e-ISSN 2280-6792

EL.LE
Vol. 15 – Num. 2 – Luglio 2026

Complexity, Accuracy, and Fluency in EFL Oral Performance of Italian Learners with Specific Learning Difficulties
Insights from the Neurodivergent Learner Corpus

Michele Daloiso
Università degli Studi di Parma, Italia

Giulia Govoni
Istituto Istruzione Superiore 'Bertrand Russell', Italia

Abstract This study investigates the complexity, accuracy, and fluency of English oral productions by 32 Italian university students with Specific Learning Difficulties (SpLDs), using data from the developing Neurodivergent Learner Corpus. Results reveal performance patterns associated with the type of course attended (intensive vs. extensive) and the type of learning difficulty. Learners with multiple SpLDs exhibited greater processing demands, reflected in longer pauses, lower grammatical accuracy, but also more successful self-corrections. By contrast, students with dyslexia spoke more fluently and maintained better grammatical control, despite less efficient self-monitoring. Findings suggest that multiple SpLDs might impose broader processing challenges affecting both fluency and accuracy. While the results are not generalisable, the study indicates that L2 difficulties in SpLD learners go beyond reading and writing and highlights the need for further corpus-based research on language teaching and learning to neurodivergent learners.

Keywords Specific learning difficulties (SpLDs). Dyslexia. Learner corpus research. English as a foreign language (EFL). Instructed second language acquisition.

Edizioni Ca'Foscari

Check for updates

Peer review

Submitted 2026-04-03
Accepted 2026-05-23
Published 2026-07-14

Open access

DOI 10.30687/ELLE/2280-6792/2026/02/001

1 Introduction

In recent years, the intersection of Educational Linguistics and inclusive education has attracted growing attention, particularly concerning English language learners with Specific Learning Difficulties (SpLDs), as they face distinct challenges in developing communicative competence.[1] While prior research has focused on L2 reading and writing difficulties in younger learner populations, less is known about how these learners perform in spoken English at the tertiary level, where linguistic and cognitive demands are intensified.

This study addresses a critical gap in the literature by examining complexity, accuracy and fluency (CAF) in the oral interlanguage of Italian-speaking English learners with SpLDs at the tertiary level. Drawing on data from the Neurodivergent Learner Corpus, a recently developed resource of transcribed and annotated oral productions, this research explores how different instructional formats and learner profiles shape L2 spoken performance. The study starts from the assumption that the CAF framework could provide new insights into the interlanguage of neurodivergent learners, as it captures not only the linguistic features but also the cognitive dynamics of L2 learner speech.

Ultimately, the study aims to contribute to more data-informed language pedagogy by informing tailored instructional strategies that address the specific needs of neurodivergent EFL learners.

2 Literature Review

This section provides an overview of the theoretical underpinnings of the study to be presented in Section 3, with a particular focus on the research gaps that the study seeks to address.

2.1 EFL Learners with SpLDs

In recent decades, interest in language education for learners with SpLDs has grown across medical, pedagogical, and linguistic domains, including Second Language Acquisition (SLA) and Educational Linguistics. In particular, English as a Foreign Language (EFL) is becoming an area of special interest, driven by both the global presence of English in educational curricula and the high prevalence of SpLDs in mainstream classes. In Italy, for example, around 6%

1 Michele Daloiso authored Sections 2.1, 3, 4.2; Giulia Govoni authored Sections 2.2., 2.3, 4.1. Sections 1 and 5 was written by both the authors.

of students have a diagnosed SpLD (MIM 2024), and nearly 20,000 university students report such characteristics (ANVUR 2022).

Research in this area draws on two psycholinguistic models: the *Linguistic Coding Differences Hypothesis* (Sparks et al. 1989) and the *Orthographic Depth Hypothesis* (Katz, Feldman 1983). The former links interlanguage development to L1 abilities, suggesting that difficulties in phonology, morphosyntax, or pragmatics may transfer to L2, while the latter highlights differences in writing systems, with shallow orthographies like Italian favouring phonological decoding and deeper ones like English relying more on visual-orthographic processing. Accordingly, learners with phonological and lexical retrieval deficits are likely to face similar challenges in L2, which can be influenced by the features of the target language. For example, studies on Italian learners aged 9-14 report slower, less accurate English word reading (Palladino et al. 2013) as well as comprehension difficulties (Bonifacci et al. 2017).

Despite growing research in this area, some gaps remain. On the one hand, the linguistic profiles of Italian EFL learners with SpLDs are still underexplored. On the other hand, university students, who face higher academic demands, remain underrepresented in this research area, as most studies focus on younger learners and lack large-scale data across age group. Additionally, descriptions of learner errors usually draw on clinical studies of L1 acquisition, which may offer useful hypotheses but require empirical validation in L2 contexts. It is also worth noting that, while SpLDs are typically associated with literacy difficulties, SLA research (e.g., Nijakowska 2010; Cappelli, Noccetti 2023; Kormos, Smith 2023) suggests that they may extend to a broader range of EFL challenges, including oral communication. Therefore, a more comprehensive account of learners' oral interlanguage is needed alongside written skills.

2.2 L2 Oral Performance Analysis: the CAF Framework

In SLA research, L2 oral proficiency is commonly conceptualised through three core dimensions: complexity, accuracy, and fluency (CAF). Introduced in the 1970s as a means of quantifying oral competence, the CAF framework has evolved into a model encompassing linguistic, psycholinguistic, and cognitive aspects of performance. Within this framework, greater complexity and accuracy are typically associated with more advanced interlanguage, while fluency reflects more efficient access to linguistic knowledge (Michel 2017).

These three dimensions are recognised as distinct yet integral components for the comprehensive assessment of L2 oral proficiency. Importantly, each component may vary according to contextual

factors, such as task features or learner profiles. While a learner's focus on one dimension may temporarily constrain their performance in the others, the three are considered cognitively interdependent (Rausch 2016). Consequently, disaggregated analysis of the CAF dimensions is essential for capturing their unique contributions to interlanguage development.

Accuracy is generally considered the most clearly definable dimension (Thewissen, Anishchanka 2022), as it refers to conformity with target language norms (Bui, Skehan 2018). However, its analysis is methodologically complex, as error types and frequency depend on task design, learner focus, and proficiency (Skehan, Foster 2012). Therefore, both holistic and analytic measures have been proposed; the former provide an overall evaluation, considering error frequency and severity (Polio 1997), whereas the latter assess discrete features. This study adopts an analytic approach, examining self-corrections, error-free clauses, errors per 100 words, and target-like use of selected forms.

Fluency has been defined in various ways (Foster 2020). Definitions include factors such as the smoothness of speech under real-time constraints (Lennon 1990) and the use of compensation strategies to manage communication breakdowns (Tarone 1980). Interestingly, pauses and hesitations do not necessarily indicate disfluency, as they may support planning or discourse organisation (Dumont 2018). Following Ellis and Barkhuizen (2005), this study measures fluency through temporal variables (e.g., speech rate, pauses) and hesitation markers (e.g., repetitions, reformulations).

Complexity is the dimension that has generated greater debate, as it includes both cognitive and linguistic aspects.

> Cognitive complexity is a relative and subjective notion. It refers to the relative difficulty with which language elements are processed during L2 performance and L2 learning, as determined in part by the learners' individual backgrounds (e.g. their aptitude, motivation, stage of L2 development, L1 background). Linguistic complexity is an important component of cognitive complexity (or difficulty), but it does not coincide with it. Linguistic complexity is an objective [property] given, independent from the learner, which refers to the intrinsic formal or semantic-functional properties of L2 elements (e.g. forms, meanings, and form-meaning mappings) or to properties of (sub-) systems of L2 elements. (Housen et al. 2012, 4)

Most SLA research has focused on linguistic complexity, typically operationalised as a quantitative measure of the number and variety of elements in a linguistic system, and the interrelations among them (Kuiken 2023). Common indices of linguistic complexity include

measures of grammatical structure (e.g., mean length of utterance, subordination and coordination indices) and lexical characteristics (e.g., lexical diversity, density, and sophistication). The current study adopts this approach by analysing both grammatical complexity (e.g., extent of subordination, diversity of verb forms, mean number of verb arguments) and lexical complexity (type-token ratio). However, interpretation of these metrics must be approached cautiously, as this traditional perspective on complexity does not fully capture phraseological, morphological, or developmental aspects (Ortega 2003).

2.3 Research Gap

Since research on EFL learners with SpLDs typically addresses linguistic challenges in isolation rather than considering the interaction of complexity, accuracy, and fluency, integrating the CAF framework could provide a more comprehensive account of their L2 production.

Regarding complexity, SpLDs may affect both cognitive and linguistic domains. The well-known deficits in phonological short-term memory (Kastamoniti et al. 2018) can limit vocabulary acquisition, syntactic development, and lexical diversity. Moreover, difficulties in working memory and serial processing (Majerus, Cowan 2016) may further constrain the use of complex structures, resulting in less elaborated output.

Accuracy may also be compromised, as the memory limitations associated with SpLDs can hinder the internalisation of grammatical rules, leading to errors in morphology, word order, and lexical selection. Oral accuracy may likewise be affected by problems in lexical retrieval and grammatical encoding.

Finally, due to limited phonological memory and slower lexical access, L2 fluency may be disrupted by disfluencies (e.g., longer pauses, hesitations), which impede spontaneous speech. Overall, fluency in learners with SpLDs may reflect a continuous negotiation between linguistic competence, cognitive constraints, and situational factors.

3 Method

This section presents an overview of the research design that guided our study. Particular attention is given to the sample composition, the key variables under investigation, and the methodological approach adopted for data collection and analysis.

3.1 Research Aims and Rationale

The present study investigates the oral English proficiency of Italian university students with SpLDs by analysing interlanguage samples within the CAF framework, with the aim of examining performance differences between learners enrolled in an intensive instructional programme and those attending an extensive programme (see Section 3.3). As all participants were preparing for a B1-level English examination, the study provides insights into the extent to which targeted pedagogical interventions may support the oral L2 performance of learners with SpLDs.

In addition, the study explores the oral L2 features of learners with dyslexia and those presenting multiple SpLDs, contributing to a deeper understanding of the linguistic and cognitive processes involved in EFL acquisition among neurodivergent learners. Accordingly, the rationale of the study is not to compare learners with SpLDs against a presumed normative benchmark, but rather to investigate the specific dynamics of their interlanguage in order to better understand the internal coherence and variability of SpLD learner profiles.

The study is guided by the following research questions:

RQ1. What are the CAF profiles of students participating in intensive and extensive instructional programmes?

RQ2. What are the CAF profiles observed among learners with dyslexia and among learners with multiple SpLDs?

RQ3. What CAF patterns characterize the oral production of learners with dyslexia and of those with multiple SpLDs across intensive and extensive instructional formats?

3.2 The Neurodivergent Learner Corpus

The interlanguage samples analysed in this study were selected from the Neurodivergent Learner Corpus, a project under development at the University of Parma coordinated by Michele Daloiso, which aims to implement an electronic corpus derived from the systematic collection of large volumes of spoken English productions by Italian-speaking young adults with SpLDs. The overarching goal of the project is to address current gaps in research on EFL teaching to learners with SpLDs (see 2.1), providing linguistic data that extend beyond the decoding of written text and focus on informants of an age group (young adults) that has been less frequently studied.

At the time of writing, the prototype version of the corpus consists of 409 transcribed and annotated oral productions, totalling 210,912 tokens and 3,956 types, with an average of 516 tokens per text (for detailed information on the corpus construction, see Daloiso, Genduso et al. 2026). The data were collected using clinical elicitation (Corder 1978), a method that employs tasks capable of eliciting a wide range of communicative features, including usages, errors, and strategies, without starting from fixed assumptions.

The tasks used to build the corpus vary in their objectives and input types to elicit different aspects of the interlanguage system (e.g. nouns, tense and aspect). The interlanguage samples examined in this research all pertain to the same task, which aims to gather information on learners' ability to describe a static scene by locating objects or people within space and detailing their physical characteristics. Specifically, the task involves an image depicting a restaurant where a murder has taken place. Learners were asked to describe the image and identify the alleged murderer, providing justification for their choice.

3.3 Instructional Context

The tasks were administered to learners with SpLDs enrolled in specific English language courses at the University of Parma. These courses, tailored to support this learner population, are offered in two formats: an extensive and an intensive programme. Upon enrolment, students complete a placement test to determine their initial proficiency and are assigned to the appropriate instructional track.

The extensive programme consists of ten units, each delivered over a two-week period, while the intensive programme comprises five units, also spanning two weeks each. Both formats aim to foster communicative competence in English at the B1 level. At the end of the course, all students complete the same oral exam, which evaluates their ability to engage in spontaneous conversation on familiar topics, describe experiences, express opinions, and narrate events.

Each teaching unit follows a five-stage structure:

1. **Let's Start**: this phase introduces the topic through visual and verbal cues, prompting learners to formulate hypotheses about the unit's content.
2. **Explore**: here the instructor focuses on listening comprehension through an audio passage in English followed by a related comprehension activity.
3. **Think**: this phase encourages metacognitive reflection on the strategies used during the learning activities.

4. **Your Turn**: this phase encourages L2 oral production through an oral task which requires the description of an image or video in English.
5. **Look Back**: here the instructor and the learners summarise the content covered in the unit and consolidate learning.

The task analysed in this study was administered during phase 4. Learners were shown an image in real time, given two minutes to observe it, and then asked to record an audio description to be uploaded on the course platform.

3.4 Participants

The study involved 32 Italian-speaking students with SpLDs, selected from a larger cohort of the 90 learners who completed the task between 2020 and 2024. Participants ranged in age from 19 to 52 (70% under 24) and included 23 females and 9 males, all of whom provided informed consent.

Participants were initially grouped by instructional programme (intensive vs extensive). Since only 16 out of 90 original learners were in the extensive track, matched pairs were created based on age, SpLD subtype, and degree programme, using data from the participants' language biographies. By the end of the programme, all participants reached B1 proficiency, although those in the extensive track required more time and support.

The sample was later analysed by SpLD subtype, distinguishing learners with dyslexia only from those with multiple SpLDs (dyslexia, dysgraphia, and/or dysorthography). Seven participants were excluded for this further analysis due to missing biographies, resulting in 11 learners in the dyslexia group and 14 in the multiple-SpLD group.

3.5 Data Preparation and Analysis

As a first step, the recorded performances were transcribed using the EUDICO Linguistic Annotator (ELAN, Version 6.9; Max Planck Institute for Psycholinguistics, The Language Archive 2024), which offers a user-friendly platform for multimodal linguistic analysis, including multi-tier annotation, customisable schemes, and advanced playback control for detailed segment analysis.

In this study, the transcription followed the guidelines of the Trinity Lancaster Corpus (Gablasova et al. 2019) and included selected prosodic and paralinguistic features such as pauses and fillers. Learners' utterances were segmented into AS-units, defined

as "a single speaker's utterance consisting of an independent clause or subclausal unit, together with any subordinate clause(s) associated with it" (Foster et al. 2000, 365).

Following transcription, the data were annotated to enrich the corpus with interpretive linguistic information. Annotation followed an adapted version of the University of Louvain framework (Granger et al. 2023), which included the following error types: phonetic, lexical, grammatical, and word order. Phonetic errors were classified into consonantal, vowel, combined (both types in a single word), and stress errors. Lexical errors included incorrect, redundant, or omitted use of nouns, verbs, adjectives, adverbs, conjunctions, and connectors. Grammatical errors covered the misuse or omission of elements such as nouns, articles, verbs, adjectives, pronouns, and possessives, as well as misclassification of word class. Word order errors included typical learner mistakes such as post-nominal adjective placement (e.g., *a t-shirt red*), literal translations influenced by Italian syntax (e.g., *the table in yellow*), and omissions affecting sentence structure.

In the final stage of data preparation, individual participant data were organised into an Excel spreadsheet structured for statistical analysis.

As regards data analysis, the following dependent variables and indicators were analysed to assess the three CAF components **[tab. 1]**.

Table 1 Indicators used to assess EFL fluency, accuracy and complexity

CAF component	Indicator	Definition
Fluency	*Speech rate*	Number of words per minute, excluding disfluencies.
	Pause density	Any silence lasting 700 milliseconds or more, calculated as total pause length divided by total speaking time.
	Mean length of run	Average number of syllables produced between pauses.
	False starts	Incomplete utterances that are abandoned or restarted.
	Repetitions	Immediate reiteration of a word or phrase.
	Reformulations	Self-repair involving the rephrasing or paraphrasing of a previously spoken item.
	Replacements	Substitution of a word or structure with an alternative during speech.

CAF component	Indicator	Definition
Accuracy	*Self-corrections*	Number of accurate self-corrections.
	Error-free clauses	Number of error-free clauses divided by the total number of clauses.
	Error density	Errors divided by the total number of words produced by the learner.
	Lexical accuracy	Number of lexical errors relative to total words produced (excluding disfluencies).
	Verbal accuracy	Proportion of target-like verbal morphology, calculated by dividing correct finite verbs by total verb phrases.[1]
	Plural accuracy	Correct plural forms divided by obligatory plural contexts.[2]
Complexity	*Verb arguments per AS-unit*	Mean number of verb arguments per AS-unit, to capture complexity at the sentence level.
	Verb arguments per verb phrase	Mean number of verb arguments per verb phrase, to highlight the internal structure of verb phrases.
	Subordination density	Number of all subordinate clauses divided by total AS-units.
	Accurate subordination density	Number of correct subordinate clauses divided by total AS-units.
	Verb form diversity	Number of distinct verb forms divided by total verb phrases.
	Lexical diversity	Type-token ratio, i.e. the number of word types divided by the total number of word tokens.

1 To account for incomplete sentences lacking necessary verbs, such cases were included in the denominator.

2 Overgeneralised plural forms were also included to capture non-standard usage.

4 Results

This section reports the findings of our study, which should be interpreted descriptively, without implying statistical significance.

4.1 Findings by Intervention Programme

This section examines the oral performance outcomes of participants grouped by intervention programme. The first group (n = 16), referred to as the Intensive Group (InG), attended a 5-unit condensed course, while the second group (n = 16), referred to as the Extensive

Group (ExG), followed a 10-unit extensive course. Detailed statistical analysis is reported in Appendix 1.

4.1.1 Fluency

The analysis of temporal variables points to distinct fluency profiles across groups. InG learners showed a higher speech rate and lower pause density compared to ExG learners, alongside fewer pauses per unit. Mean length of run was slightly higher in the InG than in the ExG. These patterns suggest faster delivery and more continuous speech in the InG, contrasted with slower but potentially more segmented production in the ExG.

Hesitation phenomena further nuance this picture. While replacements were comparable, the InG displayed higher rates of self-corrections, reformulations, repetitions, and false starts. This may indicate more active online monitoring and repair, consistent with greater processing speed. Conversely, the lower frequency of such phenomena in the ExG may reflect a more cautious and pre-planned production style.

Overall, the data suggest a trade-off between speed and planning: InG learners appear to favour more automated production with frequent self-repairs, whereas ExG learners adopt a slower, more deliberate approach with fewer overt repairs. However, substantial within-group variability, as indicated by standard deviations (SDs), cautions against generalisation.

Such tendencies may relate to the timing of the task within the respective curricula, administered in Unit 2 of the intensive programme and Unit 4 of the extensive one, which could have influenced learners' readiness. Affective or sociocultural factors may also have contributed by shaping perceived communicative pressure or opportunities for planning, although these influences cannot be isolated in our dataset.

4.1.2 Accuracy

Linguistic accuracy revealed distinct patterns across groups. The ExG showed higher target-like use of tense and plural marking than the InG. Moreover, overgeneralisation was more frequent in the InG, alongside a slightly higher error rate.

Error frequency measures reinforced this trend, as the ExG produced fewer errors per 100 words and more error-free clauses than the InG, suggesting more stable integration of grammatical knowledge. Self-correction rates were similar, but more accurate in the ExG, indicating more effective monitoring.

Overall, the ExG displays a more consistent accuracy profile. Interestingly, these results align with accounts linking extended input exposure to more stable morphosyntactic representations (Ellis, Barkhuizen 2005). Moreover, lexical accuracy patterns are consistent with research suggesting distinct trajectories for lexical and grammatical development (Bui, Skehan 2018). Greater self-correction accuracy in the ExG may reflect enhanced metalinguistic awareness fostered by sustained exposure and feedback (Ellis et al. 2006; Lantolf, Thorne 2007).

However, these patterns must be interpreted cautiously, as high SDs and mean-median discrepancies indicate substantial internal variability. From an interlanguage perspective, such variability may reflect diverse learner profiles and strategies.

4.1.3 Complexity

Linguistic complexity also revealed distinct patterns across groups. The mean number of verb arguments per AS-unit was higher in the InG than in the ExG, suggesting greater clause-level elaboration. By contrast, verb arguments per verb phrase were comparable, which indicates similar internal syntactic density.

Lexical and morphological range were slightly higher in the InG, as shown by verb form diversity and lexical diversity. At the discourse level, however, the ExG produced more AS-units, while the InG showed higher subordination and greater subordinate clause accuracy.

Overall, these results point to different complexity profiles. The higher clause-level elaboration and subordination in the InG support the idea that intensive practice fosters syntactic complexity even over shorter periods (Ellis, Barkhuizen 2005). However, similar verb phrase density suggests stability in internal structure across conditions. The ExG's higher number of AS-units indicates broader discourse expansion, whereas the InG's denser subordination reflects more compact syntactic embedding.

The lower subordinate clause accuracy observed in the ExG is consistent with the trade-off between complexity and control (Kuiken, Vedder 2008): producing more complex structures may come at the expense of precision, particularly under higher cognitive load or specific instructional conditions. By contrast, the slightly higher lexical diversity in the InG may reflect either greater vocabulary development or differences in initial proficiency.

However, descriptive results for the complexity measures revealed substantial internal variability, with SDs frequently large relative to the means. In several cases (most notably accurate subordination) the median was considerably lower than the mean, indicating skewed

distributions driven by a few higher-performing learners. The InG demonstrated higher average subordination and lexical diversity, yet both groups displayed diverse complexity profiles, which could reflect learners' individual strengths and different strategies for producing complex language.

4.2 Findings by SpLD Subtype

This section discusses the potential impact of SpLD types on L2 oral production. The sample was subdivided into two groups based on diagnostic profiles: the first group (n = 11) comprised learners diagnosed with dyslexia (hereafter DysG), while the second group (n = 14) consisted of learners with multiple learning difficulties (i.e., diagnosed with at least two of the following: dyslexia, dysgraphia, and dysorthography; hereafter MixG). Learners with unspecified SpLD diagnoses were excluded from this analysis. Detailed statistical descriptions are provided in Appendix 2.

4.2.1 Fluency

The analysis of pausing behaviour revealed only limited differences between groups. The DysG produced slightly more pauses than the MixG, which is in line with findings that adults with dyslexia show more disruptions in spontaneous speech (Pérez-Litago et al. 2025). Speech rate was marginally higher in the DysG than in the MixG, while pause density was higher in the MixG, suggesting greater time devoted to pausing.

Hesitation phenomena showed broadly similar distributions. False starts were comparable, while repetitions, replacements, and self-corrections were slightly higher in the MixG, suggesting greater reliance on compensatory strategies. Reformulations were nearly identical, which indicates similar monitoring behaviours.

Overall, both groups displayed elevated pausing and hesitation, which are consistent with the well-known cognitive demands associated with SpLDs. The faster speech rate in the DysG may reflect a compensatory focus on message delivery, whereas the MixG's higher pause density and repair phenomena may be caused by increased processing load due to co-occurring difficulties. Comparable reformulation and self-correction rates suggest similar levels of metalinguistic awareness.

These findings align with previous research showing reduced fluency in dyslexic speakers (Wiseheart, Altmann 2018) but also extend this pattern to multiple SpLD profiles. However, one should consider that group averages are influenced by higher-performing

individuals, pointing to heterogeneous strategies in managing speech production.

4.2.2 Accuracy

Accuracy measures revealed partly overlapping but distinct profiles. While target-like tense use was comparable across groups, the DysG showed higher plural accuracy than the MixG, suggesting more consistent morphosyntactic control. However, the higher SD in the DysG indicates greater variability, which is in line with evidence that inflectional marking is particularly vulnerable in this population (Helland, Kaasa 2005). Lexical accuracy was low and similar in both groups.

Error density further differentiates the groups: DysG learners showed more stable grammatical control, as they produced fewer errors and more error-free clauses. In contrast, self-correction patterns point to stronger monitoring in MixG learners, who showed greater accuracy. Nonetheless, both groups remained below 50% success, indicating that self-monitoring skills are still developing.

Overall, these findings suggest a trade-off between grammatical control and monitoring. The DysG appears to achieve greater accuracy in production, whereas the MixG shows more effective self-monitoring. This pattern is consistent with accounts of cognitive constraints in L2 learners with SpLDs, as phonological memory limitations affect inflectional control (Daloiso 2017), while working-memory constraints impact lexical retrieval (Gathercole, Baddeley 1990).

However, these patterns should be interpreted with caution, given the substantial within-group variability discussed above.

4.2.3 Complexity

Linguistic complexity revealed partly divergent patterns across groups. The MixG showed slightly greater clause-level elaboration, producing more verb arguments per AS-unit than the DysG. However, argument complexity at the verb-phrase level was comparable, suggesting similar internal organisation of predicates.

Lexical and morphological range was marginally higher in the DysG, as indicated by greater type-token ratio, as well as a higher number of AS-units. By contrast, the MixG produced more subordinate clauses, showing a tendency toward greater syntactic embedding. Interestingly, despite producing fewer subordinate clauses, the DysG showed higher accuracy, indicating more stable syntactic control. This pattern is consistent with previous findings on dyslexic learners' L2 performance (Pérez-Litago et al. 2025; Helland,

Kaasa 2005); it also highlights the trade-off between complexity and control, as increased syntactic elaboration may come at the expense of accuracy under processing constraints.

Overall, the MixG appears to favour more elaborated but less controlled constructions, whereas the DysG produces somewhat less complex output with greater accuracy in the structures attempted. These tendencies may reflect different processing demands during real-time production.

5 Discussion and Implications

This study examined CAF patterns in the EFL oral performance of Italian-speaking university learners, considering both instructional type and SpLD subtype. Although results are not generalisable and no causal claims can be made, the findings suggest tentative trends that may inform future research.

RQ1. What are the CAF profiles of students participating in intensive and extensive instructional programmes?
Our findings point to a fluency-accuracy trade-off. The InG produced faster speech with fewer pauses but more false starts and self-corrections, showing less controlled output and lower grammatical accuracy. By contrast, the ExG showed greater accuracy with fewer errors and more error-free clauses, but longer pauses and less frequent repair, suggesting more deliberate production. One tentative interpretation is that intensive exposure may favour fluency gains, whereas extensive exposure supports accuracy and control. In terms of complexity, the InG displayed greater syntactic and lexical elaboration but lower subordination accuracy, while the ExG showed more controlled use of complex structures. These patterns suggest that instructional format may shape how learners allocate their resources across CAF dimensions.

RQ2. What are the CAF profiles observed among learners with dyslexia and among learners with multiple SpLDs?
Distinct profiles emerged across subtypes. MixG learners showed greater processing load, reflected in more frequent and longer pauses, increased self-corrections, and lower accuracy in morphology and subordination. By contrast, DysG learners produced faster and more accurate speech, though with less successful self-monitoring. This may suggest that multiple SpLDs impose greater cognitive demands, affecting both fluency and accuracy, whereas dyslexic learners may prioritise fluent production while maintaining relatively stable grammatical control. Variation within groups may reflect individual differences in cognitive and strategic profiles.

RQ3. What CAF patterns characterize the oral production of learners with dyslexia and of those with multiple SpLDs across intensive and extensive instructional formats?
Across both instructional conditions, DysG learners tended to prioritise fluency, producing faster speech with fewer pauses and relatively stable accuracy, but less effective self-monitoring. MixG learners showed less fluent production but more frequent and successful self-correction, indicating stronger monitoring skills. This suggests a potential interaction between learner profile and instructional format: extensive exposure may particularly support accuracy and monitoring in MixG learners, while DysG learners' fluency appears more stable across conditions, with more limited gains in control. Overall, CAF patterns appear to reflect a dynamic interplay between cognitive constraints and instructional factors.

Existing research on SpLDs rarely applies the CAF framework to L2 oral performance. This study addresses that gap, offering new insights into an underexplored area. While the CAF framework typically highlights trade-offs among its three dimensions, findings here suggest that, for learners with SpLDs, these trade-offs could be shaped not only by interlanguage development but also by cognitive constraints specific to this population. The InG demonstrated greater fluency and syntactic complexity but lower accuracy, whereas the ExG produced more accurate yet less fluent and structurally complex speech. These findings suggest that different types of instructional exposure may influence the interplay among fluency, accuracy, and complexity in distinct ways. Future research using a pre-test/post-test design could provide empirical evidence to examine this possibility more directly.

Distinct CAF patterns between learners with dyslexia and those with multiple SpLDs further suggest the impact of cognitive profiles on oral performance. Dyslexic learners produced faster, more consistent speech, while those with multiple SpLDs showed increased pausing and hesitation, likely due to greater working memory demands that hinder simultaneous control of accuracy and complexity.

Pedagogically, these results emphasise the need for flexible, differentiated instruction that addresses specific cognitive challenges in L2 acquisition among learners with SpLDs. In light of the present findings, and subject to confirmation through studies with larger samples, conducting a needs analysis prior to instructional planning appears important, as different learner profiles may require tailored pedagogical accommodations. In our sample, learners with dyslexia tended to prioritize meaning over form, showing relatively fewer difficulties with fluency but greater challenges in accuracy and, occasionally, complexity. For these learners, intensive instruction that emphasizes strategies for monitoring and self-correcting speech

production may be particularly beneficial. In contrast, learners with multiple SpLD profiles, who experience a broader range of difficulties, including challenges related to fluency, might benefit more from extensive instructional programmes that provide targeted support for L2 oral production planning, monitoring, and self-assessment.

Taken together, our results underscore the importance of adopting a balanced perspective when assessing learners with SpLDs. Although fluency issues may be the most immediately noticeable, instructors should also consider accuracy and complexity within the CAF framework to more fully capture learners' underlying L2 competence and progress.

For a correct interpretation of the findings, readers should consider that the present study is based on a relatively small, although carefully matched, sample. This inevitably limits the statistical power of the analyses and may not fully capture the variability within the broader SpLD population. Future research should therefore seek to increase the robustness of the findings by involving larger samples and more normally distributed data.

Further studies could also expand the pedagogical implications emerging from the present research. In particular, longitudinal designs would help investigate how CAF dimensions evolve over time in relation to learners' cognitive profiles and different forms of instructional exposure. Moreover, the integration of quantitative and qualitative approaches may provide deeper insights into learners' strategies, experiences and perceptions, thereby supporting the development of more targeted and inclusive instructional practices. Finally, future research might examine the impact of different task types on CAF performance, with the aim of informing more task-sensitive approaches to language teaching and assessment.

Appendix 1 Statistical Analysis by Intervention Programme

Table 1 Descriptives of variables related to the fluency dimension between intervention programme

	Intervention Programme	Mean	Median	SD	Minimum	Maximum
Pause density	**Extensive**	0.302	0.280	0.171	0.0476	0.560
	Intensive	0.269	0.213	0.159	0.0629	0.575
Speech rate	**Extensive**	69.803	57.197	30.410	30.5202	149.643
	Intensive	73.123	76.474	22.003	31.3725	105.169
Number of pauses	**Extensive**	39.750	38.000	21.807	10	79
	Intensive	35.438	29.000	25.414	6	99
Replacements	**Extensive**	1.250	1.000	1.390	0	4
	Intensive	1.500	1.000	1.265	0	4
False starts	**Extensive**	7.063	5.000	5.272	1	18
	Intensive	8.875	8.500	5.045	1	21
Reformulations	**Extensive**	2.625	2.000	1.668	0	6
	Intensive	5.000	4.000	4.211	0	14
Number of self-corrections	**Extensive**	4.313	3.500	2.938	0	11
	Intensive	6.875	6.000	4.703	0	17
MLR	**Extensive**	8.247	4.381	9.500	1.2394	38.091
	Intensive	8.856	6.896	6.757	1.9100	26.778
Repetitions	**Extensive**	5.500	4.000	5.598	0	20
	Intensive	8.750	7.500	5.335	3	23

Table 2 Descriptives of variables related to the accuracy dimension between intervention programme

	Intervention Programme	Mean	Median	SD	Minimum	Maximum
Self-correction density	**Extensive**	0.1153	0.1000	0.0878	0.00000	0.308
	Intensive	0.1080	0.1152	0.0639	0.00000	0.211
Accurate self-correction density	**Extensive**	0.4596	0.3667	0.3337	0.00000	1.000
	Intensive	0.3764	0.3333	0.2656	0.00000	0.857
Error-free clauses	**Extensive**	0.2306	0.2550	0.1537	0.00000	0.500
	Intensive	0.1588	0.1800	0.1549	0.00000	0.620
Error density	**Extensive**	0.2451	0.2066	0.1183	0.08012	0.455
	Intensive	0.3434	0.3372	0.1132	0.15017	0.586
Lexical accuracy	**Extensive**	0.0831	0.0960	0.0434	0.00794	0.158
	Intensive	0.1069	0.1054	0.0450	0.03101	0.167
Verbal accuracy	**Extensive**	0.7688	0.8400	0.1964	0.23000	1.000
	Intensive	0.6794	0.7350	0.2432	0.00000	0.930

	Intervention Programme	Mean	Median	SD	Minimum	Maximum
Plural accuracy	Extensive	0.7636	0.9087	0.2800	0.28571	1.000
	Intensive	0.6147	0.6667	0.2792	0.00000	1.000

Table 3 Descriptives of variables related to the complexity dimension between intervention programme

	Intervention Programme	Mean	Median	SD	Minimum	Maximum
Subordination density	**Extensive**	0.132	0.1043	0.127	0.0000	0.400
	Intensive	0.281	0.2250	0.170	0.0714	0.714
Accurate subordination density	**Extensive**	0.175	0.0000	0.250	0.0000	0.667
	Intensive	0.278	0.0500	0.365	0.0000	1.000
Total AS-units	**Extensive**	17.000	16.0000	6.938	7	30
	Intensive	14.563	14.5000	5.796	5	26
TTR	**Extensive**	2.208	2.1502	0.287	1.6098	2.619
	Intensive	2.447	2.4779	0.470	1.3571	3.529
Verb form diversity	**Extensive**	0.420	0.4037	0.141	0.1538	0.714
	Intensive	0.456	0.4393	0.109	0.3125	0.647
Average VA for VP	**Extensive**	1.906	2.0000	0.259	1.3750	2.308
	Intensive	1.856	1.8062	0.561	1.0000	3.071
Average VA for AS unit	**Extensive**	2.333	2.2792	0.450	1.3750	3.000
	Intensive	2.741	2.6389	0.978	1.2632	5.111

Appendix 2 Statistical Analysis by SpLDs Subtype

Table 4 Descriptives for variables related to the fluency dimension between SpLD types

	SpLD_Type	Mean	Median	SD	Minimum	Maximum
Pause density	**Dyslexia**	0.260	0.214	0.145	0.0750	0.560
	Mixed SpLDs	0.291	0.237	0.178	0.0476	0.541
Speech rate	**Dyslexia**	74.125	58.868	34.015	36.0377	149.643
	Mixed SpLDs	72.588	70.809	19.129	49.8113	103.059
Replacements	**Dyslexia**	1.273	1	1.489	0	4
	Mixed SpLDs	1.643	1.500	1.277	0	4
False starts	**Dyslexia**	7.545	6	6.299	1	21
	Mixed SpLDs	8.214	8.000	4.336	1	18
Reformulations	**Dyslexia**	3.727	3	3.797	0	14
	Mixed SpLDs	3.857	3.500	2.248	0	9
Number of self-corrections	**Dyslexia**	5.455	4	4.824	0	16
	Mixed SpLDs	6.000	6.000	2.449	1	10
MLR	**Dyslexia**	10.412	4.184	11.813	1.8250	38.091
	Mixed SpLDs	7.863	5.824	5.731	2.7500	21.500
Repetitions	**Dyslexia**	5.727	4	4.901	0	18
	Mixed SpLDs	7.929	6.500	6.330	1	23

Table 5 Descriptives of variables related to the accuracy dimension between SpLD types

	SpLD_Type	Mean	Median	SD	Minimum	Maximum
Self-correction density	**Dyslexia**	0.1232	0.0952	0.1006	0.00000	0.308
	Mixed SpLDs	0.1123	0.1010	0.0544	0.04762	0.240
Accurate self-correction density	**Dyslexia**	0.3754	0.3125	0.3483	0.00000	1.000
	Mixed SpLDs	0.4910	0.3842	0.2255	0.00000	0.750
Error-free clauses	**Dyslexia**	0.2327	0.2000	0.1661	0.00000	0.500
	Mixed SpLDs	0.1443	0.1300	0.1284	0.00000	0.400
Error density	**Dyslexia**	0.2590	0.2211	0.1397	0.08012	0.455
	Mixed SpLDs	0.3188	0.3241	0.1217	0.13636	0.586
Lexical accuracy	**Dyslexia**	0.0861	0.0988	0.0515	0.00794	0.158
	Mixed SpLDs	0.1083	0.1016	0.0398	0.03810	0.167
Verbal accuracy	**Dyslexia**	0.7155	0.7700	0.2221	0.23000	1.000
	Mixed SpLDs	0.7050	0.8100	0.2470	0.00000	0.920
Plural accuracy	**Dyslexia**	0.7076	0.8889	0.3544	0.00000	1.000
	Mixed SpLDs	0.6422	0.6458	0.2747	0.28571	1.000

Table 6 Descriptives for variables related to the complexity dimension between SpLD types

	SpLD_Type	Mean	Median	SD	Minimum	Maximum
Subordination density	**Dyslexia**	0.173	0.158	0.131	0.000	0.364
	Mixed SpLDs	0.204	0.177	0.153	0.000	0.500
Accurate subordination density	**Dyslexia**	0.247	0.000	0.333	0.000	1.000
	Mixed SpLDs	0.174	0.000	0.256	0.000	0.667
Total AS-units	**Dyslexia**	16.455	15	6.089	10	30
	Mixed SpLDs	16.000	16.500	4.756	9	26
TTR	**Dyslexia**	2.328	2.414	0.267	2.000	2.705
	Mixed SpLDs	2.480	2.508	0.426	1.946	3.529
Verb form diversity	**Dyslexia**	0.421	0.450	0.121	0.154	0.538
	Mixed SpLDs	0.416	0.383	0.118	0.267	0.714
Average VA for VP	**Dyslexia**	1.849	2.000	0.393	1.000	2.333
	Mixed SpLDs	1.911	1.876	0.398	1.000	2.875
Average VA for AS unit	**Dyslexia**	2.372	2.167	0.705	1.375	3.636
	Mixed SpLDs	2.602	2.404	0.902	1.263	5.111

Bibliography

ANVUR (2022). *Gli studenti con disabilità e DSA nelle università italiane. Una risorsa da valorizzare*. https://www.anvur.it/wp-content/uploads/2022/06/Gli-studenti-con-disabilita%CC%80-e-DSA-nelle-universita%CC%80-italiane-Una-risorsa-da-valorizzare-ANVUR.pdf.

Bonifacci, P.; Canducci, E. et al. (2017). "English as a Foreign Language in Bilingual Language-Minority Children, Children with Dyslexia and Monolingual Typical Readers". *Dyslexia*, 23(2), 181-206. https://doi.org/10.1002/dys.1553.

Bui, G.; Skehan, P. (2018). "Complexity, Accuracy, and Fluency". *The TESOL Encyclopaedia of English Language Teaching*, 1-7.

Cappelli, G.; Noccetti, S. (eds) (2022). *A Linguistic Approach to the Study of Dyslexia*. Bristol: Multilingual Matters.

Corder, S.P. (1978). "Language Learner Language". Richards, J. (ed.), *Understanding Second and Foreign Language Learning: Issues and Approaches.* Rowley, MA: Newbury House, 71-92.

Daloiso, M. (2017). *Supporting Learners with Dyslexia in the ELT Classroom*. Oxford: OUP.

Daloiso, M.; Genduso, M. et al. (2026). "The Neurodivergent Learner Corpus: Development, Overview and Applications". *Mosaic*, 16(1), 91-108.

Dumont, A. (2018). *Fluency and Disfluency: A Corpus Study of Non-Native and Native Speaker (Dis)fluency Profiles* [PhD Dissertation]. Louvain-la-Neuve: UCL-Université Catholique de Louvain.

Ellis, R.; Barkhuizen, G. (2005). *Analysing Learner Language*. Oxford: OUP.

Ellis, R.; Loewen, S.; Erlam, R. (2006). "Implicit and Explicit Corrective Feedback and the Acquisition of L2 Grammar". *Studies in Second Language Acquisition*, 28(2), 339-68. https://doi.org/10.1017/S0272263106060141.

Foster, P. (2020). "Oral Fluency in a Second Language: A Research Agenda for the Next Ten Years". *Language Teaching*, 53(4), 446-61. https://doi.org/10.1017/S026144482000018X.

Foster, P.; Tonkyn, A.; Wigglesworth, G. (2000). "Measuring Spoken Language: A Unit for All Reasons". *Applied Linguistics*, 21(3), 354-75. https://doi.org/10.1093/applin/21.3.354.

Gablasova, D.; Brezina, V.; McEnery, T. (2019). "The Trinity Lancaster Corpus: Development, Description and Application". *International Journal of Learner Corpus Research*, 5(2), 126-58. https://doi.org/10.1075/scl.92.02gab.

Gathercole, S.E.; Baddeley, A.D. (1990). "Phonological Memory Deficits in Language Disordered Children: Is There a Causal Connection?". *Journal of Memory and Language*, 29(3), 336-60. https://psycnet.apa.org/doi/10.1016/0749-596X(90)90004-J.

Granger, S.; Paquot, M. et al. (2023). *UCLouvain Error Editor. Version 2* (UCLEEv2). Louvain: Centre for English Corpus Linguistics. http://hdl.handle.net/20.500.12279/968.

Helland, T.; Kaasa, R. (2005). "Dyslexia in English as a Second Language". *Dyslexia*, 11(1), 41-60. https://doi.org/10.1002/dys.286.

Kastamoniti, A.; Tsattalios, K. et al. (2018). "The Role of Phonological Memory in Reading Acquisition and Dyslexia: A Systematic Literature Review". *European Journal of Special Education Research*, 3(4), 278-323. https://doi.org/10.5281/zenodo.1560698.

Katz, L.; Feldman, L.B. (1983). "Relation Between Pronunciation and Recognition of Printed Words in Deep and Shallow Orthographies". *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, 9(1), 157-66. https://doi.org/10.1037//0278-7393.9.1.157.

Kormos, J.; Smith, A.M. (2023). *Teaching Languages to Students with Specific Learning Differences*. 2nd ed. Bristol: Multilingual Matters.

Kuiken, F. (2023). "Linguistic Complexity in Second Language Scquisition". *Linguistics Vanguard*, 9(s1), 83-93. https://doi.org/10.1515/lingvan-2021-0112.

Kuiken, F.; Vedder, I. (2008). "Cognitive Task Complexity and Written Output in Italian and French as a Foreign Language". *Journal of Second Language Writing*, 17(1), 48-60. https://doi.org/10.1016/j.jslw.2007.08.003.

Lantolf, J.P.; Thorne, S.L. (2007). *Sociocultural Theory and the Genesis of Second Language Development.* Oxford: OUP.

Lennon, P. (1990). "Investigating Fluency in EFL: A Quantitative Approach". *Language* Learning, 40(3), 387-417. https://doi.org/10.1111/j.1467-1770.1990.tb00669.x.

Majerus, S.; Cowan, N. (2016). "The Nature of Verbal Short-Term Impairment in Dyslexia: The Importance of Serial Order". *Frontiers in Psychology*, 7, 1-8. https://doi.org/10.3389/fpsyg.2016.01522.

Michel, M. (2017). "Complexity, Accuracy, and Fluency in L2 Production". Loewen, S; Sato, M. (eds), *The Routledge Handbook of Instructed Second Language Acquisition*. London: Routledge.

MIM (2024). *I principali dati relativi agli alunni con DSA. Aa.ss. 2021/2022 e 2022/2023*. https://www.mim.gov.it/pubblicazioni/-/asset_publisher/6Ya1FS4E4QJw/content/i-principali-dati-relativi-agli-alunni-con-dsa.

Nijakowska, J. (2010). *Dyslexia in the Foreign Language Classroom.* Bristol: Multilingual Matters.

Ortega, L. (2003). "Syntactic Complexity Measures and their Relationship to L2 Proficiency: A Research Synthesis of College-Level L2 Writing". *Applied Linguistics*, 24(4), 492-518. https://doi.org/10.1093/applin/24.4.492.

Palladino, P.; Bellagamba, I. et al. (2013). "Italian Children with Dyslexia are also Poor in Reading English Words, but Accurate in Reading English Pseudowords". *Dyslexia*, 19(3), 165-77. https://doi.org/10.1002/dys.1456.

Pérez-Litago, U.; Rojas-Guerra, J.M. et al. (2025). "Language Skills of Adults with Dyslexia in English as a Foreign Language: Proposal of a Language Spontaneity Deficit Hypothesis". *Annals of Dyslexia*, 75, 313-38. https://doi.org/10.1007/s11881-025-00326-1.

Polio, C.G. (1997). "Measures of Linguistic Accuracy in Second Language Writing Research". *Language Learning*, 47(1), 101-43. https://doi.org/10.1111/0023-8333.31997003.

Rausch, A. (2016). "The Contribution of Complexity, Accuracy and Fluency to Language for Specific Purposes". *Journal of Languages for Specific Purposes (JLSP)*, 3, 29-39.

Skehan, P.; Foster, P. (2012). "Complexity, Accuracy, Fluency and Lexis in Task-Based Performance: A Synthesis of the Ealing Research". Housen, A.; Kuiken F; Vedder, I. (eds), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*. Amsterdam: Benjamins, 199-220.

Sparks, R.; Ganschow, L.; Pohlman, J. (1989). "Linguistic Coding Deficits in Foreign Language Learners". *Annals of Dyslexia*, 39, 177-95. https://doi.org/10.1007/bf02656908.

Tarone, E. (1980). "Communication Strategies, Foreigner Talk, and Repair in Interlanguage". *Language Learning*, 30(2), 417-28. https://doi.org/10.1111/j.1467-1770.1980.tb00326.x.

Thewissen, J.; Anishchanka, A.V. (2022). "Interaction Between Grammatical Accuracy and Syntactic Complexity at Different Proficiency Levels: Insights from a Learner Corpus". Leńko-Szymańska, A.; Götz, S. (eds), *Complexity, Accuracy and Fluency in Learner Corpus Research*. Amsterdam: Benjamins, 209-40.

Wiseheart, R.; Altmann, L.J.P. (2018), "Spoken Sentence Production in College Students with Dyslexia: Working Memory and Vocabulary Effects". *International Journal of Language & Communication Disorders*, 53, 355-69. https://doi.org/10.1111/1460-6984.12353.