COMPLEXITY ACCURACY AND FLUENCY IN INFORMAL L2 LEARNING AMONG UNIVERSITY STUDENTS
Four longitudinal case studies

CAMILLA DE RISO
UNIVERSITÀ DI PAVIA

Abstract – With English increasingly moving out of the classroom and into informal settings, studies on informal second language learning (ISLL) have supplied evidence on how prolonged exposure to informal L2 input can enhance advanced receptive and productive skills in L2 learners (Sockett 2014; Sundqvist, Sylvén 2016; Arnbjörnsdóttir, Ingvarsdóttir 2018; Dressman, Sadler 2020; Toffoli *et al.* 2023). However, only recently has research addressed longitudinal development within ISLL. In this respect, the three constructs of complexity, accuracy and fluency (CAF) have provided a comprehensive framework for analysing L2 learners' acquisitional process longitudinally (Polat, Kim 2014; Kusyk 2017). CAF measures aim to assess the richness, correctness, and naturalness of learners' L2 production and use. Furthermore, the interplay between these three dimensions across time appears to echo the non-linearity of L2 development (Ferrari 2012; Lambert, Kormos 2014; Vercellotti 2015; Michel 2017; Pallotti 2021). Starting from these premises, a longitudinal study was carried out at the University of Pavia, aiming to monitor the evolution of four participants' L2 spoken production over time. CAF trajectories were observed through monologic storytelling tasks, investigating any possible relationship between L2 development and frequency, intensity and type of exposure to informal English input. Data analysis has revealed distinctively different profiles which appear to be affected by type and intensity of media input. It has further confirmed that L2 development is multifaceted, multi-layered and multidimensional in nature, as several individual factors, such as learners' identity and beliefs, also play a key role in language development.

Keywords: Informal English learning; complexity, accuracy, fluency; longitudinal study; individual differences; Italy.

1. Introduction

Over the last two decades, the landscape for second language acquisition and language learners has changed drastically, moving from near-total dependence on the knowledge, expertise and planning of others to an unprecedented level of autonomy and opportunity for self-teaching and 'picking up' new languages (Dressman 2020). Given the broad and

diversified availability of L2 input outside institutional settings, research on informal language learning is flourishing in Europe and beyond, exploring informal second language learning (ISLL) scenarios in different countries with various population groups (i.e. age, degree of instruction, language competence levels), and other social and sociolinguistic factors such as educational and translational policies (Sundqvist, Sylvén 2016; Arnbjörnsdóttir, Ingvarsdóttir 2018; De Wilde *et al.* 2020; Dressman, Sadler 2020; Muñoz, Cadierno 2021; Toffoli *et al.* 2023, among others). While findings to date substantiate the effectiveness of prolonged exposure to online and offline informal input on L2 proficiency (Peters 2018; Kusyk 2020; Ludke 2020; Pattemore, Muñoz 2022; Lyrigkou 2023), only recently has research within ISLL addressed longitudinal L2 development. A longitudinal case study design provides insights into the dynamic, individually owned process of L2 development, enabling to thoroughly monitor the developmental process of a single individual while exploring individual differences related to motivation and attitudes towards the L2. This abundance of data challenges L2 researchers seeking analytical approaches that guarantee comparability and replicability of results. In this respect, the three dimensions of complexity, accuracy and fluency (CAF) – although controversial (Housen *et al.* 2012; Michel 2017; Tavakoli, Wright 2020; Pallotti 2021) – have provided a comprehensive framework for evaluating L2 learner-users' acquisitional paths longitudinally (Polat, Kim 2014; Vercellotti 2015; Kusyk 2017; Peng *et al.* 2022).

Starting from these premises, this article describes the results of a longitudinal study carried out in a medium-sized university in Italy aiming to monitor the evolution of four participants' L2 spoken production over time. CAF trajectories were observed through monologic storytelling tasks, concurrently investigating any relationship between L2 development and frequency, intensity and type of exposure to informal English input. By positioning itself in the eclectic field of ISLL, the study moves from the assumptions that L2 development is multifaceted, multi-layered and multidimensional in nature (Housen *et al.* 2012; Peng *et al.* 2022; Suzuki, Kormos 2023), as several individual factors such as learners' identity and beliefs also play a key role in language development (Arndt 2019; Li *et al.* 2022).

In what follows, Section 2 reviews recent literature on ISLL in the European context and gives a summary overview on CAF theoretical and methodological issues. Section 3 describes the study and provides data on participants' backgrounds and types of exposure to informal English input. Results are discussed in 4 and presented in three sub-sections: (i) an overview of participant's habits when accessing informal English input, (ii) the relationship with English, attitudes and behaviours towards the L2 and (iii) the analysis of CAF trajectories in L2 oral productions through time. As

remarked in the conclusive section, language develops dynamically in non-linear fashion and it is, therefore, rather challenging to identify a shared process of CAF development, especially when multiple individual factors such as motivation, self-confidence and engagement also come into play.

2. Background of the study

2.1. Informal second language learning

ISLL is defined as a form of second language acquisition which is thought to result from leisure-oriented activities such as watching films or TV series, listening to music, reading, playing video games, and browsing social networks in the L2. These informal activities take place outside educational contexts, are independent of formal frameworks for learning or testing and are self-directed by learners themselves who typically engage in these activities out of personal interests rather than for language learning. The growing access to English outside the classroom goes hand in hand with today's language and media-saturated world – a world where globalisation and mobility have expanded language affordances and multiplied the settings where language acquisition can occur (Benson 2021; Reinhardt 2022).

Over the last 20 years, systematic research on ISLL of English has been conducted in several European countries, at first mostly in the northern area where informal contact with English starts at a very young age and is strongly rooted in everyday lives (Verspoor *et al.* 2011; Sundqvist, Sylvén 2016; Puimège, Peters 2019; Arnbjörnsdóttir, Ingvarsdóttir 2018). Due to the fact that they are traditionally subbing countries, contact with English in northern Europe is extensive among teenagers and continues throughout adulthood, with differences in preferred input sources according to age and gender (Sundqvist 2009; Peters 2018). The massive L2 exposure through media led to higher English proficiency levels and is most evident in vocabulary development, reading, listening, and speaking skills (Arnbjörnsdóttir 2018). Receptive activities like watching television and browsing the web are most popular, though listening to music and playing video games are also widespread (Sylvén, Sundqvist 2012). Because of media specificity, users tend to acquire colloquial and informal language registers, to the detriment of academic writing skills required in the educational and professional sphere (Arnbjörnsdóttir, Ingvarsdóttir 2018). More recently, studies on ISLL conducted in France and Germany (Sockett 2014; Kusyk 2017, 2020; Arndt 2019), Spain, Italy and Greece (Muñoz 2020; Pavesi, Ghia 2020; Lyrigkou 2023), called for a systematic comparison of the results. Although informal contact with English seems to occur at a later age in

traditionally dubbing countries, i.e. in secondary school or university, some trends appear to be quite generalised and shared amid learner-users coming from central and southern Europe when approaching English in extramural contexts. Among audiovisuals, TV series appear to be the most popular choice and for the most part learner-viewers rely on the support of subtitles, though preferred viewing modalities vary according to content availability and L2 proficiency. More advanced users tend to opt for bimodal subtitles or no subtitles at all whereas interlingual subtitles are preferred at lower L2 proficiency levels (Pavesi *et al.* 2019; Pujadas, Muñoz 2020; De Wilde *et al.* 2020; De Riso 2021; Ghia, Pavesi 2021; Muñoz, Cadierno 2021; see Pattemore *et al.* 2024). L2 proficiency and familiarity with extramural English also recur in the polarisation between receptive and productive skills, as learners from central and southern European countries tend to primarily engage with English informally in receptive tasks such as watching AVs, listening to music and searching the web (Kusyk 2020; De Riso 2023; Toffoli 2023). The divide in learner-users' frequency of access to receptive vs. productive input is wider than their northern European counterparts as it appears that higher degrees of linguistic self-confidence can lead to greater willingness to engage in more activities that involve the English language, including those that require L2 production (Li *et al.* 2022; Arndt 2023).

Previous research has hinted at a lack of English proficiency among Italians as self-assessed by participants themselves, as perceived by other Europeans, and as persistently indicated in European reports (see Aiello 2018). Compared to other European countries, Italy appears to fall behind in terms of citizens' competence in L2 English and extent of exposure to the language in informal settings. To date, these results have been much rebutted as new studies – including this – suggest encouraging trends in the growing incidence of informal contact with English in the country, higher self-perceived L2 competence and increased L2 identity awareness among learner-users (Cravidi 2016; Pavesi, Ghia 2020; De Riso 2023; Cicero 2023).

2.2. CAF and L2 development

The multidimensionality of language proficiency and performance can be thoroughly described by the notions of complexity, accuracy and fluency (CAF). Complexity refers to the size, elaborateness, richness and diversity of the L2 performance. Accuracy is a measure of the target-like and error-free use of the L2. Fluency is described as the ability to produce eloquent speech without undue pausing, hesitation or reformulation. Research in Second Language Acquisition (SLA) has frequently used CAF measures as dependent variables to gauge L2 performance operated by independent variables such as task complexity or repetition (Yuan, Ellis 2003; Tavakoli,

Skehan 2005; Skehan, Foster 2008; Foster, Skehan 2012; Lambert *et al.* 2017; Suzuki 2017). In recent years, studies within the field have used CAF to identify developmental change and observe longitudinal learner trajectories (Ferrari 2012; Polat, Kim 2014; Vercellotti 2015; Kusyk 2017; Michel 2017; Peng *et al.* 2022). Regardless of the considerable amount of research available on CAF (Housen, Kuiken 2009; Norris, Ortega 2009; Housen *et al.* 2012; Lambert, Kormos 2014; Pallotti 2009, 2021; Hasnain, Halder 2022 for recent reviews), the constructs' theoretical and operational definitions are still controversial topics of discussion.

Considered as the most controversial construct of the triad, the term complexity is commonly used to refer both to a structural or linguistic dimension, i.e. the elements and intricacies of a system, and to a cognitive dimension, that is the effort needed to process said system (Pallotti 2021). Since linguistic complexity itself is also multidimensional, further disambiguation is needed between lexical and grammatical complexity as both, in turn, feature different subdimensions (e.g. lexical diversity, lexical sophistication, syntactic complexity). To make way in the vast, intricate framework of complexity, several measures have been developed and critically scrutinised to generate a plausible rationale beyond their usage. To date, however, there is still a lack of consensus among researchers as the quest for omnibus measures of complexity is somewhat misguided as it "fails to adequately capture the multidimensional space of variation among complexity features" (Biber *et al.* 2023, p. 23). Unlike complexity, there is less disagreement on accuracy, defined as any deviation from a norm or native-like use of language including diverse types of errors (Housen et al. 2012). Despite the well-established tradition of adopting the 'accuracy perspective', i.e. counting and weighing errors, it can be problematic to define exactly what constitutes an error, especially when lexical, morphological, syntactic, phonological, and spelling errors are bundled together. In this respect, new studies have embraced the 'interlanguage approach' wherein accurate productions "may depend on a more proceduralised norm-oriented interlanguage system and be thus positively related to increased fluency" (Pallotti 2021, p. 205). From a cognitive perspective, fluency is connected to the proceduralisation and automatisation of a language, with functions such as planning and monitoring language production (Tavakoli, Skehan 2005). Pertaining first and foremost to spoken language, three main dimensions of fluency are usually identified: speed, breakdown and repair. Although this fluency model has been thoroughly examined and validated (Skehan et al. 2016; Tavakoli, Wright 2020; Suzuki, Kormos 2023 among others), normal dysfluency phenomena are to be expected when analysing spoken language. Whether it is a monologic/dialogic task or spontaneous conversation, it is quite natural for a

speaker's flow to be characterised by pauses, hesitators and repetitions "at points where the need to keep talking threatens to run ahead of mental planning, and the planning needs to catch up" (Biber *et al.* 2021, p. 1048). While sometimes dysfluency is so pervasive a feature that, by the standards of written language, spoken language is grammatically inchoate, it is normal for speakers to produce utterances with minor dysfluencies which do not interfere with understanding (Biber *et al.* 2021, p. 1052-53).

When measuring L2 learners' proficiency, performance and development, the three dimensions of CAF need to work together to relay exhaustive outcomes. However, earlier research on the multidimensionality of L2 development suggests that due to working memory's limited capacity for processing input and output, L2 learners have difficulties in catering simultaneously to meaning and form, and therefore need to choose which content to prioritise at the expense of the other (Ellis, Barkhuizen 2005; Vercellotti 2015; Michel 2017 for recent reviews). In terms of input, working memory extracts and temporarily stores information from both input and long-term memory. For L2 learners the load on working memory increases, because the extract-and-store processing rely considerably more on controlled processing rather than automatic. For output, L2 learners need to retrieve information from long-term memory and hold it in their short-term memory to produce messages that are meaningful, coherent, and grammatically correct. In real-time conversation, with limited planning time, the overload on working memory could prevent the simultaneous processing of comprehension (meaning) and acquisition (form), leaving L2 learners to prioritise one or the other depending on context and orientation (Ellis, Barkhuizen 2005, p. 142). To simplify, focusing on one CAF construct may result in lower performance in one or both other constructs, that is, trade-off effects. To investigate CAF and L2 development, research has analysed learners' performance by testing the predictions of two theoretical models, i.e. Skehan's (1998) Limited Attention Capacity model and Robinson's (2011) Cognition Hypothesis. Skehan's information-processing model predicts a conflicting relationship among the three CAF components where learners emphasise meaning (fluency) over form, potentially hindering further L2 development. When learners do focus on form, there is a secondary contrast between control of form (accuracy) and use of more advanced language (complexity) (Skehan 2009). Because of limited mental resources, i.e. limited attentional capacity and working memory, all language learners have these tensions during performance and thus entailing a single-source view of attention (Skehan 2015). On the other hand, researchers who reject a single-source capacity limitation accept CAF competitiveness as explained by attentional control and interference (Robinson 2003). Robison's Cognition Hypothesis (2011) expects tasks to promote either fluency or combined complexity and accuracy, a theory which supports Skehan's primary

Complexity accuracy and fluency in informal L2 learning among university students.
Four longitudinal case studies

165

trade-off but contrasts with the second. For example, simple monologic tasks are likely to foster fluency to the detriment of complexity or accuracy, whereas accuracy and complexity (but not fluency) are promoted during complex monologic tasks (Robinson 2011). Over the years, many studies have set out to contribute to the debate and despite contrasting results, researchers working in different theoretical frameworks have concluded that trade-off effects impact language performance, supporting either Skehan's or Robinson's theoretical model. Grammatical complexity has been reported to increase at the expense of fluency (Bygate 2001), and a trade-off between fluency and accuracy seems to be a particularly robust finding in the literature (Yuan, Ellis 2003; Michel *et al.* 2007; Ahmadian, Tavakoli 2011). Conversely, findings also suggest a trade-off between complexity and accuracy (Ferrari 2012).

The large variety of research designs and the plethora of CAF measures generating conflicting results call for more standardisation and a unified methodological approach in what is, in all respects, an ongoing debate. After a thorough review of CAF measures and researchers' perspectives regarding the choice of particular metrics, the present study sets out to gain preliminary insights on L2 development among a sample of Italian learner-users of English during a one-year longitudinal study.

3. The study: aims, sample, methods

The present study analyses if, how and to what extent access to several types of informal English input may affect L2 development as part of a wider study on informal contact with English (De Riso 2023). Four longitudinal case studies gauge how different access patterns to different types of informal English input may play a role in learners' L2 development and address the following research questions: i) To what extent does prolonged exposure to informal English input affect L2 development? ii) Is there any relationship between access patterns and learners' L2 proficiency? iii) Which paths do complexity, accuracy and fluency trajectories follow in elicited production tasks over a 12-month period? The study additionally provides qualitative observations about learner-users' attitudes towards the English language and how their motivations may vary in time and in relation to other factors such as increased exposure and better knowledge of the language.

From an initial large-scale questionnaire survey (*n* 605) that explored the degrees and modalities of spontaneous and naturalistic access to English by Pavia University students, 264 respondents consented to take part in the longitudinal study. With the aim of carrying out longitudinal studies, the majority of the selected cohort was made up of non-language specialists, in order to sample a wider variety of first-year undergraduates and postgraduates with a view to carrying out longitudinal studies. Early selection

criteria included self-assessed levels of competence in the language (CEFR B2 and C1) and self-reported exposure indexes calculated by multiplying frequency by length of exposure to the different types of informal input surveyed in the questionnaire (from average to high-input learner-users). 73 respondents were selected and eventually 20[1] participants confirmed their attendance. By the end of the data collection, only 4 subjects had completed the steps of the study successfully (Table 1).

Subject	Gender	Age	Course	CEFR	Exposure index
Delia	F	23	MA Biology	B2	57/140
Grace	F	20	BA Chemistry	B2	58/140
Marta	F	25	MA Biology	C1	64/140
Lea	F	20	BA Biotechnology	B2	105/140

Table 1
Personal information of the four subjects analysed as case studies.

The longitudinal study lasted throughout 2021 with one-on-one online meetings[2] scheduled on Zoom every other month for a total of six meetings per subject. Each data collection involved two main monologic oral tasks elicited through storytelling techniques, i.e. the retelling of a muted film clip,[3] and the recollection of an event from the past.[4] A total of 48 monologues were recorded, transcribed and further analysed according to 15 measures of complexity, accuracy and fluency (Table 2) to observe participants' L2 development trajectories.[5]

     Complexity included: grammatical complexity, calculated as mean length of AS-units[6] (Foster *et al.* 2000); lexical diversity, calculated using *vocd* (McKee *et al.* 2000); and lexical sophistication, measured with Lexical

---

[1] 16 females and four males distributed among different courses as follows: Biology and Biotechnology (6), Medicine (4), Communication Studies (2), Pharmacy (2), Chemistry (1), Computer Engineering (1), Environmental Engineering (1), Philosophy (1), Political Science (1), Psychology (1).

[2] Due to COVID-19 restrictions, online meetings were strongly encouraged as opposed to face-to-face.

[3] 5-7 minutes long. January: *The Wedding Date* (Kilner 2004); March: *Sex and the City: The Movie* (King 2008); May: *Four Weddings and A Funeral* (Newell 1994); July: *Runaway Bride* (Marshall 1999); September: *Made of Honor* (Weiland 2008); November: *Mamma Mia!* (Lloyd 2008).

[4] Participants were encouraged to reflect and provide a narrative on their past experiences as follows (from the first to the last meeting): your first time on holiday with friends; your first day at university; your first time surrounded by a big crowd; the first item you purchased something with your own money; your first time going to the cinema; last days of high school.

[5] Data were standardised to bring 15 CAF measures to the same scale (z-score).

[6] Analysis of Speech unit, defined as a single speaker's utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s) associated with either (Foster *et al.* 2000, p. 365).

Frequency Profile[7] (LFP) (Laufer, Nation 1995). The analysis of *vocd* is based on the probability of new vocabulary being introduced into longer and longer samples of text. The index is automatically calculated[8] through a series of type-token ratio samplings and curve fittings (McNamara *et al.* 2014), and it has been shown to be a useful measure for L2 data (McCarthy, Jarvis 2007, 2010). LFP describes the lexical context of a text in terms of frequency scores (Laufer, Nation 1995).

Complexity
Lexical
1 - Diversity: *vocd*
2 - Sophistication: Lexical Frequency Profile (LFP)
Grammatical:
3 - Length: mean length of AS-unit
Accuracy
4 - Global: Normalised error rate (per 1000 words)
Lexis:
5 - Words in L1
6 - Non-normative word formation
7 - Non-normative word choice
Syntax:
8 - Non-normative use of singular-plural forms
9 - Non-normative use of articles, pronouns, and prepositions
10 - Word order or sentence structure based on L1
Verb:
11 - Non-normative form (conjugation)
12 - Non-normative choice of tense, aspect, and mood
13 - Other (missing word)
Fluency
14 - Reformulations (normalised per 1000 words)
15 - Repeats (normalised per 1000 words)

Table 2
Selected set of measures for complexity, accuracy and fluency.

The BNC-COCA corpus was used as reference for this analysis as it offers a comprehensive sample of modern English usage. This ensures that word frequency categorisation is based on a broad linguistic dataset, enhancing the reliability of the LFP and providing insights into the text's lexical richness and sophistication. Accuracy was measured as percentage of error-free clauses. Looking at the non-normative use of the L2 is recommended (Ellis, Barkhuizen 2005), since specific measures may misinterpret learners' knowledge if learners avoid certain forms or constructions. While the study was initially set to measure speed, breakdown and repair fluency, data were collected during the pandemic when lessons in all institutions, from primary school to university, were held online. It often happened, then, that Zoom video calls would end unexpectedly or that the video or audio would lag whenever the servers were

---

[7] Automatically calculated with Compleat Web VP (frequency frameworks by Laufer *et al.* coded by Tom Cobb, available at https://www.lextutor.ca/vp/comp/).

[8] Using Coh-Metrix (Graesser *et al*.), available at http://cohmetrix-new.memphis.edu/.

overloaded. These unforeseen events compromised the reliability of speed and breakdown fluency; hence, it was decided to focus on repair fluency, i.e. repeats and reformulations.

By embracing a mixed-method approach, the study relied on semi-structured interviews to gain insights on learner-users' personal and L2 identity; every other week participants were asked to keep a journal[9] of their informal activities in the L2 and record whatever thought might have arisen from the experience.

4. Results

Each case study is analysed separately and is divided into three sections: (i) an overview of participant's habits when accessing informal English input, (ii) the relationship with English, attitudes and behaviours towards the L2 (iii) the analysis of CAF trajectories in L2 oral productions through time. Section (iii) analyses two monologic tasks (i.e. storytelling) elicited over a 12-month period, whereas section (i) and (ii) draw data from participants' questionnaire responses, semi-structured interviews, and journals.

4.1. Exposure to informal English input

Participants L2 usage profiles were representative of the informal habits surveyed in the questionnaire (De Riso 2023) insofar as they engaged in multiple informal activities while showing strong preferences for a specific type of input. More specifically, Delia preferred to watch films and TV series, Grace loved to read fantasy novels, Marta listened to podcasts and livestreams and Lea used to play video games online, read manga and watch anime. Their participation in informal activities did not vary particularly over the 12 months, though at times content consumption increased considerably whenever binge-watching (or binge-reading, binge-listening, binge-playing, etc.) occurred. Among the dominant factors leading to engaging in informal activities in English, participants mentioned the hedonic component, a search for naturalness and authenticity, the immersive experience, sociability factors and earlier availability of content in the English language.

Participants' first contacts with informal English input date back to secondary school when they first approached music, YouTube videos, books and video games recreationally. In lack of an Italian counterpart and unwilling to wait longer for an eventual translation, participants turned to

---

[9] An online Word document per participant, shared exclusively with the researcher in order to gain access to data while maintaining privacy among participants.

content in the L2 without having any language-related goal in mind. During high school they began to notice a certain ease with the English language and attributed this easiness to what they were doing out-of-the-classroom rather than their formal instruction. Nevertheless, the main reasons for accessing content in English did not vary. To date, participants' input choices are affected by habit and enjoyment rather than language learning purposes, whilst being aware of the beneficial effect that prolonged exposure to English has on their L2 competence.

Despite the significant exposure to informal English input, it should be mentioned that participants occasionally turn to Italian whenever the activity in the L2 is perceived as too demanding. For example, with AVs, dubbing (when done properly) requires less attention and results in a more relaxing experience, whereas L1 subtitling offers an aid to comprehension when dialogues are too technical. As regards reading, participants opt for Italian fantasy novels and manga whenever they get the chance: as niche genres, there is still little to no original publications (nor translations) available in Italy, thus English is 'necessarily' preferred.

4.2. Attitudes towards English

Data on individual differences, i.e. learner-user traits and characteristics that may have an impact on learning processes, behaviours and outcomes, were gathered through journal entries, semi-structured interviews and spontaneous conversations with participants. On different occasions, participants were asked to recall their first approach to informal English input and give a one-word definition of their relationship with the language. Although attitudes towards the L2 were overall positive, the participants' experiences were quite different as reflected in their motivations and beliefs.

In the case of Delia, her appreciation of the language clashed with a low self-perceived L2 competence, highlighting a complicated relationship with English. In other words, her use of the L2 is partially clouded by the pressure of other people's judgment:

> (1) "When I speak in English, I know what I want to say […] but I get nervous since I don't want to make mistakes, because I imagine that people will assume I'm not good at it […] sometimes it gets harder to transform what I'm thinking into words and sentences, and I begin to stutter."

This conflict, however, appears to be strictly related to conversation as Delia eventually admits that her insecurities might be related to a mild social anxiety rather than a lack of L2 competence:

Lingue e Linguaggi

(2) "I had one paper published, one was recently accepted plus I'm working on my thesis in English, and I don't find it hard at all […] I basically do everything in English it's just that I'm not really extroverted so that could be part of the problem too, speaking in public in general."

Similarly, Grace too underestimates her English proficiency level; however, she believes she has a natural predisposition for the language that led to a strong L2 confidence and language awareness:

(3) "When I was 13 or 14 I began to read books in English because I wanted to read them in their original language. At first it was hard I remember that I used to write all the words that I didn't know in a notebook. […] Then something just clicked […] I was able to read and speak effortlessly."

She defines her relationship with English as "a wonderful love story" and one that is meant to last:

(4) "Every time I get the chance to speak, read, listen to it I'm always very happy […] When I'm doing something and I can choose between Italian or English, I chose English. I will always choose English unless of course the original language is Italian."

That of Marta is one interesting case, as she believes that her bilingual upbringing (in Italian and Spanish) gave her an "extra boost in picking up foreign languages". According to her, "having a bilingual brain is like having multiple language microchips in your head". She describes the process of switching from one language to another as a "change input language command":

(5) "I know I'm privileged from a linguistic point of view; I know that people who grew up in bilingual environments are more inclined to learn new languages […] I consider myself extremely lucky because for me it was a rather spontaneous thing [to learn English]."

Whilst acknowledging her 'head start' in L2 learning, Marta is quite proud of her relationship with English because: "knowing other languages and cultures is both necessary and fundamental nowadays and it's something to be proud of". Quite often, Marta finds herself thinking or writing things down in English and she says that most times she hardly notices which language she is using:

(6) "The more I listen to it the more I think in English. I occasionally write my thoughts in a personal journal and sometimes I find myself writing in English, like it's natural to me"

Rather similar feelings were shared by Lea. She considers English almost as a "routinized second language" and is extremely proud of her ability to communicate effectively with both native and non-native speakers – an activity in which she engages at every chance she gets:

(7) "I simply couldn't be without English, it allowed me to meet new people and cultures from all around the world […] I use it so often almost as if it was my second language"

(8) "I was playing [a video game] online with my Norwegian friend and a friend of mine joined us, he's Italian. I was chatting in English as usual and noticed that my Italian friend was very quiet, so I began to translate things to him, I was like an interpreter. And later he thanked me, because he was really struggling to follow what we were saying and was like: "How can you talk so easily?" and I didn't know what to say it was a normal thing for me."

4.3. CAF trajectories over time

A total of 48 monologues elicited through storytelling were recorded, transcribed and further analysed according to 15 selected CAF measures (Section 3 Table 2). Each case study was analysed individually; however, few similarities can be observed across the results.

Throughout the longitudinal study, CAF trajectories presented strong variability across participants and within individual L2 development, showing a constant push and pull among measures, i.e. progression and regression. The trajectories of accuracy and fluency show that the two dimensions appear to be particularly sensitive to each other. While this result is consistent with previous literature (Yuan, Ellis 2003; Michel *et al.* 2007; Ahmadian, Tavakoli 2011), the trade-off[10] between accuracy and fluency found in the present study led to different outcomes among participants. Findings from Delia's case study (Figure 1) seem to support Skehan's primary competition between meaning and form, as grammatical complexity (mean length of AS-units) appears to increase at the expense of fluency (measured by the number of repairs).

---

[10] The present data analysis relied on descriptive statistics of a small sample of students, consequently, assumptions of trade-off effects should be made conservatively.
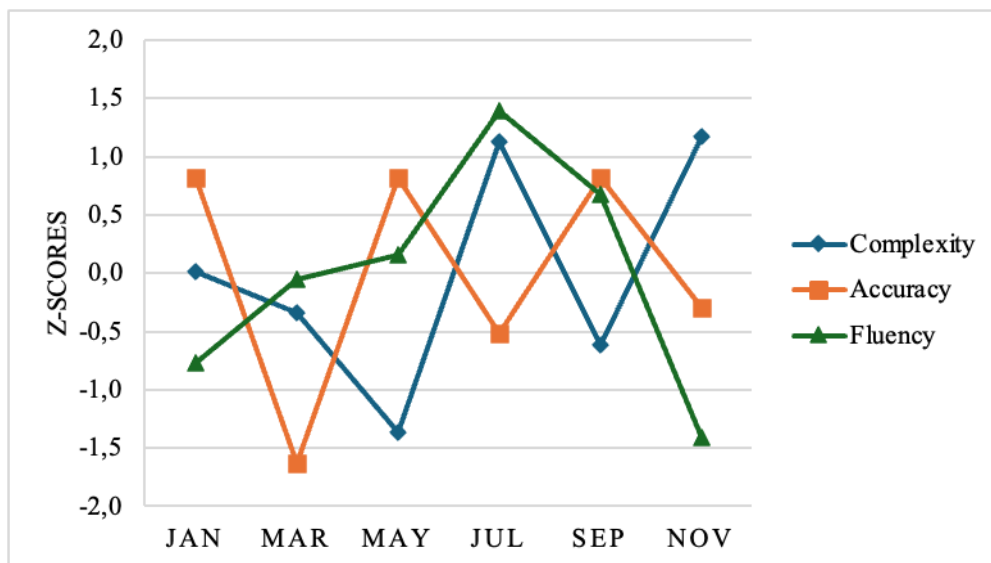
Figure 1
CAF trajectories: Delia's case study.

Conversely, the results from Marta's case study (Figure 2) show that learner-users can have higher accuracy and complexity at the expense of fluency, thus supporting Robinson's Cognition Hypothesis but refuting Skehan's secondary contrast with form.
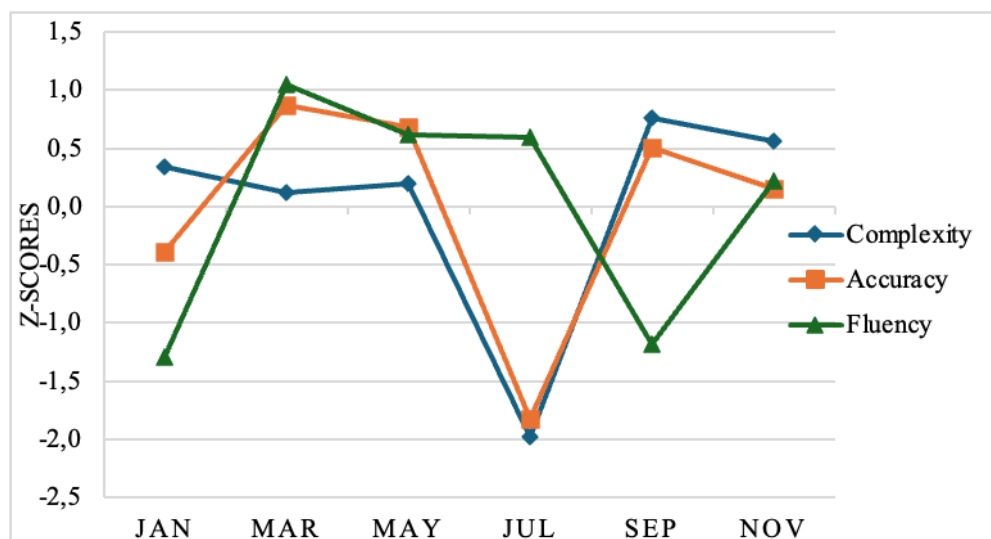

Figure 2
CAF trajectories: Marta's case study.

Even though these findings partially support competitiveness among CAF components, the fluctuations observed among participants show just how complex and dynamic L2 development is and how difficult it is to identify a shared process of CAF development.

Complexity accuracy and fluency in informal L2 learning among university students.
Four longitudinal case studies

173

The study further compared CAF trajectories with participants' habits of exposure to the different sources of informal English input. Among the several informal activities investigated in the study, extensive reading appears to be by far the most effective for L2 development in the sample. Findings from Grace's case study suggest that reading novels extensively can lead to improved complexity, i.e. more structured utterances accompanied by a diverse, sophisticated vocabulary. The following examples compare Grace's L2 oral production in March (9) and November (10):

> (9) "We were celebrating [New Year's Eve] in the square that is right in front of the castle and there was a spectacular fireworks' show it was exceptional. The fireworks started from the walls of the castle, and they were creating these beautiful, beautiful patterns in the sky and there was this huge crowd, and they were all celebrating. There were people screaming, dancing, people kissing, just having a good time in general."

> (10) "So, I basically found my footing and other than that I don't remember much of my actual exam. I do remember coming out of there and feeling very like relieved like a great burden had been taken off my shoulders, but at the same time very sad because that was the moment that I had waited for, for five years. And if you hear people talking about the, the exam they all- it always seems like this big huge thing something impossible to overcome and, and I know it gives you like so much anxiety and stress and you're so nervous and then you get there and you just blink and it's gone and you did it and you don't really know how you did it, but you got through it and you're like, 'That was it?'"

In the context of SLA, researchers agree on the value of extensive reading for L2 learning but they also warn about the possible risks of text selection. Ultimately, learners could select a text that is too simple compared to their proficiency level to the detriment of L2 learning (Arnold 2009; Grabe, Yamashita 2022). This seems to be the case of Lea who, despite reading plenty of comics and manga, did not show signs of L2 improvement throughout the longitudinal study. The following examples compare Lea's L2 oral production in March (11) and November (12):

> (11) "And during the breaks, I tried to speak with some- somebody else. To get to, to try and know them. And in the end, I met the girl that I went- I went home together on the train. To get to Pavia, I went with my, my guy friend, and when I, when I had to return, I was with him and the girl I met. So, we were all three coming back home together, and it was kind of nice."

> (12) "And then we went- we just went to high school that was like five minutes by foot for my home so not, not that far. I did my exam. It was kind of

> funny because it happened a lot of the stuff that my friends call it lucky? So,
> I was nervous at, at the beginning but it ended up being something I, I laugh
> while I- when I tell it. And when I was done I was, I was still with this
> friend of mine and we went, we went to eat outside- out."

By comparing Grace and Lea's L2 oral productions, it appears that exposure to different literary genres affects L2 development in a different way: if, on the one hand, reading prose led to more structured utterances with rich, complex vocabulary, on the other, reading comics, typically written with short sentences and simple vocabulary, resulted in a similar way of speaking.

Throughout the study, all four participants were largely exposed to informal L2 audiovisual input, and while it would be an overstatement to imply that exposure to AVs did not affect L2 development in the sample, findings differ and, in part, contradict each other. As in Kusyk (2017), it appears that watching AVs informally positively affects L2 development in the measure of lexical sophistication. However, if, in the case of Delia, extensive exposure to informal L2 AV input led to a moderately more sophisticated vocabulary, the same did not occur with Marta, whose lexical sophistication slightly worsened despite the prolonged exposure to AVs (see Appendix). It is worth mentioning that these different L2 development outcomes within exposure to the same source of informal input may have been affected by individual factors, such as motivation and media immersion. Although each participant accessed the different types of informal input extensively, some were more motivated than others and appreciated the informal activity to the point of experiencing media immersion constantly. These subjects (i.e. Grace and Delia) were the ones who ultimately showed signs of improvement compared to the others. Lastly, during the longitudinal study participants used the L2 productively by habitually interacting with both native and non-native speakers of English through face-to-face interactions, face-to-screen interactions (voice calls and video calls), text messages and chats. Although they all perceived improvements in their L2 competence following these activities, this development was not always reflected in their CAF trajectories, especially with regard to fluency.

The evolution of complexity, accuracy and fluency trajectories over time for each case study is available in the Appendix.

5. Conclusions

By positioning itself in the eclectic field of ISLL, the present study observed the evolution of CAF trajectories in L2 oral productions of four case studies over a 12-month period, aiming to investigate any potential relation among

Complexity accuracy and fluency in informal L2 learning among university students.
Four longitudinal case studies

175

frequency, intensity, and type of exposure to informal English input and L2 development. Moreover, the study explored the thoughts and feelings of L2 learner-users who accessed different sources of informal English input extensively, by investigating participants' L2 awareness in relation to themselves, others, and the informal activities they access habitually.

While prolonged exposure to informal English input arguably affects L2 development, findings show that individual factors such as motivation and patterns of access to the informal input seem to have influenced L2 competence in different ways. For example, both Grace and Lea engaged in reading extensively. Nevertheless, their L2 development had different outcomes perhaps due to the specificity of the two preferred literary genres (i.e. novels vs. comics) and participants' higher vs lower motivation, which could lead (or not lead) to media immersion. On the one hand, Grace's lexical and grammatical complexity improved greatly throughout the longitudinal study, reflecting the sophisticated structures and rich vocabulary typical of prose. She showed strong tendencies of media immersion, reading up to 5 hours a day to finish the novel she was reading. On the other, Lea used to read comics and manga every day while doing other activities, e.g. commuting on the train, while having breakfast, thus lacking media immersion and motivation to read the whole story at once. Despite extensive exposure, her L2 development lacked grammatical and lexical complexity, perhaps echoing the simple structures and minimal vocabulary of comics and manga and suggesting that less attention and motivation towards the informal activity may interfere with L2 development. Similarly, Delia and Marta were both largely exposed to informal L2 audiovisual input; however, the different modalities of access and frequency of exposure led to different outcomes in terms of lexical sophistication. By the end of the longitudinal study, Delia, who was strongly motivated and used to watch at least one film or episode per day, used moderately more sophisticated vocabulary. Marta, on the contrary, used to leave the video on while doing other activities such as cleaning or cooking and her lexical sophistication slightly worsened throughout the study. Overall, CAF trajectories presented strong variability across participants and within individual L2 development. Among CAF measures, the two dimensions of accuracy and fluency appeared to be particularly sensitive to each other and while this result is broadly consistent with a trade-off interpretation (Yuan, Ellis 2003; Michel *et al.* 2007; Ahmadian, Tavakoli 2011), findings differ and sometimes contradict each other (cf. Figure 1 and Figure 2 in Section 4.3). The fluctuations observed among the four case studies show just how complex and dynamic L2 development is and how difficult it is to identify a shared process of CAF development.

The present contribution wanted to provide better insights on informal L2 learner-users' individual profiles as well as analyse *if*, *how*, and *to what extent* access to informal English input may affect L2 development analysed

by CAF measures. However, some limitations of the data collection emerged and will need to be fully addressed in future research. The modality of data collection (recorded Zoom video calls) and the unexpected issues that came within compromised the reliability of speed and breakdown fluency, thus limiting the analysis to repair fluency, i.e. repeats and reformulations. Although the longitudinal monitoring of CAF trajectories allowed to thoroughly observe the evolution of participants' L2 oral production, the lack of pre/post L2 proficiency tests made it difficult to properly assess L2 development and identify subtle improvements, especially if participants were already proficient in the L2. Finally, results might have been somehow influenced by task design. At each data collection participants were asked to tell the same story, i.e. Test 1: four participants tell story A; Test 2: four participants tell story B, etc. Perhaps, results would have varied if each storytelling task was distributed differently among the four informants every time, ensuring that the task construct did not influence overall task performance over time i.e. Test 1: participant 1 tells story A, participant 2 tells story B, participant 3 tells story C, …; Test 2: participant 1 tells story C, participant 2 tells story A, participant 3 tells story B, etc.[11]

Bionote: Camilla De Riso is a postdoctoral researcher in Linguistics at the University of Pavia, where she earned both her MA in Theoretical, Applied and Modern Languages Linguistics (2019) and PhD in Linguistics (2023). Her doctoral thesis examined the impact of informal English input on second language acquisition through a longitudinal perspective, analysing the development of complexity, accuracy and fluency in students' L2 oral production. She has been a visiting fellow at Cardiff University (2014), Karlsruhe University of Education (2021), and the University of Malta (2022), enriching her expertise in content and language integrated learning (CLIL) and quantitative methods in applied linguistics. Her research interests include informal second language learning, audiovisual translation, language learner identities, motivations and attitudes, as well as the constructs of complexity, accuracy and fluency (CAF) in L2 development.

Author's address: camilla.deriso@unipv.it

---

[11] The author would like to thank the reviewer for this valuable insight which will contribute to improve future research.

References

Ahmadian M.J. and Tavakoli M. 2011, *The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production*, in "Language Teaching Research" 15 [1], pp. 35-59.

Aiello J. 2018, *Negotiating Englishes and English-speaking identities*, Routledge, London.

Arnbjörnsdóttir B. 2018, *Using English at University*, in Arnbjörnsdóttir B. and Ingvarsdóttir H. (Eds.), *Language Development Across the Life Span: The Impact of English on Education and Work in Iceland*, Springer International Publishing, Cham, pp. 143-162.

Arnbjörnsdóttir B. and Ingvarsdóttir H. (Eds.) 2018, *Language Development across the Life Span: The Impact of English on Education and Work in Iceland*, Springer International Publishing, Cham.

Arndt H. 2019, *Informal second language learning: The role of engagement, proficiency, attitudes, and motivation*, unpublished doctoral thesis, University of Oxford.

Arndt H. 2023, *Construction and validation of a questionnaire to study engagement in informal second language learning*, in "Studies in Second Language Acquisition", 45 [5], pp. 1456-1480.

Arnold N. 2009, *Online extensive reading and advanced foreign language learners: An evaluation study*, in "Foreign Language Annals" 42 [2], pp. 340-366.

Benson P. 2021, *Language Learning Environments*, Multilingual Matters, Bristol.

Biber D., Johansson S., Leech G., Conrad S. and Finegan E. 2021, *Longman Grammar of Spoken and Written English*, Longman, London.

Biber D., Larsson T. and Hancock G.R. 2023, *The linguistic organization of grammatical text complexity: comparing the empirical adequacy of theory-based models*, in "Corpus Linguistics and Linguistic Theory", 20 [2], pp. 347-373.

Bygate M. 2001, *Effects of task repetition on the structure and control of oral language*, in Bygate M., Skehan P. and Swain M. (Eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing*, Pearson Education Limited, pp. 23-48.

Cicero A. 2023, *Motivation and attitudes in EFL learning: A case study of Italian high school students*, unpublished doctoral thesis, University of Pavia.

Cravidi G. 2016, *Accessing English Input Outside University: A Qualitative and Quantitative Investigation*, unpublished postgraduate thesis, Università degli Studi di Pavia.

De Riso C. 2021, *English audiovisual input: a study of teenagers' habits, attitudes, and responses to subtitling versus dubbing*, in "RILA - Rassegna Italiana di Linguistica Applicata" 1 [2], pp. 245-260.

De Riso C. 2023, *Out-of-the-classroom English input and informal Second Language Acquisition: A longitudinal study of Italian university students*, unpublished doctoral thesis, University of Pavia.

De Wilde V., Brysbaert M. and Eyckmans J. 2020, *Learning English through out-of-school exposure. Which levels of language proficiency are attained and which types of input are important?*, in "Bilingualism: Language and Cognition" 23, pp. 171-185.

Dressman M. 2020, *Introduction*, in Dressman M. and Sadler R.W. (Eds.), *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken, pp. 1-12.

Dressman M. and Sadler R.W. (Eds.) 2020, *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken.

Ellis R. and Barkhuizen G. 2005, *Analysing Learner Language*, Oxford University Press, Oxford.

Ferrari S. 2012, *A longitudinal study of complexity, accuracy, and fluency variation in second language development*, in Housen A., Kuiken F. and Vedder I. (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*, John Benjamins, Amsterdam, pp. 277-298.

Foster P., Tonkyn A. and Wigglesworth G. 2000, *Measuring spoken language: a unit for all reasons*, in "Applied Linguistics" 21 [3], pp. 354-75.

Foster P. and Skehan P. 2012, Complexity, accuracy, fluency and lexis in task-based performance, in Housen A., Kuiken F. and Vedder I. (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*, John Benjamins, Amsterdam, pp. 199-220.

Ghia E. and Pavesi M. 2021, *Choosing between dubbing and subtitling in a changing landscape*, in "Lingue e Linguaggi" 46, pp. 161-177.

Grabe W. and Yamashita J. 2022, *Reading in a Second Language: Moving from Theory to Practice*, Cambridge University Press, Cambridge.

Hasnain S. and Halder S. 2022, *Intricacies of the Multifaceted Triad-Complexity, Accuracy, and Fluency: A Review of Studies on Measures of Oral Production*, in "Journal of Education" 204 [1], pp. 145-158.

Housen A. and Kuiken F. 2009, *Complexity, accuracy and fluency in second language acquisition*, in "Applied Linguistics" 30 [4], pp.461-473.

Housen A., Kuiken F. and Vedder I. (Eds.) 2012, *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*, John Benjamins, Amsterdam.

Kusyk M. 2017, *Les dynamiques du développement de l'anglais au travers d'activités informelles en ligne: une étude exploratoire auprès d'étudiants français et allemands*, unpublished doctoral thesis, University of Strasbourg/Karlsruhe University of Education.

Kusyk M. 2020, *Informal English learning in France*, in Dressman M. and Sadler R.W. (Eds.), *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken, pp. 333-348.

Lambert C. and Kormos J. 2014, *Complexity, accuracy, and fluency in task based L2 research: Toward more developmentally based measures of second language acquisition*, in "Applied Linguistics" 35 [5], pp. 606-614.

Lambert C., Kormos J. and Minn D. 2017, *Task repetition and second language speech processing*, in "Studies in Second Language Acquisition" 39 [1], pp. 167-196.

Laufer B. and Nation P. 1995, *Vocabulary size and use: Lexical richness in L2 written production*, in "Applied Linguistics" 16 [3], pp. 307-322.

Li S., Hiver P. and Papi M. 2022, *The Routledge Handbook of Second Language Acquisition and Individual Differences*, Routledge, New York.

Ludke K.M. 2020, *Songs and music*, in Dressman M. and Sadler R.W. (eds), *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken, pp. 203-213.

Lyrigkou C. 2023, *Informal second language learning and EFL learners' spoken use of discourse markers*, in Toffoli D., Sockett G. and Kusyk M. (Eds.), *Language Learning and Leisure*, De Gruyter, Berlin, pp. 91-120.

McCarthy P.M. and Jarvis S. 2007, *vocd: a theoretical and empirical evaluation*, in "Language Testing" 24 [4], pp. 459-488.

McCarthy P.M. and Jarvis S. 2010, *MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment*, in "Behavior Research Methods" 42, pp. 381-392.

McKee G.D., Malvern D. and Richards B. 2000, *Measuring vocabulary diversity using dedicated software*, in "Literary and Linguistic Computing" 15, pp. 323-37.

McNamara D.S., Graesser A.D., McCarthy P.M. and Cai Z. 2014, *Automated Evaluation of Text and Discourse with Coh-Metrix*, Cambridge University Press, Cambridge.

Michel M. 2017, *Complexity, Accuracy and Fluency (CAF)*, in Loewen S. and Sato M. (Eds.), *The Routledge Handbook of Instructed Second Language Acquisition*, Routledge, London, pp. 50-68.

Michel M.C., Kuiken F. and Vedder I. 2007, *The influence of complexity in monologic versus dialogic tasks in Dutch L2*, in "International Review of Applied Linguistics" 45 [13], pp. 241-59.

Muñoz C. 2020, *Boys like games and girls like movies. Age and gender differences in out-of-school contact with English*, in "RESLA" 33, pp. 172-202.

Muñoz C. and Cadierno T. 2021, *How do differences in exposure affect English language learning? A comparison of teenagers in two learning environments*, in "Studies in Second Language Learning and Teaching" 11 [2], pp. 185-212.

Norris J.M. and Ortega L. 2009, *Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity*, in "Applied Linguistics" 30 [4], pp. 555-578.

Pallotti G. 2009, *CAF: Defining, refining and differentiating constructs*, in "Applied Linguistics" 30 [4], pp. 590-601.

Pallotti G. 2021, *Measuring complexity, accuracy, and fluency (CAF)*, in Winke P. and Brunfaut T. (Eds.), *The Routledge Handbook of Second Language Acquisition and Language Testing*, Routledge, New York, pp. 201-210.

Pattemore A. and Muñoz C. 2022, *Captions and learnability factors in learning grammar from audio-visual input*, in "The JALT CALL Journal" 18 [1], pp. 83-109.

Pattemore A., Suárez M.D.M. and Muñoz C. 2024, *Perceptions of learning from audiovisual input and changes in L2 viewing preferences: The roles of on-screen text and proficiency*, in "ReCALL", 36 [2], pp. 135-151.

Pavesi M. and Ghia E. 2020, *Informal Contact with English: A Case Study of Italian Postgraduate Students*, Edizioni ETS, Pisa.

Pavesi M., Perego E., Ghia E. and Colletto V. 2019, *Assessing subtitling in a dubbing country: an empirical investigation on learner-viewers*, in "Studi Italiani di Linguistica Teorica e Applicata" 48 [1], pp. 5-26.

Peng H., Lowie W. and Jager S. 2022, *Unravelling the idiosyncrasy and commonality in l2 developmental processes: A time-series clustering methodology*, in "Applied Linguistics" 43 [5], pp. 891-911.

Peters E. 2018, *The effect of out-of-school exposure to English language media on learners' vocabulary knowledge*, in "International Journal of Applied Linguistics" 169 [1], pp. 142-167.

Polat B. and Kim Y. 2014, *Dynamics of complexity and accuracy: A longitudinal case study of advanced untutored development*, in "Applied Linguistics" 35 [2], pp. 184-207.

Puimège E. and Peters E. 2019, *Learners' English Vocabulary Knowledge Prior to Formal Instruction: The Role of Learner-Related and Word-Related Variables*, in "Language Learning" 69 [4], pp. 943-977.

Pujadas G. and Muñoz C. 2020, *Examining adolescent EFL learners' TV viewing comprehension through captions and subtitles*, in "Studies in Second Language Acquisition" 42 [3], pp. 551-575.

Reinhardt J. 2022, *The History of Language Learning and Teaching Beyond the Classroom*, in Reinders H., Lai C. and Sundqvist P. (Eds.), *The Routledge Handbook of Language Learning and Teaching Beyond the Classroom*, Routledge, Oxon, pp. 9-23.

Robinson P. 2003, *Attention and memory during SLA*, in Doughty C.J. and Long M.H.

(Eds.), *The Handbook of Second Language Acquisition*, Blackwell Publishing, Oxfrod, pp. 631-678.

Robinson P. 2011, *Second language task complexity, the cognition hypothesis, language learning and performance* in Robinson P. (Ed.), *Second Language Task Complexity: Researching the Cognition Hypothesis of Language Learning and Performance*, John Benjamins, Amsterdam, pp. 3-38.

Skehan P. 1998, *A Cognitive Approach to Language Learning*, Oxford University Press, Oxford.

Skehan P. 2009, *Modelling second language performance: integrating complexity, accuracy, fluency, lexis*, in "Applied Linguistics" 30 [4], pp. 1-23.

Skehan P. 2015, *Limited attentional capacity and cognition: Two hypotheses regarding second language performance on tasks*, in Bygate M. (Ed.), *Domains and directions in the development of TBLT*, John Benjamins, Amsterdam, pp. 123-155.

Skehan P. and Foster P. 2008, *Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the ealing research*, in Van Daele S, Housen A., Kuiken F., Pierrard M. and Vedder I. (Eds.), *Complexity, accuracy, and fluency in second language use, learning, and teaching*, University of Brussels Press, Brussels, pp. 207-226.

Skehan P., Foster P. and Shum S. 2016, *Ladders and snakes in second language fluency*, in "International Review of Applied Linguistics in Language Teaching" 54 [2], pp. 97-111.

Sockett G. 2014, *The Online Informal Learning of English*, Palgrave MacMillan, London.

Sundqvist P. 2009, *Extramural English matters: Out-of-school English and Its Impact on Swedish Ninth Graders' Oral Proficiency and Vocabulary*, unpublished doctoral thesis, Karlstad University Studies.

Sundqvist P. and Sylvén L.K. 2016, *Extramural English in Teaching and Learning: From Theory and Research to Practice*, Palgrave Macmillan, London.

Suzuki M. 2017, *Complexity, accuracy, and fluency measures in oral pre-task planning: A synthesis*, in "Second Language Studies" 36 [1], pp. 1-52.

Suzuki S. and Kormos J. 2023, *The multidimensionality of second language oral fluency: Interfacing cognitive fluency and utterance fluency*, in "Studies in Second Language Acquisition" 45 [1], pp. 38-64.

Sylvén L.K. and Sundqvist P. 2012, *Gaming as extramural English L2 learning and L2 proficiency among your learners*, in "ReCALL" 24 [3], pp. 302-321.

Tavakoli P. and Wright C. 2020, *Second Language Speech Fluency: From Research to Practice*, Cambridge University Press, Cambridge.

Tavakoli P. and Skehan P. 2005, *Strategic planning, task structure, and performance testing*, in Ellis R. (Ed.), *Planning and task performance in a second language*, John Benjamins, Amsterdam, pp. 239-273.

Toffoli D. 2023, *Learner profiles and ISLL trajectories*, in Toffoli D., Sockett G. and Kusyk M. (Eds.), *Language Learning and Leisure*, De Gruyter, Berlin, pp. 291-325.

Toffoli D., Sockett G. and Kusyk M. (Eds.) 2023, *Language Learning and Leisure*, De Gruyter, Berlin.

Vercellotti ML. 2015, *The development of Complexity, Accuracy, and Fluency in second language performance: A longitudinal study*, in "Applied Linguistics" 38 [1], pp. 90-111.

Verspoor M., de Bot K. and van Rein E. 2011, *English as a foreign language. The role of out-of-school language input*, in De Houwer A. and Wilton A. (Eds.), *English in Europe Today. Sociocultural and Educational Perspectives*, John Benjamins, Amsterdam, pp. 148-166.

Complexity accuracy and fluency in informal L2 learning among university students.
Four longitudinal case studies

181

Yuan F. and Ellis R. 2003, *The effects of pre- task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production*, in "Applied Linguistics" 24 [1], pp. 1-27.

Appendix

L2 development: observing CAF trajectories through time (preferred informal input source in brackets).
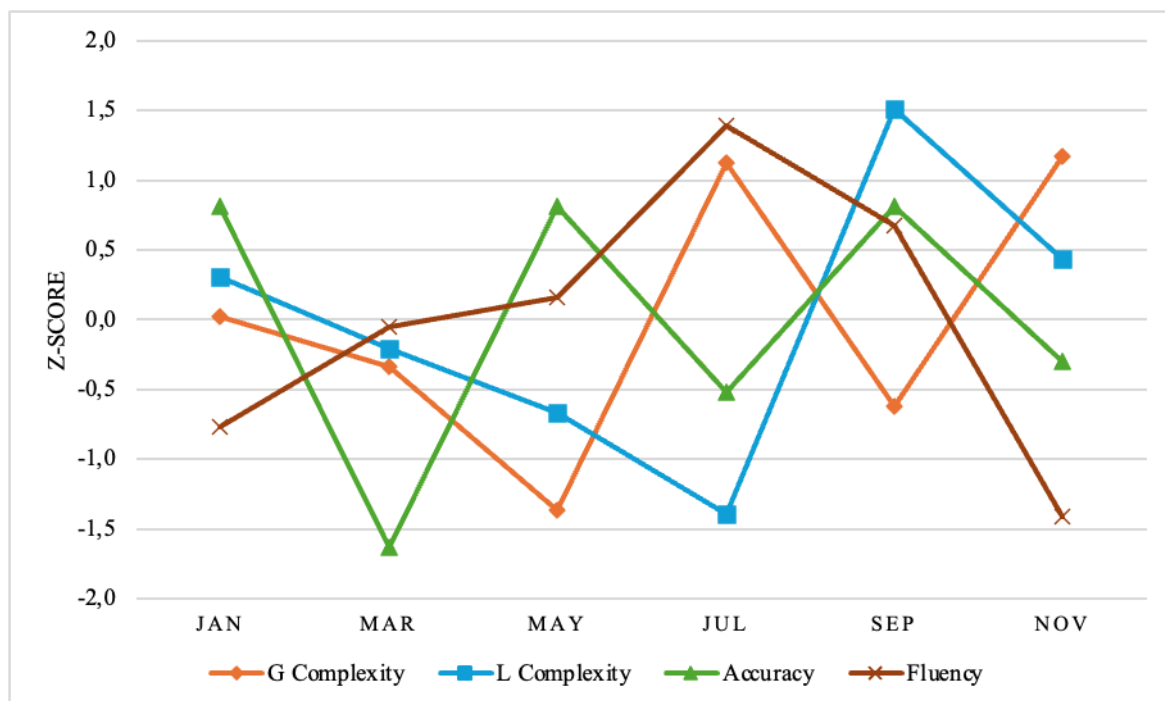
a) Findings from the case study: Delia (films and TV series)

Figure 3
CAF trajectories: Delia's case study.

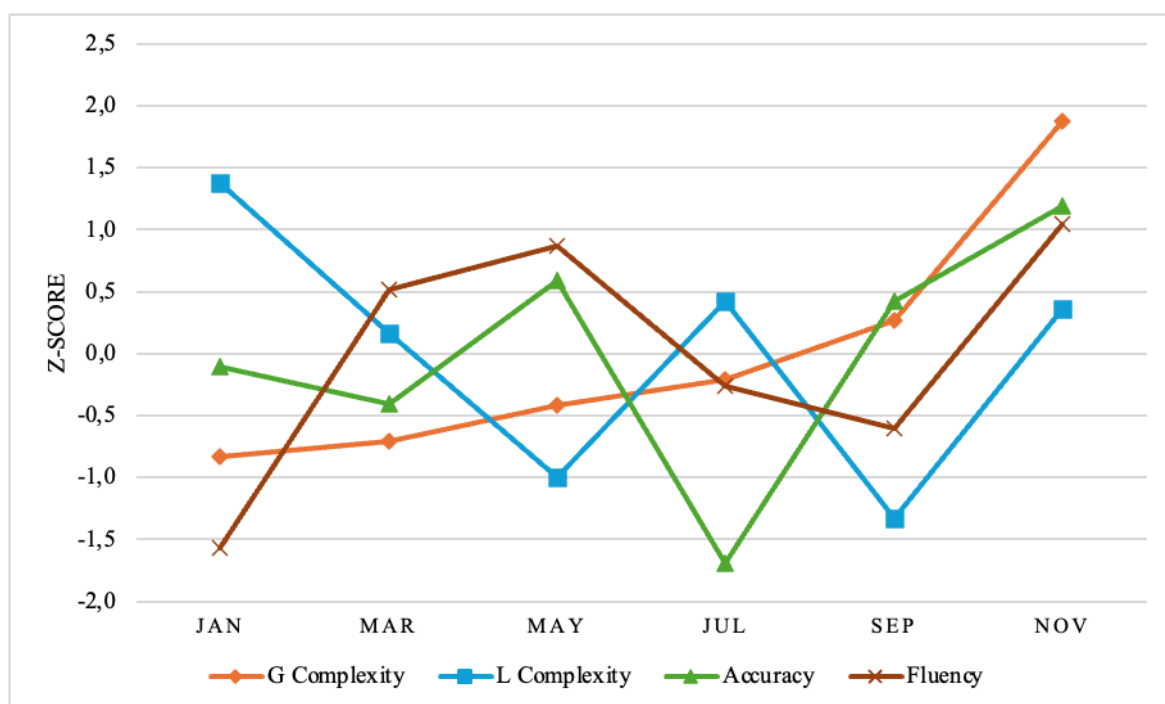*b) Findings from the case study: Grace (fantasy novels)*

Figure 4
CAF trajectories: Grace's case study.

Complexity accuracy and fluency in informal L2 learning among university students.
Four longitudinal case studies

183

c) Findings from the case study: Marta (podcasts and livestreams)
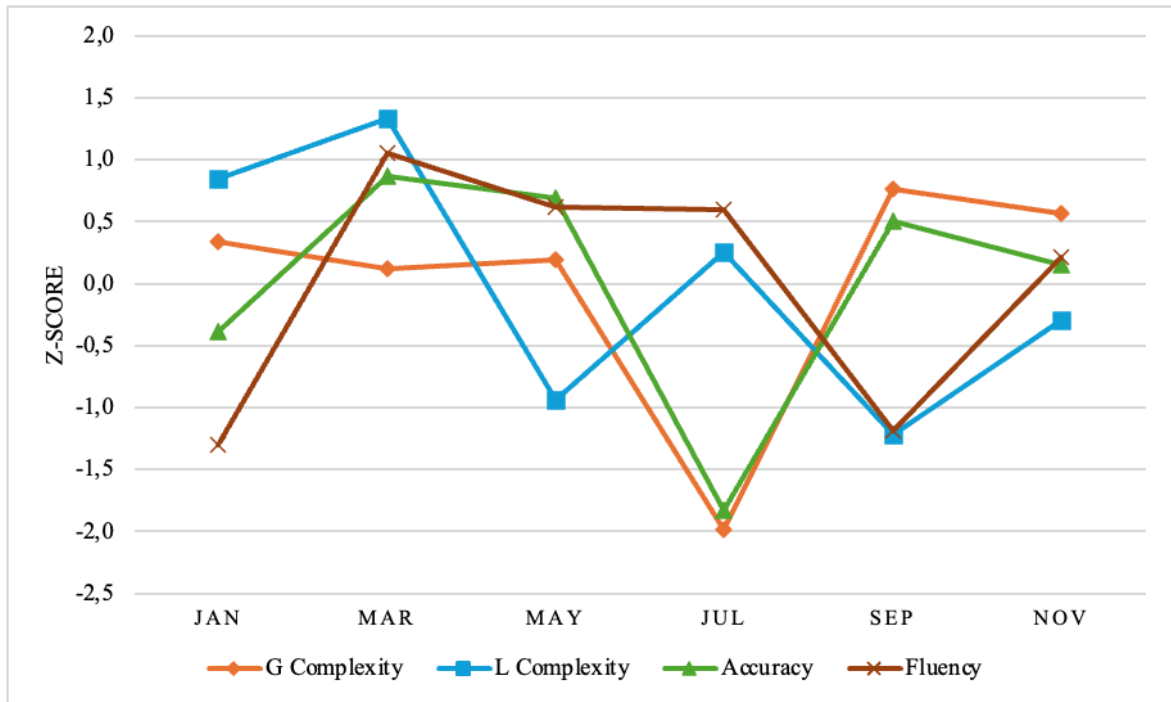

Figure 5
CAF trajectories: Marta's case study.
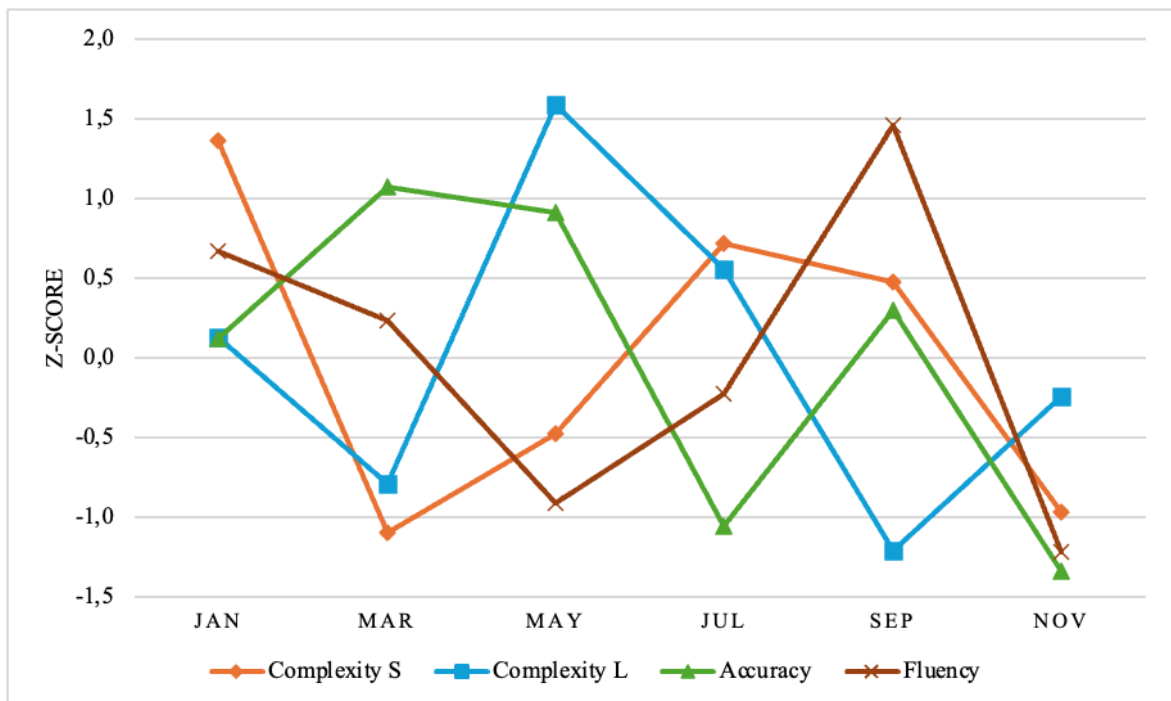
d) Findings from the case study: Lea (video games, anime and manga)

Figure 6
CAF trajectories: Lea's case study.