

# So in video-mediated communication in the Expanding Circle

Caroline Collet<sup>1</sup> | Stefan Diemer<sup>2</sup> | Marie-Louise Brunner<sup>3</sup>

<sup>1</sup> Saarland University, Germany

<sup>2</sup> Trier University of Applied Sciences, Germany

<sup>3</sup> Trier University of Applied Sciences, Germany

## Correspondence

Caroline Collet, Saarland University, Department of English, Campus A5 3, 66123 Saarbrücken, Germany.

Email: [caroline.collet@uni-saarland.de](mailto:caroline.collet@uni-saarland.de)

## Abstract

The English discourse marker *so* can mark result, inference, connection, topic development, action prompting, and emergence of incipency. In English as a Lingua Franca (ELF) contexts, it also marks hesitation, self-correction, and floor-holding. Our paper uses quantitative and qualitative discourse analytical methodology to examine *so* in a corpus of video-mediated informal conversations between ELF speakers, where it is the most frequent discourse marker. The analysis demonstrates that all functions documented in first-language English use also occur in ELF. We also document extended uses of *so* to prompt inference and to illustrate pragmatic competence, particularly in combination with other discourse markers such as *okay*.

## 1 | INTRODUCTION

This paper examines the use of the discourse marker *so* in video-mediated conversations by speakers of English as a Lingua Franca (ELF). Research in world Englishes has identified discourse marker use as a feature which potentially defines varieties, either through a specific use of existing standardized variety discourse markers or by the integration of discourse markers from other languages (Kachru, 2019; Valentine, 1991). In ELF communication, where the interaction of speakers with different language backgrounds is the norm, English discourse markers can be extended or restricted in use, and discourse markers from other languages are also used (House, 2013; Kaur, 2011). ELF shares with other world Englishes a basis in a set of locally negotiated strategies, but differs in the degree of institutionalization of those strategies (Mauranen, 2012; Widdowson, 2015). Despite this fundamental difference, consistencies in the use of discourse markers in international settings including both ELF and other world Englishes have been commented upon, for example, by Lam (2010) in her functional framework for discourse particles based on an analysis of *well* and *so* in Hong Kong English. While the respective language backgrounds influence the use of discourse

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *World Englishes* published by John Wiley & Sons Ltd

markers by ELF speakers (indicated by the use of the term 'similects' for bundles of discourse strategies that are indicative of a particular language background; Mauranen, 2012), several studies (Buysse, 2012; House, 2013) have shown that the use of English discourse markers in ELF also seems to follow general patterns transcending individual language backgrounds. As the main aim in ELF communication is mutual comprehension and successful communication, the term 'standard' is usually avoided in the ELF paradigm in favor of endonormative, spontaneous norms (Mauranen, 2012). The distinction between 'native' and 'non-native' speakers is also controversial, but for a different reason than in other world Englishes. Most ELF researchers consider ELF speakers to be users of English as second (or additional) language (Mauranen, 2012), but not learners, thus contrasting them with first-language, 'native' English speakers (or a context largely determined by 'native' speakers). Others (Jenkins, 1996) have been arguing that 'native' and 'non-native' should be replaced by 'monolingual' and 'bilingual' (Jenkins, 1996, p. 10), irrespective of whether English is the first or subsequently acquired language (so that everybody who is not monolingual would be using ELF). In order to maintain a distinction between the use of English as a first language and to accommodate both world English and ELF paradigms, in this article we will be using the terms 'first-language English' and 'ELF' speakers, respectively.

Our research motivation is thus to investigate the use of frequent English-language discourse markers in an ELF background in order to contribute to our understanding of the use of English discourse markers in world Englishes. In particular, we want to find out the quantitative differences in the use of *so* as a discourse marker across different speaker groups in both ELF and first-language English contexts. We also want to document the functions of *so* and its correlates (collocates, nonverbal, and paralinguistic co-occurrences) in an ELF setting. The study uses examples from ViMELF (Video-mediated English as a Lingua Franca conversations), a corpus of 20 Skype conversations between ELF speakers from different European countries compiled at Saarland University and Trier University of Applied Sciences. As the most frequent discourse marker in ViMELF (occurring 1,417 times compared to 941 uses of *okay*), *so* was selected as a suitable focus of our analysis. *So* has attracted considerable attention in the context of face-to-face conversations, particularly in a first-language English context, and there are several studies on its use in ELF contexts (see next section). Its use in video-mediated conversations between ELF speakers has not yet been documented, which constitutes a further motivation for the present study.

## 2 | PRIOR RESEARCH AND RESEARCH QUESTIONS

The discourse marker *so* has been extensively researched in various contexts and datasets, both written and spoken. Going back to Schiffrin (1987), discourse markers have been studied specifically in spoken conversational contexts, though initially with very small datasets, focusing on lexico-syntactical as well as conversational function. Most corpus-based studies of *so* in spoken contexts examine it as part of a larger investigation of discourse markers in a first-language English context (Aijmer, 2013; Aijmer & Simon-Vandenberg, 2011; Redeker, 2006; Schiffrin, 1987). Studies of *so* in spoken first-language English contexts were performed by Blakemore (1988), Bolden (2006, 2009), Johnson (2002), and Raymond (2004). Lam (2010) performs a comprehensive study of *so* in the Hong Kong Corpus of Spoken English, including first-language, second-language, and foreign language use. There are also several studies that take an explicit comparative approach: Buysse (2012) analyzes *so* in a corpus of Dutch speakers of English, House (2013) in ELF institutional discourse and compares it to first-language English use, Sato (2019) compares Japanese English as a foreign language (EFL) learners and first-language speakers of English (though in a written corpus), and Liu (2017) compares the use of *so* by English speakers and Chinese speakers of English. *So* has also been examined as part of a general discussion of discourse markers in learner language (Hellermann & Vergun, 2007; Müller, 2005; Trillo, 2002). We can group the documented uses of *so* into seven main functions. These are not mutually exclusive and may overlap, so that the discourse marker can fulfill more than one function:

- (1) Result and inference: *So* has generally been shown to indicate inference and to signal results (Blakemore, 1988; Buyse, 2012; Fraser, 1999; Schiffrin, 1987). Schiffrin (1987) describes *so* as a 'marker of cause and result' (Schiffrin, 1987, p. 191), distinguishing resultative, inferential, and actional meaning (in reference to the respective levels of epistemicity and justification). Blakemore (1988) and Fraser (1999) additionally find that *so* can mark a connection between two propositions made by the speaker and that this connection needs to be inferred by the listener. Fischer (2006) explains this function as the display of a self-attentive mental process.
- (2) Action prompt: Raymond (2004) presents the use of stand-alone, turn-final *so* as a prompt for the recipient to produce the next action in connection with a distinct prosodic profile, another-attentive function that we find extensively in our own study.
- (3) Topic developer and sequencer: Johnson (2002) illustrates the use of *so* as another-directed topic developer in an institutional setting, using data from police interviews. This use is similar to its role as sequencer documented by Buyse (2012), who also mentions its role as a tool for elaboration.
- (4) Connection marker: *So* as a marker of connection is first described by Howe (1991). Redeker (2006, p. 339) calls it an 'attentional cue' that occurs at a discourse transition. House (2013) also finds this use in conversations by ELF speakers in academic consultation hours, but adds that in this context, it is mainly self-attentive, contextualizing the speaker's processes of 'perception, planning understanding, and affective stance' (House, 2013, p. 63). She suggests that this allows ELF speakers to 'effectively express (inter)subjectivity and connectivity and thus advance their pragmatic competence' (House, 2013, p. 63), a notion that we will further elaborate in our data. It is similar to Sato's findings (2019, p. 26) that Japanese EFL learners 'use "so" in written language strategically as a preface to stance-taking by carefully selecting and adjusting the information to be established as common ground with the reader'.
- (5) Marker of emergence from incipency: Bolden (2006, 2009) points out that *so* is not only used to transition between themes, but also, more specifically, to launch pending or incipient actions and advance the speakers' interactional agenda, accomplishing understanding. In this use, *so* is always other-attentive, that is, directed to the recipient (Bolden, 2009, p. 996).

In addition to these functions of *so*, there are two other functions that have received attention mainly in ELF speaker or learner data.

- (6) Hesitation marker and floor-holding device: Redeker (2006) finds numerous examples where *so* is used by ELF speakers of English to mark hesitation, a function related to its role as transition marker but more attentive to the speaker in that it allows the consideration and re-organization of talk. Buyse (2012) points out that the use of *so* to mark hesitation also has an interactional function in that it allows speakers to hold the floor.
- (7) Self-correction: Buyse (2012, p. 22) finds *so* as a marker for self-correction in Dutch speakers of English, acting 'as a textual device for self-repair'. Lam (2010) also finds this feature (as 'self-editing'; Lam, 2010, p. 665) used frequently by first-language, second-language, and foreign language speakers in the Hong Kong Corpus of Spoken English.

Although these functions have been mainly researched in ELF contexts, they also occur in first-language conversations (Jucker & Ziv, 1998). An explanation for the scarcity of research in first-language data may be their connection to fluency and speech production rather than the organization of discourse – both are features that are primarily of interest in a learner context (but see Götz, 2013, on fluency in both contexts). In addition, these uses may be more pronounced in ELF data.

The present study investigates the seven uses of *so* listed here and discusses possible additional uses in the ViMELF dataset, through a combined quantitative and qualitative approach. Our research questions are as follows:

1. Are there general quantitative differences in the use of *so* as a discourse marker across different speaker groups in both ELF and first-language English contexts?
2. What are correlates of *so* (including collocates, nonverbal, and paralinguistic co-occurrences) in a video-mediated ELF setting?
3. Which documented functions does *so* fulfill in this video-mediated ELF setting, and does its position influence its function?
4. Which, if any, new functions does *so* fulfill in this video-mediated ELF setting?

We use a combined approach because, as the article will illustrate, a quantitative analysis does not provide a complete picture of discourse marker use, particularly if it is limited to lexical data. Existing studies found that specific contextual settings can influence the use of *so*, such as Johnson's (2002) study of turn-initial *so* in police interviews, which illustrates that not only genre, but also discourse marker position in the conversation turn is important. More importantly, general spoken language features such as pauses, turn-taking, and overlap, paralinguistic features such as laughter, and nonverbal features, such as gestures and facial expressions, also play an important role in discourse marker use and meaning. However, a close investigation of nonverbal and paralinguistic features is hindered by the lack of specifically annotated multimodal corpora of spoken language. While, for example, intonation and pauses have been investigated as part of general discourse marker analyses (Wennerstrom & Siegel, 2003), there are no specific investigations of these and other non-lexical features for *so* except in qualitative examples, and no annotated multimodal corpora have been examined for the interplay of discourse marker use, paralinguistic, and nonverbal features. Since ViMELF provides this information (see section 3), we have included these aspects in our analysis.

### 3 | DATA AND METHODOLOGY

The dataset used for this study is ViMELF (2018), the Corpus of Video-Mediated English as a Lingua Franca Conversations. ViMELF is a small corpus of 20 dyadic video-mediated conversations in an informal, unsupervised setting between university students from Germany, Spain, Italy, Finland, and Bulgaria, using English as a Lingua Franca. The recordings were conducted via Skype and the complete dataset comprises 12.5 hours. All conversations are first encounters between randomly selected pairs of participants. The age of participants ranges from 20 to 35 years, with an average of 23. Participants were not informed about the purpose of the conversation. They were given topic prompts to start off the conversations, which were unsupervised and developed naturally. The corpus comprises 113,677 words in the plain text version and was published in 2018 by the CASE research group at Trier University of Applied Sciences, Germany, where the corpus is also hosted. It is freely available for research, including the anonymized audio and video recordings. ViMELF was transcribed and annotated manually. The transcription includes general spoken language features such as pauses, turn-taking and overlap, paralinguistic features such as laughter, and nonverbal features, in particular gestures and facial expressions. Speakers are identified by a unique ID; abbreviations used in the examples are: FL (Forlì, Italy), HE (Helsinki, Finland), SB (Saarbrücken, Germany), SF (Sofia, Bulgaria), and ST (Santiago de Compostela, Spain). A description of transcription conventions is available on the corpus homepage (Brunner et al., 2018). Additional corpora modeled after ViMELF are being compiled: TaCoCASE (forthcoming), consists of conversations between US and British speakers, as well as between US, British, and German English speakers, and a further extension of ViMELF containing Swedish, Belgian, and French ELF speakers. The present article focuses on *so*, with 1,417 occurrences the most frequent discourse marker in the ViMELF dataset. We additionally performed a quantitative analysis of the combined use of *so* and *okay*, as preliminary results from a separate investigation of *okay* (941 occurrences in ViMELF) indicate that combinations of the two discourse markers are frequent and may fulfill a different function, as observed by Koops and Lohmann (2016).

In order to find and investigate the uses of *so* as a discourse marker, the functions of *so* were identified by means of a corpus-based discourse analysis. The dataset was annotated manually using different tags: <D = so> for use as

a discourse marker, <X = so> for all other uses. The functions were then interpreted from the context of use and tagged according to the main functions described in section 2. Coding was double-checked for consistency, with inter-coder alignment above 99%, and unclear or additional functions were marked for later qualitative analysis, feeding the results back into the tagging process. The results were then analyzed in two steps, first quantitatively and then qualitatively. The quantitative analysis started with a frequency analysis of *so* and a comparison of discourse marker use frequencies. The second step consisted of a collocation analysis of discourse marker *so* to identify co-occurrence patterns. A major challenge to quantitative corpus analysis of discourse marker data is posed by the considerable individual differences between conversations and between individual speakers. This was previously also remarked upon by Gilquin and Granger (2011), who point out the difficulty 'to distinguish between the characteristics that are typical of first-language or ELF use in general, those that are limited to certain populations and those that are only found among certain speakers.' We follow their recommendation of a combined qualitative and quantitative approach, taking into consideration both the full dataset and the separate texts. Moreover, this problem illustrates that a combined approach also needs to include a close qualitative analysis, as discourse marker use is tied to larger conversational and discourse strategies (Aijmer & Simon-Vandenberg, 2011; Brunner & Diemer, 2016). Our methodology for qualitative analysis was specifically developed for analyzing discourse elements in ViMELF. We combine an analysis of the close context with an analysis of the wider discourse unit, increasing in scope: the discourse marker's function within the turn, its role in the topical unit, its position within the conversational structure, its use in the context of the general discourse strategy, and its appropriateness in the general conversational setting. This allowed us to more precisely interpret and label the functions of *so*.

## 4 | DISTRIBUTION AND FUNCTIONS OF SO

In this section the results from the quantitative and qualitative analysis of discourse marker *so* will be presented. The distribution of *so* in ViMELF's ELF context is compared with findings from other corpora composed of first-language and ELF data.

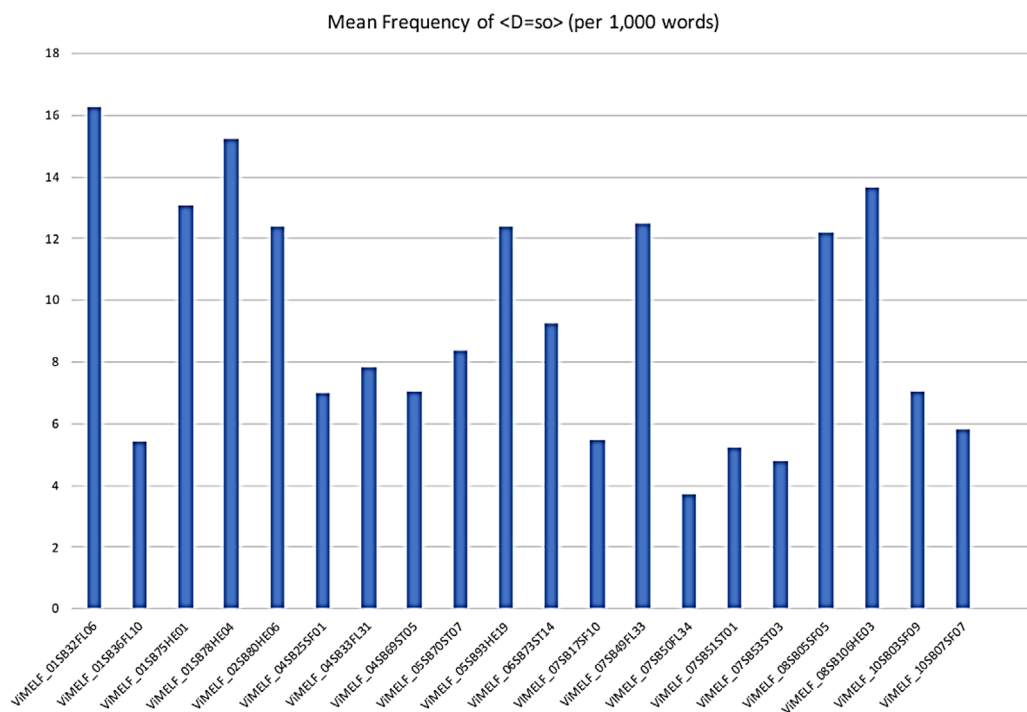
### 4.1 | Quantitative analysis and correlates of the discourse marker *so*

In order to find differences in the use of *so* as a discourse marker across different speaker groups, the distribution of *so* in ViMELF was compared to Buysse's (2012) findings on the use of discourse marker *so* in a corpus of Belgian speakers of English and his comparative analysis of first-language English speakers in the Louvain Corpus of Native English Conversation (LOCNEC, Gilquin et al., 2010), the first-language reference corpus of the Louvain International Database of Spoken English Interlanguage (LINDSEI, Gilquin et al., 2010).

Table 1 shows the total frequency of *so* in ViMELF divided into two categories: *so* with discourse marker functions and *so* with other uses. In ViMELF, *so* is used 1,417 times as a discourse marker (9.23 times per 1,000 words) and 271 times (2.01 times per 1,000 words) with other functions. Discourse marker use is considerably more frequent, with almost 84% of the cases. These results are in line with Buysse's results (2012) in his analysis of Dutch speakers of English, who, like the speakers in ViMELF, can be categorized as ELF speakers. Buysse (2012) identified a rate of 80% discourse marker function versus 18% non-discourse marker function, leaving 2% of the cases unclassified. Since

**TABLE 1** Frequency of *so* in ViMELF

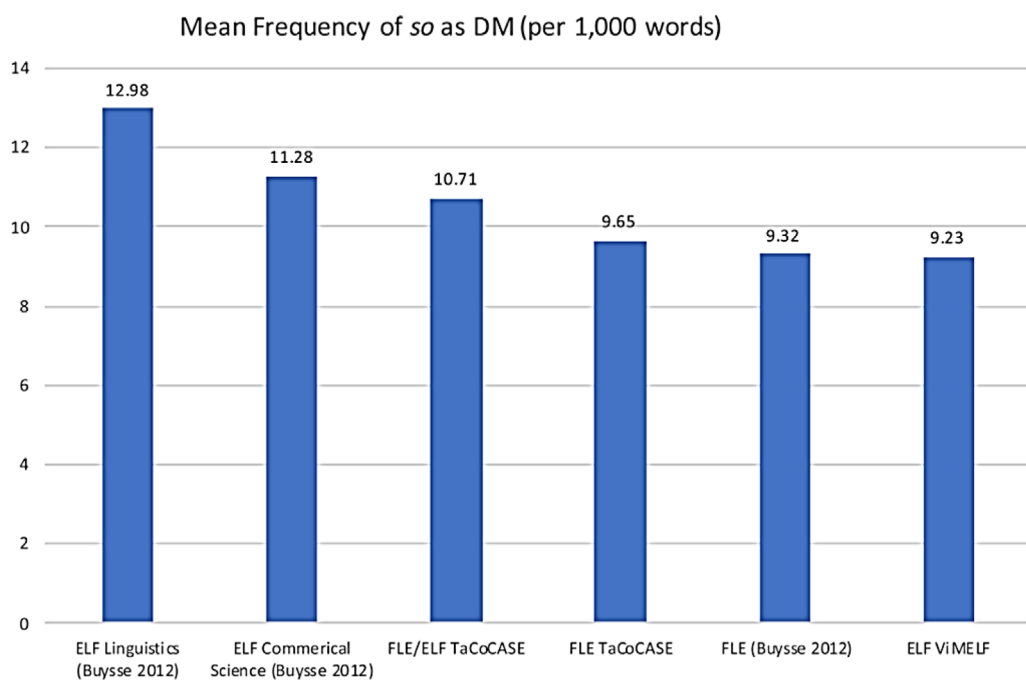
	So as discourse marker	So with other functions
Raw count	1,417	271
Frequency per 1,000 words	9.23	2.01
Percentage per category	83.95%	16.05%



**FIGURE 1** Relative frequency of *so* as discourse marker in ViMELF conversations [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

speakers, topics, and length of conversations vary within ViMELF, it is worth taking a closer look at the individual conversations and the frequency for *so* in each conversation. Figure 1 shows the frequency of *so* when used as a discourse marker for the individual conversations in ViMELF. The horizontal axis lists the individual ViMELF conversations, the vertical axis the number of uses of *so* as discourse marker normalized to 1,000 words. Relative occurrence varies considerably from conversation to conversation. While in conversation 01SB32FL06, for example, speakers use *so* 16.24 times per 1,000 words, speakers in conversation 07SB50FL34 only used it 3.71 times per 1,000 words. The mean value for *so* as discourse marker is 9.23 times per 1,000 words. All conversations contain *so* at least 3 times per 1,000 words. This considerable individual difference between discourse marker use was previously also remarked upon by Gilquin and Granger (2011) and prompted our mixed method approach in this study.

Figure 2 compares the mean frequency of *so* in ViMELF with those in the various corpora used by Buysse (2012), including the first-language reference data. Two main groups of speakers can be distinguished: ELF speakers (in ViMELF and in Buysse, 2012) and first-language English (FLE) speakers from TaCoCASE (forthcoming), which included first-language conversations and conversations between first-language and ELF speakers. ELF speakers in ViMELF use *so* less frequently (9.23 times per 1,000 words) compared to the speakers in Buysse's (2012) study (11.28 and 12.98 times per 1,000 words), though the difference is not significant. ELF speakers in ViMELF also use less *so* compared to first-language speakers of English. Indeed, ELF use of *so* in ViMELF is closer to first-language use than to the usage documented by Buysse (2012) for his Dutch speakers of English, though, again, the difference is not significant. In sum, there is no clear distinction in frequency of use between first-language and ELF use in either Buysse's data or between ViMELF and the first-language data in TaCoCASE. This result is not surprising, as previous studies also have yielded different findings in terms of frequency comparisons between first-language and ELF speakers. While Müller's (2005) study shows that German speakers of English in her data underuse *so* compared to first-language speakers of English, Buysse (2012) reaches the diametrically opposing result: his Belgian speakers of English use *so* considerably more



**FIGURE 2** Mean frequency of *so* as discourse marker by (sub)corpus [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

frequently than first-language speakers. Müller (2005, p. 251) sees the underuse due to negative L1 interference from German (with its close formal resemblance to German turn-initial *so*), while Buyse (2012, p. 29) explains the difference in use by the different text types: Müller uses an experimental setup, Buyse a more interactive interview setting, which might provide more opportunity to use the discourse marker. While these explanations can account for the different findings, the fact that both analyses reach such divergent results illustrates that purely quantitative analyses of discourse markers do not provide a complete picture of discourse marker use. The lack of a clear pattern separating first-language English and ELF use also suggests that accompanying discourse factors, such as conversational setting or context, individual language background, language competence, or idiosyncratic differences in use, play a larger role than the distinction between first-language and ELF use. This is in line with findings from ELF research and Widdowson's (2015) suggestion that pragmatic competence may not be tied to 'native' or 'non-native' speaker status.

Table 2 displays the 14 most frequent direct collocates of *so* when used as a discourse marker, using Laurence Anthony's AntConc (2019). The value of relation is given as mutual information (MI) + Log-Likelihood ( $p < 0.05$ ), as per Stubbs (1995). *So* is often found in combination with speaker IDs (SB, HE, FL) or personal pronouns like *I*, *you*, and *it*, hinting at the position of *so* in an utterance: Depending on whether the speaker's ID is in front or after *so*, it is either used at the beginning or the end of an utterance. While the results indicate that German (SB), Finnish (HE), and Italian (FL) speakers use *so* frequently at the beginning or end of an utterance, Bulgarian speakers, interestingly, use *so* in initial position much less frequently. Turn-initial ( $n = 625$ ) and turn-final uses ( $n = 228$ ) by German, Finnish, and Italian speakers account for more than 60% of all uses of *so* as a discourse marker (853 of 1,417), with a strong preference for turn-initial position. Another interesting finding is that German speakers have a stronger preference for turn-initial *so* (76% of all uses of *so* by German speakers) than Finnish (63%) or Italian (57%) speakers. Table 2 also shows that *so* frequently co-occurs with exhalations or inhalations (*h*) and the hesitation marker *uhm*, suggesting an interruption in conversational flow. *Okay* is the most frequent other discourse marker used in combination with *so*. While an

**TABLE 2** Collocates of *so* as discourse marker in ViMELF by frequency, left and right

Freq	Freq (L)	Freq (R)	MI	Collocate
498	395	103	297.452	SB
326	30	296	234.244	<i>I</i>
315	188	127	297.503	<i>yeah</i>
214	131	83	323.097	HE
209	48	161	249.429	<i>it</i>
198	61	137	257.518	<i>s</i>
191	20	171	245.358	<i>you</i>
191	130	61	255.549	<i>h</i>
177	147	30	321.378	{nods}
143	130	13	389.367	<i>okay</i>
141	99	42	318.178	FL
96	38	58	259.500	<i>Uhm</i>
93	67	26	152.600	<i>and</i>
90	36	54	206.709	<i>that</i>

**TABLE 3** Spoken feature, nonverbal and paralinguistic collocates of *so* as discourse marker in ViMELF

Freq	Freq (L)	Freq (R)	MI	Collocate
1319	637	682	4.46	Pause
434	203	231	4.20	Gesture
191	83	108	4.56	Ex- or Inhalation
131	71	60	3.95	Laughter

analysis of the complex combined occurrences of *so* and *okay* is beyond the scope of this paper, examples for the extended functionality are given in the qualitative analysis.

As the last part of our quantitative analysis, Table 3 correlates the use of discourse marker *so* with selected nonverbal and paralinguistic features. As indicated, pauses, gestures, ex- or inhalations, and laughter all significantly co-occur with *so*. The co-occurrence of pauses and *so* as discourse marker indicates that the discourse marker use disrupts normal conversational flow but at the same time allows the speaker to keep the turn and advance their interactional agenda. The co-occurrence of discourse marker *so* and gestures (such as nodding, physical stance shifts, or hand gestures) allows a more nuanced construction and negotiation of meaning. Ex- or inhalations in connection with discourse marker *so* are a sign for subtle stance shifts, while laughter allows the establishment of rapport or, again, more nuanced stance-taking in correlation with discourse marker *so*. While the quantitative analysis already illustrates relevant usage patterns of *so*, a qualitative analysis is needed in order to examine the use *so* in the conversational context.

## 4.2 | Functions of discourse marker *so* – qualitative analysis

Following our quantitative analysis, this section is based on a qualitative analysis of examples from the data to determine which functions *so* fulfills in interaction. We look at *so* from different angles, including both the transfer of typical



first-language English uses to the ELF context as well as additional uses as established in previous research (see section 2). We also take into account the different positions that *so* takes in the interaction, phrase-initial or -final, as well as the functions that *so* fulfills in these positions. Finally, we consider how *so* is used in interaction with *okay*. Some of our examples are supplemented with audio or video clips; a link to the clips is available as QR code and short link. The discourse markers under discussion in the data are marked in bold to facilitate reading.

#### 4.2.1 | Previously documented functions of discourse marker *so* in ViMELF

First, we discuss functions of *so* that have been well established in research on first-language English data and which have been found to be frequently used in ELF contexts as well, as described in section 2. Example (1) illustrates the use of *so* as a topic sequencer and connection marker. The German interlocutor SB07 is recounting the main storyline of a novel (*The Mayor of Casterbridge*) where the protagonist commits a crime and then regrets it. When the protagonist tries to make it right, though, he is somewhat inhibited as *he doesn't want others to recognise him, and to uhm, to find out about his crimes*. SB07 then uses the discourse marker *so* to sequence the topic moving to the next part and outlining the consequence *so he moves away*. *So* functions both to connect and move smoothly from one aspect to the next, building up the storyline.

##### (1) Topic sequencer & connection marker (10SB07SF07)

SB07: uhm he doesn't want others to recognise him,  
and to to uhm,  
to: find out about his crimes,  
<D = so> he moves away,  
.h and uhm his fate is also important for his ROLE,

Continuing the sequence in example (1), the German interlocutor SB07 talks about other books in example (2) and suggests that the Bulgarian interlocutor, SF07, *might also read [...] Hard Times by Dickens*. SF07 infers from SB07's statement that this is a recommendation, which is introduced by *so*: *so you recommend it*. This is an example of inference.

##### (2) Inference (10SB07SF07)

SB07: (1.1) .h you might also read uhm;  
... how is it called,  
it's called u:hm,  
Hard Times by Dickens?  
SF07: (1.1) <D = okay>,  
<D = so> you recommend it.

In our data, *so* also frequently functions as topic developer. In examples (3) to (5), *so* is used to introduce a new topic or change the direction of an already mentioned topic, moving the conversation forward. Example (3) introduces the first topic after the greeting (*so what's up?*), inviting general small talk to start the conversation. In example (4), the two interlocutors have just finished talking about *The Mayor of Casterbridge* and SB07's recommendation to read *Hard times* (see (1) and (2) above), when SF07 comes back to the first topic (that is, *The Mayor of Casterbridge*) by asking: *so u:hm, did you like\_uh, The Mayor of Casterbridge*. Example (5) puts the focus on an already mentioned aspect of a topic that was discussed previously in the conversation to make this the new main topic. The German interlocutor SB05 talked about personal family connections to New York before. The Bulgarian interlocutor refers back to this by asking *so you said you were in New York several times?* Discourse marker *so* again introduces this development.

## (3) Topic developer I (01SB36FL10)

FL10: ((laughs)) €€€ hi oh hi ()?

SB36: ((laughs)) &lt;D = so&gt; what's up?

## (4) Topic developer II (10SB07SF07)

SF07: mhm.

&lt;D = so&gt; u:hm,

&lt;did you like\_uh&gt;,

The Mayor of Casterbridge,

## (5) Topic developer III (08SB05SF05)

SF05: ... uh and &lt;D = so&gt; you said you were in New York several times?

SB05: yes I visited my cousin like °(four) times &lt;D = so&gt;°.

So in ViMELF data can also mark emergence from incipency, as shown in example (6).

## (6) Emergence from incipient action (04SB6)

SB69: [yeah],

((laughs))

... .h.. [well we could go on]?

ST05: [&lt;D = so&gt; do y- do you want to] say something?

[[((laughs))]]

SB69: [[[laughs]]]

ST05: [to end]?

SB69: ..h no I,

.. your- your last words,

.. they were,

they were just perfect, {lifts hands; palms forward} {makes swiping gesture}

to end the conver[sation].

In example (6), *so* signals emergence from incipency, as documented by Bolden (2009, p. 989), who explains this use as follows: 'By prefacing these updating questions with "so," the speakers use its function as a marker of emergence from incipency reflexively to indicate that the matter is something they had meant to raise, thus emphasizing their engagement with the addressee.' Similarly, ST05's *so*-prefaced utterance signals a shift towards the previously incipient concluding sequence of the conversation.

*So* in phrase-initial position has been well documented in research to function as an outcome signal. In example (7), *so* is used to introduce a summary of the previous description by the Bulgarian speaker who talks about spending time at home with family and being spoiled before returning to university for the exam period. SF10 sums this up by concluding *so it's like a: complete vacation, ((laughs))*.

## (7) Summing up (07SB17SF10)

SF10: I don't have to travel <X = so> >much< and all this stress I experience,  
in the big city is gone,

&lt;D = so&gt; it's like a: complete vacation, ((laughs))



**FIGURE 3** Audio – summing up (alternative access: <https://seafire.rlp.net/f/1967c643d71f4e1b8566/>)

In example (8), later in the same conversation, SB17 asks what SF10 will be doing after their talk, to which SF10 responds that she will be talking to her parents since she does not see them very often. SB17 infers that SF10 will stop revising for her exams (*so you stop- you stop (revising) for today*), which she confirms. SF10 then continues that she studied the whole day, before providing a result, introduced by *so*: *so for today, .. I'm done*.

(8) Inferred and explicit result (07SB17SF10)

SB17: [oh],

SF10: because I'm not here very often\_and ... they want me to,  
{noise in the background}  
to be with them. ((hehe))

SB17: <D = so> you stop- you stop (revising) for today.

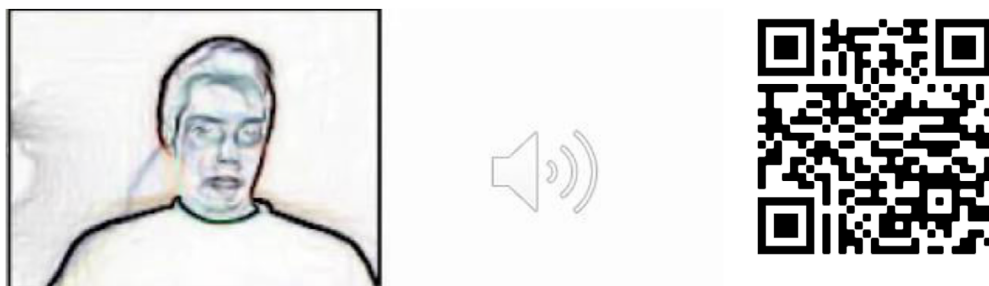
SF10: yeah.

I- I learned\_uh the whole day,  
.. <D = so> for today,  
.. I'm done. ((laughs))



**FIGURE 4** Audio – inferred and explicit result (alternative access: <https://seafire.rlp.net/f/65577e8777674932a279/>)

In initial position, *so* is frequently used to initiate the final part of an utterance or explanation in terms of introducing the results or providing a summary of what has been said. It is also often used to signal inference. Conversations in ViMELF confirm results from first-language contexts, with speakers incorporating *so* seamlessly into their interactions to: (1) indicate result and pragmatic inference derived from the context; (2) prompt the next action; (3) to develop, structure, and sequence topics; (4) to mark connections; and (5) to move on from incipency. ViMELF also contains numerous instances of *so* used in connection with hesitation and floor-holding and self-correction, that were previously documented in ELF data. In example (9), HE03 is taking nonverbal action (he looks at his computer, see Figure 5), which is supported by the following utterance: *okay let me see, {looks at his computer} ... so::: This so*, which is marked by paralinguistic lengthening, serves to hold the floor and to gain some time while trying to find the respective function on the computer.



**FIGURE 5** Video – floor-holding (alternative access: <https://seafire.rlp.net/f/235b73c78b3247b49faf/>) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

(9) Floor-holding device (08SB106HE03)

SB106: yeah .. uh you can, #00:04:30-5#  
... you can do it at this recorder?  
in the right corner?

HE03: [<D = okay>].

SB106: [there] is this audio? and video.

HE03: <D = okay> let me see, {looks at his computer}  
... <D = so>:::

SB106: if you have the same recorder,  
as I have I don't know exactly.

HE03: yeah.

So is also used for self-correction, as illustrated in example (10).

(10) Self-correction (06SB73ST14)

SB73: ((chuckles)) oh wow. ((hehe)) {lifts head & looks up}  
(1.1) <D = so> I thought you were- you would be older {lowers hand; palm up}?  
((laughs))

ST14: no I'm only [eighteen] ((heh)) ey. {looks up}

In the example, SB73 corrects her previous assumption about her conversation partner's age, prefacing her statement by discourse marker *so*.

#### 4.2.2 | New functions of discourse marker *so* in ViMELF: prompting inference and displaying pragmatic competence

The next examples illustrate uses that have not been specifically documented in either 'native' or 'non-native' speaker conversations. These are (8) prompting inference, and (9) displaying pragmatic competence.

(8) Prompting inference: *So* frequently occurs in final position in ViMELF. Turn-final *so* functions as a type of open end, signaling the end of the topic or turn, inviting the interlocutor to respond or change the topic. This use is less explicit than as 'action prompt.' Rather, there is a type of unspoken inference that can be drawn from what was said and that the interlocutor needs to read between the lines. This inference is different from that described in section 1, in which the listener has to infer the connection between two of the speaker's propositions – in this case there is no

second proposition. Examples (11) to (13) illustrate these instances of turn-final *so* prompting inference. While the first *so* in example (11) is in initial position and concludes the topic, the second *so* by ST03 appears as an afterthought to underline the argument. This turn-final *so* is used to prompt the addressee to comment on the issue or confirm the point that is made about bread not being very expensive. It also seems to imply that the interlocutor should infer the rest from the first proposition.

(11) Prompting inference I (07SB53ST03)

ST03: and .. a proper LOAF of bread would be one euro ten,  
one euro twenty,

(1.6) <D = so> it's not very expensive [definitely],

SB53: [yeah sounds], =

ST03: = everybody goes to the bakery every day <D = so>.

SB53: (2.8) yeah that would be about the same price,

The first *so* in example (12) is also a topic developer. SB75 talks about her brother's education, explaining the different components. In this context, she further develops the topic with: *so he studied in Australia as well*. The second, turn-final *so* occurs as a separate intonation unit after a short pause. It concludes the topic for SB75, providing a closing signal while also signaling a transition point (*so* as transition marker) to the interlocutor, inviting HE01 to comment or change the topic. It also expresses a degree of hesitation, as SB75 does not seem to have anything to add, making it more open-ended and open to interpretation than the *so* in example (9).

(12) Prompting inference II (01SB75HE01)

SB75: [<D = so>] he studied in Australia as well,

.. half a year,

..h and u:h .. but you don't have both things on the same time. #00:04:30-9#

... <D = so>.

HE01: ah <D = okay> because u:h,

here in Finland people usually;

usually work at the same time as they're studying,

In example (13), the first *so* is other-attentive and initiates a new topic: *so, you've never been to Italy?* The second, turn-final *so* concludes the turn. The Italian speaker, FL06, has just asked whether the German interlocutor has ever been to Italy. SB32 responds with *unfortunately not* and elaborates that there was an opportunity to go to Venice but that that did not work out due to a cancellation. SB32 finishes the statement (*they cancelled the trip*) with *so* which functions as a signal to the interlocutor to infer the obvious outcome, namely that they did not go to Italy, which reinforces the original statement that SB32 has never been to Italy. It also signals a transition point to the interlocutor, who responds partially in overlap with *oh, that's bad*.

(13) Prompting inference III (01SB32FL06)

FL06: <D = so>,

you've never been to Italy?

SB32: unfortunately not.

[I would really love to that],

FL06: [oh is that bad],

SB32: in uhm three years ago I was in Croatia,

.. uhm and we had the chance to go by ferry to Venice,

.. but yeah they cancelled the trip [<D = so>].

FL06: [oh],  
that's bad, ((giggles))

Turn-final *so* is frequently employed. It provides a prompt for the interlocutor to comment or change the topic, functioning as a topic close and a transition marker. It is also often used to signal to the interlocutor that they should draw their own conclusions from what is said and infer what remains unsaid. In other words, *so* is working as a type of signal for an unspoken inference that remains to be made.

*So* in ViMELF is also used by speakers repeatedly and in connection with other discourse markers in order to display pragmatic competence. This is related to the self-attentive display of the speaker's thought processes mentioned by House (2013); but rather than just advancing their local pragmatic agenda, speakers can also display their overarching pragmatic competence by using multiple discourse markers or discourse marker combinations in an authentic fashion. The combination of *so* and *okay* is particularly frequent, and their uses in these combinations supplement and build on each other. The functions may also partially overlap. Example (14) illustrates this complex use.

(14) Interaction of *so* and *okay* (08SB106HE03)

HE03: ((laughs)) <D = okay> <D = so> that works.  
((laughs)) apparently.  
<D = okay>.  
<D = so> uhm very good.  
<D = so> we'll just I'm just gonna save it then,  
in the way they wanted to- they wanted to have it.  
<D = so> uhm,  
<D = okay> <D = so> there goes my,  
recording again.  
<D = okay> uhm,  
<D = so>:,  
<D = so> w- we were just to talk about general stuff I guess?

SB106: uhm,;  
yeah but we also have to talk about one special topic?

HE03: <D = okay>?  
yeah.

SB106: uhm about our favourite TV series,



**FIGURE 6** Video – pragmatic competence (alternative access: <https://seafire.rlp.net/f/df46e5bb41314a17bbc0/>) [Colour figure can be viewed at wileyonlinelibrary.com]

The first *okay* and *so* combination in the conversation between the German speaker SB 106 and the Finnish speaker HE03 serves to wrap up previous utterances. While the *okay* is a backward-oriented deictic expression, confirming the previous statement, the *so* is forward moving and concludes the interlude on the recording modalities. HE03 then is slightly stuck as he does not seem to know exactly how to continue, using an *okay so* combination and then an additional *so* as hesitation and floor-keeping devices. Simultaneously, they also fulfill additional functions. The second *okay* in the interaction serves to link and transition from the previous to the next utterance. We can see House's (2013) self-attentive display of the speaker's thought processes in the two instances of *so* that follow. They act as a focus on the speaker's own utterance and a type of 'self-prompt' (House, 2013, p. 63). Both the floor-keeping and the speaker-supportive functions are typical of ELF interaction and here are highly successful strategies for negotiating the interactional development. The following *so uhm, okay so there goes my, recording again. okay uhm, so; so w-we were just to talk about general stuff I guess?* continues with a total of six discourse markers (two *okay* and four *so*) which also serve speaker-supportive as well as floor-keeping functions. In the last part, *okay uhm, so; so w-we were just to talk about general stuff*, the *okay* also serves to signal a pre-close. This is resolved with the statement about just having to talk about general topics which is initiated by *so* as a topic introduction. The last *okay* in the interaction is a confirmation check. This may have been caused by a processing delay, that is, the speaker needed a bit of time to understand, or because he was not sure or did not remember at first. HE03 then answers his own check by confirming this with *yeah*. This complex, interweaving pattern of discourse markers helps speakers to negotiate the interaction and allows them to display pragmatic competence on a larger scale by clustering and nuancing discourse markers in their utterances. At the same time, the two discourse markers are also used in a complementary way as they serve to provide a more complete connection between units of discourse which is both backward and forward looking. This combined use with extended meaning also illustrates Koops and Lohmann's (2016) observation that discourse marker combinations may fulfill different functions than either of their components or the sum of their components.

As we have shown, *so* as a discourse marker fulfills various functions in ViMELF. In addition to the seven main uses previously documented in first-language and ELF data, *so* can function as an invitation to make an inference from an unspoken proposition, the interlocutor having to draw their own conclusions and deciding whether to provide a comment or initiate a topic change. Speakers can also display their general pragmatic competence to successfully negotiate the interactional development by using discourse marker *so* or combinations of *so* and *okay*.

## 5 | CONCLUSION

This paper illustrated the use of *so* by ELF speakers in ViMELF, a corpus of informal video-mediated conversations. The four research questions in this study were:

1. Are there general quantitative differences in the use of *so* as a discourse marker across different speaker groups in both ELF and first-language English contexts?
2. What are correlates of *so* (including collocates, nonverbal, and paralinguistic co-occurrences) in a video-mediated ELF setting?
3. Which documented functions does *so* fulfill in this video-mediated ELF setting, and does its position influence its function?
4. Which, if any, new functions does *so* fulfill in this video-mediated ELF setting?

In response to research question 1, our quantitative analysis shows that *so* is the most frequent discourse marker used, but that its use varies considerably between conversations. A comparison with both first-language English and other ELF data shows that the frequency of the discourse marker *so* is not significantly higher or lower in one type of data. The difference in use may rather depend on conversational setting or context, individual language background, language competence, or idiosyncratic differences. Discourse marker use, as one indicator of pragmatic competence,

may thus not be tied to 'native' or 'non-native' speaker status. Concerning research question 2, *so* was found to frequently interact with other discourse markers such as *okay* as part of a sequence of two (or more) discourse markers. This interplay is similar to what Koops and Lohmann (2016) observed in their data. The quantitative correlate analysis also shows significant correlation between *so* and features such as intonation and pauses and nonverbal elements such as gestures and paralinguistic, particularly laughter. This complex use of *so* as an element of an embodied performance (Goodwin & Goodwin, 2000), with paralinguistic aspects (such as intonation, pauses, pitch, vowel length, and laughter) and nonverbal aspects (in particular gestures and facial expressions), demonstrates that its pragmatic meaning is created only in the situational context. A qualitative analysis of *so* in ViMELF is thus necessary in order to illustrate its full functionality and meaning. The qualitative analysis of functions as per research question 3 shows that all seven previously documented main functions, (1) result and inference, (2) action prompt, (3) topic developer and sequencer, (4) connection marker, (5) marker of emergence from incipency, (6) hesitation marker and floor-holding device, and (7) self-correction can also be found in the video-mediated ELF data. Research question 4 can be answered affirmatively. Two additional uses of *so* were found in ViMELF: prompting inference and displaying pragmatic competence. In total, we have thus documented nine different uses of *so* as discourse marker in our data.

In sum, the use of *so* in spoken ELF Skype conversations illustrates uses established in first-language English research while at the same time also showing the considerable complexity and flexibility inherent in lingua franca communication, further broadening the complex range of functions it fulfills in interaction. Extended flexibility in the use of discourse markers is thus a common feature of ELF and other world Englishes. The very high flexibility in discourse marker use may, however, also be a differentiator between ELF and other world Englishes, though this would need to be investigated further.

## ACKNOWLEDGMENTS

Open access funding enabled and organized by Projekt DEAL.

## REFERENCES

- Aijmer, K. (2013). *Understanding pragmatic markers*. Edinburgh: Edinburgh University Press.
- Aijmer, K., & Simon-Vandenberg, A. M. (2011). Pragmatic markers. *Discursive Pragmatics*, 8, 223–247.
- Anthony, L. (2019). *AntConc*. Retrieved from <https://www.laurenceanthony.net/software/antconc/>
- Blakemore, D. (1988). 'So' as a constraint on relevance. In R. Kempson (Ed.), *Mental representations: The interface between language and reality* (pp. 183–195). New York: Cambridge University Press.
- Bolden, G. (2006). Little words that matter: Discourse markers 'so' and 'oh' and the doing of other-attentiveness in social interaction. *Journal of Communication*, 56, 661–688.
- Bolden, G. (2009). Implementing incipient actions: The discourse marker 'so' in English conversation. *Journal of Pragmatics*, 41, 974–998.
- Brunner, M.-L., & Diemer, S. (2016). 'Mhm, ... okay so u:h, maybe we should start with this topic' –Conversation starts in English as a Lingua Franca (EFL) Skype talks. IWODA'16 - Fourth international workshop on discourse analysis, Santiago de Compostela.
- Brunner, M.-L., Collet, C., Diemer, S., & Schmidt, S. (2018). *The CASE project*. Birkenfeld: Trier University of Applied Sciences. Retrieved from <http://umwelt-campus.de/case>.
- Buysse, L. (2012). *So* as multifunctional discourse marker in native and learner speech. *Journal of Pragmatics*, 44, 1764–1782.
- Fischer, K. (2006). Frames, constructions, and invariant meaning. The functional polysemy of discourse particles. In K. Fischer (Ed.), *Approaches to discourse particles* (pp. 427–447). Amsterdam: Elsevier.
- Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics*, 31, 931–952.
- Gilquin, G., & Granger, S. (2011). The use of discourse markers in corpora of native and learner speech: From aggregate to individual data. Paper presented at the *Corpus Linguistics Conference 2011*, Birmingham.
- Gilquin, G., De Cock, S., & Granger, S. (2010). *The Louvain International database of spoken English interlanguage*. Louvain-la-Neuve: Presses Universitaires de Louvain.
- Goodwin, M. H., & Goodwin, C. (2000). Emotion within situated activity. In N. Budwig, I. Č. Užgiris, & J. V. Wertsch (Eds.), *Communication: An arena of development* (pp. 33–53). Westport, CT: Greenwood.
- Götz, S. (2013). *Fluency in native and nonnative English speech*. Amsterdam: John Benjamins.
- Hellermann, J., & Vergun, A. (2007). Language which is not taught: The discourse marker use of beginning adult learners of English. *Journal of Pragmatics*, 39, 157–179.



- 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.
- Howe, L. (1991). *Topic change in conversation* (Unpublished PhD thesis). University of Kansas, Kansas.
- Jenkins, J. (1996). Native speaker, non-native speaker and English as a Foreign Language: Time for a change. *IATEFL Newsletter*, 131, 10–11.
- Johnson, A. (2002). 'So...?': Pragmatic implications of so-prefaced questions in formal police interviews. In J. Cotterill (Ed.), *Language in the legal process* (pp. 91–110). New York: Palgrave Macmillan.
- Jucker, A. H., & Ziv, Y. (Eds.). (1998). *Discourse markers: Descriptions and theory*. Amsterdam: John Benjamins.
- Kachru, Y. (2019). Speaking and writing in world Englishes. In C. L. Nelson, Z. G. Proshina, & D. R. Davis (Eds.), *The handbook of world Englishes* (2nd ed., pp. 349–367). Malden, MA: Blackwell.
- Kaur, J. (2011). Raising explicitness through self-repair in English as a lingua franca. *Journal of Pragmatics*, 43, 2704–2715.
- Koops, C., & Lohmann, A. (2016). Discourse marker sequencing and grammaticalization. In G. Kaltenböck, E. Keizer, & A. Lohmann (Eds.), *Outside the clause: Form and function of extra-clausal constituents* (pp. 417–446). Amsterdam: John Benjamins.
- Lam, P. W. (2010). Toward a functional framework for discourse particles: A comparison of *well* and *so*. *Text & Talk*, 30, 657–677.
- Liu, B. (2017). The use of discourse markers *but* and *so* by native English speakers and Chinese speakers of English. *Pragmatics*, 27, 479–506.
- Mauranen, A. (2012). *Exploring ELF: Academic English shaped by non-native speakers*. Cambridge: Cambridge University Press.
- Müller, S. (2005). *Discourse markers in native and non-native English discourse*. Amsterdam: John Benjamins.
- Raymond, G. (2004). Prompting action: The stand-alone 'so' in ordinary conversation. *Research on Language and Social Interaction*, 37, 185–281.
- Redeker, G. (2006). Discourse markers as attentional cues at discourse transitions. In K. Fischer (Ed.), *Approaches to discourse particles* (pp. 339–358). Amsterdam: Elsevier.
- Sato, S. (2019). A corpus-based analysis of *so* in written discourse: A comparison between L1 English speakers and Japanese EFL learners. *Applied Pragmatics*, 1, 26–45.
- Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.
- Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of Language*, 2, 23–55.
- TaCoCASE. (forthcoming). *Transatlantic component of the CASE project*. Saarbrücken & Trier: Saarland University & Trier University of Applied Sciences.
- Trillo, J. R. (2002). The pragmatic fossilization of discourse markers in non-native speakers of English. *Journal of Pragmatics*, 34, 769–784.
- Valentine, T. M. (1991). Getting the message across: Discourse markers in Indian English. *World Englishes*, 10, 325–334.
- ViMELF. (2018). *Corpus of video-mediated English as lingua franca conversations*. Birkenfeld: Trier University of Applied Sciences. Retrieved from <http://umwelt-campus.de/case>
- Wennerstrom, A., & Siegel, A. F. (2003). Keeping the floor in multiparty conversations: Intonation, syntax, and pause. *Discourse Processes*, 36, 77–107.
- Widdowson, H. G. (2015). ELF and the pragmatics of language variation. *Journal of English as a Lingua Franca*, 4, 359–372.

**How to cite this article:** Collet C, Diemer S, Brunner M-L. *So* in video-mediated communication in the Expanding Circle. *World Englishes*. 2021;40:594–610. <https://doi.org/10.1111/weng.12543>