frontiers | Frontiers in Communication

TYPE Original Research
PUBLISHED 07 January 2026
DOI 10.3389/fcomm.2025.1532750

OPEN ACCESS

EDITED BY
Eliane Lorenz,
University of Giessen, Germany

REVIEWED BY
Tugba Elif Toprak Yildiz,
Saarland University, Germany
Ahmed Seddik Al-Wahy,
Ain Shams University, Egypt

*CORRESPONDENCE
Víctor Parra-Guinaldo
✉ parraguinaldo.v@gust.edu.kw

RECEIVED 22 November 2024
REVISED 16 November 2025
ACCEPTED 24 November 2025
PUBLISHED 07 January 2026

CITATION
Parra-Guinaldo V (2026) Emergent academic English as a lingua franca in the UAE: in-depth analysis of the ZAEBUC-50 corpus. *Front. Commun.* 10:1532750. doi: 10.3389/fcomm.2025.1532750

Emergent academic English as a lingua franca in the UAE: in-depth analysis of the ZAEBUC-50 corpus

Víctor Parra-Guinaldo*

English Department, Gulf University for Science and Technology, West Mishref, Kuwait

Comprehensive morpho-syntactic analyses in the description of English as a lingua franca (ELF) are to date not very common, especially in the Gulf region, and the present paper attempts to remedy this lacuna. Following Parra-Guinaldo and Lanteigne's (2020) study of morpho-syntactic features of transactional ELF and their classification of linguistic variants into processes and categories, this study provides a qualitative analysis of the morphology and syntax of a selection of 50 writing samples produced by first-year students of English contained in the recently compiled Zayed Arabic-English Bilingual Undergraduate Corpus (ZAEBUC) (Habash and Palfreyman, 2022). The study examines grammatical features by integrating insights from generative grammar and usage-based linguistics and situates them within the context of previous lexico-grammar studies. Based on novel uses of the language identified in the data, the paper posits the emergence of a new variety of ELF within the Gulf region (Gulf English) in that some of the linguistic variants found in the study seem *a priori* particular to this region. Important observations include the *sui generis* use of *generic* forms, *morphological reanalysis*, *anticipatory* 3rd person singular *-s*, *phantom pronouns*, and *intruding constituents*. Not only have these processes been identified and classified within the corpus, but plausible motivations behind these have also been hypothesized.

KEYWORDS

emergent academic English as a lingua franca, Gulf English, morpho-syntactic features, ZAEBUC-50, morphological reanalysis, phantom pronoun, intruding constituent

1 Introduction

Globalization has intensified the spread of English worldwide so much so that, as has long been acknowledged, "[a]n inexorable trend in the use of global English is that fewer interactions now involve a native-speaker" (Graddol, 2006, p. 87). The British Council (2013) reports that "English is the dominant international language of the 21st century" and recognizes that "it has increasingly become the operating system for the global conversation" (p. 5). In a more recent publication, Melitz (2018) examines the global dominance of English and asserts that English is spoken as a lingua franca by roughly three times as many people as native speakers. Consequently, English is most commonly used as "an additionally acquired language system that serves as a means of communication between speakers of different first languages" (Seidlhofer, 2001, p. 146). In effect, English has become the de facto *lingua franca* of science, international business, education, information technology, and popular culture around the world. For the purpose of this paper, English as a lingua franca (henceforth ELF) will entail the use of English as a medium of communication by individuals whose first language is not English, as in Seidlhofer's definition but, crucially, it will include communicative events where all participants share the same first language (L1), as may be the case in higher education institutions in the UAE.

Although a relatively recent field of study, ELF research has garnered considerable interest since the turn of the millennium. We have witnessed the creation of a few large-scale ELF corpora, such as the VOICE corpus (the Vienna Oxford International Corpus of English) (VOICE, 2021), compiled by Seidlhofer and her team from 2001; the ELFA corpus (English as a *Lingua Franca* in Academic Settings) (ELFA, 2008), started by Mauranen and her colleagues in 2003 and completed in 2008, and its companion, the WrELFA corpus (Written English as a *Lingua Franca* in Academic Settings) (WrELFA, 2015); and the ACE corpus (Asian Corpus of English) (ACE, 2020), compiled by Kirkpatrick and his team from 2009 to 2014. Other than these one million-word databases, an increasing number of smaller projects are surfacing as well (cf. Cogo, 2005; Dewey, 2003; as reported in Cogo and Dewey, 2006), such as The Brazil Corpus of English (Brazil Corpus of English, n.d.), active since 2010. One of these smaller corpora is The Zayed Arabic-English Bilingual Undergraduate Corpus (ZAEBUC), put together by David M. Palfreyman and Nizar Habash in 2019 (Habash and Palfreyman, 2022). This bilingual writer corpus comprises short essays written by first-year Emirati students, some written in Arabic and some in English. A total of 388 English essays (about 88,000 words) constitute the English part of the corpus, of which ZAEBUC-50 is a sub-corpus containing 50 of its shortest essays, selected by the author of this paper for the purpose of the present study. ZAEBUC is significant not only because it is one of very few bilingual corpora in existence, but because it provides valuable data for ELF scholars interested in an understudied region, namely, the Middle East and North Africa (MENA).

In their state-of-the-art paper, Jenkins et al. (2011) recognized a decade ago that "two clear geographical strands are emerging in ELF research: mainland European (Seidlhofer et al., 2006; Seidlhofer, 2010) and East Asia/ASEAN (Kirkpatrick, 2010; Baker, 2011)" and added that the presence of research carried out in Latin America is "lukewarm" and mostly focused on "wider rather than local issues" (p. 285). Since then, research in this region is experiencing a surge in activity, as evidenced by the release of a dedicated volume, Morán Panero et al. (2024), within the *Developments in English as a Lingua Franca* (DELF) series. The MENA region, although slowly gaining momentum, continues to lag somewhat behind in regard to research on ELF. Investigations into this region address a range of issues [see for example, Hillman et al. (2020) for a scoping review of World Englishes in the MENA region; for recent research specifically on the UAE, see for example Hopkyns et al. (2020), Siemund et al. (2020), Leimgruber et al. (2022), or Lorenz (2022)]. However, linguistic studies are still scarce, were it not for a few notable exceptions, such as are Zoghbor's (2009) comparative study between the *Lingua Franca* Core (LFC) and the phonology of Modern Standard Arabic, Fussell's (2011) brief but insightful account of distinctive linguistic features of what he terms "Gulf English," Boyle's (2011) lexico-grammatical analysis of excerpts from the *Gulf News*, a local newspaper written in English, and more recently Parra-Guinaldo and Lanteigne (2020), who analyzed the morpho-syntactic features of ELF in the Dubai/Sharjah metropolitan area within the domain of transactional communication.

The adoption of English as a *lingua franca* in the United Arab Emirates, and particularly in Dubai, across all levels of society has brought about an interesting linguistic paradigm, which ELF researchers interested in this region ought to consider as a unique opportunity for investigation. Thus, it is the aim of the present paper to contribute to this pursuit by investigating salient features in the morphology and syntax of the writing produced by first-year university students at Zayed University, a private university with campuses in Dubai and Abu Dhabi, in the UAE. The present analysis has adopted a descriptive and explanatory perspective and thus it has sought to not only identify and categorize linguistic features that may characterize the nature of the language under investigation, but, more importantly and guided by a usage-based linguistics framework, to hypothesize about plausible motivations behind the identified linguistic phenomena; for example, known factors, such as creativity and pragmatic competence, instability, and simplification, and newly proposed ones, such as *morphological reanalysis*, *phantom pronouns*, and *intruding constituents*. Therefore, the unique contribution of the present study is twofold: it expands the investigation of linguistic features in ELF into an underrepresented geographical area, the MENA region, and, more importantly, it provides a deeper understanding of the underlying mechanisms at play. This new approach is significant given that most publications of ELF grammar have predominantly focused on the identification of features, rather than on motivating factors (Ranta, 2018).

In this research, I draw upon the generative framework of sentence structure, focusing on the three fundamental layers: the complementizer phrase (CP), the tense phrase (TP), and the verb phrase (VP). This theoretical approach provides a lens through which to examine and understand the linguistic phenomena that occur when writers of ELF produce non-standard English constructions. Breaking sentences into these layers helps reveal the basic mechanisms and patterns behind linguistic variations. This conceptual foundation forms the basis for addressing the following research questions as regards the writing of first-year university students in the present study:

- RQ1: Are there any non-standard features or non-standard uses within the CP layer?
- RQ2: Are there any non-standard features or non-standard uses within the TP layer?
- RQ3: Are there any non-standard features or non-standard uses within the VP layer?

2 The MENA region and English as a lingua franca

A handful of studies have surfaced over the past few decades describing the morphology and syntax of ELF as manifested in the MENA region. One of the earliest studies, Diab (1996), analyses the writing of Lebanese university students and lists the following: agreement (verbal and nominal), articles, prepositions, word order, omission of copula, overuse of the conjunction *and*, and issues with word choice or meaning, all surmised to be errors due to language interference. Likewise, Fussell (2011) attributes a series of distinguishing features to "language transfer,[1] where Arabic functions

1 Crystal (2008) defines language transfer as "the influence of a person's first language on the language being acquired" (491).

as a 'mother-tongue' substratum": the use of the 'dummy object'[2] ("the city which I live in *it*"), a variant use of both subject and object pronouns, a preference for the masculine form, the use of the gerund form following *for* in purposive clauses ("I sometimes go to Muscat *for* shopping with my friends"), the plural *-s* with non-countable nouns and concomitant article *a* in "countable syntactic environments" ("What *a* bad luck"), and preference for *-ing* with stative verbs ("The wadi [seasonal dry river bed] *is coming* from the mountains to my house") (pp. 27–30). Additionally, "a wide range of lexical items have been directly borrowed from Arabic into" what he tentatively terms "Gulf English" (p. 30). The influence of Arabic on English, this time specifically in the UAE, has also been acknowledged by Boyle (2011), for example, in the use of –'*s* genitive with inanimate objects. He further makes a case for the instability in the system of transitivity and observes that the plural inflection is marked in the first element of a noun compound. Al_Surmi (2018) reaches the same conclusion regarding the stage English is in as used in the Arab world based on some distinctive lexico-grammatical features, in particular the use of some Muslim religious words and the frequent use of the modal *will*.

One more contribution to the description of morpho-syntactic features of ELF in the UAE is Parra-Guinaldo and Lanteigne (2020). This in-depth study analyses naturally occurring spoken interactions in the domain of transactional communication. It is suggested that the idiosyncratic nature of some of the features identified may be indicative of an emergent ELF in the area of Dubai/Sharjah. Some of the most significant features in the study are: processes of omission (conjunctions, auxiliary DO in interrogatives, direct or indirect objects), insertion (subject pronoun *you* in imperatives, *anticipatory it*), and substitution (certain pronouns, prepositions, and verbs); other features were related to tense (base form, *-ing* form), word order (left dislocation and reverse order), negation, number, and concord. A couple of examples from Parra-Guinaldo and Lanteigne (2020) relevant to the present study are: "*It*'s ok for you the *blade*?" [Is the blade ok with you?] (anticipatory *it*) and "I *no* know what you *say*." [I do not understand what you are saying] (negation and *bare* form of the verb).

In sum, a review of the aforementioned sources reveals several common and unique patterns. All but one identify issues with number agreement, with Fussell (2011) highlighting instances where plural nouns appear alongside the singular indefinite determiner *a* and Boyle (2011) observing the occurrence of plurals marked in the first element of noun compounds. Most sources point to the influence of language transfer as a major influence on these morpho-syntactic features. An overuse of particular forms was also noted, with Fussell (2011) and Parra-Guinaldo and Lanteigne (2020) both reporting an overuse of the *-ing* form, while Boyle (2011) found an overuse of the particle *to* instead. Unique features identified by Parra-Guinaldo and Lanteigne (2020), which relied on oral data, included the omission of *wh*-words, the insertion of subject pronouns with imperatives, the substitution of pronouns with the word *madam*, and the use of the bare form of the verb. For further clarification, Table 1 provides a summary of these findings, detailing the study, specific features, participants involved, and the motivations adduced by the scholars.

TABLE 1 Studies on morpho-syntactic features of ELF in the MENA region.

Study	Features	Subjects	Motivation
Diab (1996)	Errors (agreement, articles, prepositions, number, syntactic errors)	Lebanese university students	Language transfer
Fussell (2011)	Variants (dummy object, sub. & obj. pronouns, purposive gerund)	Oman college students	Language transfer
Boyle (2011)	ELF features (inanimate '-`s', *overuse of to clause*, unstable intransitivity, number)	News outlet	World language divergence
Al_Surmi (2018)	Distinctive features (adverbial position, overuse of modal *will*)	Bahrain news outlet corpus	Distinctive features
Parra-Guinaldo and Lanteigne (2020)	Tense (present progressive, will + *-ing*, BE + bare form), word order (adj. with trans. verb, reverse order with BE, noun + adj.)	1st year university students	Transfer/ idiosyncratic variants

By comparing these features with the findings of the present study, this work seeks to identify any emerging features that may offer novel insights into the communicative strategies within the context of emergent academic[3] ELF writing in the MENA region.

3 Methodology

3.1 X-bar theory

The generative framework, and in particular X-bar theory, adopted in this study follows the approach outlined in a previous publication (Parra-Guinaldo and Lanteigne, 2020). The

2 As rightly noted by one of the reviewers, the pronoun *it* in constructions such as *the city which I live in it* is more accurately described as a resumptive pronoun, rather than a dummy object; the latter typically referring to non-referential uses of *it* (e.g., *It is strange that*...). I fully agree with this clarification. The term *dummy object* is used here following Fussell's (2011) terminology, and this usage is retained to reflect his original framing.

3 While the students' output may not yet reflect fully developed academic English, the data was collected in an academic setting (produced by university students in an instructional context) and reflects early-stage academic engagement.

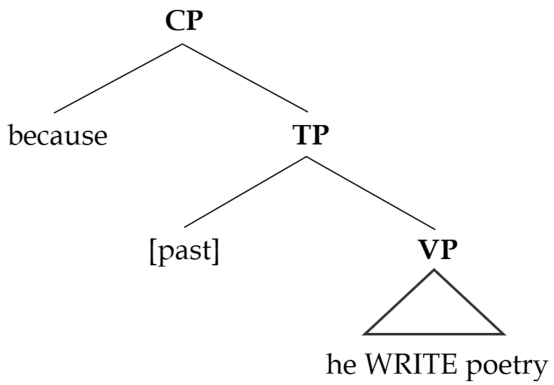

FIGURE 1
Basic structure of a sentence based on X-bar theory.

classification of tokens[4] into syntactic layers or domains is useful, because it helps explain the hierarchical nature of language and their functional properties. In a generative framework, the clausal structure is divided into the complementizer phrase or CP (the outer layer), the tense phrase or TP (the intermediate layer), and the verb phrase or VP (the inner layer). The CP corresponds with the discourse domain and is where illocutionary force is encoded (an example of complementizers is elements joining clauses, such as conjunctions); the TP is the layer carrying grammatical information (as opposed to lexical), such as negation, tense, and aspect; finally, the VP is the layer containing information about the verb, i.e., its arguments (subject, direct and indirect object). Accordingly, the tokens found in ZAEBUC-50, a total of 584, have been categorized consistent with the three structural layers commonly used in generative grammar: 10 tokens within the CP, 131 within the TP, and 388 within the VP. Additionally, 36 instances of non-standard repetition and 19 instances of non-standard word order were identified; these two categories can potentially occur at any of the three clausal levels, so I describe them separately under the label *cross-layered categories*. Figure 1 depicts the basic structure of a sentence according to X-bar theory, where the CP layer is occupied by the conjunction *because*, the TP by the grammatical information conveying past tense, and the VP comprising the base form of the verb (represented here with capital letters) and its main arguments (subject *I* and direct object *poetry*).

X-bar theory is a basic, but powerful, tenet of generative theory, which facilitates the analysis of grammatical elements within the sentence and ensures reliability and consistency. For example, Parra-Guinaldo and Lanteigne (2020) identify an underutilized CP, and this impoverished layer was interpreted as evidence of syntactic simplification, which is one of the characteristics of ELF transactional communication. Similarly, the present study employs the generative framework, particularly the X-bar model, offering a principled means of classifying linguistic deviations according to their syntactic locus, whether at the CP, TP, or VP level. This approach is particularly useful in the description and explanation of such phenomena as the phantom pronoun and intruding constituents, where structural reassignments are reflected within the VP layer. By situating these phenomena within a layered architecture, the study articulates a theoretically grounded explanation of their underlying structure and motivation.

3.2 A note on usage-based linguistics

Generally speaking, and with no further implications other than its basic premise, the motivations herein proposed for the processes identified in this study can be said to fit within a usage-based linguistics framework, whereby

> language is conceptualized as a cognitive resource constructed and continuously developing on the basis of analyses of the frequency and distribution of form-meaning pairings in the input experienced during usage events [and where] systematic patterns of language are not determined "top-down", as rules conforming to hard-wired universal principles, but rather emerge "bottom-up," on the basis of variable, socially contextualized, individual experience. (Hall, 2018, p. 75)

In practical terms, the usage-based approach is interpreted here as the language user's affordances to establish *ad hoc* communicative strategies. Rather than attempting to comply with normative standards, the students seem to resort to a variety of available linguistic mechanisms that deviate from standard English, but facilitate and even enhance the conveyance of meaning. The usage-based approach informs the analytical process by recognizing the value of actual language use in shaping linguistic phenomena. While this approach is not explicitly developed in the subsequent analysis, it is referenced here to foreground students' adaptive language strategies, which remain a consistent undercurrent throughout and may be explored more fully in future work.

4 The data

4.1 The source for this study, ZAEBUC

Corpora of English texts written by Arabic speakers in the Middle East are scarce, so a recent contribution, such as ZAEBUC (see Habash and Palfreyman (2022) for a detailed description of the corpus), is an indispensable source for the present study.

ZAEBUC is a bilingual annotated corpus comprising short essays written by first-year university students at Zayed University, a public university in the United Arab Emirates. These essays were collected during the fall semester of 2019 across several campuses (Abu Dhabi and Dubai's female and male campuses). The total number of essays in English is 388 (about 88,000 words) and 214 in Arabic (33,000 words). It is a unique corpus in that it focuses on bilingual writers, but rather than being a parallel corpus of texts with their translations, it contains samples of students' writing both in English and Arabic. The corpus is annotated for parts of speech and lemmas, but it also includes the following meta-data features: ID, gender, school

4 The term "token" refers to instances of non-standard linguistic forms.

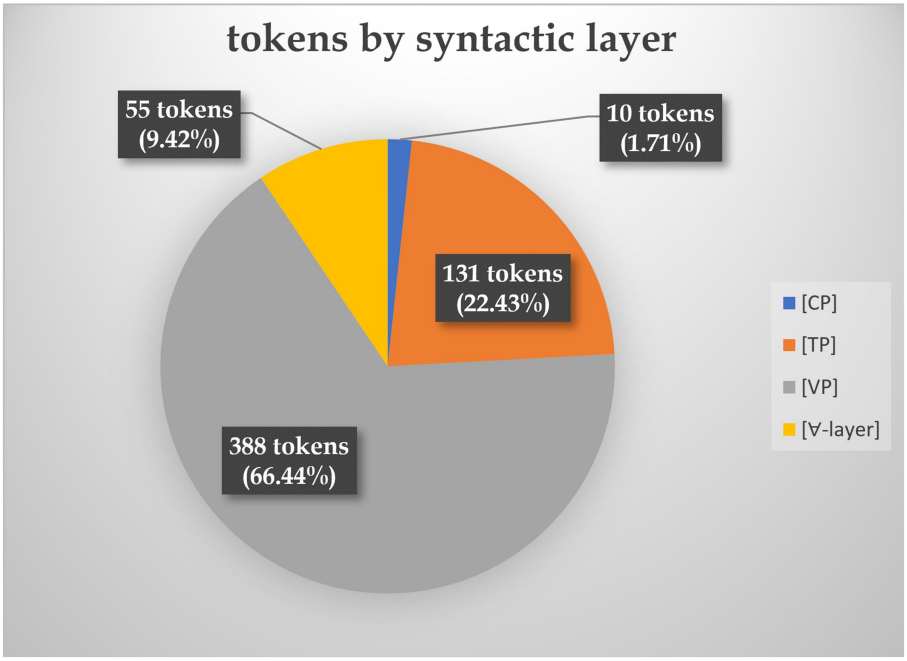

FIGURE 2
Distribution of tokens by syntactic layer.

type, school language, address, campus, course, section, handwritten, date added, date last saved, and topic.

All essays were written for ENG-140, English Composition I, a course that introduces students to academic reading and writing strategies and practice. The student catalog at Zayed University indicates that the students should strive in this course for correct grammar, rhetorical modes and so on; therefore, the course can be considered normative-dependent, assuming Standard English as the norm, rather than encouraging the use of ELF, for which there is no mention at all (ZU catalog, 2021–2022).

4.2 About the sub-corpus, ZAEBUC-50

Out of a total of 388 essays written in English in the main corpus, 50[5] of the shortest essays were selected for the sub-corpus used in the present study, with the assumption that students who are not as productive in their writing are the ones deviating more from the standard, and therefore, generate a greater number of tokens. A cursory examination of larger pieces seemed to corroborate this initial conjecture. The 50 texts that made it into the sub-corpus will be referred to as ZAEBUC-50, and throughout the data analysis and discussion, I may refer to these texts by this name or simply by "the corpus" or "this corpus." ZAEBUC-50 came to a total of 5,638 words; the mean (average) of words per text is 112.76, the mode (or most frequent number) is 135, 109, and 139 words per text (3 times each), and finally, the minimum number of words per text is 7, and the maximum 172.

5 Data analysis and discussion

While the study includes the frequency of instances for each linguistic feature examined, the primary focus remains on offering a nuanced interpretation and contextual understanding of these features, consistent with the qualitative nature of the study. The current paper is intended as a synopsis of a large data-intensive study on ELF as observed in written samples of English produced by student speakers of Arabic in the ZAEBUC-50 sub-corpus. Although there were numerous examples of the processes that are suggested as operative in the writing produced by these students, space restrictions of this paper require that the number of examples be limited to those that adequately explain the linguistic phenomena herein discussed. Additional examples of the processes and phenomena proposed here may be found in a more comprehensive article which is planned for publication at a future date.

5.1 Distribution of tokens

Figure 2 illustrates the distribution of tokens according to syntactic layer. Most of the tokens identified occur within the domain of the verbal phrase, with 388 (or 66.44%) of a total of 584 tokens. Second in importance, although by a distant margin, is the tense

5 The decision to analyze 50 texts was based not only on the fact that shorter texts tend to exhibit a higher proportion of deviant forms than longer texts, but they are on average 112.76 words each, and to the researcher this seemed a manageable amount to carry out such an extensive analysis.

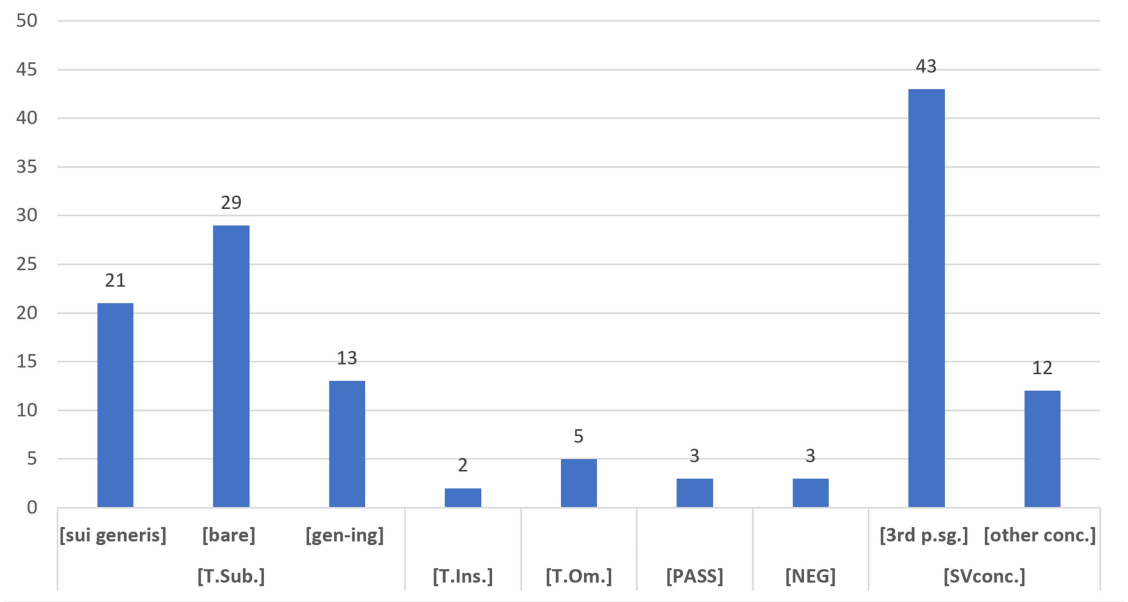

FIGURE 3
Distribution of tokens within the TP layer.

phrase layers, with 131 (or 22.43%) tokens. The complementizer phrase contains only 1.71% (10 tokens) of all tokens. Finally, and indicated by the symbol ∀, which stands for "for all" or "for any," to borrow a useful mathematical sign, is a segment representing tokens that may affect any or all of the three layers; these cross-layered categories [∀-layer], including repetition and word order, account for almost 10% of all tokens.

The description of the complementizer layer or phrase is straightforward; it mainly contains omissions leading to asyndeton,[6] which characterizes a more direct style, with juxtaposed rather than subordinated clauses. As for the other two layers, the verb phrase and the tense phrase, a more detailed illustration would be useful. The two main issues within the TP layer, as depicted in Figure 3, deal with concord, including the 3rd person singular *-s* (3rd p.sg. *-s*, henceforth), and substitutions of tense, as when for example a bare form or the *-ing* form take the place of other tenses. For the reader's convenience, an Appendix has been compiled to present a typology of the analyzed features, including the number of tokens by category, providing a structured reference for the codes employed in the analysis.

As for the VP, not only does this layer contain most of the tokens, but the range of tokens is quite diverse. Thus, Figure 4 shows 12 main categories, among which issues with determiners, number, and prepositions are the most numerous. The following sections will describe each of these categories in turn.

5.2 [CP] The complementizer layer (10/584 tokens or 1.71%)

Out of a total of 10 instances within the complementizer layer or phrase (CP), 6 omissions, 3 insertions, and 1 substitution were identified. Asyndeton occurs when the student omits a coordinating conjunction and clauses are juxtaposed rather than linked by elements in the surface structure.

The non-standard insertion of complementizers, with only three occurrences, is not a common feature in the corpus. On two occasions, two different complementizers seem to be competing for a more felicitous structure. For example, when the student writes (1) "that the most important development in our country is *that to* make rocketship to make fake moon and to see if people can life there or not." [192479],[7] she may be replacing the complementizer *that* with *to* in a process known in ELF research as *repair*.[8] Another student simply inserts a seemingly unnecessary complementizer *that* to introduce a clause with a gerund, (2) "one of the most important affect is *that* know*ing*

6 Asyndeton is the omission of conjunctions, leads to clauses being juxtaposed rather than linked by elements in the surface structure, as in Julius Ceasar's celebrated aphorism "Veni, vidi, vici" *I came, I saw, I conquered*.

7 The examples presented in this paper are in line with The Chicago Manual of Style. Quotations from research participants are enclosed in quotation marks, and the standardized versions, where applicable, are provided in square brackets. Additionally, each example is followed by a number in square brackets, which corresponds to the text identification number from the original corpus. Elements relevant to the discussion are italicized by the author to emphasize their significance.

8 In discourse analysis, the attempt an L2 language user makes to rectify a real or perceived deficiency, whether self-initiated or other-initiated, is commonly known as *repair*.

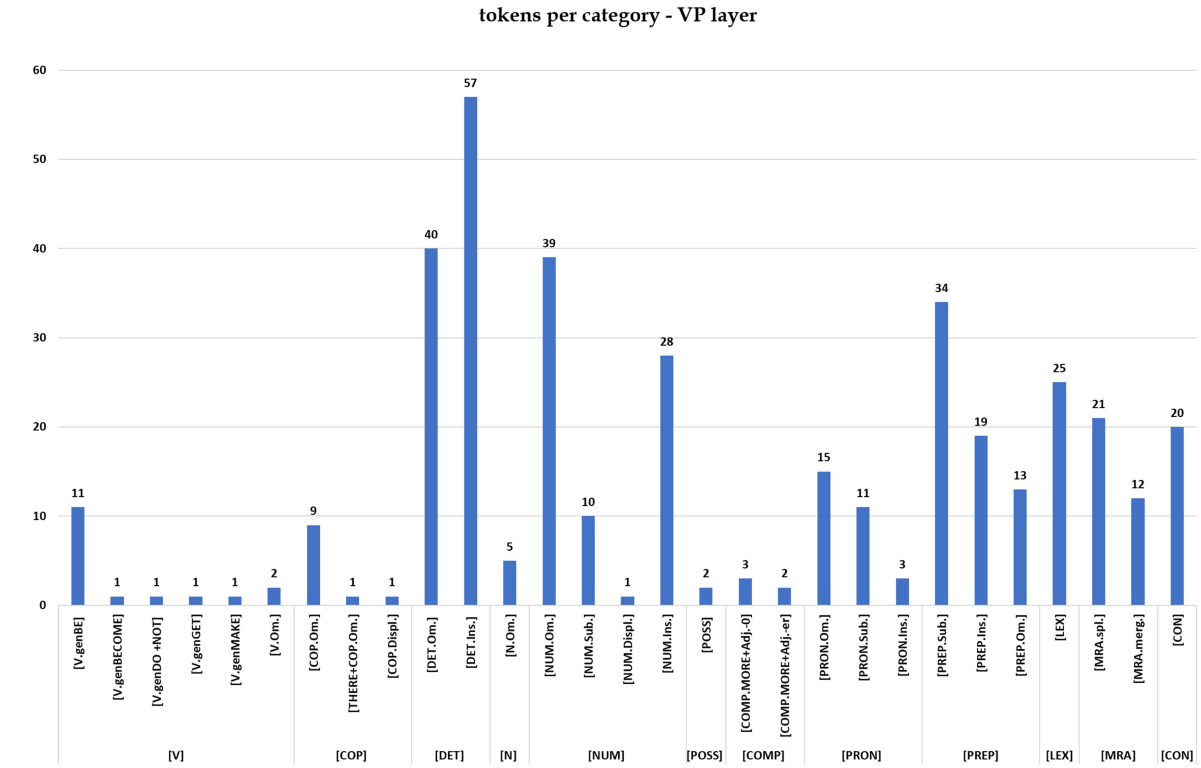

FIGURE 4
Distribution of tokens within the VP layer.

people all over the world." [18992]. This case of *fluctuating grammar*[9] may be motivated by the initial choice of a construction the student is familiar with, *it's important that*, only to *self-repair* with a more relevant construction. Fluctuation in second language (L2) can also be explained by Mauranen's (2012) observation "that L2 speakers' second languages, especially language forms, are not as deeply 'entrenched' in their minds as their L1's" (p. 4, cited in Ranta, 2018, p. 245).

The only instance of a possible complementizer substitution, (3) "What is also good about it is *the* you also can know news" [89953], may simply be a misspelling mistake, *the* for *that*. An alternative explanation may be that the student is (subconsciously) aware of the need for a complementizer, but not completely familiar with the complementizer itself, resorts to *lexical approximation*, *the* being the closest element to *that* in the student's mental lexicon. Likewise, Mauranen (2009, 2012; cited in Wang and Kaatari, 2021) found that ELF users may resort to approximation of conventional forms resulting in linguistic variability.

Along with asyndetic syntax, which in and of itself denotes absence of linking constituents, the paucity of tokens in the complementizer layer is an indication that this is underutilized in the data under investigation. Similarly, Parra-Guinaldo and Lanteigne (2020) claim the presence of an impoverished CP in spoken transactional communication in the same region (Dubai/Sharjah), sometimes due to left-dislocated elements[10] competing for the argument position or more commonly to the number of omissions in this layer of the syntactic structure of the sentence.

5.3 [TP] The tense layer (131/584 tokens or 22.43%)

In the present study, Negation [NEG], Subject-Verb Concord [SVconc], and Passives [PASS] are three types of non-standard variants within the Tense layer of syntactic structure. Issues related to tense in the traditional sense (i.e., time reference) are considered first and they are categorized as Tense *proper* or [T]; non-standard variants under the other three categories are described thereafter. The rationale behind placing Verb [V] and Copula [COP] under the Verb layer is to draw a distinction between issues related specifically to grammar, such as the choice of tense, and issues to do with the choice of verb (or the decision to use or not a verb in the first place), which are more closely related with the student's mental lexicon, and the relation to its arguments (whether a verb is transitive or intransitive). By itself, category [T] *proper* (including omission, insertion, and substitution of Tense) accounts for 70 occurrences, or roughly 50% of all occurrences in [TP], but if we add the other three categories, [NEG], [SVconc], and

9 In this paper, *fluctuating grammar* refers to the apparent capricious choice of grammatical elements, as when the ELF writer is not sure what the right element is to introduce a complement clause, whether the subordinate conjunction *that* or the infinitive marker *to*.

10 Left-dislocated elements are constituents, such as noun phrases or clauses, that are placed at the beginning of the sentence to express emphasis, as in "*That* I did not break" (although other things may have gotten broken along the way).

[PASS], then the sum of occurrences of [T] adds up to 131, which is equivalent to close to a quarter of all occurrences in ZAEBUC-50.

It is hard to disassociate Tense from Mood or Aspect,[11] since these functions may all be carried by the same form. While the focus in the present analysis is on the temporal function of the verb phrase, a few instances of non-standard use of modals have also been addressed [for further insights into the emergent use of modal auxiliaries in ELF, see for example, Laitinen (2020)].

5.3.1 [T] Tense *proper*

The issues found in the corpus regarding Tense *proper* are varied. [T] *proper* comprises about half of all [T] occurrences, of which the majority have to do with the student choosing a tense that would be unexpected in Standard English, but there are also a few instances of omission and insertion. The three occurrences to do with the passive voice are assigned a subcategory of its own, since these may fall under any of these mentioned types of [T] occurrences.

5.3.1.1 [T.Sub] *sui generis*. Substitution of tense

The substitution of Tense forms refers to the situation where the student opts for a tense form different from what would be expected in Standard English. This type of Tense substitution is referred to as [T.Sub] *sui generis*, since the student has at her[12] disposal a repertoire of tense forms *of her own* to express a variety of temporal expressions, which may or may not be different from that of a native speaker's. Following are sentences produced by students representing instances of tense substitution *sui generis* as identified in the corpus.

5.3.1.1.1 Simple past instead of present. Example (4) illustrates how the student alternates between the past and the present tense to describe what seems to be a general statement about the benefits of using social media.

> (4) It also has a lot of benefits in it, for example knowing about the news and how everything is going on and its also esair for everyone because people *had* to get out to buy some newspapers and it takes a lot of time and probebly half of the pepole *were* lazy. [89953]

General conclusions are drawn by inductive reasoning, and it may very well be the case here that the student resorts to the past tense, "people *had* to get out to buy some newspapers," to indicate what seems to be past events from which to draw a general statement, "knowing about the news and how everything is going on." Thus, the apparent fluctuation of tense (as manifested in the surface structure) in this instance seems justified, since it is the result of formally encoding the student's contextual argumentation. Consequently, it is argued that this is a good example of pragmatic competence. This interpretation aligns with Young's (2011), whereby pragmatic competence is defined "as the ability to negotiate meaning in a flexible, adaptive manner and to co-construct a communicative act" (cited in Taguchi and Ishihara, 2018).

5.3.1.1.2 [T.Sub.B] bare form of the verb. The most frequent type of Tense substitution, with a total of 29 occurrences, involves the use of the bare form of the verb or [b], that is, a verb with no morphology,[13] to express a variety of tenses. For example, in (5) "Moreover *in the past* people usually *go* out if they *want* to by anything" [193451], where the adverbial *in the past* clearly establishes an action in the past, but the verbs *go* and *want* display zero morphology. It is precisely the use of the adverbial that renders inflectional morphology unnecessary; hence, the use of the bare form.

5.3.1.2 [T.Sub.gen-*ing*] bare form instead of gerund

The use of *-ing* in ELF has drawn the attention of some scholars; for example, Ranta (2006, 2009) observes that the *-ing* form is used more extensively than in Standard English and that "in addition to the standard-like uses, it may also have its own peculiar function in ELF based on its 'attention-catching' form" (2006, p. 114) and rejects the idea that L1 interference is the only, or even the main, motivator. In all, I found 13 instances where an *-ing* form was used in place of another expected form. For example, (6) "thy take her tima and *beaing* lazy at *studing* thy canit foucas at studying" [189349] [They take their time and are lazy at studying]. The presence of other *-ing* forms, *studing* [studying] twice, may have influenced the student's choice. In this other sentence, (7) "Fainly i think should be caerfull and do not *using* you phon" [189349], the use of an *-ing* form seems a bit odd following an auxiliary plus a negative marker, DO + NOT,[14] but it is not an isolated case; this construction is found in two sentences produced by different writers, (8) "they dont *have in* save streerts for bicycles" [190745], where I assume "have in" means [having], and (9) "they do not *playing* sports" [194708]. Presumably, the construction DO NOT is used as a default negating marker (not just the negative marker *not* in itself, but both this and the auxiliary, so a two-element construction) regardless of the verbal form it is negating. One of the non-standard uses of the *-ing* form identified in Ranta (2006) is the reference to a point in the past, of which not a single instance was found in the present corpus, but she provides a convincing explanation, I believe, to the question "what, then, could explain the 'attractiveness' of the *-ing* form in L2 use?" (p. 111); she proposes that a new function is at play, expressivity or clarity of expression, by means of "prominence and salience in the speaker's utterance" (p. 112). In the present data, the preference for the *-ing* form *in lieu of* other temporal forms has been observed, and whether prominence of

11 The verb in "The kid *had been running* all day before he sprained his ankle," carries a past tense (an earlier event), an indicative mood stating a fact, and a continuous aspect (marking the ongoing nature of the action in the past before a second event, in a more recent past, happened).

12 The decision to use feminine as the generic gender for this paper was mainly guided by the fact that all but 2 of the essays analyzed were produced by female students. In any case, whether gender plays a role in the use of non-standard forms is subject for further research.

13 The terms "no morphology" and "zero morphology" in generative grammar refer to the absence of overt, or visible, morphological markers. In such cases, grammatical features, such as tense or number, are conveyed covertly or via adjuncts (adverbials) and other constituents.

14 An element presented in all capital letters, such as DO, represent its bare form, which can then be morphologically realized in different ways depending on the grammatical context. Thus, DO could take the form of *did* when expressed in the past tense.

form or prominence of function, ELF users seem to gravitate toward a trust-worthy, *quasi* do-it-all verbal form.

5.3.1.3 [T.Ins.] Insertion of T(ense)

Under this subcategory, cases of non-standard grammatical items were included. When the student writes (10) "Social media has alot of effects so. we should *to* be carefull about our selfes" [191587], it may be the case that two available T forms expressing deontic modality are competing, the result being a hybrid construction SHOULD + TO + BE.

5.3.2 [PASS] Passive

Two issues were found with the passive voice, either part of the morphology is absent, as in (11) "people can be *affect* their social media and socicty" [289115], where the inflectional suffix *-ed* of the past participle and the preposition introducing the agent *by* are absent, or the passive voice is used where in Standard English it would not be, as in (12) "Also it *can been effect* their future" rather than "It can also affect their future" [289115]. These two types of non-standard uses of the passive are found in Björkman (2008), even though, as she demonstrates, cases of "deviant passive voice," as she calls them, are "rare in spoken university registers," a statement already made by Biber (2006, as cited in Björkman, 2008, p. 113).

5.3.3 [NEG] Negation

Only one occurrence of negation by means of a particle NO preceding the main verb was found, (13) "that can let the parents *no* take a bad idea about social media." [168718]. This strategy is well known in English as a foreign language, as well as in ELF, but the other two occurrences of non-standard negation may prove more enlightening, (14) "they will *dosent* live their child life" [190122] and (15) "they *dont* have in save streerts for bicycles" [190745], with an understood [they are not having safe streets], where Auxiliary DO (whether in the singular or the plural form) is used generically for the purpose of negation, obviating negation on auxiliary WILL or auxiliary BE. As previously pointed out, the construction DO + NOT seems to be used as a default marker for negation regardless of the verbal form it negates: a modal verb, verbal forms ending in *-ing*, a simple future, or a simple present.

5.3.4 [SVconc] Subject-verb concord

Issues regarding the agreement between the subject and the verb are included in the category [SVconc], where "conc" stands for concord.

5.3.4.1 [3rd p.sg.] Third person singular *-s*

Most of the occurrences of this type of variant reflect the omission of *-s*, 37 out of 43 occurrences of all [SVconc] cases or 90%. The rest of the tokens include insertion, insertion and displacement, and displacement.

5.3.4.1.1 Insertion of *-s*. On two occasions, *-s* is added to the main verb even though the subject it is in concord with is a plural noun, N.pl + V-*s*, as in (16) "*the actions* of these individuals on social media *affects* the society entirely" [83847]. Notice the distance between the head of the noun phrase and the verb in both sentences. I would argue that whenever the syntax becomes complex (for example, when phrases are embedded within other phrases), the student resorts to whatever tools she has at her disposal to convey her thought fully despite this complexity (i.e., to alleviate the *cognitive load* of keeping track of long-distance dependencies; for valuable insights into syntactic complexity in L2 writing, see Jiang et al., 2019). In practice, this may mean that the language user will opt for *generic (or multi-purpose) forms*, such as the bare form of a verb [V.b], the *-ing* form to indicate a variety of tense settings, and so on. In this case, it seems that when the noun phrase (NP) is complex (i.e., contains more than a single binary branch) and the verb is not adjacent to the head (N), the solution is to resort to a default subject, which I assume is an implicit *it*. We can call this implicit pronoun *resumptive it*, since it "repeats or in some way recapitulates the meaning of a prior element" (Crystal, 2008). The use of resumptive pronouns is not uncommon in informal Standard English, as when someone says "Mary, I know *her*" (also from Crystal, 2008, p. 415; emphasis added), but the reinterpretation of the subject obviating number is rather innovative. In her study of English as a European lingua franca, Breiteneder (2005) refers to the principle of proximity to explain the "overprovision of the *-s* marker" (p. 16) where the verb appears in concord with a noun in closer proximity with the verb. This principle of proximity is defined as "the tendency for the verb to agree with a noun which is closer to the verb […] but which is not the head of the subject phrase" (Biber et al., 1999, p. 189, as quoted in Breiteneder, 2005, p. 15). Similarly, the principle of proximity could explain the presence of an *-s* marking given an implicit resumptive pronoun in proximity with the verb, with an important difference; in this case, the proximal subject is implicit and therefore it does not show in the surface structure.

5.3.4.1.2 Displacement of *-s*. An anticipatory 3rd p.sg. *-s* is placed on the agreeing subject rather than on the verb, as in (17) "overall the negativ ways *its take* tima from the people" [189349]. Although I found no literature on this phenomenon, I have noticed it anecdotally on numerous occasions in interactions between ELF users from the MENA region. Further investigation into the use of the anticipatory 3rd p.sg. *-s* is needed.

5.3.4.1.3 Omission of *-s*. As was mentioned above, the omission of 3rd p.sg. *-s* is by far more common than insertion or displacement. A common example entails the presence of the pronoun *it*, as in (18) "it *have* a good affect" [188847], but nouns also appear, (19) "*The government* spend alot of money" [294378]. It is tempting to think that the student is treating *the government* as a group noun agreeing with the verb in the plural, as is the case in Standard British English, but the student is most likely not aware of this type of plural. Interestingly, not even a resumptive *it* triggers the 3rd p.sg. *-s*; for example, (20) "*The social media it have* many affect for the people" [188847]. This sentence shows that the student may be treating *social media* as a noun in the singular rather than the plural of [medium] (although it is also possible that *it* is a resumptive *it* that obviates number, as explained above).

5.3.4.2 [SVconc\3rd p.sg.] Subject-verb concord other than third person singular *-s*

Other than issues with 3rd p.sg. *-s*, 12 occurrences were found in which the subject does not agree with the verb as would be expected in Standard English. Borrowing from mathematics the symbol "\," which stands for "remove from a set," the formula heading this subcategory reads as "instances of subject-verb concord, excluding 3rd p.sg. concord instances."

A point in case is (21) "things that *we wasent* nknow it" [192218], where a plural pronoun is in dual co-reference relationship first with an auxiliary BE marked in the singular and (concurrently?) with the main verb KNOW with no inflection. Like the use of DO + NOT as a clausal negative marker (as opposed to a nominal negative marker, such as "*no* one"), BE may be used as a generic auxiliary to negate the main verb KNOW; assuming a generic WASN'T with a singular feature as a default number would explain why a plural pronoun *we* would be paired with *wasn't*. Although Arabic can be considered a copula-less language, the copula does appear in some restricted syntactic environments, one of which being a special negative construction "lays-/las-," as in "*Laysa* tilmiðan" [He *is not* a student] (Ferguson, 1971). It follows that the student may be mirroring this special construction to negate a lexical verb. The constructions in L1 and L2 are not identical, so we can infer that the student is *repurposing* a construction from her L1 to achieve a similar, but not identical, function; i.e., clausal negation via a negative marker. If this assessment is correct, this phenomenon may be considered a case of language transfer.

5.4 [VP] The verb layer (388/584 tokens or 66.44%)

5.4.1 [V] The verb

The category Verb [V] includes occurrences considered non-standard in English involving the choice of verb or the presence or absence of a verb. Specifically, this section deals with the lexical aspect of the verb, rather than its grammatical morphology and function. Therefore, this category excludes the morphological agreement between subject and verb (whether 3rd p.sg. or otherwise) and morphology indicating tense (or voice). A variety of verb choices are noticeable in the corpus; in particular, the use of BE as a generic verb, with more than half the occurrences of this type. Generic verbs may appear by themselves, in the company of another verb, with or without its own subject.

5.4.1.1 [V.gen*BE*] Generic BE

In (22) "Always we have to be careful about what we post on it, who's people I'm following, *is* I'm on a right place or?" [192397], the student resorts to using a generic BE *is* before constructing an interrogative. This may serve an *interrogative marker* signaling the force of the clause (i.e., interrogative polarity item). The last clause in (22) could have been an indirect question, a subordinate clause along with the previous ones, relative clauses. Instead, the student ended up creating a direct interrogative (grammatically independent, but semantically dependent).

Generic BE seems to aid in the construction of verbal phrases, particularly the ones with complex constructions (an NP with more than a single N as head or a VP with a tense other than the present). A common occurrence is the presence of BE between the subject and the main verb, as in (23) "Social media *is* affect alot about the people" [119150]. BE seems to be a *default wild card*, taking a temporary place for the verb, before the more complex part of the sentence, the predicate, is construed. A default wild card in this context could therefore be defined as a grammatical element which, by its generic nature, can take on the role of any other grammatical element within its same grammatical category. In this instance, BE functions as a place holder for any other verb. Verbs other than BE were also used in the present corpus as generic verbs, such as BECOME, DO, GET, and MAKE. The 'overuse' of certain verbs of high semantic generality was already identified by Seidlhofer (2004), but the place holding function of the default wild cards described here seems an innovative resource at the language user's disposal.

5.4.2 [COP] Copula

In traditional grammar, *copula* refers to the verb BE when this is interpreted either as a lexical (even if with minimal semantic value) or as a connective verb (Crystal, 2008; but see Payne (2011) for a different view). Since the grammatical features of BE have been discussed above, I will focus here on the semantic value of the verb. The absence of copula is a common phenomenon in languages around the world and it is no stranger to informal English or regional varieties of English either, as in "You the man!" or in "And your mom, how is she? *She fine* too." As Ferguson (1971) puts it, "all natural languages have grammatical machinery for equational clauses, but the details vary considerably from one language to another," and suggests "two main types of language as far as copula clauses": Type A, which includes languages with "a copula in all normal neutral equational clauses" and where "the absence of the copula is limited to certain set expressions" and Type B, "normally" with "no copula in equational clauses […] when both members of the clause (subject and complement) are present, the clause is timeless or unmarked present in time, the complement is attributive (i.e., adjectival rather than nominal), and the subject is third person" (pp. 141–142). Arabic thus can be classified as a Type B language, for which the use of explicit copula is restricted to certain syntactic environments or constructions.

5.4.2.1 [COP.Om.] Omission of a copula

The copula was absent in 12 instances in ZAEBUC-50. For example, (24) "and thagt really good and helpful" [89953] [and that is really good and helpful] and (25) "some people famous in a good way but some people *are* reverse" [119150], where the copula was omitted in the first clause of the coordinated construction, but not the second. Instability in the use of certain elements is noticeable in the data; this is particularly evident when the same element is present in one clause, but absent in the next (especially when coordinated).

5.4.2.2 [COP.Displ.] Displacement of a copula

Although it is hard to know what the student meant exactly in (26) "Some of us take it a life and a job now and it those who'*s* get affected the most" [219833], a possible interpretation could be [Some of us take a life and a job and it *is* those who *are*/*get* affected the most]. It looks like BE was first omitted in *it those*, then it surfaced *in those who's*, and finally competed against another copular verb, GET. From the data analyzed so far, a modest, but nonetheless conspicuous element, *–s*, whether a contracted form of BE, as in (26), or the 3rd person singular form of the present simple, as in previous examples, is not only multifunctional (e.g., interrogative marker), but flexible in so far as its landing site is not fixed. Floating elements are known in Standard English, such as the quantifier *all*, which can take different positions in the sentence. I would argue then that to the language user, albeit subconsciously, *-s* is in fact a floating element with important functions but flexible placement, i.e., a floating *-s*.

5.4.3 [DET] Determiner

The description of the use of the determiner in ZAEBUC-50 is rather straightforward (see for example, Al-Najjar (2014) for an alternative, more nuanced approached to the differences between the determiner system in Arabic vis-à-vis English). First of all, this category rendered the largest number of tokens, 97/584 or 16.61%. Essentially, the non-standard use of the determiner consists of either using it when it is not expected in Standard English, as in (27) "It has both negative and postive effects to individuals and *the* society." [83847] or omitting it when it would be (28) "*UAE goverment* make the first step to mars, send two emirate to mars" [286381]. The determiners in question include the definite article *the* and mainly the indefinite article allomorphs *a* and *an*. The determiner was omitted 40/97 times or 41.24% and it was inserted 57/97 times or 58.76%. Arabic uses the determiner extensively, at a much higher degree than English, and yet the distribution of omissions versus insertions in the corpus is not considerable.

5.4.4 [NUM] Number

The non-standard marking of number is one of the most frequent categories in the corpus. Half of the instances (39/78 tokens) found entail the absence of plural marking on the noun [NUM.Om.]; the other half consists of 10 occurrences of substitution [NUM.Sub.], one occurrence of displacement [NUM.Dis.], and 28 occurrences of insertion [NUM.Ins.].

5.4.4.1 [NUM.Displ.] cross-segmental metathesis

The morphology indicating plural seems to have gone through a process of *cross-segmental metathesis*,[15] although I suspect the motivation is not necessarily phonetic, but syntactic (it can also be interpreted as a combination of omission and subsequential insertion). An instance of number displacement is when the plural marking surfaces on the adjective qualifying the noun, rather than on the noun itself, as in (29) "in a diffrent*s* way" [190122] [in different ways]. Previous examples have been adduced in support of the concept of flexibility in the selection of landing sites for either grammatical features or lexical items. Moreover, a (semantically) plural noun introduced by the indefinite determiner *a* surfaces again.

5.4.5 [POSS] Possessive

Two occurrences of non-standard possessive forms were found. The periphrastic possessive (or analytical genitive) construction in place of the expected inflected one (or synthetic genitive), (30) "This is very important for *the level of the country*" [91113] [for the country's level]. Arabic has a variety of strategies to express possession depending on dialectal variation or certain semantic or syntactic conditions, including the use of possessive markers (*hagg*, *tab'*, and *maal*, among others) (Guella, 2011), but none of these resemble the possessive forms in Standard English. ZAEBUC-50 only produced two occurrences, so attempting to elucidate the language users' choices would be premature.

5.4.6 [PRON] Pronouns

Of all 29 occurrences of non-standard use of the pronoun, pronouns are absent on 15 occasions, 11 alternative forms are used, and 3 occurrences of resumptive pronouns account for pronoun insertions.

5.4.6.1 [PRON.Om.] Omission of a pronoun

Expletive pronouns are sometimes absent, as in (31) "Addiction is a bad thing *can* easily ruin someone's personal life" [186196], where one would expect expletive *it* before the verb *can*. Restrictive (or defining) relative pronouns are absent, as in (32) "ther is to much people *are* famous" [293472] [there are too many people who are famous], although word order could have produced a more straightforward construction, [Too many people are famous]. The pronoun is also omitted when its referent is in subject position, as in (33) "Fainly i think should be caerfull" [189349] [Finally, I think *you* should], with a second-person pronoun absent. An interesting omission of the pronoun in direct object position is when this is hinted at by showing up in a prepositional phrase complementing the same verb, as in (34) "this is *affecting* a lot *in or society*" [189799]. This last case merits further investigation, since it occurs on several occasions in the corpus. This phenomenon, which I will name phantom pronoun,[16] refers to the absence of a direct pronoun from its expected position, the complement of the verb V + $NP_{D.O.}$, while it surfaces, in an oblique fashion, embedded into a prepositional phrase (where the PP functions as an adverbial). This PP thus carries two functions, that of a conventional adverbial and, more significantly, that of a complement to the transitive verb. For example, in (34) the expected direct object of the verb *affecting* is *or society* [our society], but it is instead reinterpreted as a prepositional phrase [in our society].

5.4.6.2 [PRON.Ins.] Resumptive pronoun[17]

This section presents an interesting use of the pronoun, which I would like to refer to as resumptive pronoun, whereby an expletive pronoun, whether singular *it* or plural *they*, re-establishes subjectivity based on the writer's reinterpretation (see Stein and Wright (1995) for insights into a different interpretation of the subject), resulting in the main sentence being partitioned into two simpler clauses with *semantic replication* and *grammatical dissociation*. For example, in (35) "but some time these news *it* can be false" [292703], expletive *it* refers to *these news* even though it is likely that the student treats *news* as a plural noun, judging by the plural demonstrative. Rather than a more direct reading, [but sometimes this news can be false], the student splits the sentence into two constructions, a phrase *but sometimes*

15 *Cross-segmental metathesis* is used here to refer to a grammatical element, such as plural *-s*, that is marked on a morpheme other than the expected one, so that "in different way**s**" becomes "in a different**s** way." Therefore, plural marking has crossed the segmental boundary of the noun to eventually land within the segmental boundary of the adjective.

16 C.f. Figure 5 in section 6.5 below for a graphical representation of this phenomenon.

17 I use the term resumptive pronoun here in a specialized sense to refer to a phenomenon that, in my interpretation, goes beyond pronominal reinforcement. Specifically, the pronoun not only co-occurs with its antecedent but also plays a structural role in reanalyzing the utterance as two distinct propositions. This usage is inspired by generative grammar traditions, notably Ross (1967), though the phenomenon discussed here diverges in key respects.

these news with a missing predicate, and a clause *it can be false*, which picks up the subject from the previous phrase, albeit with a different pronoun, and makes the same statement.

5.4.7 [PREP] Preposition

The redundant use of prepositions is described under the [REP] repetition section. Here we focus on the substitution, insertion, and omission of prepositions as contrasted with standard English. Substitutions alone account for half of the tokens found, with insertions accounting for 28% of the remaining tokens, and omission for slightly over 20%.

5.4.7.1 [PREP.Sub.] Substitution of a preposition

Some of the substituted prepositions are: *to* instead of *on*, *for* instead of *on*, *about* instead of *with*, *in* instead of *as*, and *in* instead of *on*, as for example in (36) "It has both negative and postive effects *to* individuals and the society" [83847], a common substitution in the corpus. The substitution of prepositions is a prevalent phenomenon not only within the MENA region (Diab, 1996; Parra-Guinaldo and Lanteigne, 2020), but also elsewhere.

5.4.7.2 [PREP.Ins.] insertion of a preposition

A frequent occurrence is the insertion of the preposition *to* with the verb *to affect*, as in (37) "Social media *affect to* the pepole in work" [189349]. Another instance of a "redundant" preposition (the use of scare quotes here is intended to convey my own reticence to use a term prevalent in the ELF literature) is (38) "its easy to *cotact with* my family in KSA"[18] [194708]. This use of the preposition has been referred to as "redundant preposition" (Seidlhofer, 2004; Cogo and Dewey, 2011), but I would argue that it is "redundant" only from the point of view of Standard English and its expected collation; if instead we assume the ELF user has reanalyzed the grammatical features of the verb, rendering this intransitive (more economical, or cognitively less challenging, than a transitive verb), then the preposition gains a new and important role, that of linking the verb with its argument (the *patient*, or entity affected by the action of the verb).

5.4.7.3 [PREP.Om.] Omission of a preposition

If the verb *to affect* appears with a preposition, as we have just seen, the noun *effect* shows up without a preposition, the reverse of what we would expect in Standard English, as in (39) "this is can be very dangerous and ofcourse cause sides *affects* our society" [189442], [side effects *on* our society].

5.4.8 [MRA] Morphological reanalysis

Morphological reanalysis entails the reinterpretation of morphological boundaries resulting in newly structured lexical items. For example, the coalescence of two separate lexical items into a single lexical item, which I here refer to as *merging*, or the breaking down of a single lexical item into separate morphologically analyzable lexical items, *splitting*. Although reanalysis is a common phenomenon in all languages at all linguistic levels (phonological, morphological, and syntactic), the occurrences identified in ZAEBUC-50 can be characterized as *innovative*, since they are all different from current reanalyzed lexical items in Standard English, and as *expressive* in that the writer's pragmatic perspective can be elucidated. A commonly used example of reanalysis from a diachronic perspective in Standard English is *a napron* ⇒ *an apron*, from Old French "naperon" *small table-cloth*. "The formation of a new lexical item through a wrong analysis of an existing word boundary," such as this, has been referred to as *metanalysis* (Crystal, 2008) in historical linguistics. Other examples include coexisting pairs, such as *afire* and *on fire*, *aboard* and *on board*. Following are a few examples from the corpus subcategorized into splitting and merging morphological reanalysis.

5.4.8.1 [MRA.spl.] Splitting morphological reanalysis

About half of all occurrences under this subcategory are reanalyzed items found several times throughout the corpus, as in example (40), where the adverb *also* is rendered as *all so* (3 tokens), "in the UAE *all sow* in Fujairah" [186798].

5.4.8.2 [MRA.merg.] Merging morphological reanalysis

By far, the most frequent occurrence of merging morphological reanalysis (or *merging MRA*, for short) is *a lot*, reanalyzed as one word, *alot*. This construction alone is present 11 out of a total of 12 times, as in (41) "the education changed *alot*" [191115], and *everyday* being the remaining occurrence, (42) "between 4–6 h *everyday*" [219833]. Although prescriptively incorrect in Standard English, the latter example is quite frequent, even in TV ads or billboards (I recall multiple examples over the years).

What merging tells us is that the student is reinterpreting the construction as a linguistic unit, with a single meaning, probably replaceable with another single one-word item, for example, *alot* can be substituted with *much*. Another example is *everyday*, which could be replaced with *daily*. In Standard English, we also have constructions with identifiable parts that can appear separate (split) or together (merged) depending on their function, and so we have *overall* as an adverb and *over all* as a combination of a preposition and a pronoun and, as just mentioned, *everyday* as an adjective or *every day* as an adverbial phrase.

As to the motivations behind either splitting or merging, one can only speculate. Splitting may be a mechanism resulting from the literal analysis of constructions, and so rather than retrieving a single lexical item (with a heavier cognitive load) from the language user's lexicon, separate items are retrieved resulting in a simple composite construction. Similarly, in the use of idioms and other formulaic constructions, Wang and Kaatari (2021) acknowledge, "Unlike L1 speakers, for whom those 'prefabricated' sequences are stored and retrieved from memory as whole entities at the time of use in accordance with the 'idiom principle'[19] (Sinclair, 1991), L2 speakers are said to rely primarily on 'analytical processing' (Wray, 2002) (or the 'open-choice' principle in Sinclair's (1991) terminology"). After all, idioms, among other features which are regarded as 'the most typically English', "turn out to be non-essential for

18 One reviewer rightly observes that this is a case of cross-linguistic transfer, as the Arabic equivalent of "contact" typically co-occurs with the preposition ma'a *with*, which likely influenced the student's phrasing in English.

19 The principle of idiom, according to Sinclair (1991), "is that a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments [...] it may reflect a natural tendency to economy of effort" (110).

mutual understanding" in ELF communication (Seidlhofer, 2001, p. 149), and therefore opting for more literal interpretations is a logical consequence. If this analysis is correct, and the ELF user tends to split (and interpret constructions literally) more than merge, this would explain why in our data splitting morphological reanalysis (or *splitting MRA*, for short) is twice as common as merging, especially when considering that almost all instances of the *merge* operation refer to the same lexical item, *alot*. Conversely, merging involves the combination of separate lexical items into one; this may seem contradictory at first, but notice that the two elements being merged in 11 of 12 instances of splitting MRA are one determiner and one noun, so cognitive processing is not as much of an issue for the student. More data would be necessary to corroborate this tentative interpretation.

5.4.9 [CON] Conversion

Conversion is a word-formation process whereby the word class of a lexical item shifts to a different word class without any change in its morphology; for example, the word *taste* can be used either as a verb or a noun and the word *empty* as an adjective or a verb depending on context (Crystal, 2008). The occurrences found in this corpus are noteworthy for several reasons; for example, the majority of occurrences, about 70%, involve a noun taking over the role of other word classes, such as adjective or verb, but other combinations are also possible. Another interesting feature is that converted nouns can take the inflectional morphology of a noun even while functioning as a verb as well as taking on verbal inflection.

The derivational morphology attached to the verb does not prevent it from carrying its expected grammatical features. It is not necessarily a case of absorption though, since the transposed feature has no grammatical utility. On the other hand, it was noticed before that the student has a proclivity toward the use of nouns, so it is possible that the student retrieves from her lexicon a noun form and plugs it into the verb slot. In other words, given a reduced lexicon, or inventory of lexical items, some of these items must take the role of other items belonging to the same lexical set. For instance, (43) "SO in UAE need *to develoment* the streets for baicycles" [190745] [to develop] or (44) "UAE do alot for the commenty to be confortabl and to make everthing *to developements* contry" [191115] [*to develop* the country], where the noun is used as a verb or (45) "after the big *developming* they make" [286381], [the big development they made], where the verb *-ing* is used as a noun.

5.5 Cross-layered categories (55/584 tokens or 9.42%)

5.5.1 [REP] Repetition

Numerous are the researchers who have paid attention to repetition in communicative events (Cappuzzo (2015) mentions the following: Schnelby (1994) on self-repetition, Simpson (1994) on other-repetition, Murata (1995) on allo-repetition) and have described the various techniques speakers use to promote intelligibility and enhance communication. The instances of repetition in ZAEBUC-50 have been categorized on the basis of the type of element being repeated, whether a single lexical item, such as a preposition (the use of two prepositions, one of which being seemingly superfluous), or a prepositional phrase (the superfluous use of a prepositional phrase where adverbial information has already been provided by other means).

Two different students use the combination of *in my opinion* followed by *I think*, as in (46) "*In my opinon think* socail media" [89953]. It may be the case that the student finds the starter phrase complex, in the sense that it veers away from the simple S + V construction. The student starts out with a frequently-used construction, one she is very familiar with, but then reverts to a default S + V construction.

5.5.2 [w.o.] Word order

This category contains examples of word order that deviates from Standard English. Except for one, all occurrences involve the repositioning of the adverbial, for example (47) "[…] you *also* can know news" [89953], where the adverb *also* is placed in an unexpected position. Arabic exhibits a certain degree of flexibility in terms of adverb placement, although it often has a post-verbal position. Cross-linguistic interference may therefore not be a factor in this case, but then the question of what factors might be influencing this phenomenon remains.

5.5.3 [w.o.intrud.] Intruding constituents

When a transitive verb contains more than just its direct object (whether this surfaces as such or not), for example an adverbial, the student seems to prioritize the relationship between the verb and the adverbial, relegating the noun phrase to a later merger operation. I will refer to this as the *intruding constituent phenomenon*, whereby the adverbial (or possibly another constituent) *intrudes into* the V_{trans} + $NP_{D.O.}$ merge operation. Thus, recall (34) "this is affecting *a lot* in or society"[20] [189799] [this is affecting our society *a lot*], where the adverbial (the quantifier *a lot*), in this case *the intruder*, has merged directly with the verb, and in a later merger operation the direct object (the NP *or society* [our society]) surfaces as a prepositional phrase on a later merging operation. Section 5.4.6.1 addressed the concept of the phantom pronoun, a related phenomenon, demonstrating how the noun is indirectly expressed via an adverbial.

6 Key insights from the discussion

In the present section, I will highlight some of the observations and hypothesized motivations behind the linguistic processes identified in the corpus under investigation.

6.1 Instability

In developing an ELF perspective in pedagogy, Jenkins et al. (2011) posit that "generating […] an understanding among learners and teachers of the inherent variability (even *instability*) of human language in general and English more specifically" (p. 306, my italics) is necessary. As a matter of fact, the data show how the same element may be present in one part of the sentence and absent in another, as in coordinated constructions (see 26). Another example of instability is the placement of morphological elements, such as the 3rd p.sg. *-s*, which may appear sometimes attached to the verb, as expected, but

20 Likely a case of cross-linguistic transfer from Arabic, where *affect* typically co-occurs with the preposition *ma'a* [with].

more interestingly, it may also appear attached to the preceding subject; this anticipatory 3rd p.sg. form has been referred to as floating *-s* in the present study, since the grammatical information (inflectional morphology) is carried by the subject in anticipation of the verbal morphology (see 18). This phenomenon seems recurrent not only in the data, but also in my own observations during ELF interactions with Arab speakers. One more instance of instability in the selection of landing sites for grammatical features is the case of the plural suffix *-s*, which occurs not only in the expected position (attached to the noun), but also in neighboring elements (see 29).

6.2 Simplification and the *resumptive pronoun*

As part of a process of simplification, a sentence with a complex noun phrase is split such that the original NP becomes cut off from the sentence and a resumptive pronoun, with less computational load, takes on the role of the subject. This resumptive pronoun does not necessarily agree with the original noun in number, since this will be dependent on the semantic reinterpretation of the subject by the language user (see 35). Although Fussell's (2011) 'dummy object' likewise accounts for the non-standard use of a pronoun, the main difference between the 'dummy object' and the resumptive pronoun is that the former is coreferential with a restrictive relative pronoun, whereas the latter replaces a complete sentence. What they both seem to have in common is the need for meaning reinforcement.

6.3 Generic forms as *default wildcards*

The use of generic verbs in ELF is a well-known phenomenon. Mauranen (2015), for example, identified a few frequency verbs of communication, such as GO, ASK, SAY, TELL, PUT, BEGIN, and ADD, and other non-communicative generic verbs, such as GIVE, TAKE, and SHOW. But, whereas this phenomenon mainly concerns the frequency of use of these verbs, the generic verbs adduced from the present data bear specialized functions: a temporary verb (of less or generalized semantic content) in anticipation of a complex VP (see 23) or a verbal base carrying a negative marker (see 21). Generic verbs with these functions have been referred to as *default wildcards* in this study. The use of default wildcards, along with other phenomena such as the bare form of the verb or the overuse of the *-ing* form, simplify the selection of grammatical forms and thus provides evidence for syntactic simplification as a characteristic of ELF, as proposed in Parra-Guinaldo and Lanteigne (2020).

6.4 *Morphological reanalysis* or MRA

ELF users in ZAEBUC-50 resort to the reassessed demarcation of morphological boundaries to convey an accurate picture (in the language user's mind) of the reconceptualization of lexical items. As a result, separate elements in Standard English appear as a single lexical item, in a process the researcher has coined merging MRA (see 41), or single elements in Standard English result in separate morphologically analyzable lexical items under a process of splitting MRA (see 40). The second of these processes is by far more frequent and it entails the literal analysis of the construction. As has been observed in studies of formulaicity and phraseology, ELF users produce in many cases approximation of idiomatic constructions due to their proclivity to analyze their meaning literally rather than idiomatically (Wang and Kaatari, 2021; see also Kecskés (2007) for a detailed account of formulaic language in ELF). This propensity toward compositionality and "semantically transparent language" (Kecskés, 2007) may be the principal motivator of the splitting type of morphological reanalysis here proposed, as manifested in the corpus.

6.5 *Phantom pronouns* and *intruding constituents*

A phantom pronoun may occur in combination with an intruding constituent (see 34), but not necessarily, as in (49) "Social media has advantage and dis advantage that can effect in our society." [119813]. In the present corpus, it is an adverbial that intrudes into the link between the transitive verb and its object. In generative grammar parlance, one would say that a *merge* operation between a transitive verb and its complement, the direct object, precedes (and therefore sits lower in the tree) the *merge* operation between the VP[V_{trans} + $C_{DirectObject}$] and an adjunct (in this case, the adverbial). Indeed, this is the expected order of elements in Standard English, except to indicate emphasis or when using poetic license. In Figure 5, (1) represents a verbal phrase, including a transitive verb and its complement (a direct object) and an adjunct (an adverbial) in Standard English, and (2) and (3) represent a process assumed to occur in ELF, based on the data in ZAEBUC-50, where (2)

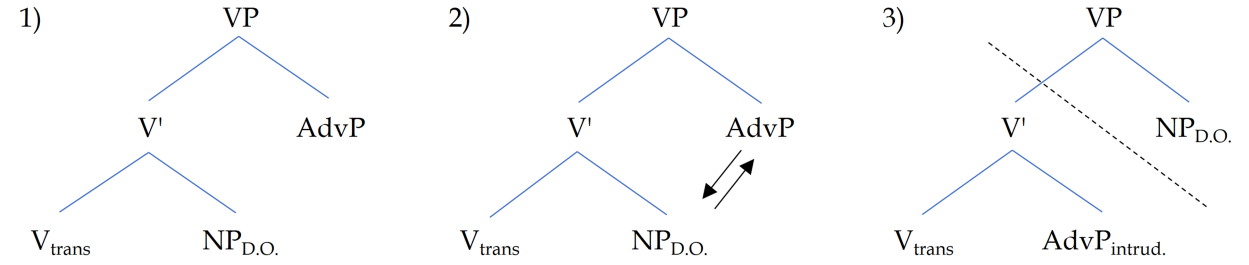

FIGURE 5
Syntactic trees representing **(A)** transitivity in Standard English, **(B)** transposition of phrases in ELF, and **(C)** resulting structure with an intruding constituent in ELF.

identifies the transposition of phrases between the NP and the AdvP and (3) depicts the resulting syntactic structure, whereby the AdvP emerges as an intruding constituent under the control of the transitive verb.

If syntactic simplification is responsible for the intruding constituent phenomenon, it follows that merging a single transitive verb with an adjunct carries less computational load than merging a combined VP[V_{trans} + $C_{DirectObject}$] with an adjunct. The noun (direct object) is left for a later operation, as if scaffolding contentful items (nouns and verbs) in consecutive, but crucially not tightly linked operations (the syntactic laxity represented by a broken line in Figure 5C). This phenomenon is not unlike the breaking up of complex sentences into separate clauses linked through apposition, as we saw before; the main difference being that, whereas the latter occurs at the clausal level, the former does so at the phrasal level. A related phenomenon is that of the phantom pronoun, whereby the complement of a transitive verb becomes implicit in the sentence, that is, it disappears from the surface structure as such, only to reappear under the guise of an adverbial. Again, dealing with a single contentful item is syntactically preferable than dealing with two. Notice that it is the transposition of functions, as shown in Figure 5B, that would enable the adverbial to assume the nominal value (the direct object feature) of the original noun. Boyle (2011) suggests that "instability in the system of transitivity gives the ELF user scope for creativity" (p. 153), and although he was referring specifically to whether a verb is treated as transitive or intransitive by the ELF user, the phenomena of the phantom pronoun and the intruding constituent not only qualify as creative, but possibly unique to ELF, and possibly unique to ELF for whom L1 is Arabic, that is, ELF of the MENA region. We are clearly in need of further investigation to lend support to this claim.

6.6 *-ing* and the preference for nouns

Creativity does not stop at the *sui generis* use of transitive verb constructions, as can be surmised from the description and commentary of the data presented thus far. Quite the contrary, creativity comes in all shapes and sizes, so to speak. In regard to the so-called "extended" use of the progressive, Ranta (2006) posits that "it [is] difficult to believe that the reason behind such use of the progressive resides in mere L1 interference, target language input or teaching related factors" (p. 110) and adds that this view "fails to acknowledge the way L2 speakers actually make use of the language for their own purposes" (p. 95); finally, she concludes that "the source of the 'attractiveness' of the progressive resides in the grammatical form itself" (p. 112), since it "is actually used for the very purpose of gaining explicitness and expressivity in L2 communication" (p. 114). Two main uses of the *-ing* suffix have been observed in ZAEBUC-50: (1) a generic temporal (or multi-purpose) form to express different instances of tense and aspect (see 2 and 6) the nominalization of verbs. Creativity, and in particular, expressivity and clarity of expression, may very well motivate the ELF user to the "extended," or rather *sui generis*, use of the *-ing* morph, but we ought to take Ranta's proposition a step further and ask: what is it then about the *-ing* form that makes it "attractive"? what is so "explicit" and "expressive" about the *-ing* form? The answer to this, I suggest, is the capacity that *-ing* has to carry out both verbal and nominal functions on the one hand and the special value assigned to nouns as default forms on the other. The gerund is the key element here, which by its own nature shares the characteristics of both verb and noun. Thus, when the ELF user is confronted with a complex (or not well-known) construction, she may resort to a default, multi-purpose form. Sentence (42) above provides interesting supporting evidence for this explanation, where the student, confronted with the complex or unfamiliar "development," resorts to a simplex or familiar "develop" and the nominalizer, plus the generic, *-ing*, producing a new, if not creative, noun *developming*.

7 Concluding remarks

In this paper, we have presented a comprehensive analysis of the morphology and syntax of the language produced by first-year students at a private university in the UAE as recorded in ZAEBUC-50, a small corpus of short essays written in English. This study contributes to the description of morpho-syntactic features of ELF in what, up to now, has been an understudied geographic strand, the MENA region, thus filling a persistent gap in the literature. The present study is both descriptive and explanatory in nature; first, salient features have been identified, codified, and categorized according to phrase structure layers and then, following a usage-based linguistics framework, the motivations behind the phenomena identified have been hypothesized. Some of the most significant findings in the study include: (1) an under-utilized CP, due to two main factors, the narrative nature of the language, and therefore devoid of interrogatives, and the language style, characterized by asyndeton, with juxtaposed rather than subordinated clauses; (2) most of the features identified in the TP are issues with SV concord and special uses of the tense, such as the predominant use of bare verb forms and *-ing* forms; and (3) the most common features in the VP are variant uses of the determiner, the preposition, and number. Overall, this paper confirms the presence of some of the features already identified in previous studies of ELF in general, but more interestingly those particular to the Gulf region (or possibly the MENA region at large). Notably, novel phenomena in the morphology and syntax of the ELF under investigation have been identified and their motivations explained, such as morphological reanalysis, the case of a phantom pronoun and certain intruding constituents, and the particular use of *-ing* forms. These findings though should be interpreted with caution given the limitations of this study, such as the limited number of tokens, but it is hoped they encourage future studies to corroborate, or disprove, the phenomena presented here as they occur in other instances of ELF in the region. After all, and as Seidlhofer (2004) pointed out, providing "comprehensive and reliable descriptions of salient features of ELF [should provide the] basis for an eventual codification" (p. 215).

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the participants or participants legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

VP-G: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author declares that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomm.2025.1532750/full#supplementary-material

References

ACE. (2020). The Asian Corpus of English. Director: Andy Kirkpatrick; Researchers: Wang Lixun, John Patkin, Sophiann Subhan. https://corpus.eduhk.hk/ace/index.html (Accessed April 19, 2024).

Al_Surmi, M. (2018). Postcolonial English varieties in the Arab world: a preliminary study. *Asian Englishes* 21, 1–14. doi: 10.1080/13488678.2018.1464094

Al-Najjar, M. (2014). A contrastive study on the syntactic structure of standard English and standard Arabic determiner phrase. *J. College Lang.* 28, 149–158. Available online at: https://jcolang.uobaghdad.edu.iq/index.php/JCL/article/view/77

Baker, W. (2011). "Global cultures and identities: refocusing the aims of ELT in Asia through intercultural awareness" in Innovating EFL education in Asia. eds. T. Muller, S. Herder, J. Adamson and P. S. Brown (Basingstoke, Hampshire, UK: Palgrave Macmillan).

Biber, D. (2006). University language: a corpus-based study of spoken and written registers. Amsterdam (Netherlands) and Philadelphia (USA): John Benjamins.

Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E. (1999). Longman grammar of spoken and written English. Harlow, Essex Longman.

Björkman, B. (2008). English as the lingua franca of engineering: the morphosyntax of academic speech events. *Nord. J. Engl. Stud.* 7, 103–122. doi: 10.35360/njes.103

Boyle, R. (2011). Patterns of change in English as a lingua franca in the UAE. *Int. J. Appl. Linguist.* 21, 143–161. doi: 10.1111/j.1473-4192.2010.00262.x

Brazil Corpus of English (n.d.). Available online at: https://www.elfbrasil.com/ (Accessed March 18, 2025).

Breiteneder, A. (2005). The naturalness of English as a European lingua franca: the case of the 'third person -s'. *Vienna English Work. Pap.* 14, 3–26. Available online at: http://www.univie.ac.at/Anglistik/Views0502ALL.pdf

Cappuzzo, B. (2015). Allo-repetition in academic settings. Cooperation, understanding co-construction and knowledge negotiation in the medical section of the ELFA corpus. *ESP Across Cult.* 12, 33–54.

Cogo, A. (2005). The expression of identity in intercultural communication: the case of English as a lingua franca. *Paper given at BAAL conference*. Bristol.

Cogo, A., and Dewey, M. (2006). Intercultural communication. *Symposium given at IATEFL conference*. Harrogate.

Cogo, A., and Dewey, M. (2011). Analysing English as a lingua Franca. A corpus-driven investigation. London and New York: Continuum.

Crystal, D. (2008). A dictionary of linguistics and phonetics. 6th Edn. Malden, Massachusetts (USA) and Oxford (UK): Blackwell Publishing.

Dewey, M. (2003). Codifying lingua franca English. *Paper given at IATEFL conference*. Brighton.

Diab, N. (1996). The transfer of Arabic in the English writings of Lebanese students. *ESP São Paulo* 18, 71–83.

ELFA. (2008). The corpus of English as a Lingua Franca in academic settings. Director: Anna Mauranen. Available online at: http://www.helsinki.fi/elfa (Accessed April 19, 2024).

Ferguson, C. A. (1971). Absence of copula and the notion of simplicity. In D. Hymes (Ed.), Pidginization and creolization of languages. *Cambridge: Cambridge University Press*, 141–150

Fussell, B. (2011). The local flavour of English in the Gulf: in a matter of fifty years, English has attained its own distinctive features in the Gulf. *English Today* 27, 26–32. doi: 10.1017/S0266078411000502

Graddol, D. (2006). English next. London: The British Council. Algiers: GERFLINT.

Guella, N. (2011). Possessive constructions in Arabic: a cross-dialectal study. Algiers: GERFLINT.

Habash, N., and Palfreyman, D. (2022). ZAEBUC: an annotated Arabic-English bilingual writer corpus. In *Proceedings of the 13th conference on language resources and evaluation (LREC 2022)*, pp. 79–88, Marseille.

Hall, C. J. (2018). "Cognitive perspectives on English as a lingua franca" in The Routledge handbook of English as a lingua franca. eds. J. Jenkins, W. Baker and M. Dewey (London & New York: Routledge Taylor & Francis Group).

Hillman, S., Selvi, A. F., and Yazan, B. (2020). A scoping review of world Englishes in the Middle East and North Africa. *World Englishes* 40, 159–175. doi: 10.1111/weng.12505

Hopkyns, S., Zoghbor, W., and Hassall, P. J. (2020). The use of English and linguistic hybridity among Emirati millennials. *World Englishes* 40, 176–190. doi: 10.1111/weng.12506

Jenkins, J., Cogo, A., and Dewey, M. (2011). Review of developments in research into English as a lingua franca. *Lang. Teach.* 44, 281–315. doi: 10.1017/S0261444811000115

Jiang, J., Bi, P., and Liu, H. (2019). Syntactic complexity development in the writings of EFL learners. Insights from a dependency syntactically-annotated corpus. *J. Second. Lang. Writ.* 46:100666. doi: 10.1016/j.jslw.2019.100666

Kecskés, I. (2007). "Formulaic language in English lingua franca" in Explorations in pragmatics: linguistic, cognitive and intercultural aspects. eds. I. Kecskés and L. Horn (Berlin & New York: Gruyter).

Kirkpatrick, A. (2010). English as a lingua franca in ASEAN. A multilingual model: Hong Kong University Press.

Laitinen, M. (2020). Empirical perspectives on English as a lingua franca (ELF) grammar. *World Englishes* 39, 427–442. doi: 10.1111/weng.12482

Leimgruber, J. R. E., Al-Issa, A., Lorenz, E., and Siemund, P. (2022). Managing and investing in hybrid identities in the globalized United Arab Emirates. *J. Lang. Identity Educ.* 23, 955–972. doi: 10.1080/15348458.2022.2070849

Lorenz, E. (2022). "We use English but not like all the time like": discourse marker *like* in UAE English. *Front. Commun.* 7:778036. doi: 10.3389/fcomm.2022.778036

Mauranen, A. (2009). Chunking in ELF: expressions for managing interaction. *Intercult. Pragmat.* 6, 217–233. doi: 10.1515/IPRG.2009.012

Mauranen, A. (2012). Exploring ELF. Academic English shaped by non-native speakers. Cambridge: Hong Kong: Cambridge University Press.

Mauranen, A. (2015). "English as a global lingua franca: changing language in changing global academia" in Exploring ELF in Japanese academic and business contexts. ed. K. Murata (London: John Benjamins).

Melitz, J. (2018). English as a lingua franca: facts, benefits and costs. *World Econ.* 41, 1750–1774. doi: 10.1111/twec.12643

Morán Panero, S., Martínez-Sánchez, M. M., and Ronzón Montiel, G. J. (2024). English as a lingua Franca in Latin America: educational perspectives.: Berlin, Boston: De Gruyter Mouton.

Murata, K. (1995). Repetitions: a cross-cultural study. *World Englishes* 14, 343–356. doi: 10.1111/j.1467-971X.1995.tb00078.x

Parra-Guinaldo, V., and Lanteigne, B. (2020). "Morpho-syntactic features of English as a lingua franca in Dubai and Sharjah" in Multilingual global cities. Singapore, Hong Kong, Dubai. eds. P. Siemund and J. R. E. Leimgruber (London & New York: Routledge).

Payne, T. E. (2011). Understanding English grammar. Cambridge Cambridge University Press.

Ranta, E. (2006). The 'attractive' progressive: why use the-*ing* form in English as a lingua Franca? *Nord. J. Engl. Stud.* 5, 95–116. doi: 10.35360/njes.13

Ranta, E. (2009). "Syntactic features in spoken ELF–learner language or spoken grammar" in English as a lingua franca. Studies and findings. eds. A. Mauranen and E. Ranta (Castle upon Tyne: Cambridge Scholars Publishing).

Ranta, E. (2018). "Grammar in ELF" in The Routledge handbook of English as a lingua franca. eds. J. Jenkins, W. Baker and M. Dewey (London:: Routledge).

Ross, J. R. (1967). Constraints on variables in syntax (Doctoral dissertation). Cambridge, Massachusetts: Massachusetts Institute of Technology.

Schnelby, C. (1994). "Repetition and failed conversation in the theatre of the absurd" in Repetition in discourse. Interdisciplinary perspectives. ed. B. Johnstone (Norwood, NJ: Ablex).

Seidlhofer, B. (2001). Closing a conceptual gap: the case for a description of English as a lingua franca. *Int. J. Appl. Linguist.* 11, 133–158. doi: 10.1111/1473-4192.00011

Seidlhofer, B. (2004). Research perspectives on teaching English as a lingua franca. *Annu. Rev. Appl. Linguist.* 24, 209–239. doi: 10.1017/S0267190504000145

Seidlhofer, B. (2010). "Lingua franca English in Europe" in The Routledge handbook of world Englishes. ed. A. Kirkpatrick (London & New York: Routledge).

Seidlhofer, B., Breiteneder, A., and Pitzl, M.-L. (2006). "English as a lingua franca in Europe: Challenges for applied linguistics". *Annu. Rev. Appl. Linguist.* 26, 1–34. doi: 10.1017/S026719050600002X

Siemund, P., Al-Issa, A., and Leimgruber, J. R. E. (2020). Multilingualism and the role of English in the United Arab Emirates. *World Englishes* 40, 191–204. doi: 10.1111/weng.12507

Simpson, J. M. (1994). Regularized intonation in conversational repetition. *Adv. Discourse Process.* 48, 41–41.

Sinclair, J. (1991). Corpus, concordance and collocation. Oxford: Oxford University Press.

D. Stein and S. Wright (Eds.) (1995). Subjectivity and subjectivisation. Linguistic perspectives. Cambridge: Cambridge University Press.

Taguchi, N., and Ishihara, N. (2018). The pragmatics of English as a lingua franca: research and pedagogy in the era of globalization. *Annu. Rev. Appl. Linguist.* 38, 80–101. doi: 10.1017/S0267190518000028

The British Council (2013). The English effect: The impact of English, what it's worth to the UK and why it matters to the world. London: The British Council.

VOICE (2021). The Vienna-Oxford International Corpus of English (version VOICE 3.0 Online). Available online at: https://voice3.acdh.oeaw.ac.at (Accessed April 19, 2024).

Wang, Y., and Kaatari, H. (2021). Phraseological patterns of SAY in academic ELF communication. *J. Engl. Acad. Purp.* 54:101046. doi: 10.1016/j.jeap.2021.101046

Wray, A. (2002). Formulaic Language and the Lexicon. Cambridge: Cambridge University Press..

WrELFA (2015). The Corpus of Written English as a Lingua Franca in Academic Settings. Director: Anna Mauranen. Compilation manager: Ray Carey. Available online at: http://www.helsinki.fi/elfa (Accessed April 19, 2024).

Young, R. (2011). "Interactional competence in language learning, teaching, and testing" in Handbook of research in language learning and teaching. ed. E. Hinkel (New York, NY: Routledge), 426–443.

Zoghbor, W. (2009). The implications of the LFC for the Arab context. *IATEFL Pronunc. Special Interest Group Newsletter*, Available online at: https://hdl.handle.net/2381/9232

ZU catalog (2021–2022). Available online at: https://www.zu.ac.ae/main/files/contents/docs/zu-catalog-21-22.pdf (Accessed July 15, 2024).