"We Use English But Not Like All the Time Like"—Discourse Marker *Like* in UAE English

Eliane Lorenz[1,2]*

[1]Department of English, English Linguistics, Justus-Liebig-University Giessen, Giessen, Germany, [2]Department of Teacher Education, English and Foreign Languages, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

The United Arab Emirates (UAE) is characterized by extensive language contact. Although Arabic is the official language, practically all communication in general as well as in higher education, in particular, takes place in English. The current study reports from the larger project *Language, Attitudes, and Repertoires in the Emirates* (LARES, 2019–2021) and investigates the use of English as a lingua franca (ELF) among university students in Sharjah, one of the seven sovereign emirates of the UAE. A spoken corpus based on 58 semi-structured interviews is used to examine the use of the discourse marker *like*. It has been shown to be a ubiquitous feature of English no longer confined to American English and occurs frequently in the corpus. It doubtlessly is a prominent discourse marker in the type of English spoken among the heterogeneous group of multicultural university students considered here. Although a large individual variation with respect to normalized frequencies of *like* can be observed, none of the social variables (i.e., gender, citizenship, L1, year of birth, number of languages, college, self-assessed proficiency in English, and English usage score) included in the analysis account for this variability. Instead, I argue that *like* as a discourse marker is part of the English repertoire of all students and appears to be even more frequently used than in other English varieties. This supports previous research arguing for an intensification of language change in ELF contexts as well as high individual variation as a characteristic of multilingual ELF users.

Keywords: discourse marker *like*, English as a lingua franca, spoken corpus, United Arab Emirates, varieties of English

INTRODUCTION

The current study is located in the United Arab Emirates (UAE) and investigates the use of the discourse marker *like* among university students in Sharjah. The focus is on the frequency of use as well as influence pertaining to social variables. First, in the two introductory sections, the status of English as a lingua franca (ELF) in the UAE is presented (*The United Arab Emirates and English as a Lingua Franca*), and a brief overview of the use of the discourse marker *like* in different varieties of English is given (*Discourse Marker Like*), which leads to the two research questions guiding this study. Second, in the *Methodology* section, the *Language, Attitudes, and Repertoires in the Emirates* (LARES, 2019–2021) project is introduced (*The LARES Project*), together with background information of the participants (*LARES Participants and Social Variables*) and the corpus as well as the data coding and analysis steps (*LARES Corpus: Data Coding and Analysis*). Third, in the *Results* section, the frequency of the discourse marker *like* is given (*Frequency of Like*) and contrasted

with the social variables of the participants (*Like Versus Social Variables*). In addition, the most frequent *like*-users are presented in some detail (*Most Frequent Like-Users*). Fourth, in the *Discussion* section, the current findings are discussed in light of previous research, in particular with respect to the overall frequency of *like* (*Frequency of Like*) and in conjunction with the social background information of the speakers (*Like Versus Social Variables*). Finally, the paper concludes with a short summary and outlook section.

The United Arab Emirates and English as a Lingua Franca

The UAE, located in the eastern part of the Arabian Peninsula bordering Saudi Arabia, Oman, and Qatar, constitutes a federation of seven sovereign emirates (Abu Dhabi, Ajman, Dubai, Fujairah, Ras al-Khaimah, Sharjah, and Umm al-Quwain). Even though the current study is set in Sharjah, it would be imprecise to neglect the other emirates because of their close geographical and political proximity. The seven emirates share the same constitution, they are strictly speaking not separated by borders, and the inhabitants have the same nationality (Siemund et al., 2020). For instance, it is not at all uncommon to live and work or study in two different emirates (Parra-Guinaldo and Lanteigne, 2021). Dubai is certainly the most famous of all sheikhdoms; however, Sharjah, along with the other emirates, shows a comparable development (Davidson, 2005; Madichie and Madichie, 2013; Siemund et al., 2020).

The metropolitan area of Sharjah is particularly fascinating for a (socio-)linguistic study because it is characterized by large-scale language contact and has undergone an interesting language development. Originally developing from small fishing villages (Pacione 2005; Siemund et al., 2020), the UAE has experienced unprecedented growth, mainly because of large-scale immigration. Ahmad (2016, p. 31) argued that the six Gulf Cooperation Council countries, to which the UAE belongs, "became an extremely attractive destination for skilled and unskilled labor from within the Arab World and beyond," particularly due to the oil industry. The UAE only gained independence from Great Britain in 1971, until which it had been part of a British protectorate, and has ever since been economically on the rise and gaining in popularity mainly because of the discovery of oil in the mid-twentieth century (Fussell, 2011). Another driving factor was the establishment and development of the tourism industry (Boyle, 2012; Leimgruber and Siemund, 2021). Whereas in 1971 population figures were below 300,000, they have, 50 years later in 2021, reached over 10 million, which is more than a 33-fold increase.[1]

Moreover, what is particularly intriguing about the UAE is that the nonnational population greatly outnumbers the local Emiratis in all seven emirates. For example, there are only about 10%–15% Emirati nationals living in Dubai (Government of Dubai, 2019; Dubai Population, 2020) and approximately 12%

Emirati inhabitants in Sharjah (Sharjah Population, 2020). The nonnational population represents a diverse multilingual and multicultural group, with South Asians being the largest (approximately 60% of the entire population in the UAE) (Al-Issa, 2021). Thus, the UAE is what Vertovec (2007) called a society characterized by super-diversity (see also Hopkyns, 2021). This high share of expatriates in the UAE has essentially helped to guarantee its fast economic growth, as most of those coming to the UAE are not refugees but economic migrants (Al-Issa, 2021).

This population distribution ultimately creates a complex linguistic landscape (Boyle, 2011) and necessarily results in intense language contact at all levels of society (Siemund et al., 2020; Siemund and Leimgruber, 2021; Siemund, 2022). Al-Issa (2021) reported that there are more than 100 languages represented in the UAE, which are, apart from Arabic and English (which are listed as "principal languages" on *Ethnologue*[2]), for example, Bengali, French, Farsi, Hindi, Malayalam, Pashto, Punjabi, Somali, Tagalog, Telugu, and Urdu. Arabic is the official language of the UAE, but in order to work or live in this country, it is strictly speaking not necessary to have a command of Arabic (Al-Issa, 2021). Instead, it is competence in English, which ensures successful communication and secures job opportunities. For instance, anecdotal evidence presented in Hopkyns (2017, 2021) showed that not being sufficiently proficient in English "could be seen as a linguistic disability," even for daily tasks such as going shopping (Hopkyns, 2021, p. 253). Apart from the Arab expatriates, other migrants rarely use Arabic but heavily rely on English as the language of communication. English has also become more important among Emirati citizens, especially among the younger generations, as many families employ nannies with whom they communicate in English (Hopkyns, 2021). The important and ubiquitous role of English can also be supported with a quote from one of the participants of the current study, who is an Emirati citizen:

> (1) *Because as I said we live in a very diverse place I think English is one of the connecting languages that we have* (f8)[3]

Not only are there many different languages present, but there are also different native and non-native varieties of English spoken in the UAE (Parra-Guinaldo and Lanteigne, 2021; Thomas, 2021). Many of the numerous expatriates grew up in countries where English is at least one of the official or national languages (such as India or the Philippines) or come from countries where English is the majority language, such as the United Kingdom or the United States (Hopkyns, 2017). Therefore, many citizens have been in contact with English

---

[1]These numbers were taken from https://www.worldometers.info/world-population/united-arab-emirates-population/, accessed July 28, 2021.

[2]See https://www.ethnologue.com/country/AE.

[3]This quote comes from the LARES corpus, which will be presented in *LARES Corpus: Data Coding and Analysis*. The speakers remain anonymous and were numbered randomly. The only information visible from the ID is their gender. The letter "f" stands for female and "m" for male. All further quotes are also taken from the LARES corpus.

during school education in a foreign country, often following an English curriculum. Others have studied English as a foreign language (EFL) in school. In the UAE, English is introduced as an obligatory school subject already from the early school years onwards; in addition, there are many private schools that have English as their medium of instruction, and practically all higher education takes place in English (Hopkyns, 2017; Al-Issa, 2021; Thomas, 2021). Some Emirati families even decide to send their children to private instead of government schools where the medium of instruction is English (Ahmed, 2021). This is motivated because in order to be admitted to a university, an entry test documenting sufficient English proficiency is usually a requirement (Ahmed, 2021; Al-Issa, 2021). Interestingly, similar placement tests for Arabic do not exist (Ahmed, 2021). It clearly follows that whoever wants to pursue a university degree in the UAE needs to know English (but not necessarily Arabic).

Typically, English is associated with modernity and internationalism, used for business and education, and with this, it "dominates everyday public life and, to a lesser extent, private life too" (Hopkyns, 2017, p. 40; see also Thomas, 2021, for a recent overview of English users and the use of English in the UAE). Arabic, however, is largely confined to the home and family context as well as to practicing religion (Al-Issa, 2021). Without any doubt, English has developed into the lingua franca (Theodoropoulou, 2021; Thomas, 2021), and it has replaced Arabic in many domains (Fussell, 2011; Al-Issa, 2021). ELF could be defined as a vehicular language used by speakers who do not share the same language (Filppula et al., 2017; Mauranen, 2012, 2017). As has been explained above, the result is a complex contact situation of diverse multilingual speakers. It is precisely in such situations of multilingual contact that new varieties of English will emerge (Mair, 2021).

Some have tentatively argued that a new variety of English has emerged or will emerge in the UAE, referred to as either "Gulf English" (Fussell, 2011, p. 31) or "UAE English" (Boyle, 2012, p. 321). Gulf English, clearly not confined to the UAE but to be found in the larger Gulf region, is said to have emerged because of language contact between speakers of Arabic and expatriates who speak different varieties of English (Fussel, 2011). According to Fussel (2011), in the early 2000s, this variety was still at an initial stage, moving towards Schneider's (2007) phase three of the Dynamic Model, i.e., the nativization phase. Boyle hypothesized shortly after Fussel (2011) that the English found in the UAE "should in time become 'UAE English', a variety, perhaps, with a distinct South Asian flavour" (2012, p. 321). Others, however, remark approximately 10 years later that a "local norm of Dubai or Gulf English" has not yet emerged and that "it remains to be seen if it will ever develop" (Leimgruber and Siemund, 2021, p. 1; see also; Ahmed, 2021; Siemund et al., 2021).

As there is to date a lack of research investigating the use of this lingua franca and its status as a new English variety (Siemund et al., 2020), the proposed study sets out to examine the use of ELF in the UAE. There are a number of recent linguistic studies based on Dubai, Sharjah, or the UAE in general (Randall and Samimi, 2010; Boyle, 2011; O'Neill, 2014; Thomas, 2016; Cook, 2017; Piller, 2017; Parra-Guinaldo and Lanteigne, 2021). However, most of these are based on a limited number of participants or specific subgroups (such as female students, see O'Neill, 2014; or police officers, see Randall and Samimi, 2010) or on a small set of linguistics examples that do not yet qualify to formulate generalizations (Boyle, 2011; Parra-Guinaldo and Lanteigne, 2021). The current study is of course also based on a limited number of participants, representing only one group of the entire population (see *Like Versus Social Variables*); yet this research tries to add a puzzle piece to the emerging picture and to contribute to studies investigating varieties of English.

More specifically, this study examines the use of the discourse marker *like*. In general, *like* has received much scholarly attention (see *Discourse Marker Like*), and it is clearly undergoing frequency shifts in present-day Englishes (D'Arcy, 2017). It has mainly been studied in native Englishes, but much less research focuses on non-native speakers of English or ELF varieties (see, for example, Diskin-Holdaway, 2021 or; Rüdiger, 2021). Yet studies investigating language change in ELF varieties identified accelerated grammatical language change (Laitinen, 2020). The observed frequency shifts in present-day Englishes may be understood as language change of a pragmatic phenomenon. Therefore, the current research aims to add to this latter context by investigating the discourse marker *like* used in the English spoken in the UAE. The following subsection introduces the discourse marker *like* and its use in different varieties of English.

Discourse Marker *Like*

This section focuses on an extremely versatile and multifunctional word form in English. The four-letter word *like* has been shown to appear with (at least) 12 different functions. D'Arcy's (2017) account of this word form, which is arguably the most comprehensive, distinguishes between the "unremarkable" functions as a verb (*I like ice cream*), adjective (*they are as like as twin brothers*), noun (*such as fishing or the like*), preposition (*a difficulty like this*), conjunction (*like I said*), complementizer (*it feels like a bit too much*), and suffix (*an Earth-like planet*). D'Arcy (2017) further listed some more notable functions of *like* as an approximative adverb (*it took like 3 hours*), a sentence adverb,[4] a quotative use (*and I was like . . .*), a discourse marker (*like they accepted it*), and as a discourse particle (*he was like falling asleep*).

For the current study, only the last two uses of *like* are of interest and will be focused on in the following sections. D'Arcy (2017, p. 14) stated that "as a discourse marker, *like* encodes textual relations by relating the current utterance to prior discourse" and that it "signals exemplification, illustration, elaboration, or clarification." She further points out that its use increased, particularly so in the second half of the last century, and that it is "widely attested across varieties of English in speech

---

[4]*Like* as a sentence adverb seems to be restricted to some dialects of English, particularly to English dialects in Ireland (D'Arcy 2017, pp. 12–13). D'Arcy (2017, p. 13) provides some examples of this use, for instance, "You'd hit the mud on the bottom *like*."

materials" (D'Arcy, 2017, p. 14). In opposition, D'Arcy (2017, p. 15) explained that particularly as a particle, *like* "signals subjective information" and establishes "common ground, solidarity, or intimacy" between the speech partners. It is this latter use of *like* that is prominently associated with young and female speakers, driving the observed frequency shift mentioned above. Examples (2) and (3) exemplify these two uses respectively. Notice that as a discourse marker, *like* occupies the clause initial position, yet as a particle, it occurs clause internally.

(2) *Mostly the US. Like all the states in the US.* (f23) [discourse marker]

(3) *It wasn't like a challenge but still it was something new for me.* (m4) [discourse particle]

Other than D'Arcy (2017), the current study follows Schweinberger, who does not distinguish between discourse marker and discourse particle uses but who refers to both types using the label "discourse marker" (2014, p. 52). The remaining discussion also refers to both types (which are further subclassified according to their position, see *LARES Corpus: Data Coding and Analysis*) as discourse marker *like*. This is justified because both marker and particle uses (as defined by D'Arcy 2017) of *like* share a number of features outlined below.

An important property of discourse markers is their optionality. This means that they are not required for a sentence to be grammatical (Fuller, 2003). Moreover, particularly in interviews, the use of discourse markers fulfills a stylistic role by creating a rather casual style. This seems to be reinforced by their co-occurrence with other markers; for example, *well* and *let's see* and could be understood as an "interactional tool" to establish a common basis between the interlocutors (Fuller, 2003, p. 372). Fuller (2003, p. 370) further argued that *like* is "pragmatically useful" for creating closeness, placing focus on something, or implying approximation, which are typical contexts of (personal and somewhat informal) interviews. Even though *like* (among other discourse markers) is a way to accommodate planning and continuation in spontaneous speech (Hasselgren 2002; Wolk et al., 2021), it is frequently considered as something negative. Rüdiger (2021, p. 1) even talked about "public language stigmatization" and presents a number of quite strong, negative attitudes towards the use of *like* (Rüdiger, 2021, p. 2).

A number of recent publications focus on the use of the discourse marker *like* (e.g., Schweinberger, 2014; D'Arcy, 2017; Diskin, 2017; Gabrys, 2017; Corrigan and Diskin, 2019; Corrigan and Diskin, 2019), most likely because it is such a prominent or salient feature of English across its different varieties and due to its frequent use (Schweinberger, 2014; Corrigan and Diskin, 2019; Leuckert and Rüdiger, 2021). Yet it can also undergo frequency shifts (D'Arcy, 2017), and its frequency of use differs across English varieties, as impressively demonstrated in Schweinberger's (2014) comprehensive study. To provide only a selection, Schweinberger (2014, p. 185, 379) reported frequencies of the discourse marker *like* ranging from 0.49 per

one thousand words (ptw) for British English, 1.51 ptw for Indian English, 2.18 ptw for New Zealand English, and 2.23 ptw for Philippine English, up to 4.38 ptw for Canadian English. These differences underline the importance of investigating the use of *like* in other English varieties.

Whereas many studies focus on native speakers of English, fewer studies target non-native speakers of English (such as L2 or foreign language users) or ELF varieties (however, there seems to be an increasing interest in analyzing discourse markers among second or foreign language learners, see, for example, Gilquin, 2016). Keeping in mind the frequency differences across English varieties discovered in Schweinberger (2014) and the fact that, typically, the use of discourse markers such as *like* are not normally taught to foreign language learners in schools (Mukherjee and Rohrbach, 2006, p. 216; Rüdiger, 2021, p. 2; Wolk et al., 2021, p. 10), the use of *like* among non-native speakers proves particularly insightful. Moreover, non-native speakers of English are often found to use discourse markers less frequently and differently than native speakers (Liao, 2009; Gilquin, 2016). Mukherjee and Rohrbach (2006) provided an explanation for this. They argue that discourse markers are among the more challenging elements when learning a foreign language, and thus, they are typically acquired relatively late (Mukherjee and Rohrbach, 2006, p. 213). Yet the (correct) use of discourse markers contributes to sound more native-like or, in other words, more natural or idiomatic (Wolk et al., 2021, p. 9; see also Liao (2009)). The following selected findings of studies about discourse markers in general as well as *like* in particular underline that more research focusing on non-native speakers of English is needed.

Hasselgren (2002), for instance, discovered that higher proficiency, or rather higher fluency in English, resulted in a more target-like use of what she called "smallwords." The discourse marker *like* was among the smallwords investigated in her study comparing native speakers of English with Norwegian learners of English, grouped into more fluent and less fluent users of English (Hasselgren, 2002). *Like* appeared among those smallwords, which were acquired comparably late and thus require a certain level of proficiency. Similarly, in a study on pragmatic markers, Neary-Sundquist (2014) found that with increasing English proficiency of Korean and Chinese L2 learners, the use of these pragmatic markers increased as well. This conclusion was based on overall frequencies, and no specific mention of the discourse marker *like* was made, even though *like* had been part of the analysis. Gilquin (2016) also underlined the importance of proficiency, but in particular language skills acquired through naturalistic language exposure. More precisely, Gilquin (2016, p. 216) noted that for non-native learners to acquire the use of discourse markers, "exposure to naturalistic speech outside the classroom" is particularly important. Moreover, she found that native speakers of English (United Kingdom) used the discourse marker *like* approximately 3.5 times more frequently than non-native speakers of English (various L1s) (Gilquin, 2016, p. 220). In addition, she reported a statistically significant difference between those foreign language learners who had spent some time in an English-speaking country versus those who had not, with the

former using *like* more frequently than the latter (Gilquin, 2016, p. 221). She further exemplified that with the increasing length of stay (particularly so after 10 months), the frequency of *like* increased (Gilquin 2016, p. 227). This underlines the importance of exposure to language use outside of normative EFL contexts, where the (over)use of discourse markers would rather be discouraged, in order to acquire this discourse marker. Another striking finding was the observation that some foreign language learner groups used *like* relatively frequently, whereas others barely produced it at all. Among the former were the Polish, Dutch, and Swedish, as well as the German and Spanish learners of English, who had at least some (naturalistic) exposure to English in their respective countries of origin, for instance, *via* media or the internet. Chinese, French, and Italian speakers of English, who were assumed to have more limited access to naturalistic English language, were among the latter group. Gilquin (2016) further predicted that English as a second language (ESL) speakers used discourse markers more frequently than EFL speakers. This could be confirmed with *like* occurring more frequently in the ESL than in the EFL data (Gilquin, 2016, pp. 240–241).

Müller (2005) also identified frequency differences between native and non-native speakers of English, namely, that the native speakers of English (US) used the discourse marker *like* more frequently than the foreign language learners of English (L1 German). More precisely, all American speakers used the discourse marker *like* at least once, whereas only less than 60% of the German learners of English had at least one occurrence of *like* in their utterances (Müller, 2005, p. 230). Furthermore, she found some age-related tendencies, namely, that the younger German learners of English used the discourse marker *like* more frequently than those in the middle-age group (Müller, 2005, p. 232). The speaker relationship is also shown to have an effect. Among friends, *like* was used more often than among strangers (Müller, 2005, p. 233). In addition, Müller showed that interaction in English in informal situations as well as using English as the primary means of communication at least occasionally had boosting effects on the use of discourse marker *like* (Müller, 2005, p. 239). Finally, the influence of American English was prominent in her study. She showed that those German learners of English who had spent time abroad in the United States had higher rates of *like* than those who had been to the United Kingdom (Müller, 2005, p. 239).

In a study by Wolk et al. (2021), *like* was the fifth most frequently used discourse marker, but it generally appeared relatively infrequently (Wolk et al., 2021, p. 23). Their findings are based on English major university students with German, Spanish, Bulgarian, or Japanese as native languages. They further noticed that the L1 background influenced the use of discourse markers. For example, *like* appeared most frequently among the Spanish students studying in Madrid (Wolk et al., 2021). Moreover, they could also show that length of English instruction positively correlated with the frequency of discourse marker usage (Wolk et al., 2021, p. 30).

Interestingly, Diskin-Holdaway (2021) could not attest to the differences in frequencies of *like* uses among Irish (L1) and Chinese or Polish (L2) speakers of English. Moreover, neither proficiency in English nor length of stay in Ireland was shown to influence the use of this discourse marker (Diskin-Holdway, 2021). A difference between the L1 and L2 speakers, however, was detected in the positioning of *like* within the clause. The former showed higher frequencies of clause-final *like* (Diskin-Holdaway, 2021). Yet one additional crucial finding identified by Diskin-Holdaway (2021) was that the use of *like* as a discourse marker was characterized by a high individual variation. Similarly, Liao (2009) also found a large individual variation among the six Chinese teaching assistants living in the United States. At first sight, this may be interpreted as a gender effect, namely, that, as is sometimes reported, female speakers use discourse markers more frequently than their male peers (Liao, 2009, p. 1321). Yet in Liao's (2009) study, one of the female speakers showed considerably lower rates than the male speakers, underlining that a simple generalization would be imprecise. Furthermore, *like* appeared more frequently in personal interviews than in discussions. Thus, the register turned out to be more conclusive than gender (Liao, 2009, p. 1326). Finally, she remarked that L2 speakers should not be considered as homogenous groups but rather analyzed as individuals with distinct and complex identities (Liao, 2009, p. 1326).

A high level of individual variation was also detected in Rüdiger (2021), even though *like* was overall found to be a prominent feature of the young and educated Korean speakers of English investigated in her study. These findings are based on personal interviews conducted between the author (a young female German speaker of English) and one Korean speaker (i.e., the setting is quite comparable to the setting of the current study, see *The LARES Project*). She identified that, on average, each speaker used approximately eight instances of *like* per one thousand words (Rüdiger, 2021, p. 7).[5] Strikingly, four speakers did not use *like* as a discourse marker at all, whereas one speaker used it as frequently as 70 times per one thousand words (Rüdiger, 2021, p. 8). Even though the speaker variation was quite large, Rüdiger (2021, p. 8) found effects of time spent in an English-speaking country and self-reported proficiency in English but no effect with respect to gender. Those who had spent time abroad and those who reported having higher proficiency in English used *like* more frequently (Rüdiger, 2021, 9).

In summary, the most important variables argued to have an impact on the use of the discourse marker *like* are proficiency in English, input to (naturalistic) English language use, length of stay in an English-speaking country, the status of English (native speaker versus ESL versus EFL), age, and register or the context of language use, including formal versus informal speech. Moreover, a special role may be assigned to influence from American English, and the L1 of the English learners might also have an influence on the use of *like* as a discourse marker. Finally, the effect of gender remains inconclusive, as some studies

---

[5]Rüdiger (2021) used the same classification as D'Arcy (2017). This means that she differentiated between discourse marker and discourse particle uses. The frequency reported here represents the sum of both uses.

acknowledge the role of gender (with higher frequencies for female speakers), whereas others do not find differences between female and male speakers.

Based on the preceding discussion on the use of the discourse marker *like* among different (mainly L2) users of English with at times conflicting results, the current study seeks to answer the following two research questions, shifting the focus from L2 English to ELF users.

RQ1: How frequent is the discourse marker *like* in university student oral interviews conducted in Sharjah and how does its use differ in comparison to other varieties of English?

RQ2: What is the influence of social variables (e.g., gender, citizenship, and year of birth) on the frequency of use?

Before answering these research questions by presenting the results in *Results*, the next section will outline the methodology of the study.

METHODOLOGY

The following three subsections 1) briefly describe the larger project this study is part of, 2) present the participants as well as the social variables relevant to the analysis, and 3) introduce the spoken corpus as well as explain the coding and the subsequent corpus analysis.

The LARES Project

The study employs a subsample of a spoken corpus that consists of semi-structured interviews, approximately 30 min each. In total, 116 students attending the American University of Sharjah participated in the interviews, out of which 58 randomly selected interviews (50%) make up the corpus used in the current study. The participants come from a variety of linguistic backgrounds and include both Emirati and non-Emirati populations. As indicated earlier, some of the interviewees live in Sharjah and others in neighboring emirates, for instance, Dubai. The interviews were conducted by three young, female researchers in March and April 2019 as part of a larger project on LARES (2019–2021). The author of this study was one of the interviewers. Two of the interviewers have a German background, and one has an Iranian background but grew up in Germany. Each interview included questions targeting family background, educational history, specific language biographies, and attitudes towards English, Arabic, and other languages in the students' repertoires. Prior to participating in the interviews, the students completed a comprehensive online survey that was specifically developed for the project LARES based on Siemund et al. (2014) and Leimgruber et al. (2018).[6] The

TABLE 1 | Overview of participants (citizenship and L1).

Citizenship	L1 Arabic	L1 English	Other L1	Total
Arab expatriate	12	8	—	20
Emirati	6	6	—	12
Other	—	4	3	7
South Asian	—	18	1	19
Total	18	36	4	58

survey consisted of questions and agreement statements concerning the demographic, educational, and socioeconomic background, as well as language use and language attitudes (see Siemund et al., 2020; Siemund and Leimgruber, 2021).[7] With this, it is possible to outline and assess the migration history, educational background, language history, and attitudes towards English and Arabic. This detailed information nicely complements the spoken corpus and does not only allow investigating the use of *like* but makes it possible to correlate it with different social (non-linguistic) variables.

A limitation of the dataset used in the current study is that only university students' language production is included. This necessarily restricts any claims made further below to this particular population. Moreover, only one specific genre (i.e., one-to-one interviews) of one specific English variety (English spoken in the UAE) at one specific point in time (2019) is considered here. Ideally, including "a wide array of genres" of several ELF varieties, potentially even at different times or with different generations, would be a broader basis for this kind of investigation (Laitinen, 2020, p. 430). As such, it is not possible to provide a language change perspective per se. Instead, this study can only document the use of *like* from a synchronic perspective and in relation to other studies.

LARES Participants and Social Variables

The current corpus includes 58 semi-structured interviews, conducted with female ($n = 27$) and male ($n = 31$) university students. Their mean age is 20.2 (SD = 1.5) and ranges from 17 to 24. Fifteen different citizenships are represented in this sample, grouped into four distinct groups, namely, Emirati, Arab expatriate, South Asian, and other.[8] The students are part of all four colleges at the American University of Sharjah, that is, Architecture, Art, and Design ($n = 3$); Arts and Sciences ($n = 14$); Engineering ($n = 28$); and business administration ($n = 13$).

In the online survey, the students were asked to rank the languages they speak, starting with the language they are most proficient in. Thirty-six students ranked English first, followed by 18 who ranked Arabic first. Four students reported another language to be their most proficient or dominant language.

---

[6]As part of LARES, 692 students completed the online questionnaire. Part of the questionnaire was to signal a willingness to additionally participate in a semi-structured interview. From those indicating their interest, 116 students finally participated in the semi-structured interview. These interviews make up the LARES corpus.

[7]The sociolinguistic variables applied in this study reflect Anglo-American/European settings, i.e., gender, citizenship, and language background. By choosing these variables, comparability with previous research and other sociolinguistics studies can be assured.

[8]The 15 citizenships are Argentina, Bahrain, Bangladesh, Egypt, India, Iraq, Jordan, Lebanon, Morocco, Pakistan, Russia, Sri Lanka, Tunisia, the UAE, and the USA.

TABLE 2 | Number of participants per number of languages.

Number of languages	Total
2	9
3	20
4	19
5	7
6	3

The language ranked the highest will henceforth be referred to as L1. Note that this does not necessarily have to be the native language or the language acquired chronologically first. This is visible from **Table 1**, where citizenship and the respective L1 are presented. In the online survey, 14 of the Emiratis or Arab expatriates indicated English to be their strongest language (L1), yet in the interviews, they reported Arabic to be their native language.

In both the online survey and during the interviews, the students were asked how many languages they knew or had some proficiency in. Particularly during the interviews, the interviewers stressed that all languages counted, i.e., also those in which the participants had a relatively low proficiency such as a foreign language studied for some years in school a while ago. When comparing the responses from the survey with those of the interviews, it is quite striking that 27 students reported knowing a higher number of languages during the interviews. At times, they remembered halfway through the conversation that they had learned French in school for some years, for instance. Here, the numbers from the interviews are used, as they seem to better reflect the multilingual repertoires of the students. The majority of the current sample has at least some proficiency in either three or four languages (*n* = 39). The overall distribution can be found in **Table 2**.

In the online survey, the students were asked to self-assess their proficiency in English. Using a scale from one (mastery) to six (beginner), they should assign a score to listening/ understanding, speaking fluency, reading proficiency, and writing proficiency separately. The proficiency score used in the current study is the resulting mean of the four individual measures.

In addition to these rather traditional sociolinguistics variables, a new variable was created. This variable is based on the interview data and is called "English usage score." The main motivation for this analysis step was that, as will become apparent in the *Results* section, the previously described social variables turned out not to be significant predictor variables of the distribution of the discourse marker *like*. Moreover, previous research has shown that learners who are exposed to "naturalistic English" seem to use the discourse marker *like* more frequently (Gilquin, 2016, p. 244). Since all participants are students enrolled in an American University, hence, the medium of instruction is English, and they are all clearly exposed to English on a regular basis. However, there may be differences in English usage outside of university as well as stemming from their school education. For instance, the use of English (social) media may show dissimilarities. In addition, some students may have attended

schools with English as the educational language, whereas others may have attended schools using other instructional languages such as Arabic.

In order to get a better understanding, each interview was coded for seven sub-variables, which are "English TV/movies," "English-speaking country," "Length of stay in an English-speaking country," "Language of instruction in school," "School system," "Pick one language to keep,"[9] and "Best language." After coding, two variables were dropped, namely, "English TV/movies" and "Pick one language to keep." The former was removed, as practically everyone reported watching movies or TV in English, and more specific details (such as "frequently" or "only occasionally") were impossible to determine from the interviewees' responses. Two students did not specifically mention that they watched films or TV in English. Yet they did not deny it either but rather stated that they enjoyed Hindi and Arabic movies a lot. Thus, it can be assumed that all students are in some way or another exposed to English *via* movies or TV, and this variable could therefore be discarded. Unfortunately, in ten interviews, the question as to which language they would keep if they had to choose one was not asked. To avoid the reduction of the dataset, this variable was also removed from the English usage score. Finally, four of the remaining sub-variables were coded as two, as they belonged to the same category (see the explanation of the final variables below).

The refinement of the English usage score is thus composed of three different measures, "English-speaking country" (i.e., number of times traveled to an English-speaking country: 0, never; 1, once or twice/infrequently; 2, regularly, frequently, often), "Language of instructions" (i.e., school education received in: 0, Arabic or another language than English; 1, English and another language; 2, English only), and "Best language" (i.e., the language they feel most comfortable with: 0, Arabic or another language than English; 1, English and another language; 2, English). The resulting English usage score is the sum of the three measures ranging from zero to six.

LARES Corpus: Data Coding and Analysis

The recordings of 58 interviews that represent 50% of the entire dataset were transcribed by one person and checked by another one. XML tags were used to distinguish between the interviewee's (iwe) and interviewer's (iwr) utterances. For the current study, exclusively, the interviewee data were analyzed. This part of the corpus contains 139,630-word tokens of orthographically transcribed speech, with an average of 2,407-word tokens per interview file. Strikingly, *like* turned out to be the third most frequent word form (*n* = 3,937), which clearly supports the importance of investigating its usage.

All instances of *like* were extracted from the LARES corpus using the concordance program AntConc (Anthony, 2018). In a second step, all hits were manually annotated. In the first round of coding, the discourse marker *like* was differentiated from other

---

[9]One of the questions the interviewees were asked was "If you have to choose only one of the languages you know, which one would it be?"

uses of *like*, such as verbs, comparative prepositions/complementizers, nouns, suffix, quotative *like*, or repetitions, following D'Arcy (2017, pp. 3–13) and Schweinberger (2014, pp. 140–143). In the second round of coding, the initial decision was checked and, if necessary, corrected; and each instance of a discourse marker was further categorized depending on the clausal position. This coding step was based on Schweinberger's (2014, pp. 146–149) coding scheme. Thus, the current study distinguishes between clause-initial (INI), clause-medial (MED), clause-final (FIN), and non-clausal (NON) *like*. Examples (4) to (7) represent each type.

(4) *Like English is easy to communicate but not every local knows English.* (f20) [INI]

(5) *I mean um my academic career is like mostly English* (m3) [MED]

(6) *[...] but we took like English as a course like.* (f14) [FIN]

(7) *Okay Arabic is for me it's the like the its more complex than any of the others.* (f22) [NON]

The absolute frequencies of the discourse marker *like* were then, in order to ensure comparability across the interview files, normalized to the basis of 1,000 words. As a next step, each social variable was investigated separately to assess its relation with the frequency of *like*. For this, the variables identified as influencing the use of discourse markers in general or *like* in particular (see *Discourse Marker Like*) and those assessed *via* the online survey (see *Discourse Marker Like* and *LARES Participants and Social Variables*) were used. Following this monofactorial analysis, a generalized linear regression analysis was run.

RESULTS

The following three subsections present the results of the corpus analysis. First, the overall (absolute) frequency of *like* is discussed. Second, the uses of the discourse marker *like* versus social variables are exhibited based on the normalized frequencies, and third, the most frequent *like*-users are examined as a separate cohort.

Frequency of *Like*

Research question 1 asked how frequently the discourse marker *like* appeared in the spoken corpus and how this differed in comparison to other varieties of English. On the whole, *like* is the third most frequently used word in the interviewees' utterances ($n = 3,937$), with 2,951 (75%) uses as a discourse marker and 986 (25%) other uses.[10] The classification of the discourse marker uses shows that clause-initial ($n = 1,466$, 50%) and clause-medial ($n = 1,206$, 41%) make up the largest part, whereas non-clausal ($n = 235$, 8%) and particularly clause-final ($n = 44$, 1%) uses are

relatively infrequent. The mean frequency per 1,000 words (ptw) across the entire corpus is 19.5 (median: 16.0), yet the relatively high SD of 14.75 shows that the individual variation among the speakers is comparably large. The lowest frequency is 0.51 ptw, and the highest is 55.14 ptw. The visualization across 11 intervals (see **Figure 1**) shows that lower frequencies (from 1 to 20 or perhaps even until 30 discourse marker uses ptw) are more often represented in the dataset than higher frequencies (from 30 to 55 ptw).

A closer look at the social background variables in combination with the normalized frequencies of the discourse marker *like* provides a more detailed picture.

Like Versus Social Variables

To answer the second research question, namely, how the social background of the speakers influences the use of *like*, seven different social variables will be juxtaposed with the normalized frequencies of the discourse marker. These are gender, citizenship, L1, year of birth, number of languages, college, self-assessed proficiency in English, and the English usage score. For all statistical tests, the free software R (R Core Team, 2020) is used.

Even though the mean frequency of the female students (21.1) is slightly higher than the mean frequency of their male peers (18.0), the Wilcoxon test did not return a statistically significant difference ($W = 468.5$, $p = 0.22$). This suggests that there is no difference between male and female speakers with respect to the frequency (ptw) of the discourse marker *like* in the current sample.
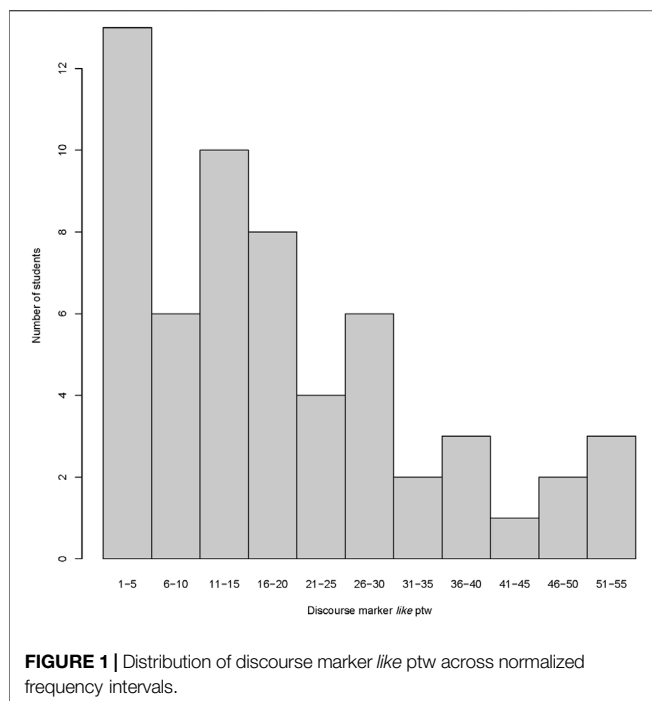
A similar result was obtained for the Kruskal–Wallis test when testing citizenship (Emirati, Arab expatriate, and South Asian)[11] versus normalized frequency of discourse marker *like*. No statistical significant difference was attested (H (3) = 2.5, $p = 0.47$), confirmed by nonsignificant post-hoc tests.

An equally nonsignificant difference resulted from the Wilcoxon test ($W = 376.5$, $p = 0.17$) when comparing the means of the discourse marker use with respect to L1 English and Arabic. This means that citizenship is also not a good predictor when explaining frequency differences in the current sample.

When correlating the age of the participants (range from 17 to 24) with the normalized frequencies of *like*, a very small negative correlation coefficient was obtained ($r = -0.09$); yet this very weak negative correlation did not reach statistical significance ($p = 0.53$). Perhaps the age range of the participants is too small to find a correlation. A less homogeneous group may indeed return a significant correlation of age and frequency of *like*.

---

[10]The pronouns *I* and *it* are the two most frequent words in the LARES corpus, followed by *like*, *in*, *the*, and *to*.

[11]For the monofactorial statistical analysis, only Emirati, Arab expatriate, and South Asian students were considered. Those with another citizenship had to be excluded, as this group consisted of four students only. The same procedure was followed for L1 as well as college. Those who reported another L1 than English or Arabic were not considered for the statistical analysis, because only four students belonged to this group. The college of Architecture, Art, and Design ($n = 3$) was excluded, and only the colleges of arts and sciences, business administration, and engineering were featured in the statistical analysis.

FIGURE 1 | Distribution of discourse marker *like* ptw across normalized frequency intervals.

The number of languages in the students' repertoires was also correlated with the normalized frequencies of the discourse marker *like*. The correlation coefficient is positive even if very small (0.11), and it does again not reach statistical significance ($p = 0.20$). Once more, no statistically significant difference of a background variable, here the number of languages, can be attested.

The Kruskal–Wallis test was used to compare the mean frequencies of *like* among the students attending the different colleges including the Colleges of Arts and Sciences, Business Administration, and Engineering. The results show that there is no statistically significant difference across the three colleges based on the Kruskal–Wallis test statistics ($H (2) = 2.70$, $p = 0.26$) as well as following post-hoc tests.

A correlation analysis of the self-assessed English proficiency versus the normalized frequencies of *like* returned a very small negative correlation ($r = -0.04$), which is not statistically significant ($p = 0.77$). The following Wilcoxon test based on two groups (low versus high proficiency) derived *via* the median split (1.625) does not return a statistically significant difference between the two groups ($W = 408.5$, $p = 0.22$) either.

Finally—and as initially indicated because all preceding variables returned no statistically significant differences—the English usage score was correlated with the normalized frequencies of the discourse marker *like*. The resulting correlation coefficient is once again very small ($r = -0.11$), and this negative correlation is also not statistically significant ($p = 0.42$). As before, a Wilcoxon test comparing those who have a low English usage score (0–3) versus those with a higher score (4–6) does not reach statistical significance either ($W = 342$, $p = 0.20$).

What this means is that none of the social variables considered in this study turn out to statistically significantly differ with respect to the use of the discourse marker *like*. This was ultimately confirmed with a generalized linear regression analysis, using a Poisson regression. The social variables did not statistically significantly contribute to explaining the variance in the frequency of the discourse marker *like*.[12] A logical next step is then to have a closer look at those students who have the highest ratios of *like* used as a discourse marker and to investigate if these students share specific characteristics that could not be identified with the preceding analyses.

Most Frequent *Like*-Users

As an extension of research question 2, the ten most frequent *like* users were separately analyzed in order to identify features they share. This additional analysis step may allow further conclusions as to which social variables are particularly strongly associated with high frequencies of *like*. The ten speakers with the highest ratios of *like* (between 33 and 55 ptw) form a relatively heterogeneous group. A close inspection of the characteristics they have in common reveals that they present the entire range of all possible variable manifestations. Five male and female speakers are represented, some reported to know only two or three languages, others know four or more, and there are Emirati, Arab expatriate, and South Asian students present. The only citizenship group missing is other, but we have to keep in mind that in the entire sample, there were only four students with citizenship other than the three mentioned here. One interesting indication may perhaps be the college associated with the students. All but one attend the college of engineering (one is part of the college of arts and sciences). Yet it has to be acknowledged that nearly 50% of all students are part of engineering, which clearly increases the likelihood of appearing among this subsample as well and may thus be a sampling condition instead of a finding. Moreover, more of these ten students reported English to be their L1 ($n = 6$), and fewer ranked Arabic first ($n = 4$). With this ratio of 3:2, there are fewer English L1 speakers than in the overall sample (which has a ratio of 2:1). Half of the participants indicated that they have very high competencies in English (a self-assessed proficiency score of 1.25 or lower), whereas the other half rated their English skills slightly lower (four scored two or lower, and one scored 3.75). The mean score, however, is only marginally, if at all, higher among the ten students (1.65) in comparison to the entire sample (1.69), and it is in general relatively high. Even though so far self-assessed proficiency in Arabic has not played a role, it had been considered for this final analysis. The scoring procedure was the same as for English (i.e., from 1 to 6), and among the ten students, there are some with high skills in Arabic (1) up to relatively low skills (3.75), in addition to two students who indicated that they do not speak any Arabic.

---

[12]Two separate generalized linear regressions were fit. The first included all main effects, and the second allowed two-way interactions of all variables. *Via* model selection (backward stepwise) based on *p*-values (threshold 0.05) (see, for example, Gries, 2021, p. 366), all predictors (interactions as well as main effects) had to be dropped, as they did not contribute statistically significantly to the models. No predictors remained.

All in all, no specific pattern could be identified, which at first may seem even disappointing. Yet, and this will be argued for in the next sections, this may indeed produce some interesting implications. Arguably, there is quite a bit of variation across the students, yet none of the variables considered seem to have a particularly strong association with the use of the discourse marker *like*. In the following discussion, these results will be looked at in relation to findings from other studies investigating this discourse marker.

DISCUSSION

The results indicate that *like* used as a discourse marker is a prominent feature among the students investigated here. The first subsection of the discussion considers the overall frequency of *like* in the semi-structured interviews and discusses this in light of the earlier studies investigating different varieties and different speakers of English. The second subsection looks at *like* in combination with the social background variables of the students.

Frequency of *Like*

The overall frequency of the discourse marker *like* in the current study is surprisingly high (mean = 19.50 ptw; SD = 14.75), particularly so when compared to the frequencies found in Schweinberger (2014). He reported mean frequencies for British, Indian, New Zealand, Philippine, and Canadian English ranging from 0.45 to 4.39 instances of *like* per one thousand words (Schweinberger, 2014, p. 185). This may be partly due to the specific genre used in the current study, namely, semi-structured, personal, and relatively informal interviews, as opposed to Schweinberger (2014) who relied on the International Corpus of English (ICE), which represents various types of spoken language. This genre-related argument can be supported with the results discussed in Fuller (2003) investigating American English. She noticed that *like* appeared relatively frequently in her study (11.6 ptw). These findings were also based on interviews. In interviews, the use of discourse markers may be pragmatically useful, particularly in interviews perceived as relatively informal and personal, even though Fuller (2003) thought it was remarkable that the speakers in her study used *like* with such a high frequency. The reason for her surprise was the presumed stigmatization of *like* and its association with "a lack of intelligence" (Fuller, 2003, p. 369). This, however, could not be confirmed in Fuller's (2003) research with young speakers of US English. Yet the frequencies reported in her study are still considerably lower compared to the mean frequency found in the LARES corpus.

In a study equally based on personal one-to-one interviews, Rüdiger's (2021) investigation of Korean English found a mean frequency of *like* used as a discourse marker of approximately eight per one thousand words. However, similar to the current study, she noticed high rates of internal variation with one speaker having a frequency of 70 uses of *like* per one thousand words. This is even more extreme than what was found in the LARES corpus, where the most frequent *like*-user had a frequency of 55.14 uses per one thousand words.

Two potential explanations may be feasible to interpret the high speaker variability as well as the overall high frequency of *like*. On the one hand, ELF may be particularly prone to show variability across speakers of one speech community, irrespective of their linguistic or social background (see *Like Versus Social Variables* for more details). On the other hand, from a language change perspective, such a generally high frequency of *like* may also be explained with the specific setting in which English is used. Users of ELF have a multilingual background and comparable multilingual contexts have been shown to accelerate language change (see, for example, Laitinen, 2020, p. 428). The use of *like* as a discourse marker has seen a recent increase in varieties of English (D'Arcy, 2017, pp. 14–15), and the multilingual context present in the UAE may be responsible for an even greater frequency increase. In addition, influence through (digital) media, particularly from the United States, may further advance this development (for more information about the influence of the United States, see the following section).

Like Versus Social Variables

The most important finding is that the social background of the speakers cannot explain the variability identified in the use of the discourse maker *like*. The statistical analysis presented in *Like Versus Social Variables* demonstrated that the female students did not use the discourse marker *like* significantly more frequently than their male peers, even though the mean frequency of the female students (21.1 ptw) was slightly higher than that of the male students (18.0 ptw). Both Fuller (2003) and D'Arcy (2007) identified a female speaker lead among speakers of American English, yet Schweinberger (2014, p. 393) presented a more diverse picture and argued for "variety-specific" differences with respect to gender. Moreover, Rüdiger (2021) could not identify differences with respect to gender either. Perhaps, in our modern and globalized era, it may be time to focus more on other variables than on the binary variable gender or to approach it as less binary. Cleary, it is a very convenient variable, relatively straightforward to code and include in research. Yet it may not necessarily be a reliable indicator of discourse marker usage (or other markers, for that matter). Perhaps a more fine-grained or scalar category, potentially not with respect to gender but perhaps a personality marker instead (i.e., in relation to the big five personality traits), would shed more light on uses of *like* as a discourse marker.

The slight tendency that younger students used the discourse marker *like* more often than older students would further substantiate that the use of *like* is increasing in the English of the UAE students investigated here. However, this needs to be taken with great caution, as the negative correlation was extremely small and did not reach statistical significance. This may be due to the small age range of the students and further research including older (and also younger) groups of people may further substantiate this observation.

Furthermore, this study did not identify statistically significant differences in the frequency of *like* pertaining to

citizenship, L1, number of languages, or college. A few words pertaining to citizenship are in order. Similar to gender, citizenship as a definite category may not necessarily reflect the multilingual and multicultural identities of these students. Perhaps the participants could be better classified as global citizens having a "modern global identity" (Fuller, 2020, p. 167). Many of them have moved multiple times across countries or even continents, and they converse on a regular basis with many different individuals coming from diverse backgrounds. Yet a relative closeness towards especially American English but also British English cannot be denied. Many of the students have either followed an American or a British school curriculum, they all attend an American university, and quite a number of students associate their English either with American or British English or aim at using one of the two varieties. Particularly because of their affiliation with an American university, influence from the United States could be argued to be relatively prominent, which, in turn, could have an effect on the use of the discourse marker *like* as shown by Müller (2005). Nevertheless, parallels may be drawn to what Fuller (2020, p. 167) means when she talks about "a modern global identity that is not linked to any particular nationality" where English is "used to convey a cosmopolitan connotation." Even though Fuller (2020) argued this to be true for English in Germany—and this context may be understood as distinctly different from the UAE context—it is still worthwhile extending this to the current study.

Thus, in light of the literature review introduced above, it should be acknowledged that L2 learners are not straightforwardly comparable to the ELF users considered in the current study. The main reason is that the LARES participants are advanced speakers of English who use this language during their studies and mostly also outside of the university. Hence, proficiency, which had been identified as a predictor variable of usage rates of *like* (see Hasselgren, 2002; Neary-Sundquist, 2014), is necessarily high among all interviewees (which was already visible in the self-reported proficiency ratings). It may thus be less surprising that the variable self-assessed proficiency did not turn out as a significant predictor. Including further speakers with more variability in relation to proficiency in English may further corroborate the claim that proficiency impacts the use of the discourse marker *like*. Moreover, as indicated in Gilquin (2016), it is access to naturalistic language input outside of the English language classroom that particularly boosts the acquisition and use of discourse markers (see also Liao, 2009, p. 1314). More than half of the participants reported that English was their dominant language. Even the remaining students, whose dominant language is either Arabic or another language, can be assumed to have access to English on a regular basis, first because of their studies (English is the exclusive medium of instruction) in addition to media consumption and interaction with peers and partly even within the home. English truly plays an important role in the lives of these young students. They are

in fact confident speakers and may also identify with this language. The following two quotes, taken from the LARES interviews, underline this.

(8) *So like I said earlier when it comes to like English it's very it's very much like a tool.* (m28)

(9) Answer to the questions of which language to keep, if only one could be kept: *English.* [Interviewer: *Why*?] *Coz like, the other languages* [English, Hindi, Urdu, Arabic, Farsi, Russian] *are like small parts of my life, but like my life runs in English.* (m12)

English is understood as a tool, it dominates the lives of the students, and it may even replace the native language in terms of daily use and importance. Employment aspirations and the prospective importance of English with respect to future careers are certainly two of the driving factors. Nevertheless, it is imperative to acknowledge that other languages, particularly Arabic, also play a role (Thomas, 2021). Yet as student m12 admits in (9), these languages compete on a different level with English and typically are of secondary importance. To be clear, this is true for the particular population considered here and may not necessarily be generalized to other groups residing in the UAE (or even students attending another university).

Moreover, what has to be acknowledged is that the discourse marker use of *like* seems to be a relatively prominent feature of the English repertoires of the UAE students. In a sense, these speakers could therefore be seen to be somewhat comparable to the Korean English speakers analyzed in Rüdiger (2021). Clearly, Korea and the UAE are two quite different geographical locations and social realities, but perhaps the younger generations in a globalized world are not that different anymore. As argued above, young people and their access to media may make it relevant to assess the global context in addition to or instead of simply regarding citizenship or country of residence. However, the higher rates of *like* among the UAE students in comparison to the Korean speakers in Rüdiger's (2021) study may indeed hint at higher English proficiency and more frequent English use of the former.

Finally, even though many speakers considered here made use of *like* as a discourse marker relatively frequently, this was not true for all speakers. The observed individual variation might be the result of the particular multilingual setting investigated. One characteristic of ELF encounters is that speakers with various linguistic backgrounds are in contact. Mauranen describes this context as a speech situation "[w]ith vast numbers of similects coming into contact with each other" (2017, p. 230). What this means is that different Englishes are in contact, which itself are influenced by the L1 or L1s of each person (at least in the case of non-native speakers of English) in addition to the specific English acquisition setting (Mauranen, 2017, p. 227). Because of this heterogeneity, Mauranen (2017, p. 230) assumes a substantial amount of variability in ELF uses. The present study provides evidence for the high variability of *like* among highly proficient

English users in the UAE. Whether this holds true for other features of the English spoken in the UAE remains to be seen in future research.

CONCLUSION AND OUTLOOK

This study contributes to the understanding of the use of the discourse marker *like* in ELF, more precisely in the English spoken by university students in the UAE. This is important for world Englishes at large, as ELF has so far not been in focus with respect to the use of discourse markers.

Like used as a discourse marker appeared to be a frequent and prominent feature in the semi-structured, personal, and relatively informal interviews between young female researchers and young multilingual and multicultural university students in Sharjah. The social variables included in this study (i.e., gender, citizenship, L1, year of birth, number of languages, college, self-assessed proficiency in English, and English usage score) returned no statistically significant differences. Even though variability among the students was relatively high, all participants used the discourse marker *like* at least once. Therefore, it could be argued that *like* presents a prominent marker of the English spoken by this elite and educated group of speakers. The relatively high frequency of *like* might support the hypothesis of accelerated language change in multilingual settings such as the UAE, where different Englishes of speakers with various L1s are in contact. Moreover, high individual variability among the heterogeneous speakers might be an additional characteristic of ELF users.

What this study cannot provide is a broader view of other groups living in the UAE. Future studies focusing on speakers other than university students may find out how *like* is distributed more generally in the English spoken in the UAE. This study is thus only a starting point for understanding the English used in the UAE, and further research, including studies on other linguistic features, is urgently needed.

Finally, Siemund (2022) argued that even though English in the UAE can be considered a lingua franca, it may in fact rather be categorized as a second language. The findings of the current study may perhaps offer modest support of this claim, particularly so for young university students,

demonstrated by the high rates or frequent use of *like* as a discourse marker.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because of the agreement signed in the IRB form. Requests to access the datasets should be directed to EL, eliane.lorenz@anglistik.uni-giessen.de.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Office of Research, American University of Sharjah, UAE, *via* the IRB Application Form. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

The author was involved in the data collection and confirms being the sole contributor of this article.

ACKNOWLEDGMENTS

REFERENCES

Ahmad, R. (2016). Expatriate Languages in Kuwait: Tension between Public and Private Domains. *J. Arabian Stud.* 6 (1), 29–52. doi:10.1080/21534764.2016.1192767

Ahmed, K. (2021). "Linguistic and Semiotic Landscapes of Dubai," in *The Multilingual Global Cities: Singapore, Hong Kong, Dubai*. Editors P. Siemund and J. R. E. Leimgruber (Singapore: Routledge), 185–202.

Al-Issa, A. (2021). "Multilingualism, Language Management, and Social Diversity in the United Arab Emirates," in *Multilingual Global Cities: Singapore, Hong Kong, Dubai*. Editors P. Siemund and J. R. E. Leimgruber (Singapore: Routledge), 116–130.

Anthony, L. (2018). *AntConc (Version 3.4.4.0) [Computer Software]*. Tokyo, Japan: Waseda.

Boyle, R. (2012). Language Contact in the United Arab Emirates. *World Englishes* 31 (3), 312–330. doi:10.1111/j.1467-971x.2012.01749.x

Boyle, R. (2011). Patterns of Change in English as a Lingua Franca in the UAE. *Int. J. Appl. Linguistics* 21 (2), 143–161. doi:10.1111/j.1473-4192.2010.00262.x

Cook, W. R. A. (2017). More Vision Than Renaissance: Arabic as a Language of Science in the UAE. *Lang. Pol.* 16, 385–406. doi:10.1007/s10993-016-9413-3

Corrigan, K. P., and Diskin, C. (2019). 'Northmen, Southmen, Comrades All'? the Adoption of Discourse like by Migrants north and South of the Irish Border. *Lang. Soc.* 49, 745–773. doi:10.1017/s0047404519000800

D'Arcy, A. (2017). *Discourse-Pragmatic Variation in ContextEight Hundred Years of LIKE*. Amsterdam: Benjamins.

D'Arcy, A. (2007). Like and Language Ideology: Disentangling Fact from Fiction. *Am. Speech* 82, 386–419.

Davidson, C. M. (2005). *The United Arab Emirates: A Study in Survival*. Boulder, CO: Lynne Rienner.

Diskin, C. (2017). The Use of the Discourse-Pragmatic Marker 'like' by Native and Non-native Speakers of English in Ireland. *J. Pragmatics* 120, 144–157. doi:10.1016/j.pragma.2017.08.004

Diskin-Holdaway, C. (2021). You Know and like Among Migrants in Ireland and Australia. *World Englishes*. doi:10.1111/weng.12541

Dubai Population (2020). Dubai Population 2020. Available at: https://worldpopulationreview.com/world-cities/dubai-population (Accessed September 1, 2020).

Filppula, M., Klemola, J., Mauranen, A., and Vetchinnikova, S. (2017). "Changing English: Global and Local Perspectives," in *Changing English. Global and Local Perspectives*. Editors M. Filppula, J. Klemola, A. Mauranen, and S. Vetchinnikova (Berlin/Boston: DeGruyter Mouton), xi–xiii. doi:10.1515/9783110429657-204

Fuller, J. M. (2020). "English in the German-Speaking World: Immigration and Integration," in *English in the German-Speaking World*. Editor R. Hickey (Cambridge: Cambridge University Press), 165–184.

Fuller, J. M. (2003). Use of the Discourse Marker like in Interviews. *J. Sociolinguistics* 7 (3), 365–377. doi:10.1111/1467-9481.00229

Fussell, B. (2011). The Local Flavour of English in the Gulf. *English Today* 27 (4), 26–32. doi:10.1017/s0266078411000502

Gabrys, M. (2017). "Like as a Discourse Marker in Different Varieties of English. A Contrastive Corpus-Based Study,". Master's Thesis (Louvain: Université catholique de Louvain). Available at: https://dial.uclouvain.be/memoire/ucl/fr/object/thesis%3A11925/datastream/PDF_01/view (Accessed January 20, 2021).

Gilquin, G. (2016). "Discourse Markers in L2 English," in *New Approaches to English Linguistics: Building Bridges*. Editors O. Timofeeva, A.-C. Gardner, A. Honkapohja, and S. Chevalier (Amsterdam: Benjamins), 213–249. doi:10.1075/slcs.177.09gil

Government of Dubai (2019). Population and Vital Statistics. Available at: https://www.dsc.gov.ae/en-us/Themes/Pages/Population-and-Vital-Statistics.aspx?Theme=42 (Accessed September 1, 2020).

Gries, S. Th. (2021). *Statistics for Linguistics with R. A Practical Introduction*. 3rd Ed. Berlin: De Gruyter Mouton.

Hasselgren, A. (2002). "Learner Corpora and Language Testing," in *Computer Learner Corpora, Second Language Acquisition, and Foreign Language Teaching*. Editors S. Granger, J. Hung, and S. Petch-Tyson (Amsterdam: Benjamins), 143–173. doi:10.1075/lllt.6.11has

Hopkyns, S. (2017). "A Conflict of Desires: Global English and its Effects on Cultural Identity in the United Arab Emirates,". Dissertation (Leicester: University of Leicester).

Hopkyns, S. (2021). "Multilingualism and Linguistic Hybridity in Dubai," in *Multilingual Global Cities: Singapore, Hong Kong, Dubai*. Editors P. Siemund and J. R. E. Leimgruber (Singapore: Routledge), 248–264.

Laitinen, M. (2020). Empirical Perspectives on English as a Lingua Franca (ELF) Grammar. *World Englishes* 39 (3), 427–442. doi:10.1111/weng.12482

Leimgruber, J. R. E., Siemund, P., and Terassa, L. (2018). Singaporean Students' Language Repertoires and Attitudes Revisited. *World Englishes* 37 (2), 282–306. doi:10.1111/weng.12292

Leimgruber, J. R. E., and Siemund, P. (2021). "The Multilingual Ecologies of Singapore, Hong Kong, and Dubai," in *Multilingual Global Cities: Singapore, Hong Kong, Dubai*. Editors P. Siemund and J. R. E. Leimgruber (Singapore: Routledge), 1–15.

Leuckert, S., and Rüdiger, S. (2021). Discourse Markers and World Englishes. *World Englishes*. doi:10.1111/weng.12535

Liao, S. (2009). Variation in the Use of Discourse Markers by Chinese Teaching Assistants in the US. *J. Pragmatics* 41, 1313–1328. doi:10.1016/j.pragma.2008.09.026

Madichie, N. O., and Madichie, L. (2013). City Brand Challenge 101: Sharjah in a Globalised UAE Context. *Ijbg* 11 (1), 63–85. doi:10.1504/ijbg.2013.055316

Mair, C. (2021). "World Englishes: From Methodological Nationalism to a Global Perspective," in *The Bloomsbury Handbook of World Englishes. Volume 1*. Editors B. Schneider and T. Heyd (London: ParadigmsBloomsbury), 27–45.

Mauranen, A. (2017). "A Glimpse of ELF," in *Changing English. Global and Local Perspectives*. Editors M. Filppula, J. Klemola, A. Mauranen, and S. Vetchinnikova (Berlin/Boston: DeGruyter Mouton), 223–254. doi:10.1515/9783110429657-013

Mauranen, A. (2012). *Exploring ELF. Academic English Shaped by Non-native Speakers*. Cambridge: Cambridge University Press.

Mukherjee, J., and Rohrbach, J.-M. (2006). "Rethinking Applied Corpus Linguistics from a Language-Pedagogical Perspective: New Departures in Learner Corpus Research," in *Planing, Gluing and Painting Corpora: Inside the Applied Corpus Linguist's Workshop*. Editors B. Kettemann and G. Marko (Frankfurt am Main: Peter Lang), 205–232.

Müller, S. (2005). *Discourse Markers in Native and Non-native English Discourse*. Amsterdam: John Benjamins.

Neary-Sundquist, C. (2014). The Use of Pragmatic Markers across Proficiency Levels in Second Language Speech. *Ssllt* 4 (4), 637–663. doi:10.14746/ssllt.2014.4.4.4

O'Neill, G. T. (2014). Just a Natural Move towards English: Gulf Youth Attitudes towards Arabic and English Literacy. *Learn. Teach. Higher Educ. Gulf Perspect.* 11 (1), 1–21.

Pacione, M. (2005). Dubai. *Cities* 22 (3), 255–265. doi:10.1016/j.cities.2005.02.001

Parra-Guinaldo, V., and Lanteigne, B. (2021). "Morpho-syntactic Features of Transactional ELF in Dubai/Sharjah," in *Multilingual Global Cities: Singapore, Hong Kong, Dubai*. Editors P. Siemund and J. R. E. Leimgruber (Singapore: Routledge), 303–320.

Piller, I. (2017). "Dubai: Language in the Ethnocratic, Corporate and Mobile City," in *Urban Sociolinguistics: The City as a Linguistic Process and Experience*. Editors D. Smakman and P. Heinrich (Abingdon: Routledge), 77–94. doi:10.4324/9781315514659-7

P. Siemund and J. R. E. Leimgruber (Editors) (2021). *Multilingual Global Cities: Singapore, Hong Kong* (Dubai. Singapore: Routledge).

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.R-project.org/.

Randall, M., and Samimi, M. A. (2010). The Status of English in Dubai. *English Today* 26 (1), 43–50. doi:10.1017/s0266078409990617

Rüdiger, S. (2021). Like in Korean English Speech. *World Englishes* 40, 548–561. doi:10.1111/weng.12540

Schneider, E. W. (2007). *Postcolonial English. Varieties Around the World*. Cambridge: Cambridge University Press.

Schweinberger, M. (2014). "The Discourse Marker LIKE: A Corpus-Based Analysis of Selected Varieties of English,". Dissertation (Hamburg: University of Hamburg).

Sharjah Population (2020). Sharjah Population 2020. Available at: https://worldpopulationreview.com/world-cities/sharjah-population (Accessed September 1, 2020).

Siemund, P., Al-Issa, A., and Al-Issa, 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

Siemund, P., Al-Issa, A., Rahbari, S., and Leimgruber, J. R. E. (2021). "Multilingualism and the Role of English in the United Arab Emirates, with Views from Singapore and Hong Kong," in *Research Developments in World Englishes*. Editor A. Onysko (London: Bloomsbury), 95–120. doi:10.5040/9781350167087.ch-006

Siemund, P. (2022). *Multilingual Development: English in a Global Context*. Cambridge: Cambridge University Press.

Siemund, P., Schulz, M. E., and Schweinberger, M. (2014). Studying the Linguistic Ecology of Singapore: A Comparison of College and University Students. *World Englishes* 33 (3), 340–362. doi:10.1111/weng.12094

Theodoropoulou, I. (2021). "Socio-historical Multilingualism and Language Policies in Dubai," in *Multilingual Global Cities: Singapore, Hong Kong, Dubai*. Editors P. Siemund and J. R. E. Leimgruber (Singapore: Routledge), 36–80.

Thomas, S. (2021). English in the United Arab Emirates. *World Englishes*. doi:10.1111/weng.12560

Thomas, S. (2016). "The Case of the 'Innocuous' Middle-Class Migrant Employee: English Language Use and Attitudes in Dubai, United Arab Emirates,". Dissertation (West Lafayette (IN): Purdue University).

Vertovec, S. (2007). Super-diversity and its Implications. *Ethnic Racial Stud.* 30 (6), 1024–1054. doi:10.1080/01419870701599465

Wolk, C., Götz, S., and Jäschke, K. (2021). Possibilities and Drawbacks of Using an Online Application for Semi-automatic Corpus Analysis to Investigate

Discourse Markers and Alternative Fluency Variables. *Corpus Pragmatics* 5, 7–36. doi:10.1007/s41701-019-00072-x

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.