PAPER

You know and *I think* in English(es) in Zanzibar

Susanne Mohr

Norwegian University of Science and Technology, Norway

Correspondence
Susanne Mohr, Department of Language and Literature, Norwegian University of Science and Technology, Norway.
Email: susanne.mohr@ntnu.no

Abstract

This paper provides an analysis of discourse markers in English(es) spoken by Zanzibaris, using data from sociolinguistic interviews. The focus is on *you know* and *I think*. In English as a lingua franca (ELF), both have been shown to be used differently than by 'native' speakers, especially marking subjectivity or speaker-centeredness. The data show similarities with ELF communication, that is the markers express subjectivity in highlighting or evaluating part of the discourse, and convey speaker-centeredness as 'fumbling devices'. Thus, the paper raises the question as to whether these and similar sociolinguistic interviews applied for the documentation of New Englishes should be analyzed as lingua franca communication.

1 | INTRODUCTION

Discourse markers, their forms and functions, have received increased attention in the past decades. This is reflected by a considerable heterogeneity of names – 'discourse particles', 'discourse connectives', 'pragmatic markers' to name but a few – as well as approaches taken to the study of discourse markers. Generally, they can be defined as 'sequentially dependent elements which bracket units of talk' (Schiffrin, 1987, p. 31). Regarding the functional level, discourse markers have also been shown to be versatile, indicating discourse relations and regulating interpersonal relationships or politeness, for instance.

In light of this growing body of research, investigating discourse markers from a variety-oriented perspective is called for (Aijmer, 2013). Importantly, this investigation should take sociolinguistic factors, such as types of speakers and text types, into consideration (Aijmer, 2002, 2013; Wilson et al., 2017; Mazzon, 2019). The present paper aims at contributing to this description by providing an analysis of discourse markers in the as yet little-described English(es)[1] spoken by Zanzibaris, critically considering the situational context of data collection. The data stem from sociolinguistic interviews conducted on Unguja Island of Zanzibar. The analytical focus is on the markers *you know* and *I think*. The

first has been described as a 'shared knowledge indicator' (Leech & Svartvik, 1975) or an expression of the speaker's uncertainty regarding the addressee's attitude/degree of linguistic precision achieved (Holmes & Stubbe, 1995). The latter has usually been analyzed as an epistemic stance marker (Aijmer, 1997), which might also function as a hedge or booster. In ELF communication, both markers have been shown to be used differently than by 'native' speakers, namely marking subjectivity rather than intersubjectivity; especially *you know* as a prefab drawing attention to the speaker (House, 2009, 2013). The Zanzibari data show similarities with these functions described for ELF discourse, that is expressing subjectivity in highlighting or evaluating part of the discourse, and conveying speaker-centeredness (House, 2009). Ultimately, the paper thus raises the (methodological) question as to whether the interviews at hand, as well as similar sociolinguistic interviews, should not be analyzed as instances of a certain second language variety but rather as lingua franca communication.

2 | DISCOURSE MARKERS AND WORLD ENGLISHES

As outlined by Maschler and Schiffrin (2015, p. 203), the way in which a researcher identifies discourse markers is a consequence of their general approach to language. They identify three broad approaches to the description and analysis of discourse markers, namely Schiffrin's (1987, 2006) sociolinguistic perspective, Fraser's (1990, 2009) pragmatic approach, and Maschler's (1994, 2009) interactional linguistics perspective. Given the fundamentally different conceptualizations of communication and meaning making underlying them, they are distinctly different with regard to the linguistic items they consider for analysis, as well as the functions they attribute to them within the discourse and larger communicative context. Within a sociolinguistic framework, discourse markers are compared to indexicals and are claimed to create meaning at the discourse level (Schiffrin, 1987, pp. 318, 322–325). In this vein, discourse markers are viewed as deriving meaning from the context and possessing meaning potential rather than fixed meaning (Norén & Linell, 2007). Hence, the communicative force of a discourse marker is on the one hand created by its fixed semantic core meaning, and the syntactic positioning of the marker at a specific slot in the discourse on the other (Schiffrin, 1987, pp. 317–318). Discourse markers have a lot of social meaning potential, as they 'indexically refer to a number of "speech act" features such as the speaker, […] social identities and the speech event itself' (Aijmer, 2013, p. 14). This sociolinguistic view is similar to Maschler's (1994) interactional perspective, in which discourse markers must have a metalinguistic interpretation in a particular context, fulfilling the process of 'metalanguaging'. Fraser's pragmatic approach downplays discourse markers' social-interactional functions (Maschler & Schiffrin, 2015, p. 193), which is why it is not considered in detail in this paper, as it is set within a sociolinguistic framework.

In this article, discourse markers are conceptualized as surface phenomena, which at the same time reflect a speaker's mental processes (Aijmer, 2013, p. 4). Thus, they are overt indicators of a speaker's metalinguistic awareness (Verschueren, 1999). On the other hand, discourse markers function as contextualization cues similar to prosody or co-speech gestures, organizing speech for the addressee (Aijmer, 2013). Structurally, discourse markers usually occur at transitions in the discourse, when the addressee needs to be made aware that a new activity starts or that the speaker takes on a new role (Aijmer, 2013, p. 7). Formally, discourse markers can be of different types, for example adverbs, parenthetical clauses, and so on (Mazzon, 2019). However, Maschler (2009) shows that from a participant's perspective, discourse markers form a category of their own similar to other, more established categories.

Finally, one rather important aspect of the speech event that is particularly relevant to the study of discourse markers is stance, of which there are two types. Affective stance refers to an attitude, feeling, or emotional intensity, while epistemic stance denotes knowledge or beliefs regarding some focus of concern, including degrees of certainty of that knowledge (Ochs, 1996, p. 410). In this vein, discourse markers allow speakers to take up (dis)alignment with what is being said.

2.1 │ Previous research

There is a sizeable amount of literature on discourse markers in different varieties of English. Some of these studies are cross-varietal, like Aijmer's (2013) analysis of a number of discourse markers in a variational pragmatics framework, or Mazzon's (2019) corpus investigation of *I'm afraid* across 20 varieties of English. Many studies are concerned with L1 varieties, like Aijmer (2013) focusing on British English, and considering American, Australian, New Zealand, and Singapore English, or a study by Amador-Moreno et al. (2015) of discourse markers in Irish English. Analyses of 'Outer Circle Englishes' are rarer and largely focus on specific varieties, notably on Singapore English, such as Tan (2010) on *right* and Leimgruber (2016) on *bah*, and Nigerian English, for instance Ogoanah (2011) on *as in*, Fuchs et al. (2013) on *even* and *still*, and Unuabonah and Oladipupo (2021) on bilingual markers. While a few studies on discourse markers in varieties as yet neglected have been conducted recently, for example Wilson et al. (2017) on question tags in Trinidadian English and Westphal (2020) on question tags in Philippine English, the description of discourse markers in Outer and specifically Expanding Circle Englishes is a desideratum that remains to be addressed. Several of the aforementioned studies focus on substrate influence regarding discourse markers, or 'bilingual markers'. Thus, Leimgruber (2016) adds to the literature on discourse particles derived from Southern Min, Malay, and Cantonese in colloquial Singapore English by describing the previously undocumented *bah*, possibly derived from Mandarin *ba*. The Mandarin and the English particle are syntactically identical, and the meanings of accommodation, advice, and tentativeness are found in both languages (Leimgruber, 2016, p. 90). In a similar vein, Unuabonah and Oladipupo (2021) describe six bilingual markers borrowed from the three major indigenous Nigerian languages – that is, Yoruba, Igbo, and Hausa – into Nigerian English. Overall, the discourse markers are shown to have (inter)personal rather than textual functions, such as agreement-seeking (Unuabonah & Oladipupo, 2021).

Other studies, like that by Ogoanah (2011), focus on the use of English discourse markers in a specific variety, new structural patterns of use, and functions of the investigated discourse markers. Thus, Ogoanah (2011) shows that *as in*, even though it is classified as obsolete in British English by the *Oxford English Dictionary*, is frequently used in Nigerian English and shares features with other discourse markers, such as *you know* or *I mean*. In Nigerian usage, *as in* encodes procedural meaning and signals in various ways the speaker's desire to get the hearer to recognize that a crucial aspect of meaning is communicated (Ogoanah, 2011, p. 200). Fuchs et al. (2013) describe that *even* is used significantly more frequently in Nigerian English than in British English. Functionally, it has a wider range of uses. Further, *still*, expresses promises and predictions in Nigerian English. These usages mirror the meanings of equivalents in several indigenous languages (Fuchs et al., 2013, p. 123). Finally, studies like Wilson et al. (2017) demonstrate both the use of discourse markers that are shared with other Englishes, such as *(all) right* and *you know* on the one hand, as well as particular Trinidadian markers such as *nah* on the other. In general, they find that invariant question tags are by far more frequent than variant tags in Trinidadian English. Mostly, they have speaker-centered functions, although some addressee-related functions are attested (Wilson et al., 2017, p. 739). This emphasizes the importance of speaker-centeredness in discourse marker usage in New Englishes; an issue that is also stressed in the description of discourse markers in ELF discourse.

2.2 │ *You know* and *I think*

In their function of expressing epistemic or interpersonal stance (Tan, 2010; Mazzon, 2019), '[discourse markers] are geared to serve relational and face-work needs and the implementation of politeness-impoliteness strategies' (Mazzon, 2019, p. 594). Mazzon (2019) mentions the example of *like*, either signaling a) hesitation, b) uncertainty thus fulfilling a subjective, speaker-related function, or c) softening a potentially face-threatening turn hence fulfilling an intersubjective, relational function by taking into account the hearer's face needs. *You know* and *I think* were chosen for analysis here because they have both been analyzed as subjectivity markers, signifying speaker-centeredness in

ELF (House, 2009, 2013; Baumgarten & House, 2010), but also fulfilling intersubjective functions such as indicating shared knowledge (Leech & Svartvik, 1975) when used by 'native' speakers. The analysis of these discourse markers might thus shed light on the participants' attitudes towards the interview situation, and their own language competence, that is linguistic precision of the message achieved (Holmes & Stubbe, 1995), which non-'native' speakers are concerned with frequently in their own discourse production (House, 1999). In this way, they can be helpful tools in determining variety status(es) of English(es) in Zanzibar. *You know* as a shared knowledge indicator can signal confidence in the existence of common information and fulfill the interpersonal function of engaging the listener (Stenström, 1994). Östman (1981) suggests its analysis on three levels: coherence (as in attention-getting), politeness (as in camaraderie), and utterance (as in construing common background knowledge). In this vein, Holmes and Stubbe (1995) also suggest that it might express a speaker's uncertainty regarding the addressee's attitude towards the message. Jucker and Smith (1998) analyze it as a presentation marker, making the addressee recognize the implications of an utterance. This versatility regarding functionality is summarized well by Sankoff et al. (1997), who demonstrate the differences in individual preferences concerning *you know*.

House (2009) emphasizes that, in her data, *you know* is used by speakers who have spent considerable time in an English-speaking country or have generally had ample opportunity to use the language. In a way, it is hence an indication of familiarity with English. She further claims that the intersubjective function of *you know* seems to be secondary in ELF talk (House, 2009, p. 191), instead helping speakers 'own the[ir] thinking and verbalization process' (House, 2013, p. 59). This is similar to discourse markers marking speaker stance like *I think*, which usually expresses prototypical meanings, like cogitation, relating to the speaker in ELF (Baumgarten & House, 2010). Structurally, House (2009, p. 179) found that *you know* occurs most frequently in sentence medial position. In the vicinity of conjunctions like *but* or *because*, it functions as a reinforcing or highlighting strategy in her data. She also finds that it is used when the speaker is momentarily incoherent, hesitates, fumbles for the right formulation or tries to repair a misstep; however, it is never used solely as a marker of fumbling (House, 2009, p. 187).

I think and *I don't know* are the most frequent *I + verb* collocations in American and British corpus data (Baumgarten & House, 2010). Aijmer (1997) distinguishes between a prototypical meaning of *I think*, signifying 'cogitation' and three epistemic meanings – belief, opinion and subjective evaluation – which derive from the context. *I think* can hence express a speaker's degree of belief in, opinion on, or evaluation of a proposition (Scheibmann, 2001). The prototypical meaning of *I think* seems most prominent in ELF discourse. In ELF communication, *I think* has been described to express a distinctly speaker-oriented, subjective perspective (Baumgarten & House, 2010; House, 2013). Indeed, it is one of the most common means to encode a subjective view, and ELF discourse seems 'especially fragile and susceptible to expressions of subjectivity and stance' (Baumgarten & House, 2010, pp. 1193, 1197). In particular, it can have a starting point function for a speaker's perspective, but also focus the completion of a sequence of utterances or a turn, or function as a coping mechanism to deal with the demand of ad hoc language production (Baumgarten & House, 2010, p. 1189). Functions of *I think* specific to ELF discourse according to Baumgarten and House (2010, p. 1192) are for instance expressing agreement, drawing conclusions, contradicting, sharing personal experience, and conceding. Thus, *I think* seems more versatile in ELF discourse than in L1 English. Structurally, *I think* is frequently used as a main clause in complement clause constructions in ELF. In Baumgarten and House's (2010, p. 1190) study, non-'native' English speakers showed a much higher ratio of *that*-omission than 'native' speakers in this environment. Another construction, a verbal routine in sentences like *it's mostly the older people*, **I think** , *who still go outside*, was more frequent in the 'native' speaker data.

The present paper investigates the use of the two discourse markers *you know* and *I think* in English(es) spoken by Zanzibaris. It aims at answering the question as to how the markers are used, that is, formally in terms of their position in a sentence and functionally. Specifically this latter analysis, combined with a comparison of the discourse markers' functions in ELF and certain Outer Circle Englishes, will also shed light on questions concerning the variety status of English(es) spoken by Zanzibaris. As mentioned in the introduction, the analysis ultimately enables a reflection on the variety type of English elicited using sociolinguistic interviews.

3 | DATA AND METHODOLOGY

Zanzibar is a semi-autonomous region of Tanzania. After the mainland's independence in 1963, the Zanzibar archipelago merged with the newly independent republic to form the United Republic of Tanzania in 1964. Thus, Tanzanian legislation, including that on language use, is largely applicable in Zanzibar too, which means that Kiswahili, the lingua franca of large parts of East Africa, is the national language, and English a de facto national working language (Eberhard et al., 2020). English is used in many official matters, for example in court, and in education, where it is taught as a subject in primary school and used as medium of instruction in secondary education and at university level (Mohr & Ochieng, 2017). However, language competence seems to be rather low and it has been argued that English is rather a foreign than a second language (Schmied, 1990).[2] Descriptions of English in Tanzania have so far neglected English(es) spoken by Zanzibaris though. First, small-scale analyses (Mohr, 2019b, 2021) emphasize the importance of English for Zanzibaris working in tourism. It is one of the largest economic sectors in Zanzibar besides agriculture, employing more than 50 per cent of the population (Keshodkar, 2013).

The data used for this study stem from sociolinguistic interviews conducted for a larger project investigating language learning and use in the tourist space of Zanzibar. While the data collected for the project are of different types, namely sociolinguistic interviews, recordings of guided tours, observational data, photographs of the linguistic landscape, and attitudinal data from so-called Q-sorts referring to a card sorting task (Mohr, 2019a, 2020), the analysis presented here considers the sociolinguistic interviews only. They were semi-structured, one-on-one interviews conducted by myself in a location of the participants' choosing in order to provide as comfortable and secure an environment as possible. Several of these interviews could not be audio-recorded because the participants felt more comfortable with me taking written notes. Yet other interviewees were not originally from Zanzibar, as the tourist sector attracts many skilled laborers from mainland Tanzania and other parts of the world. For this analysis, the audio-recorded data, approximately 4.5 hours, collected among indigenous Zanzibaris, are focused on. Table 1 provides an overview of the participants and their socio-demographic background.

All participants are male. This is representative of the Zanzibari tourist industry in that most jobs are fulfilled by men, as the majority of Zanzibaris are Muslim and many women do not want to or are prohibited from working in tourism where they could meet male strangers, may not be allowed to wear head scarves, and so on (Keshodkar, 2013). The participants are in their late 20s to late 30s; an average age could not be calculated as two of them did not share their exact age. The L1 of all participants is Kiswahili, only Mikidadi mentioned both Kiswahili and English. As shown in Table 1, all of them speak at least one foreign language, thus catering for the highly linguistically diverse

TABLE 1 Socio-demographics of the participants

Pseudonym	Education	Occupation	Languages apart from L1
Abdalla	University degree in tourism	Tour guide, university lecturer	English, Italian, Spanish, French
Hamadi	Primary school, part of secondary school	Hotel staff (waiter, assistant manager)	English, Arabic, French, Italian, Spanish
Hussein	University degree in environmental planning and management	Reservation and reception manager (hotel)	English
Maburuki	University degree in tourism	Tour guide, teacher in madrasa school	English, Arabic, German, Czech
Mikidadi	University degree in computer science	Hotel manager	Arabic, Hindi
Saburi	Degree in education	Hotel staff (receptionist)	English, Arabic, Spanish

tourist space (Mohr, 2019b). English is the most important of these languages for their professional lives as they unanimously claimed that English is the language they use most frequently with tourists. They acquired English in different ways, ranging from formal education in school over private language classes to acquisition 'at the grassroots' level, that is in interaction (Mohr, 2021). Given the often very complex language biographies of the participants, the age of English onset is difficult to determine exactly. However, all of them started learning the language when they were school children. All participants are relatively highly educated, holding a college or university degree. Hamadi, who only finished primary school, is the only exception. This is accounted for by the fact that the interviews were held in English, which remains a minority language of the elite in many postcolonial countries in general and in East Africa in particular (Bwenge, 2012). Thus, while I collected data from participants who were less competent in English, these data were not collected in interviews and are not considered for the analysis at hand. This hints at the fact that English in Zanzibar seems to belong to the Expanding Circle, despite its official status in the public sphere.

4 | RESULTS

In the following, the frequency of the chosen discourse markers is analyzed, also providing an overview of where in an utterance the markers occur. Subsequently, their functions within the discourse are outlined.

4.1 | Frequency of *you know* and *I think*

Generally, *you know* occurred more frequently than *I think*. Table 2 provides an overview of the discourse markers' frequencies including their position in a sentence. As illustrated, the occurrence of the discourse markers with respect to sentence position differs remarkably. While the distribution of *you know* is more or less balanced, *I think* occurs by far most frequently in sentence-initial position. It is noteworthy that much of the data analyzed in previous studies, for instance by House (2009), is conversational, dialogic data in which the beginnings and ends of turns are relatively easily made out. The data analyzed here stem from sociolinguistic interviews which, usually, are conducted with the intention of collecting as much spoken data from a participant as possible and hence being rather monologic. Thus, the beginnings and ends of turns could not be considered here and it was sometimes difficult to unequivocally determine whether a discourse marker occurred at the beginning, in the middle, or at the end of a sentence. This is shown in example (1) from the interview with Mikidadi (see the appendix for the transcription conventions).

(1) Mikidadi: even swahili have its depth <interviewer> mhm </interviewer> **you know** , the way my grandfather speak or people who are old enough to be my my parents

In this utterance, *you know* could be utterance-medial, not considering the backchannel *mhm* by the interviewer as a separate turn. If the backchannel were considered, *you know* might indeed be utterance-initial. In this case, it was however classified as utterance-final, given that Mikidadi made a brief pause after the discourse marker. Generally, pauses and sometimes other prosodic features were taken into account in order to determine the position of discourse markers in a sentence. Thus, no examples had to be discarded.

TABLE 2 Frequency of the chosen discourse markers in the data, including their position in a sentence

	Initial	Medial	Final	Total
you know	33.8% (89)	40.7% (107)	25.5% (67)	100% (263)
I think	70% (35)	24% (12)	6% (3)	100% (50)

Further, it should be mentioned that repetitions of the discourse markers were counted multiple times, that is as many times as the discourse marker was repeated. Equally, immediate repetitions were counted multiple times for the same position in the sentence (see example (2), and also (15)). In other repetitions, the discourse markers do not immediately follow each other but were still counted for the same sentence position as the sentence was begun twice, as in example (3).

(2) Abdalla: cultural limitations and **you know you know** plus with the religion concept and thoughts
(3) Hussein: **you know** most of the **y- you know** most of the zanzibari now they are they they see that uum hotels are the opportunities

Another issue worth mentioning here is that the usage of both discourse markers varied considerably between different participants. Thus, the interview with Mikidadi accounts for more than half of all instances of *you know* (N = 144) and *I think* (N = 31) in the data, while Maburuki, whose interview is only 10 minutes shorter than Mikidadi's, uses the markers considerably less, (eight tokens of *you know* and two of *I think*). Mikidadi even uttered sentences with several occurrences of *you know* shortly following each other (example (4)).

(4) Mikidadi: yeah uh coz uh people understand you're of a different culture and [...] **you know** you are you are y- y- (**you know**) people **you know** i mean expect you to be [...] [**you know**] slightly different in your mannerism and **you know** so it's it's it's okay

This difference might be related to House's (2009) finding that speakers who have spent time in an English-speaking country and are very familiar with the language, use discourse markers more frequently than those who have not. Mikidadi had lived in the UK for more than a decade before he returned to Zanzibar so it is not surprising that he uses both discourse markers analyzed here most. There are two caveats to this, however. Firstly, Shaw et al. (2009) show that in their ELF data the frequency of some discourse markers, most notably *I think*, decreased despite increasing competence in English, due to accommodation of a conceived American youth language. Secondly, the relation between frequency of discourse markers and familiarity with the English language is difficult to determine here, simply because familiarity with English is difficult to define. It definitely does not correlate with formal education in Zanzibar as Hamadi, the only participant who has not finished secondary school or university, uses both discourse markers more frequently than other participants with a degree in higher education, such as Maburuki. English is often acquired informally by Zanzibaris and is used in grassroots contexts (Mohr, 2021), which might also further the acquisition of discourse phenomena like discourse markers. In this data set, the use of discourse markers seems rather related to the personality of a participant: the more outgoing and likely they were to talk without interruption from the interviewer, the more likely they were to use either discourse marker. Thus, discourse marker usage seems, to some extent, related to self-confidence in the Zanzibari data.

In a similar vein, considerable differences can be observed between frequencies with regard to the participants' use of the discourse markers in different positions within a sentence: Hamadi uses *you know* most frequently in sentence-initial position (63.6% of his tokens), while Abdalla uses it most frequently in medial position (70.3% of his tokens). Altogether, the heterogeneity of discourse marker usage among the Zanzibaris is a noteworthy result.

4.2 | Functions of *you know* and *I think*

4.2.1 | You know

The functions of *you know* and *I think* are closely linked to their position in a sentence. While the distribution of *you know* is, as mentioned above, relatively balanced, the discourse marker occurs most often in medial position. This is in line

with House's (2009) findings, although the differences between positions of the discourse marker in her data were more pronounced. In sentence-medial position, *you know* has been argued to be a warning of re-phrasing or further explanation (House, 2009, p. 173). Examples (5) and (6) illustrate this usage in the data from Zanzibar.

(5) Abdalla: things to say in english is like pride **you know** is like […] oh this is this is really going to study
(6) Hamadi: so it's hard to explain **you know** so hard

In example (5), Abdalla further specifies why saying things in English is a sort of pride for him. In (6), *you know* introduces the repetition of the adjective *hard* and thus emphasizes the adjective. Further, there are several instances in the data in which *you know* is used as a fumbling device sentence-medially, as shown in (7) and (8).

(7) Abdalla: don't know how to describe uuh **you know** uh just only few personal details
(8) Mikidadi: to me it was all about **you know** if i felt if if i went if i if i did IT it was part of engineering

In (7), Abdalla is clearly searching for words, as is emphasized by the hesitation marker *uh* before and after *you know*. In example (8), there is no hesitation marker, but the frequent repetition of the word *if* by Mikidadi shows that he is lost for words. This function of *you know* and it expressing insecurity with regard to the utterance is generally emphasized by the combination of the discourse marker with hesitation markers (*uh, um*) and other discourse markers like *yeah, well, I think, I mean* (see also example (3)). This occurs with 50 tokens (19%) of all instances of *you know*. This personal function of *you know* is hence relatively frequent. Similar to House's (2009) finding that the discourse marker is never solely used as a fumbling device, it fulfills other functions like expressing the speaker's opinion in (8) as well.

The support of discourse markers with the ad hoc planning and processing of speech is less frequent when *you know* is used in sentence-initial position. One instance is shown in example (9), however.

(9) Mikidadi: and **y- you know** it's uh , but i understand

Here, even the discourse marker itself is stammered, which emphasizes its function as a fumbling device, and it is followed by a hesitation marker and a pause. Usually, *you know* has been found to secure the attention of the addressee when used sentence initially (House, 2009, p. 173). In the data from Zanzibar, this function is less frequent due to the data collection situation outlined before. Thus, the participants rarely had to secure the interviewer's attention as they already had it. Overall, there are 14 instances (15.7% of all sentence-initial occurrences) of *you know* in turn-initial position, that is even though the participant had already been given the right to speak. This is shown in example (10).

(10) Hamadi: **you know** in zanzibar we don't have much this

Here, Hamadi comments on the state of education in Zanzibar. *You know* seems to signify 'I want you to know', thus engaging the listener (Stenström, 1994) and fulfilling an interpersonal function. Further, this discourse marker can be analyzed as a presentation marker making the listener realize the implications of the utterance (Jucker & Smith, 1998). This is indeed its function in many sentence-initial, albeit not turn-initial positions, where *you know* introduces a new topic. This is stressed by the combination of *you know* with the conjunctions *and* (N = 23, 8.7%) and *but* (N = 6, 2.3%). Example (11) provides an instance of *you know* on its own, introducing English as a concept opposed to Kiswahili. Example (12) shows *you know* in combination with *and*. In this sentence, the second token of *you know* introduces the aspect of religion as separate from culture.

(11) Hamadi: swahili , most of the time because uuuh **you know** some people here , they like much to speak english but they don't understand (it)

(12) Abdalla: obstacles of uh women involvement in tourism is uh . **you know** cultural . things […] and . **you know** you know plus with the religion concept

In a way, all of these uses of *you know* are also explanations of whatever preceded in the utterance. They are hence similar in their function of establishing shared knowledge.

Sentence finally, *you know* rather indicates that the speaker already assumes shared knowledge on the part of the listener (Leech & Svartvik, 1975), an interpersonal function. *You know* might however also signal uncertainty on the part of the speaker in ELF (House, 2009), a rather personal, subjective function. Both are observed in the Zanzibari data.

(13) Abdalla: yeah but many of them they refuse this […] they better stay home […] because you are thinking more on your (xx) cultural orientations more than tangible benefits that they can take you further **you know**
(14) Mikidadi: uuuh uh uh the they're called the nollywood uh nigerian […] movies and swahili movies **you know**

In example (13), Abdalla appeals to the listener's knowledge that cultural orientation and tangible benefits are in opposition to each other (the discussion is on women refusing to work in tourism). In (14), Mikidadi seems rather uncertain of his statement, he does not know how to phrase his utterance, possibly because he is unsure whether the interviewer is familiar with the term *Nollywood*. Altogether, *you know* can be observed to fulfill both personal (such as rephrasing, supporting searching for words) and interpersonal functions (as in establishing shared knowledge, securing the interlocutor's attention) in the Zanzibari data. The personal functions of the discourse marker are slightly more frequent.

4.2.2 | I think

As shown in Table 2, *I think* is much less frequent than *you know* in the data. However, there are a number of interesting findings relating to the discourse marker's use. One issue that should be mentioned here is that one participant, Hamadi, consistently used the variant *I thinks*. Given that he was the only one to do so, this is probably an idiosyncrasy rather than a characteristically Zanzibari usage though. The functions of *I think* are, like those of *you know*, related to their position within a sentence. In sentence-initial position, where *I think* occurred most frequently, the discourse marker often marks the speaker's opinion or subjective evaluation (Aijmer, 1997; Scheibmann, 2001). This is illustrated in examples (15), (16), and (17).

(15) Abdalla: **i think i think** is something like they can- we find like they cannot pronounce well how we greeting
(16) Hussein: **i think** if you go to hotel 1 and hotel 2[3] you'll get uuuh good feedbacks
(17) Saburi: aah so here there's something maybe you haven't noticed it , […] someone like you […] when we call them, […] they can use any language because of your presence […] but when they are not around some- s- someone like them , yeah **i think** that they cannot

Example (15) also demonstrates the starting point function for a speaker's perspective that has been reported for ELF discourse (Baumgarten & House, 2010), as well as the ELF-specific function of sharing personal experience. Thus, *I think* seems to fulfill a rather subjective function in these cases, which is emphasized by the repetition of the discourse marker in (15). This might be an indication of the speaker's ongoing ad hoc planning process, which *I think* has also been suggested to mark in ELF (Baumgarten & House, 2010, p. 1189). Generally, 30 per cent (N = 15) of all instances of *I think* are combined with a hesitation marker or another discourse marker like *yeah*, *well*, which underlines its function as a fumbling device. This number is higher than the 19 per cent observed for *you know* and suggests that this function is more important for *I think*. Example (16) is similar but exhibits another function of the discourse marker that has been reported to be specific to ELF: displaying speaker knowledge (Baumgarten & House, 2010, p. 1192). In

the conversation, Hussein reports on his knowledge about two hotels and then advises the interviewer to go there to collect data. Example (16) is hence an indirect way of giving advice and fulfills an interpersonal function. Example (17) is similar in that Saburi displays his superior knowledge of language usage by starting the utterance with 'maybe you haven't noticed it'.

In sentence-medial position, the functions of *I think* are relatively similar to those in initial position. Thus, the discourse marker might express the speaker's opinion or evaluate what is said, and support the ad hoc planning process. This is illustrated in (18), where the 'searching for words'-function is emphasized by the hesitation marker *uh*. In (19), an example of *I think* expressing the prototypical function of the discourse marker, that is, cogitation (Aijmer, 1997) in a very literal sense, is provided. However, this was the only token of the discourse marker with this meaning in the data.

(18) Mikidadi: same as uh uh **i think** it's the same as arabic you know
(19) Mikidadi: uh in my thoughts **i i think** in english language uh

I think in sentence-final position is similar in function to the instances observed in other positions, that is expressing the speaker's opinion or an evaluation of what is expressed in the utterance (example (20)). Its usage hence suggests that it is indeed 'one of the most common means of encoding a subjective perspective in the discourse' (Baumgarten & House, 2010, p. 1194). It might, however, also be used to express epistemic stance, as shown in example (21) where Abdalla shares his knowledge on when English is introduced in school in Zanzibar. In this example, he seems unsure about the certainty of this information though.

(20) Mikidadi: so people don't even attempt to to speak it in case they come across […] and being a laughing stock **i think**
(21) Abdalla: they started from form one **i think**

With respect to focusing the completion of a sequence of utterances or a turn (Baumgarten & House, 2010), *I think* indeed fulfills this function in the data, but rather in sentence-initial position with the discourse marker introducing the concluding sentence of a turn. This is shown in example (22), which is the concluding statement of Mikidadi's account of where and how he acquired English.

(22) Mikidadi: **i think** i got exposure at an early age […] to the english language

Altogether, these results are in line with the claim that in ELF discourse, *I think* expresses a distinctly speaker-oriented, subjective perspective (Baumgarten & House, 2010; House, 2013), although it might be involved in intersubjective functions like giving (indirect) advice, too. This insinuates that the data analyzed here indeed exemplifies ELF discourse.

4.3 | Summary of results

As shown here, the discourse markers *you know* and *I think* are both used in English(es) spoken by Zanzibaris, albeit with different frequencies as compared to each other and with significant inter-participant variation. *You know* is much more frequent than *I think*, and equally seems to fulfill more various functions, which have all been reported in the literature. The functions fulfilled by both discourse markers are support with ad hoc utterance planning and processing ('fumbling device') (House, 2009, 2013) shown for example in (7), (11), and (15), as well as in the expression of uncertainty about the utterance. The latter seems to refer to the formulation of the utterance in the case of *you know* (Holmes & Stubbe, 1995) and to the information provided by the utterance with *I think*, that means epistemic stance

(Aijmer, 1997). Moreover, this function is more frequent with *I think*. Two functions that seem related are that of a presentation marker in the case of *you know* (Jucker & Smith, 1998) illustrated in example (10) and providing a starting point for the speaker's opinion in the case of *I think* (Baumgarten & House, 2010), shown in example (15). Functions only fulfilled by *you know* are a warning of re-phrasing or repetition demonstrated for example in (5) and securing the attention of the addressee (House, 2009) shown in (10), as well as assuming and especially establishing shared knowledge (Leech & Svartvik, 1975), as for instance in (11). To some extent, the latter is related to a function that is distinctive of *I think*, namely displaying speaker knowledge. However, *I think* fulfills a subjective rather than an intersubjective function in that the discourse marker expresses affective stance, that is the speaker's opinion or a subjective evaluation as for instance in (15). Altogether, both discourse markers analyzed here fulfill mainly personal, that is subjective, functions and thus emphasize the speaker-centeredness of English(es) spoken by Zanzibaris, which has been reported to be typical of New Englishes (Wilson et al., 2017) and ELF (House, 2009; Baumgarten & House, 2010) alike.

5 | CONCLUSION

The results obtained by the analysis of the data from Zanzibar are intriguing and show similarities with analyses of discourse markers in other varieties of English, besides illustrating the speaker-centeredness of the discourse. Particularly Zanzibari patterns or discourse markers could not be observed, the two tokens of *I thinks* in the interview with Hamadi hardly qualify for distinctively Zanzibari usage. Kiswahili discourse markers like *sasa* 'now', *basi* 'enough', *haya* 'ok', could not be observed either, despite the fact that all Zanzibari participants knew that I speak Kiswahili and we even discussed Kiswahili language practices in large parts of the interviews. However, this might be an effect of my position as a (white) European researcher. Due to the omnipresence of European tourists in the space under investigation, I suspected that my outsider status might be less important. This idea was disproven by the data presented here and has been attempted to be remedied by applying another method in a subsequent fieldwork phase (Mohr, 2020). Still, it is likely that interviews with Zanzibari or mainland Tanzanian interlocutors would yield different results and this is an interesting future line of research. In the present study, the significant inter-participant variation and heterogeneity of discourse marker usage observed suggest no fixed or even emerging norms with regard to one distinctively Zanzibari variety of English. The speaker-centeredness emphasized by the usage of discourse markers as expressions of subjectivity, functioning as fumbling devices, or expressions of the speaker's opinion, for instance, is not indicative of whether English(es) spoken by Zanzibaris might be part of the Outer or Expanding Circle, if one wanted to place them in the Kachruvian model. Whether making this decision is in fact desirable in view of the inapplicability of the model to more and more varieties of English nowadays (see Buschfeld, 2013) is controversial. Regardless, the considerable inter-participant variation with respect to the discourse markers investigated insinuates that there is no homogeneous variety of English spoken by Zanzibaris. This might also be due to the very dynamic nature of the tourist space in which English(es) are used in Zanzibar, namely in grassroots contexts, which are usually only minimally regulated (Mohr, 2021). Further, English(es) are often acquired in informal contexts as well, which might also have an influence on the observed heterogeneity with regard to discourse marker usage.

Two issues further suggest that the data analyzed here and much of the discourse to be encountered in English on the archipelago is in fact ELF communication. Firstly, my presence as a researcher, the 'observer's paradox' mentioned above, has certainly had an influence on the data. Thus, the interview situation clearly fulfills the criteria for ELF to emerge, that is, a group communicating in English although none of them are 'native' English speakers (Seidlhofer, 2001). While this is often ignored in analyses of New Englishes, I would like to argue that it needs to be taken into account. Specifically in a very dynamic community of practice like the one observed here, the interview situation with two non-'native' speakers as interlocutors should be considered more thoroughly. Certainly, even in ELF communication, speakers bring features of 'their' variety to the situation (Mauranen, 2012), which ultimately remains one of lingua franca discourse however. Secondly, the situations in which English is used in Zanzibar are also almost exclusively lingua franca situations, namely in tourism. While English is used as medium of instruction in secondary and tertiary

education, Abdalla mentioned explicitly that 'it is difficult to just go straight with the language' and that code-switching to Kiswahili is the norm. Thus, most communication in English in Zanzibar takes place with tourists from all over the world (Mohr, 2019b) and clearly qualifies as ELF discourse. An interesting future avenue of research in this regard would be a comparison of the interviews analyzed here with focus group interviews with Zanzibaris and tourists. In the dynamic communities of practice of ELF communication that operate on little common ground of expectable communicative behavior (Baumgarten & House, 2010), the expression of subjectivity (by discourse markers) is particularly important, given that it helps the speaker position themselves in the discourse and increases their sense of community (House, 2013, p. 59). Thus, the investigation of the discourse markers analyzed here is only a starting point in order to find out more about the speaker community/ies in Zanzibar and the ways in which they linguistically create cohesion.

NOTES

[1] While, according to a traditional 'sociolinguistics of distribution', languages and varieties are defined according to national territories, a more contemporary 'sociolinguistics of mobility' acknowledges the heterogeneity of language varieties spoken within one country (Blommaert, 2010). Acknowledging the importance of speaker styles as indexical of identities in our globalized world (Eckert, 2008), this paper uses the term 'English(es) spoken by Zanzibaris' instead of the singular 'English'.

[2] For a detailed overview of the dynamics of English within its multilingual context in Tanzania see Bwenge (2012), for instance.

[3] The type of madrasa Maburuki mentioned is a school for Muslim children, where they are acquainted with the Quran and the principles of Muslim faith.

[4] The names of the hotels have been anonymized.

REFERENCES

Aijmer, K. (1997). *I think*—An English modal particle. In T. Swan & O. Jansen Westvik (Eds.), *Modality in the Germanic languages* (pp. 1–47). Berlin: Mouton de Gruyter.

Aijmer, K. (2002). *English discourse particles: Evidence from a corpus*. Amsterdam: John Benjamins.

Aijmer, K. (2013). *Understanding pragmatic markers*. Edinburgh: Edinburgh University Press.

Amador-Moreno, C. P., McCafferty, K., & Vaughan, E. (2015). *Pragmatic markers in Irish English*. Amsterdam: John Benjamins.

Baumgarten, N., & House, J. (2010). *I think* and *I don't know* in English as Lingua Franca and native English discourse. *Journal of Pragmatics*, *42*, 1148–1200.

Blommaert, J. (2010). *The sociolinguistics of globalization*. Cambridge: Cambridge University Press.

Buschfeld, S. (2013). *English in Cyprus or Cyprus English: An empirical investigation of variety status*. Amsterdam: John Benjamins.

Bwenge, C. (2012). English in Tanzania: A linguistic cultural perspective. *International Journal of Language, Translation and Intercultural Communication*, *1*, 167–182.

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2020). *Ethnologue: Languages of the world* (23rd ed.). Dallas: SIL International. Retrieved from: http://www.ethnologue.com/

Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, *12*, 453–476.

Fraser, B. (1990). An approach to discourse markers. *Journal of Pragmatics*, *14*, 383–395.

Fraser, B. (2009). An account of discourse markers. *International Review of Pragmatics*, *1*, 293–320.

Fuchs, R., Gut, U., & Soneye, T. (2013). We just don't even know: The usage of the pragmatic focus particles *even* and *still* in Nigerian English. *English World-Wide*, *34*, 123–145.

Holmes, J., & Stubbe, M. (1995). *You know*, *eh* and other 'exasperating expressions': An analysis of social and stylistic variation in the use of pragmatic devices in a sample of New Zealand English. *Language and Communication*, *15*, 63–88.

House, J. (1999). Misunderstanding in intercultural communication: Interactions in English as a Lingua Franca and the myth of mutual intelligibility. In C. Gnutzmann (Ed.), *Teaching and learning English as a global language* (pp. 73–89). Tübingen: Stauffenburg.

House, J. (2009). Subjectivity in English as Lingua Franca discourse: The case of *you know*. *Intercultural Pragmatics*, *6*, 171–193.

House, J. (2013). Developing pragmatic competence in English as a Lingua Franca: Using discourse markers to express (inter)subjectivity and connectivity. *Journal of Pragmatics*, *59*, 57–67.

Jucker, A., & Smith, S. (1998). *And people just you know like 'wow'*: Discourse markers as negotiating strategies. In A. Jucker & Y. Ziv (Eds.), *Discourse markers: Description and theory* (pp. 171–201). Amsterdam: John Benjamins.

Keshodkar, A. (2013). *Tourism and social change in post-socialist Zanzibar*. Lanham, MD: Lexington Books.

Leech, G., & Svartvik, J. (1975). *A communicative grammar of English*. London: Longman.

Leimgruber, J. (2016). *Bah* in Singapore English. *World Englishes*, *35*, 78–97.

Maschler, Y. (1994). Metalanguaging and discourse markers in bilingual conversation. *Language in Society*, *23*, 325–366.

Maschler, Y. (2009). *Metalanguage in interaction: Hebrew discourse markers*. Amsterdam: John Benjamins.

Maschler, Y., & Schiffrin, D. (2015). Discourse markers: Language, meaning and context. In D. Tannen, H. E. Hamilton, & D. Schiffrin (Eds.), *The handbook of discourse analysis* (pp. 189–221). Hoboken: Wiley-Blackwell.

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

Mazzon, G. (2019). Variation in the expression of stance across varieties of English. *World Englishes*, *38*, 593–605.

Mohr, S. (2019a). Assembling concourse material and selecting Q-samples on the sociolinguistics of tourism discourse in Zanzibar. *Operant Subjectivity*, *41*, 1–18.

Mohr, S. (2019b) Motivations for language choices in tourist-host interactions in Zanzibar. Paper presented at the *24th Conference of the International Association of World Englishes*, Limerick, Ireland, 20 June 2019.

Mohr, S. (2020). The 'I' in sociolinguistics: The role of subjectivity in ethnographic fieldwork. *The Mouth: Critical Studies in Language, Culture and Society*, *6*, 101–118.

Mohr, S. (2021). English language learning trajectories among Zanzibaris working in tourism. In C. Meierkord & E. W. Schneider (Eds.), *World Englishes at the grassroots* (pp. 70–90). Edinburgh: Edinburgh University Press.

Mohr, S., & Ochieng, D. (2017). Language usage in everyday life and in education: Current attitudes towards English in Tanzania. *English Today*, *33*(4), 12–18.

Norén, K., & Linell, P. (2007). Meaning potentials and the interaction between lexis and contexts: An empirical substantiation. *Pragmatics*, *17*, 387–416.

Ochs, E. (1996). Linguistic resources for socializing humanity. In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 407–437). New York: Cambridge University Press.

Ogoanah, F. N. (2011). The pragmatic roles of *as in* in Nigerian English usage. *World Englishes*, *30*, 200–210.

Östman, J.-O. (1981). *You know: A discourse-functional approach*. Amsterdam: John Benjamins.

Sankoff, G., Thibault, P., Nagy, N., Blondeau, H., Fonollosa, M.-B., & Gagnon, L. (1997). Variation in the use of discourse markers in a language contact situation. *Language Variation and Change*, *9*, 191–218.

Scheibmann, J. (2001). Local patterns of subjectivity in person and verb type in American English conversation. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 61–89). Amsterdam: John Benjamins.

Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.

Schiffrin, D. (2006). Discourse marker research and theory: Revisiting *and*. In K. Fischer (Ed.), *Approaches to discourse particles* (pp. 315–338). North Holland: Elsevier.

Schmied, J. (1990). Language use, attitudes, performance and sociolinguistic background: A study of English in Kenya, Tanzania and Zambia. *English World-Wide*, *11*, 217–238.

Seidlhofer, B. (2001). Closing a conceptual gap: The case for a description of English as a Lingua Franca. *Annual Review of Applied Linguistics*, *24*, 209–239.

Shaw, P., Caudery, T., & Petersen, M. (2009). Students on exchange in Scandinavia: Motivation, interaction, ELF development. In A. Mauranen & E. Ranta (Eds.), *English as a Lingua Franca: Studies and findings* (pp. 178–199). Newcastle: Cambridge Scholars.

Stenström, A.-B. (1994). *An introduction to spoken interaction*. London: Longman.

Tan, A. (2010). *Right* in Singapore English. *World Englishes*, *29*, 234–256.

Unuabonah, F. O., & Oladipupo, R. O. (2021). Bilingual pragmatic markers in Nigerian English. *World Englishes*, *40*, 390–406.

Verschueren, J. (1999). *Understanding pragmatics*. London: Edward Arnold.

Westphal, M. (2020). Question tags in Philippine English. *Corpus Pragmatics*, *4*. https://doi.org/10.1007/s41701-020-00078-w

Wilson, G., Westphal, M., Hartmann, J., & Deuber, D. (2017). The use of question tags in different text types of Trinidadian English. *World Englishes*, *36*, 726–743.

APPENDIX: TRANSCRIPTION CONVENTIONS

(text)	-	uncertain transcription
@	-	laughter
,	-	brief pause (2-3 sec.)
.	-	longer pause (3-4 sec.)
[…]	-	speech left out in the example