition, although LAP produced longer texts in the seventh and eighth tasks, LAI produced longer texts in the third, fourth, fifth, and tenth tasks. Finally, in a general comparison, LAI produced more words than other participants in tasks 3, 4, 5, 6, and 10.

Regarding the number of T-units, LAI produced more T-units than HAI in tasks 3, 4, 5, 6, 9, and 10. Furthermore, HAP produced more T-units than LAP in tasks 1, 2, 3, 4, and 7. Moreover, LAI produced more T-units than LAP in tasks 2, 3, 4, 5, 6, and 10. Furthermore, HAI produced more T-units than HAP in tasks 2, 3, 5, 6, 8, and 10. Ultimately, in a comparison, LAI produced more T-units than other participants in tasks 3, 4, 5, 6, and 10.

Considering the number of clauses, LAI produced more clauses than HAI in tasks 3, 4, 5, 6, 8, 9, and 10. In addition, HAP produced more clauses than LAP in tasks 1, 4, 5, 6, 7, 8, and 9. Furthermore, LAI produced more clauses than LAP in tasks 3, 4, 5, 6, 8, 9, and 10. In addition, HAI produced more clauses than HAP in tasks 2, 3, 5, 6, and 10. In addition, HAP produced more clauses than HAI in tasks 1, 4, 7, 8, and 9. Finally, in a comparison, LAI produced more clauses than other participants in tasks 3, 4, 5, 6, 8, 9, and 10.



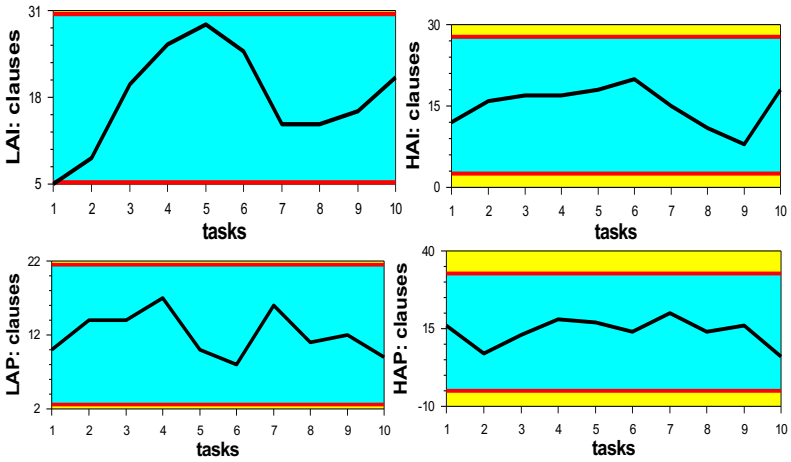
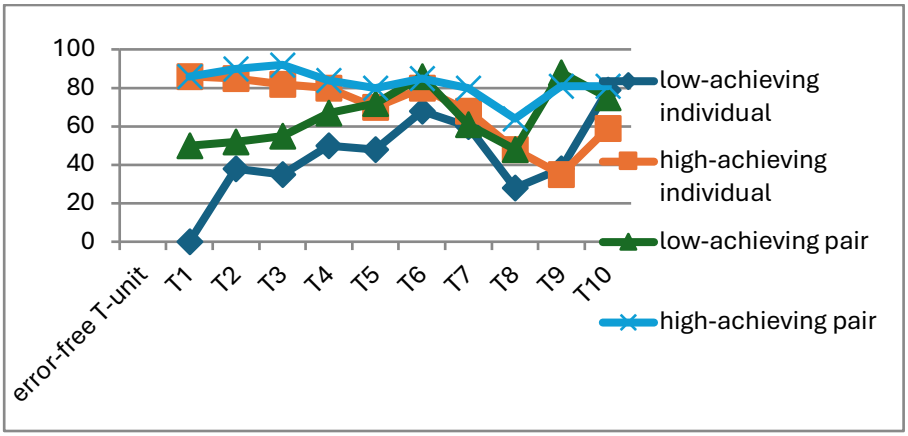


Figure 2. The change point analyses of the participants' performances in terms of fluency measures

Figure 2 shows the change point analysis (conducted by Change-Point Analyzer) of the participants' performances in terms of fluency measures. As is evident in this figure, LAI experienced a sudden change in terms of the number of words in task 2, LAP experienced a sudden change in terms of the number of T-units in task 3, and HAP experienced a sudden change in terms of the number of T-units in task 7. Nevertheless, the participants did not experience a sudden change in terms of the number of clauses.



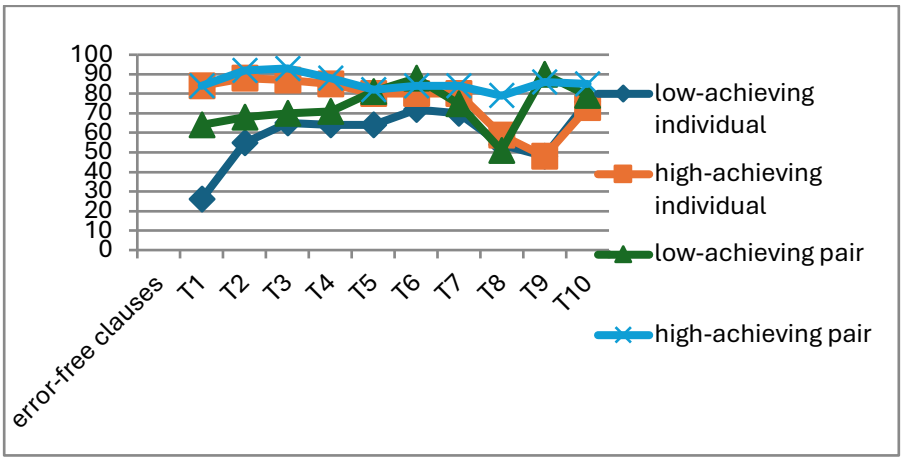


Figure 3. The participants' developmental paths in terms of accuracy measures

Figure 3 represents the developmental paths of the participants in terms of accuracy measures. Considering error-free T-units, HAI produced more error-free T-units than LAI in tasks 1, 2, 3, 4, 5, 6, 7, and 8. In addition, HAP produced more error-free T-units than LAP in tasks 1, 2, 3, 4, 5, 7, 8, and 10. Furthermore, LAP produced more error-free T-units than LAI in tasks 1, 2, 3, 4, 5, 6, 7, 8, and 9. Moreover, HAP produced more error-free T-units than HAI in tasks 2, 3, 4, 5, 6, 7, 8, 9, and 10. Finally, in a general comparison, HAP produced more error-free T-units than other participants in tasks 2, 3, 4, 5, 6, 7, 8, and 10.

Regarding error-free clauses, HAI produced more error-free clauses than LAI in tasks 1, 2, 3, 4, 5, 6, 7, and 8. Moreover, HAP produced more error-free clauses than LAP in tasks 1, 2, 3, 4, 5, 7, 8, and 10. In addition, LAP produced more error-free clauses than LAI in tasks 1, 2, 3, 4, 5, 6, 7, 9, and 10. Moreover, HAP produced more error-free clauses than HAI in tasks 2, 3, 4, 5, 6, 7, 8, 9, and 10. Ultimately, in a general comparison, HAP produced more error-free clauses than other participants in tasks 2, 3, 4, 5, 7, 8, and 10.

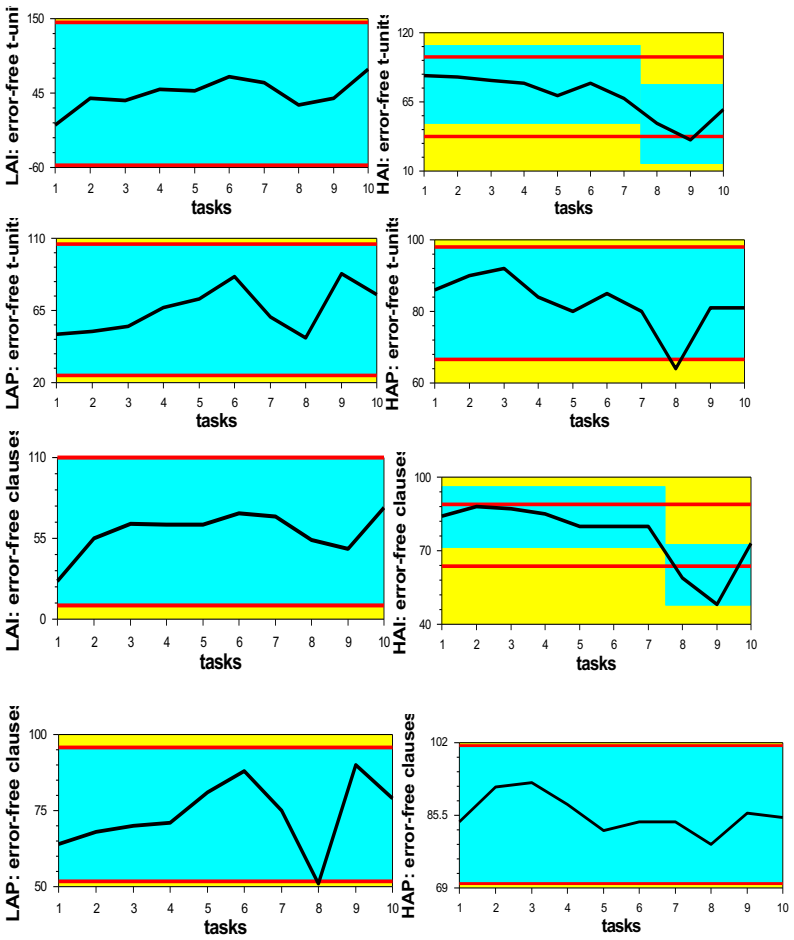


Figure 4 represents the change point analysis of the participants' performances considering accuracy measures. As this figure indicates, no sudden changes appeared in the performances of LAI, LAP, and HAP regarding the number of error-free T-units. However, HAI experienced a sudden change considering the number of error-free T-units at task 7. Furthermore, with regard to the number of error-free clauses, just HAI experienced a sudden change at task 7.

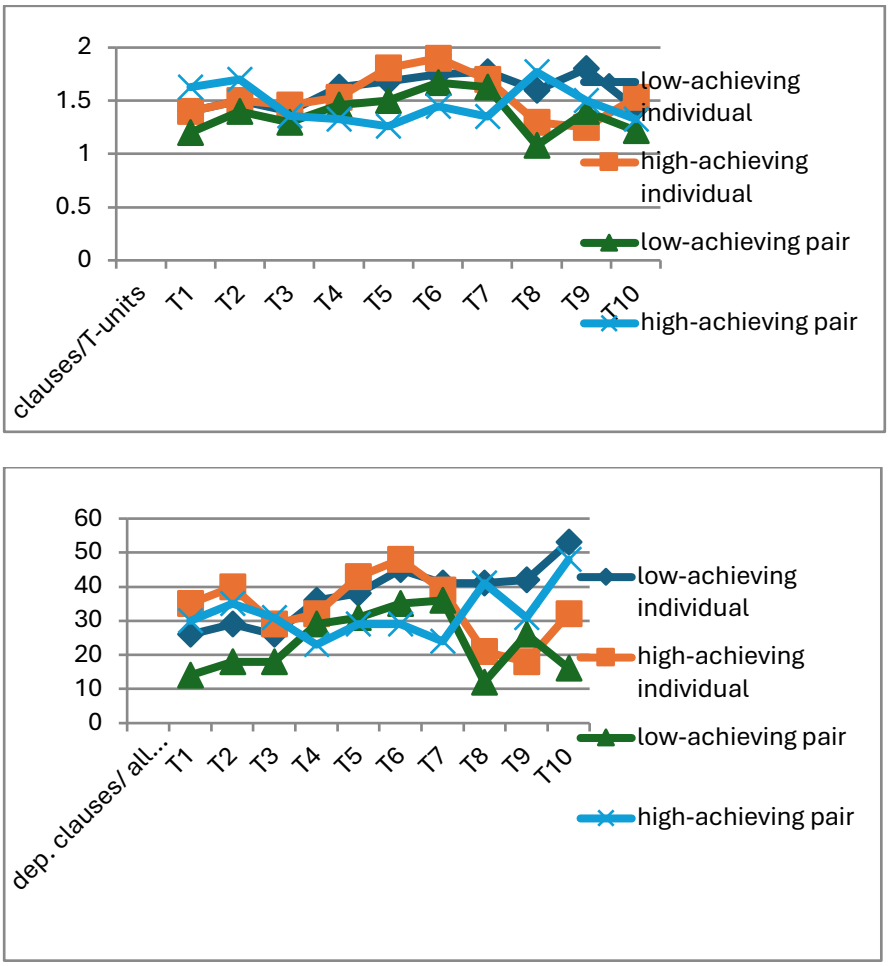


Figure 5. The participants' developmental paths in terms of complexity measures

Figure 5 shows the developmental paths of the participants in terms of complexity. Considering the performance of the learners regarding the number of clauses to the T-units, HAI outperformed HAP in tasks 3, 5, 6, and 10. Moreover, HAP outperformed LAP in tasks 1, 2, 3, 8, 9, and 10. In addition, LAI produced more clauses to T-units than LAP in tasks 2, 3, 4, 5, 6, 7, 8, 9, and 10. Moreover, HAI produced more clauses to T-units than HAP in tasks 3, 4, 5, 6, 7, and 10. Ultimately, in a general comparison, HAI produced more proportions of clauses to T-units than other participants in tasks 3, 5, 6, and 10.

Regarding the number of dependent clauses to all clauses, HAI produced a greater proportion of dependent clauses than LAI in tasks 1, 2, 3, 5, and 6. In addition, HAP outperformed LAP in tasks 1, 2, 3, 8, 9, and 10. Moreover, LAI outperformed LAP in tasks 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10. In addition, HAI outperformed HAP in tasks 1, 2, 4, 5, 6, and 7. Ultimately, in a general comparison, HAI produced a greater proportion of dependent clauses to all clauses than other participants in tasks 1, 2, 5, and 6.

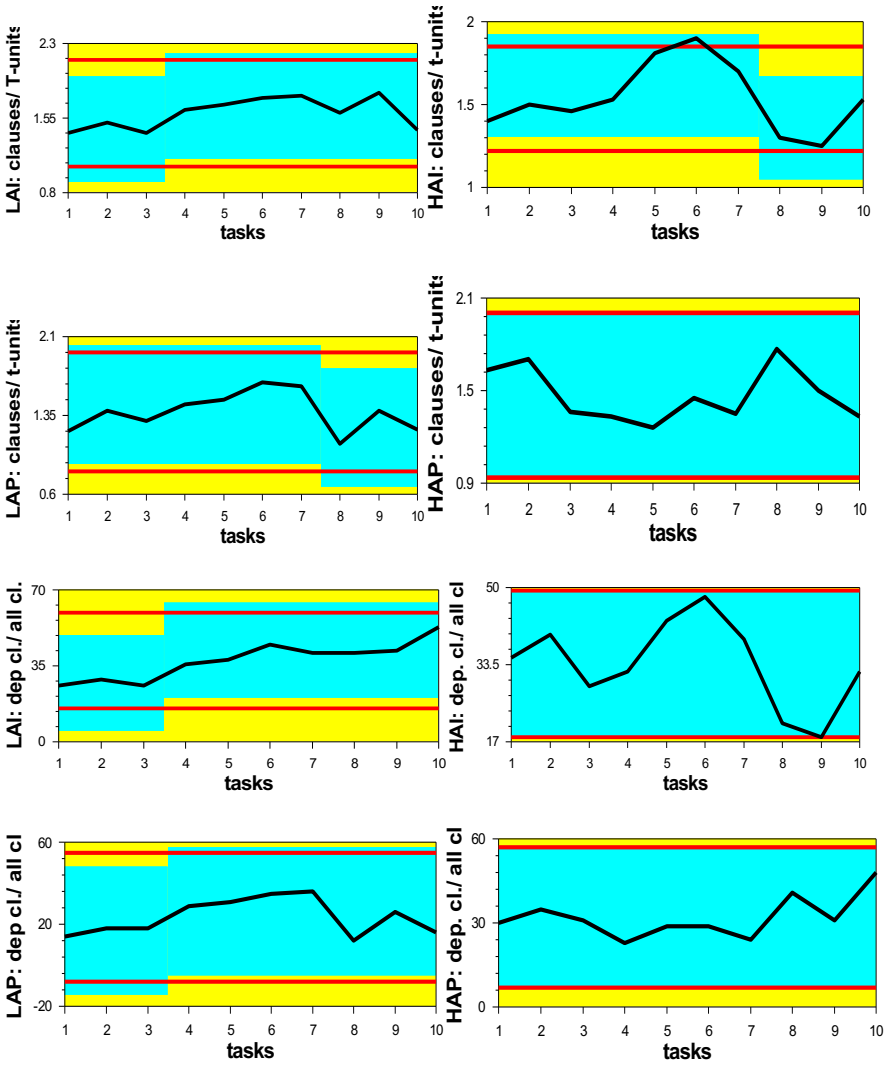


Figure 6. The change point analyses of the participants' performances in terms of complexity measures

Figure 6 represents the change point analyses of the participants' performances in terms of complexity measures. As this figure reveals, with regard to the proportion of clauses to T-units, LAI underwent a sudden change at task 3. Moreover, HAI and LAP experienced a sudden change at task 7. Considering the proportion of dependent clauses to all clauses, both LAI and LAP experienced a sudden change at task 3 while other participants did not experience a sudden change.

Discussion

Considering all the graphs, one will notice that the development of each participant's writing performance during the ten tasks was nearly similar in most of the subcategories of each of the CAF measures. In other words, the shapes of the graphs displaying the development of each participant's writing performance considering the number of words, the number of T-units, and the number of clauses were nearly similar. Concerning the complexity feature, the graphs depicting the development of each participant's performances in terms of the proportion of clauses to T-units and the proportion of dependent clauses to all clauses were somehow similar. As for the accuracy feature, each participant had somehow similar performances in terms of error-free T-units and error-free clauses during the ten successive tasks.

Therefore, it might be proposed that the measures of each of the CAF features consistently estimated their intended feature (Dobao, 2012; Norris & Ortega, 2009; Wigglesworth & Storch, 2009). In other words, the number of words, T-units, and clauses estimated the same feature (i.e., fluency of productions) since each participant's (i.e., LAI, HAI, LAP, and HAP) performances were similar in these three measures. In the same vein, the proportion of clauses to T-units and the proportion of dependent clauses to all clauses estimated the same feature (i.e., complexity of productions). Similarly, measures of error-free T-units and error-free clauses estimated the same feature (i.e., accuracy of productions).

As it was evident in the graphs, there were oscillations in the performances of all participants (i.e., LAI, HAI, LAP, and HAP) in terms of all production features (i.e., CAF) during the ten tasks. In other words, the participants progressed and regressed during the ten tasks. This finding lent support to Dynamic Systems Theory, which suggests that the process of development is nonlinear. Accordingly, a dynamic system changes with time. Therefore, the language production aspects, comprising a dynamic system, change over time, too. In other words, variability in the performance of learners is a norm (Larsen-Freeman, 2009). Moreover, in development, there are unpredictable regressions, progressions, and jumps (Larsen-Freeman, 1997; Van Dijk et al. 2024), and different learners move on diverging paths to second language acquisition (Larsen-Freeman, 2006).

Learners experienced distinct paths of development in each of the CAF measures. Even the learners who had the same writing experience (i.e., those who wrote the tasks individually or those who did so in pairs) showed distinct developmental patterns. In other words, LAI and HAI, who similarly composed individually, showed distinct paths of development in each of the CAF measures. Likewise, LAP and HAP, who composed in pairs, presented varying paths of development in CAF measures over time. The results of change point analyses also supported the different developmental paths in learners; each individual or pair's experience of a sudden change was unique. The sudden changes in each of the CAF features occurred at different tasks. Thus, learners' diverging developmental paths can be attributed to their different physical, affective, and cognitive selves and unique learning experiences; they respond differently to the learning context (Larsen–Freeman & Cameron, 2008; Van Dijk et al., 2024). Furthermore, this difference could also be accounted for by referring to the effect of their different writing proficiency level.

The participants' (i.e., LAI, HAI, LAP, and HAP) developmental paths in each of the CAF measures showed more variability in the early phases of performance. However, during the final tasks, the levels of variability decreased, and more settlement was observed in the learners' performances. This is in line with Spoelman and Verspoor's (2010) observation which noticed more variability in the individual's performances in the beginning tasks than the final ones. It should also be emphasized that although more settlement was observed during the final tasks' performances, variability in performance was still present; it was due to the concept explicates that variability is

the innate feature of a self-organizing system (Larsen-Freeman & Cameron, 2008; van Dijk et al. 2024).

There were also frequent instances of occasions in which when the participants progressed in terms of one of the CAF features in a task, they regressed in the other feature in the same task. For instance, although LAI progressed in terms of fluency at task 3, he regressed in terms of accuracy and complexity at the same task. This finding is in line with the Trade-Off Hypothesis (Skehan, 1998) which suggests that learners cannot manage to concentrate on all aspect of language at once because of the limited information processing capacity; if they pay attention to one feature of language production (e.g., each of the CAF features), they would lose attending on the other features. This finding is also in line with the Dynamic Systems Theory, which claims that the interaction of different aspects of language may be complex and non-linear (Van Dijk et al., 2024).

Concerning the performances in terms of fluency feature, the comparison of the graphs showed that the writing performance graphs of the participants who wrote texts individually (i.e., LAI and HAI) indicated more fluency in most tasks than those who wrote tasks collaboratively (i.e., LAP and HAP). This finding is in line with the literature on the effect of collaboration on the fluency of the productions (e.g., Dobao, 2012; Peng, 2024; Storch, 2005; Tavakoli & Rezazadeh, 2014; Villarreal & Gil-Sarratea, 2020). Learners performing individually composed longer paragraphs than those performing in pairs since the latter required more time to discuss different points of view, to reach a consensus on the content and form of their writings. On the contrary, the learners writing individually could not limit their ideas and wrote more detailed texts.

Furthermore, HAP produced more accurate texts than LAI and HAI. This is in line with the tenets of the Sociocultural Theory, which explicates that learning is the product of meaningful social interactions among the community members (Lantolf & Pohner, 2014). This finding is also in line with the findings of Davison (2024), Peng (2024), Mujtaba (2021), Villarreal and Gil-Sarratea (2020), Mazdayasna and Zaini (2015), Storch (2005), Dobao (2012), and Tavakoli and Rezazadeh (2014). This finding can be related to the assumption of growth in the development (de Bot, 2008). Accordingly, growth in the developmental process occurs through resources (e.g., time, motivation, feedback, and attention) which are all limited. Feedback and motivation can be created through collaboration. Thus, individuals who engage in collaborative performances can use resources in their collaborative endeavors; they produce texts with higher quality (e.g., more accurate texts) than those writing individually. On the other hand, this finding is consistent with the Output Hypothesis (Swain, 1995), claiming that interaction can improve grammatical performance. Moreover, as Storch (2019) pointed out, collaboration provides the learners with the opportunity to discuss, criticize, and explain the language used and offer more accurate texts. However, this finding is inconsistent with that of Shehadeh's (2011) study which showed that collaborative writing groups did not outperform individual writing groups in terms of accuracy. This difference can be accounted for by referring to the difference in the level of proficiency of learners in the studies. In Shehadeh's (2011) study, learners were at a low proficiency level and could not assist each other in terms of accuracy.

On the other hand, the reason why LAP did not outperform HAI in terms of accuracy may be related to the learners' level of proficiency. In line with Shehadeh's (2011) idea, LAP could not assist each other with grammatical accuracy due to their low proficiency level in English.

With regard to the development of learners in terms of complexity, the learners writing individually (i.e., LAI and HAI) outperformed the pairs. This finding is consistent with the cross-sectional study of Peng (2024) and Villarreal and Gil-Sarratea (2020). However, Storch (2005) noticed that learners composing collaboratively had more complex productions than learners composing individually. This discrepancy in findings may be due to two reasons. First, Storch (ibid) examined ESL learners

while in this study, EFL learners are studied. Second, in Storch's (ibid) cross-sectional study, descriptive statistics (and no inferential statistics) were provided.

Considering the participants' writing proficiency level, in each of the CAF measures, more stable patterns of development were observed in the productions of individuals and pairs with a higher proficiency level. However, the performances of the individuals and pairs with low levels of proficiency showed more regressions and progressions. Thus, learners' proficiency level may be a factor affecting the degree of variability in developmental paths. This finding is consistent with the findings of Spoelman and Verspoor (2010), showing that the development of complexity in advanced-level texts was relatively stable over time. This finding can also be interpreted in light of Dynamic Systems Theory which claims that the development of learners is affected by their initial state (van Dijk et al., 2024). On the other hand, higher proficiency learners who wrote individually and in pairs produced more accurate and more complex texts in most tasks than their lower proficiency counterparts. In addition, the higher proficiency pair produced more fluent texts than the lower proficiency pair. This finding is consistent with Liu et al. (2025) and Larsen-Freeman's (2006) studies which showed that learners' writing proficiency measures were boosted with the improvement in learners' proficiency level, and with Dao and McDonough (2018), Namkung and Kim (2024), Zabihi and Ghahramanzadeh,(2022), and Qiu and Lo's (2017) studies which showed that higher proficiency learners deal with more language engagement. Moreover, cross-sectional studies showed that learners' level of proficiency was effective in increasing the accuracy (Soleimani & Mahdavipour, 2014), syntactic complexity (Kim et al., 2024), lexical complexity (Liu et al., 2025), accuracy and fluency (Sánchez & Sunesson, 2023); Panagopoulos, 2024), and fluency (Panagopoulos, 2024).

Conclusion

This study was a novel investigation of how and when individuals and pairs with low and high levels of language proficiency developed in terms of writing proficiency measures over time. The participants had unique developmental paths with oscillations. Even the participants writing in similar contexts (those writing individually or those writing in pairs) showed different developmental patterns. Moreover, trade-offs between the components of writing proficiency at each data point showed that neither individuals nor pairs could focus on all aspects of the language simultaneously. In addition, the participants in the pair writing groups produced more accurate and more complex texts than those writing individually over time. However, the participants who wrote individually outperformed in terms of fluency. Finally, it was shown that proficiency could be considered a factor affecting the participants' different developmental paths in CAF measures.

Several theoretical and pedagogical implications may be derived from the findings of the study. From a theoretical perspective, this study was an examination of the Sociocultural Theory (Vygotsky, 1978), Dynamic Systems Theory (Thelen & Smith, 1994), and Trade-Off Hypothesis (Skehan, 2009). From a pedagogical perspective, the findings provided further empirical evidence of the uniqueness of the developmental paths of each learner and the usefulness of both collaborative and individual writing to enhance EFL learners' writing. Material developers and teachers can design both individual and collaborative writing activities and tasks to facilitate learning to write. They should also be informed of the trade-offs between the components of writing proficiency to devote separate sections to practice each of the CAF measures.

This study suffers from some limitations. Although this qualitative investigation is a case study, more cases could be examined. In addition, the performance of each case in each task could be compared with the performance of a group of learners.

Some suggestions for interested researchers in this field are provided. More studies can compare the development process in individual and collaborative groups of different proficiency-level learners. Second, further studies may compare learners composing collaboratively and those composing in groups with more members in terms of the development of writing proficiency. Finally, the development of CAF can be tracked in the learners' oral productions.

References

Akbarzadeh, R., Saeidi, M., & Chehreh, M. (2014). The effect of oral interactive feedback on the accuracy and complexity of EFL learners' writing performance: Uptake and retention. *Iranian Journal of Language Teaching Research, 2*(2), 105-126. https://doi.org/10.30466/ijltr.2014.20416

Ajideh, P., Leitner, G. & Yazdi-Amirkhiz, S. Y. (2016). The influence of collaboration on individual writing quality: The case of Iranian vs. Malaysian college students. *Journal of English Language Teaching and Learning, 17*, 1-24.

Ameri-Golestan, A. & Dousti, E. (2015). Collaborative vs. individual task planning and Iranian EFL learners' writing performance. *International Journal of Foreign Language Teaching and Research, 12*(3), 25-40.

Ary, D., Jacobs, L. C. & Sorensen, C. (2010). *Introduction to research in education* (8th ed.). Thomson.

Baba, K. & Nitta, R. (2014). Phase transitions in development of writing fluency from a complex dynamic systems perspective. *Language Learning, 64*(1), 1–35. https://doi.org/10.1111/lang.12033

Beers, S. F., & Nagy, W. E. (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? *Reading and Writing, 22*(2), 185-200.

Bot, K., D., Lowie, W. & Verspoor, M., H. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and cognition, 10*(1), 7–21. https://doi.org/10.1017/S1366728906002732

Cao, Z. (2021). Improving young learners' L2 writing accuracy with written corrective feedback: A case study. *English Teaching and Learning, 45*(4), 375-376. https://doi.org/10.1007/s42321-020-00071-1

Christiansen, M. H., & Arnon, I. (2017). More than words: The role of multiword sequences in language learning and use. *Topics in Cognitive Science, 9*(3), 542–551. https://doi.org/10.1111/tops.12274

Crossley, S. A. & McNamara, D. S. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing, 26*, 66–79. https://doi.org/10.32601/ejal.460961

Dao, P., & McDonough, K. (2018). Effect of proficiency on Vietnamese EFL learners' engagement in peer interaction. *International Journal of Educational Research, 88,* 60–72. http://dx.doi.org/10.1016/j.ijer.2018.01.008

Davison, I. (2024). The effects of completing collaborative or independent writing on the development of language use in individual writing. *Journal of Second Language Writing, 65*, 1-13. https://doi.org/10.1016/j.jslw.2024.101128

de Bot, K. (2008). Second language development as a dynamic process. *The Modern Language Journal, 92*(2), 166–178. https://doi.org/10.1111/j.1540- 4781.2008.00712.x

Dobao, A. F. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing, 21*, 40–58. https://doi.org/10.1016/j.jslw.2011.12.002

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics, 30*(4), 474–509. https://doi.org/10.1093/applin/amp042

Ellis, R. & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, *26*, 59–84. https://doi.org/10.1017/S0272263104026130

Housen, A. & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics, 30*(4), 461–473. https://doi.org/10.1093/applin/amp048

Housen, A., Kuiken, F. & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. *Language Learning and Language Teaching, 32*, 1-20.

Hsu, H. (2025). The effect of collaborative prewriting on L2 collaborative writing production and individual L2 writing development. *International Review of Applied Linguistics in Language Teaching, 63*(1), 499-533. https://doi.org/10.1515/iral-2023-0043

Kim, M., Kim, Y., & Kang, S. (2024). The dynamics of changes in linguistic complexity and writing scores in timed argumentative writing among beginning-level EFL learners. *International Review of Applied Linguistics in Language Teaching,* 1-35. https://doi.org/10.1515/iral-2023-0287

Lantolf, J. P. & Poehner, M. E. (2014). *Sociocultural theory and the pedagogical imperative in L2 education.* Routledge.

Larsen–Freeman, D. & Cameron, L. (2008). Research methodology on language development from a complex systems perspective. *The Modern Language Journal, 92*, 200-213. https://doi.org/10.1111/j.1540-4781.2008.00714.x

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics, 18,* 141–165. https://doi.org/10.1093/applin/18.2.141

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics, 27*(4), 590-619. https://doi.org/10.1093/applin/aml029

Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics, 30*(4), 579–589. https://doi.org/10.1093/applin/amp043

Larsen-Freeman, L. (2020). Complex dynamic systems theory. In B. VanPatten, G. Keating, & S. Wulff (Eds.) *Theories in second language acquisition* (pp. 248–270). Routledge. https://doi.org/10.4324/9780429503986-11.

Leeser, M. J. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research, 8,* 1, 55–81. https://doi.org/10.1191/1362168804lr137oa

Li, S. (2014). The interface between feedback type, L2 proficiency, and the nature of the linguistic target. *Language Teaching Research, 18*, 373–396. https://doi.org/10.1177/1362168813510384

Liu, J., Wu, M. & Liu, H. (2025). Investigating the dynamic relationship of lexical and syntactic complexity in L2 writing across proficiency levels: A CDST-inspired study. *System, 131*, 103678. https://doi.org/10.1016/j.system.2025.103678

Long, M. H. (1983). Linguistic and conversational adjustments to non-native speakers. *Studies in Second Language Acquisition, 5,* 177–194. http://dx.doi.org/10.1017/S0272263100004848

Lu, X. (2011). A Corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly, 45*(1), 36-62. https://doi.org/10.5054/tq.2011.240859

Mazdayasna, G. & Zaini, A. (2015). The effect of collaborative prewriting discussions on L2 writing development and learners' identity construction. *Iranian Journal of Applied Linguistics, 18*(2), 141-164. https://doi.org/10.18869/acadpub.ijal.18.2.141

Mujtaba, S. M., Reynolds, B. L., Parkash, R., & Singh, M. K. M. (2021). Individual and collaborative processing of written corrective feedback affects second language writing accuracy and revision. *Assessing Writing, 50*, 100566. http://dx.doi.org/10.1016/j.asw.2021.100566

Namkung, Y. & Kim, Y. (2024). Learner engagement in collaborative writing: The effects of SCMC mode, interlocutor familiarity, L2 proficiency, and task repetition. *System, 121*, 103251. https://doi.org/10.1016/j.system.2024.103251

Niu, R., Jiang, L., & Deng, Y. (2018). Effect of proficiency pairing on L2 learners' language learning and scaffolding in collaborative writing. *Asia Pacific Education Researcher, 27*, 187–195. http://dx.doi.org/10.1007/s40299-018-0377-2

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30*(4), 555–578. https://doi.org/10.1093/applin/amp044

Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics, 30*(4), 590–601. https://doi.org/10.1093/applin/amp045

Panagopoulos, P., Andria., Mikros, G. & Varlokosta, S. (2024). Writing in L2 Greek: Exploring the effect of L2 proficiency and learning context on complexity, accuracy, and fluency. *Journal of Second Language Writing, 64*, 101111. https://doi.org/10.1016/j.jslw.2024.101111.

Peng, K., X. (2024). Beyond accuracy gains: Investigating the impact of individual and collaborative feedback processing on L2 writing development. *Assessing Writing, 61*, 100876. https://doi.org/10.1016/j.asw.2024.100876

Polat, B. & Kim, Y. (2014). Dynamics of complexity and accuracy: A longitudinal case study of advanced untutored development. *Applied Linguistics, 35*(2), 184–207. https://doi.org/10.1093/applin/amt013

Qiu, X., & Lo, Y. Y. (2017). Content familiarity, task repetition and Chinese EFL learners' engagement in second language use. *Language Teaching Research, 21*(6), 681–698. https://doi.org/10.1177/1362168816684368

Sánchez, L. & Sunesson, M. (2023). Grasping the effects of storyline complexity, task structure and proficiency in narrative writing performance. *Journal of Second Language Writing, 60*, 100986. https://doi.org/10.1016/j.jslw.2023.100986.

Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing, 20*, 286–305. https://doi.org/10.1016/j.jslw.2011.05.010

Skehan, P. (2009). Modeling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510–532. https://doi.org/10.1093/applin/amp047

Soleimani, H. & Mahdavipour, M. (2014). The effect of variations in integrated writing tasks and proficiency level on features of written discourse generated by Iranian EFL learners. *The Journal of Teaching Language Skills, 6*(2), 131-159. https://doi.org/10.22099/jtls.2014.2060

Spoelman, M. & Verspoor, M. (2010). Dynamic patterns in development of accuracy and complexity: A longitudinal case study in the acquisition of Finnish. *Applied Linguistics, 31*(4), 532–553. https://doi.org/10.1093/applin/amq001

Storch, N. & Aldosari, A. (2012). Pairing learners in pair work activity. *Language Teaching Research, 17*(1), 31– 48. https://doi.org/10.1177/1362168812457530

Storch, N. (2005). Collaborative writing: Product, process and students' reflections. *Journal of Second Language Writing, 14*, 153–173. https://doi.org/10.1016/j.jslw.2005.05.002

Storch, N. (2019). Collaborative writing. *Language Teaching, 52*(1), 40–59.

Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *Canadian Modern Language Review, 50*(1), 158–164. https://doi.org/10.1017/S0261444818000320

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (eds.). *Principles and practice in applied linguistics: Studies in honor of H., G., Widdowson* (pp.125-144). Oxford University Press.

Tavakoli, M. & Rezazadeh, M. (2014). Individual and collaborative planning conditions: Effects on fluency, complexity and accuracy in L2 argumentative writing. *The Journal of Teaching Language Skills, 5*(4), 85-110. https://doi.org/10.22099/jtls.2014.1857

Thelen, E. & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. MIT Press. https://doi.org/10.7551/mitpress/2524.001.0001

van Dijk, M., & van Geert, P. (2007). Wobbles, humps and sudden jumps: A case study of continuity, discontinuity and variability in early language development. *Infant and Child Development, 16*, 7–33. https://doi.org/10.1002/icd.506

van Dijk, M., Lowie, W., Smit, N., Verspoor, M., & van Geert, P. (2024). Complex dynamic systems theory as a foundation for process-oriented research on second language development. *Second Language Research,* 1-15. https://doi.org/10.1177/02676583241246739.

Villarreal, I. & Gil-Sarratea, N. (2020). The effect of collaborative writing in an EFL secondary setting. *Language Teaching Research*, 24(6), 874-897. https://doi.org/10.1177/1362168819829017

Vyatkina, N. (2012). The development of second language writing complexity in groups and individuals: A longitudinal learner corpus study. *The Modern Language Journal, 96*(4), 576-598. https://doi.org/10.1111/j.1540-4781.2012.01401.x

Vygotsky, L. S. (1978). *Mind in society: The development of higher mental process*. Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

Watanabe, Y. & Swain, M. (2008). Perception of learner proficiency: Its impact on the interaction between an ESL learner and her higher and lower proficiency partners. *Language Awareness, 17*(2), 115-130. https://doi.org/10.1080/09658410802146651

Wigglesworth, G. & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, *26*(3) 445-466. https://doi.org/10.1177/0265532209104670

Zabihi, R. & Ghahramanzadeh, B. (2022). Proficiency pairing and engagement with language in L2 peer interaction. *System, 105*, 102725. https://doi.org/10.1016/j.system.2022.102725

Zhang, M. (2018). Collaborative writing in the EFL classroom: The effects of L1 and L2 use. *System, 76*, 1–12. https://doi.org/10.1016/j.system.2018.04.009

Sahar Zahed Alavi holds a Ph.D. in Teaching English as a Foreign Language (TEFL) and is currently an Assistant Professor of TESL in the Department of Foreign Languages at the University of Bojnord, Iran. She has authored multiple scholarly articles and books, with research interests focusing on second language writing and language testing.