www.paedagogische-horizonte.at
Pädagogische Horizonte, 8(2), 2024
https://doi.org/10.17883/pa-ho-2024-02

PÄDAGOGISCHE HORIZONTE

EFL interaction in face-to-face versus synchronous computer-mediated communication settings

A post-secondary mixed-methods study

Gudrun Keplinger[a], Harald Spann[b], Thomas Wagner[b1]

[a]Private Pädagogische Hochschule der Diözese Linz, [b]Pädagogische Hochschule Oberösterreich
gudrun.keplinger@ph-linz.at; https://doi.org/10.17883/pa-ho-2024-02-02

SUBMITTED 09 JAN 2024 REVISED 08 MAI 2024 ACCEPTED 11 MAI 2024

Interaction, a central construct in EFL teaching (Huth, 2021), has gained momentum in post-COVID online interactional contexts, in particular for face-to-face (FTF) versus synchronous computer mediated communication (SCMC, Aubrey & Philpott, 2023; Berglund, 2009; Kim, 2014; Loewen & Sato, 2018; Ziegler, 2016; Zou & Jin, 2021). The present exploratory study investigated these two contexts with regard to the interactional components willingness to communicate (WTC), communicative self-confidence (CSC), and frequency of language use (FE) in Austrian student teachers, partly replicating Darasawang and Reinders (2021). Using a parallel convergent mixed-methods design (Creswell & Plano Clark, 2006), quantitative and qualitative data were collected from 58 student teachers after exposure to both contexts in two CLIL-based courses. An online survey with 21 items explored in how far interactional context affected WTC, CSC, and FE quantitatively. The qualitative study, using two open questions for each component, explored students' perceptions and explanations of potential interactional differences. Quantitative data were analysed using cumulative link mixed models, qualitative data underwent computer-assisted (MAXQDA) content analysis. Data triangulation revealed partly inconsistent results. On the one hand, there were overall negative effects for SCMC, with explanations including contextual and linguistic factors for CSC and pedagogical reasons for FE. Likewise, students reported higher WTC in FTF, and argued for this with interactional and technology-related reasons. On the other hand, CSC remained inconclusive as to its anxiety component. We interpreted these results with regard to current challenges of post-COVID EFL teacher education in Austria.

KEIYWORDS: EFL teacher education, willingness to communicate,
perceived communicative self-confidence, frequency of use of English

1 The current project was funded by the University Colleges of Linz Diocese and Upper Austria. We are grateful to the participants in our study as well as two anonymous reviewers for their constructive feedback on earlier versions of the manuscript. The authors have no conflicts of interest to disclose. Correspondence concerning this article should be addressed to Gudrun Keplinger, Private University of Education, Linz Diocese, Salesianumweg 3, 4020 Linz, Austria. Email: gudrun.keplinger@ph-linz.at.

1. Introduction

Ever since the rise of the Interaction Approach (Loewen & Sato, 2018; Mackey, 1999) and Sociocultural Theory (Lantolf & Thorne, 2006, 2007), the construct interaction has had a central place in the context of EFL teaching and learning (Ellis & Shintani, 2014; Hall, 2010; Huth, 2021; Smit et al., 2021). That is why one major challenge that teachers are confronted with is providing opportunities for and facilitating interaction in the language classroom in order to provide an environment conducive to learning. This central place, widely accepted and thoroughly researched with respect to classical face-to-face environments (FTF), has now gained momentum in SCMC language learning research, too (Aubrey & Philpott, 2023; Berglund, 2009; Junn, 2021; Kim, 2014; Qiu & Bui, 2022; Skidmore, 2023; York et al., 2020; Ziegler, 2016; Zou et al., 2021, for an overview see Loewen & Sato, 2018), in particular after the COVID-19 pandemic. In fact, SCMC environments have repeatedly been acknowledged as a meaningful complement, or even alternative, to FTF instruction (Yu, 2022). More research on the potential impact of SCMC environments on interaction is needed, however (Dao, Duong & Nguyen, 2021), especially in post-secondary and tertiary education. The present explorative study is trying to fill this gap. We build on Keplinger et al. (2021), who reported perceived deficits in student teachers' perception of EFL peer interaction in SCMC and suggest exploring this type of interaction in more detail. Using a tertiary Content and Language Integrated Learning (CLIL) setting (Lasagabaster & Doiz, 2016; Dalton-Puffer et al., 2022), in which dialogic teaching and the interactive language use for the negotiation for meaning is central (Coyle et al., 2010; Spratt, 2017), the present mixed-methods study explores the potential influence of interactional context (FTF versus SCMC) with regard to the three interaction components willingness to communicate (WTC), communicative self-confidence (CSC), and frequency of language use (FE). These three constructs were selected in order to operationalise and quantify a measure for interaction (Darasawang & Reinders, 2021).

The empirical study was driven by three research questions. Question 1 asked in how far interactional context affected the three interaction components. Question 2 explored in how far students identified potential differences between the two interactional contexts. And question 3 explored reasons students would put forward for such potential differences. While the quantitative study pursued research question 1, the qualitative study was concerned with all three questions.

2. Theoretical Background

In this section, we will first summarise recent finding about interaction in SCMC contexts and discuss its particular relevance in CLIL settings. We will then briefly report how we operationalised interaction in terms of WTC, CSC, and FE. Each of

these concepts will first be discussed in turn and then synthesised for the research questions and hypotheses.

2.1 Interaction in SCMC and CLIL

For a long time, research focusing on interaction in SCMC mainly looked into characteristics of written discourse (Dao et al., 2021), and only recently, due to the technological development brought about by the pandemic, video-chats have been receiving more attention (Hung & Higgins, 2016; Jung & Lee, 2015; Lenkaitis, 2020; Shih, 2014; Wigham & Chanier, 2015; Yanguas, 2010; 2012). The main focus of these studies, however, has been on general learners' language skills and knowledge rather than (prospective) educators in the field, who are supposed to develop their skills not only in language proficiency courses but also in content courses taught in the target language.

In these CLIL settings, negotiation for meaning, being "at the heart of the interaction hypothesis" (Loewen & Sato, 2018), has a central place (Coyle et al., 2010), as facilitating engagement with resources, other learners, and the teacher have proved to be elements of effective pedagogy (Badertscher & Bieri, 2009). Effective, in this context, does not only refer to the development of language proficiency but also to the acquisition of deep knowledge, i.e., knowledge that can be used in new situations (Weinert, 2001), which is regarded as crucial in initial teacher education (Flores, 2020).

2.2 Willingness to Communicate

Willingness to communicate (WTC) in SLA emerged in the 1990s as a concept strongly associated with L2 communication and interaction, which, in turn, could facilitate successful L2 acquisition (Clément et al., 2003; MacIntyre, Clément, Dörnyei, & Noels, 1998). It can be described as an L2 learner's readiness, intention, and volition to enter into interaction or to remain silent given free choice (MacIntyre, 2007). The original scale to measure WTC was developed in 1991 with respect to pre-defined communication contexts and three receiver types (strangers, acquaintances, and friends, cf. Ayer-Glassey & MacIntyre, 2019).

WTC has a static as well as a situational trait, the latter referring to an individual's intention to involve themselves in particular communicative situations (Cao & Philp, 2006), such as teaching context (Aubrey & Philpott, 2023), social status (Dörnyei & Kormos, 2000), and positive reinforcement of peers (MacIntyre, Baker, Clément, & Conrod, 2001). These factors impact on a learner's immediate desire and confidence to engage in interaction and help account for learners taking advantage of affordances to communicate as they arise, for example, volunteering

to respond to a question that the teacher has posed to the whole class (Ducker, 2022). In other words, this layer of WTC is regarded as being context-dependent.

2.3 Communicative Self-Confidence

Perceived communicative (or communication) self-confidence (CSC), also referred to as L2 confidence or linguistic self-confidence (Clément, Dörnyei, & Noels, 1994; Noels et al., 1996), has been defined as "the overall belief in being able to communicate in the L2 in an adaptive and efficient manner" (MacIntyre et al., 1998, p. 551). Along with factors such as mood, motivation, or anxiety, CSC was already part of Krashen's (1981) hypothesis of the affective filter, and, as an affective factor in L2 individual differences research, it has been thoroughly examined since the 1990s (Clément et al., 1994; Lee & Lee, 2019; Noels et al., 1996, Khajavy et al., 2016).

On the one hand, CSC has often been treated as a distinct construct, either static or dynamic, but separated from anxiety. On the other hand, it has repeatedly been viewed as some sort of composite variable, a combination of perceived competence and lack of anxiety, for instance (Clément & Kruidenier, 1985; MacIntyre et al., 2003; Mulyono & Saskia, 2021), or essentially as an immediate precursor of L2 WTC (for a discussion see MacIntyre et al., 1998, and Loewen & Sato, 2018).

A strong link between CSC and WTC has been reported in numerous studies over the last three decades, also in relation to motivation, language proficiency (Darasawang & Reinders, 2021), and, most importantly here, interactional context, too (Léger & Storch, 2009). And it was Léger and Storch, who, almost 15 years ago, emphasized the complex interplay between CSC, anxiety, and a learner's behaviour in certain interactional contexts, such as FTF or SCMC.

With regard to CSC and interactional contexts in post-secondary language learning, there is ample evidence for a whole range of factors CSC is associated with (Moratinos-Johnston, Juan-Garau, & Salazar-Noguera, 2019 for an overview). Lee (2019), for instance, could show that CSC was positively associated with L2 oral communicative proficiency irrespective of interactional context. As for WTC, Fallah (2014), in a quantitative study with English-major university students, revealed significant structural equation paths from CSC to WTC. In a similar vein, Lee and Hsieh (2019), in a quantitative study with undergraduate students, found that CSC was positively associated with WTC in both FTF and SCMC communicative settings. Aoyama and Takahashi (2020), Fatima et al. (2020), and Mulyiono and Saskia (2021), also in quantitative studies, found the same effect.

Lee and Hsieh (2019), however, could show that a lack of anxiety was positively associated with WTC in FTF but not in SCMC. In another study, Lee and Lee (2019) showed that a positive association between CSC and WTC was limited to exposure outside the regular classroom, and that CSC was positively associated with WTC in digital settings only once learners had previously engaged in what

they called cross-cultural online communication. A positive association does thus appear to depend on earlier experiences with online teaching and learning.

There are also qualitative data on CSC and WTC. Cao and Philp (2006), for instance, in a study with eight adult learners of English in a language school ESL program, showed in their interview data that these adult learners identified both CSC and medium of communication, amongst others, as a strong influence on their WTC. Overall, it seems to remain unclear if CSC is generally affected by interactional context, and in how far such potential effects are mediated by anxiety, in particular in post-secondary educational contexts. That is why the present study explicitly explores CSC in both contexts.

2.4 Frequency of English Use

Even though frequency of language use (FE) cannot be defined as a construct per se, it can be argued that language output necessarily involves language use, and production practice is necessary to develop fluency and automaticity (DeKeyser, 2001; Lyster & Sato, 2013; DeKeyser, 2017), abilities which are of utmost importance for language teachers. Hence, it was important for us to see whether any differences between the two settings could be determined.

Moreover, we are addressing an issue which has been reported as a major limitation in many WTC studies, which, as highlighted by Ducker (2022), have yielded mixed results with regard to correlation between observable communication and situated moment-to-moment ratings of WTC. This is why the aspect of frequency of English use was included in our study, as we were hoping to gain a better understanding of the WTC-talk relationship, though limited to self-perceived ratings of the participants.

3. The Empirical Studies

The present study employed an exploratory mixed-methods convergent parallel design (Creswell & Plano Clark, 2006; Creswell et al., 2002; Riazi & Candlin, 2014). This is illustrated in Figure 1.

FIG. 1 *Exploratory Mixed-Methods Convergent Parallel Design*

Quantitative data	Parallel Analysis		
		Triangulation	Interpretation
Qualitative data	Parallel Analysis		

The rationale behind this choice was the assumption that parallelly triangulating quantitative and qualitative data would produce insights not gleaned from one of these fields alone. Thus, in a pre-meditated fashion, quantitative and qualitative studies were devised, data were collected parallelly, analysed individually first,

and then triangulated in order to arrive at appropriate conclusions and suggest avenues for further research. The goal of the quantitative part was to analyse the quantitative within-subject questionnaire data and trace potential effects of the delivery mode by means of cumulative link mixed regression modelling. The aim of the qualitative part was to explore potential effects in more detail by using concept-driven categories, derived from the open questions in the questionnaire, and data-driven categories which were developed through inductive coding. The triangulation of both types of data was intended to complement and contextualise individual findings, guide the discussion of potential didactic implications, and finally illuminate avenues for future research.

The synchronous computer-mediated communication (SCMC) tool of choice was the proprietary cloud-based collaborative VoIP application *Zoom* (Zoom Video Communications Inc., 2022), since at the time it was the service that all institutes involved in the present study had subscribed to. Although Zoom offers secure session recordings without recourse to third-party services and software, the sessions under investigation were not recorded, as this kind of data collection was not part of the research design.

3.1 The Quantitative Study

3.1.1 Sample

For the quantitative empirical study, 59 EFL student teachers were recruited from a joint-honours-degree programme, delivered by two University Colleges as well as one University in an Austrian educational cluster. From 59 data sets, one case had to be discarded due to data not missing at random. There were no ethnicity issues interfering with data collection. Power analyses (Champely, 2020; Kohl, 2023) suggested that for 80% power and a significance level of $\alpha = .05$ the minimum sample size would be 54. Additionally, the within-subject design with random effects would ensure a minimum of Type II errors. All 58 participants took part anonymously, voluntarily, without any financial renumeration and with explicit consent. Data collection and analysis procedures were approved and funded by both University Colleges involved. Their ethical and institutional guidelines regarding the rights of research participants, in keeping with the APA Ethics Code Standard (American Psychological Association, 2017) as well as the guidelines from the British Association for Applied Linguistics (BAAL, 2021) were adhered to at all times.

3.1.2 Materials and Procedure

Two undergraduate courses in the academic year 2022-2023 provided the backdrop for class delivery mode. One of them, called *Teaching Language Skills I*[2] was delivered online, in synchronous computer-mediated communication mode (SCMC),

2 For details on these two courses see https://osf.io/tyx3b

due to the COVID 19-pandemic, whereas the other one, *Teaching Language Skills II*, could take place in a face-to-face setting (FTF). Potential effects of delivery mode on students' language behaviour were elicited using three constructs.

The first construct was willingness to communicate (WTC), measured through a scale adapted from MacIntyre et al. (2001), Léger and Storch (2009), as well as Darasawang & Reinders (2021). Adaptions mainly concerned minor wording alterations; the term *teacher*, for instance, was changed into *lecturer* in order to fit our data elicitation context. The fully labelled bipolar scale consisted of four items on a 5-point Likert scale ranging from 1 (*very unwilling*) to 5 (*very willing*). Thus, high values would reflect a high willingness to communicate.

The second one was communicative self-confidence (CSC), consisting of eleven items on a fully labelled bipolar 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*), adapted from Ryan (2009) and Darasawang and Reinders (2021). Items were worded both positively and negatively in order to control acquiescence response bias (Kreitchmann et al., 2019). For composite sum score aggregation, the four negative items were reversed for the analysis, however, so that high values would reflect consent. For a critical review of this established practice see Sonderen et al. (2013).

The final construct assessed was frequency of English (FE), based on Darasawang and Reinders (2021). Their seven items were complemented with one additional question in order to differentiate frequency of English usage with peers versus lecturers. Thus, eight items on a fully labelled unipolar 5-point Likert scale ranging from 1 (*never*) to 5 (*always*) were used. All items had a positive wording so that high values would reflect frequent use of English[3].

Data collection was done via an online survey platform (*SoSciSurvey*, Leiner, 2019). Both the instructions and all scales were administered in English. In the introduction, participants were informed that this survey revolved around the three constructs in both SCMC and FTF classes with explicit reference to the two courses. The introduction then stated that it would take approximately 15 minutes to finish the survey, and that it was vital that they replied to all items truthfully.

3.1.3 Data Coding and Analysis

Individual Likert ratings as well as aggregated composite sum scores constituted the dependent variables. Delivery mode, SCMC versus FTF, formed the dichotomous independent variable. First, the three scales were examined for internal consistency (Cronbach's α), normality (Shapiro-Wilk-test), and homogeneity of variance (Levene test, extended Brown-Forsythe test).

Then, in order to examine effect of delivery mode on the unaggregated Likert ratings (Carifio & Perla, 2007; Lantz, 2013; Taylor et al., 2023), cumulative link mixed models (Christensen, 2019) were fitted to the data. Analysing the original

3 A complete list of all the items can be found in Table A1 at https://osf.io/tyx3b/

ordinal data as repeated measures in mixed models has a number of advantages (Cunnings & Finlayson, 2015; Linck & Cunnings, 2015; Taylor et al., 2023). First of all, our observations (Likert responses) are not independent, so traditional ANOVA approaches would be inappropriate. Second, Likert ratings should probably not be treated as metric data (Lantz, 2013; Liddell & Kruschke, 2018; Janda & Endresen, 2017), and, apart from unrealistic estimates, both Type I and Type II errors tend to occur more often. Third, the mixed models are less sensitive to error distributions, sphericity, or unequal variance. Fourth, mixed models can control typical repeated measure random effects (Baayen et al., 2008). And finally, typical Likert biases such as social desirability and acquiescence can at least partly be controlled by allowing random response intercepts and slopes for participants and items (Kreitchmann et al., 2019). Random effects for participants are particularly useful, since we also statistically control potential personal rating preferences across the items, typically known as 'pockets' of inhomogeneity in a repeated-measures research design. Accounting for such random effects in turn improves the accuracy of fixed effect estimation.

Data were statistically analysed using the R environment for statistical computing (R Core Team, 2022). The cumulative link mixed models, also known as the proportional odds models, were fitted with Laplace approximation (Christensen, 2019, version 2019.12.10, cf. Olsson, 2022). Main regression effects from the mono-factorial mixed models are reported based on Type-II Wald χ^2-tests, as implemented in the *RVAideMemoire* packages (Fox & Weisberg, 2019). Partial effects, along with their statistics and Wald-based *p*-values, were derived from the model summaries. Significance levels are reported in categories of $p < .05$, $p < .01$, and $p < .001$, or in full, once $\alpha > .05$. Effect size interpretation followed Plonsky and Oswald (2014), with small effects for Cohen's approximations around .25, medium sizes between .26 and .59, and large effect sizes ≥ .60. The survey, the data and the statistical analyses from the present quantitative study are available via the free repository OSF at https://osf.io/tyx3b/

3.1.4 Results

The composite sum scores of the three constructs were first examined for their distribution, variance, and internal consistency. Table 1 reports conditional means, medians, standard deviations, and raw Cronbach's α values for each of the three scales (WTC, CSC, FE) in each of the two types of instruction, face-to-face (FTF) versus online (SCMC).

TAB. 1 *Conditional Means, Medians, Standard Deviations, and Cronbach's α for the Three Scales*

		Delivery Mode							
		Face-to-Face (FTF)				Online (SCMC)			
		M	*Mdn*	*SD*	α	*M*	*Mdn*	*SD*	α
Scale	WTC	16.29	17.0	3.77	.88, 95% CI [.78, .93]	13.26	13.5	3.57	.72, 95% CI [.53, .81]
	CSC	43.02	45.5	6.68	.85, 95% CI [.79, .89]	44.59	43.5	6.29	.81, 95% CI [.74, .85]
	FE	34.26	35.0	4.22	.64, 95% CI [.47, .75]	32.95	34.0	4.71	.70, 95% CI [.51, .80]

Note. The raw Cronbach's α values with bootstrapped confidence intervals based on 1000 samples were calculated here, since some items did not show equal variances. Apart from the online CSC and WTC scores, subscales are not distributed normally ($p < .05$).

As we can see in Table 1, both mean and median values of the three constructs are generally rather similar across types of instruction, except for WTC, which is markedly higher in the live teaching (FTF). We can also see that the subscales come with acceptable raw Cronbach's α values (≥ .70), except for face-to-face FE (α = .64).

We then explored potential effects of delivery mode on the original Likert ratings on each of the three constructs. Therefore, cumulative link mixed models by Laplace approximation with random effects for subject and item were fitted to these scales. For CSC and FE, final models included random intercepts and slopes for both participants and items; for WTC, random slopes for items had to be dropped, since it could not define the variance-covariance matrix of the parameters and hence did not converge properly.

As far as CSC is concerned, the model did not produce a significant main effect for delivery mode (Type II analysis of deviance, $\chi^2(1) = 1.36$, $p = .24$), with a negative but small to medium effect size of Cohen's $d = -0.29$) only. For FE, a corresponding cumulative link mixed model did not find a significant main effect for delivery mode on FE either (Type II analysis of deviance, $\chi^2(1) = 0.62$, $p = .43$), again with a negative but rather small effect size of Cohen's $d = -0.32$. For willingness to communicate (WTC), however, the cumulative link mixed model found a significant main effect for instruction ($\chi^2(1) = 15.05$, $p < .001$), with a strong effect size of Cohen's $d = -1.11$. The resulting model for WTC is summarised in Table 2 (next page).

We can see in Table 2 that the predictor level 'SCMC classes' comes with a negative logit, an odds ratio below 1.00, and a considerable effect size of Cohen's $d = -1.01$ The corresponding odds ratio of 0.13 means that the odds of getting a higher willingness to communicate rating in SCMC instruction is 0.13 times that of FTF instruction, or 87% lower than for the face-to-face variant; in other words, this model would predict that participants perceive their willingness to communicate in FTF instruction far better than they do for online environments.

TAB. 2 *Coefficients and Test Statistics from a Cumulative Link Mixed Model (Laplace Approximation) for Willingness to Communicate by Delivery Mode, with Random Intercepts and Slopes for Participants as well as Random Intercepts for Items*

effects	estimate		Cohen's *d* approx.[a]	*SE*	*z*	95% *CI* odds ratios		*p*
	logits	odds ratios				*LL*	*UL*	
fixed effects								
threshold coefficients								
1 | 2	-5.81	0.00	-3.21	0.53	-10.89	0.00	0.01	< .001 ***
2 | 3	-3.44	0.03	-1.90	0.44	-7.82	0.01	0.08	< .001 ***
3 | 4	-1.84	0.16	-1.01	0.41	-4.45	0.07	0.36	< .001 ***
4 | 5	0.20	1.22	0.11	0.40	0.51	0.56	2.67	0.61
predictor								
SCMC classes	-2.01	0.13	-1.11	0.42	-4.81	0.06	0.30	< .001 ***

random effects	variance	*SD*	corr.
ID intercept τ	6.16	2.48	
online instruction	4.79	2.19	-0.71
Item intercept τ	0.07	0.27	

Note. Number of partial effects =1, number of observations =464, total *N* =58, *ICC* =0.59, marginal R^2 =0.11, conditional R^2 =0.64, pseudo-R^2 (McFadden) =0.09, Cox and Snell (ML) =0.22, *CI* =confidence interval, *LL* =lower limit, *UL* =upper limit

[a] Cohen's *d* effect size approximations were calculated using (*log*(odds ratios)x√3)/π, Sánchez-Meca et al., 2003). Condition number of the Hessian, measuring the empirical identifiability of the model, is < 0.001. Maximum absolute gradient of the log-likelihood function with respect to the parameters is 0.002. Scale and nominal tests on a parallel cumulative link model suggest that the proportional odds assumption holds.

In sum, the quantitative data suggest that delivery mode does not substantially affect the students' perceived self-communicative competence or frequency of English usage. The negative coefficients and effect sizes from these regression model suggest a merely mild negative trend for SCMC; it did significantly decrease, however, the students' willingness to communicate in the foreign language.

3.2 The Qualitative Study

3.2.1 Sample

The same 58 students from the quantitative study also took part in the qualitative study.

3.2.2 Materials and Procedure

The qualitative data, too, were collected using SoSciSurvey (Leiner, 2019). After the introduction and the quantitative Likert survey, seven prompts elicited open replies, as illustrated in Table 3.

TAB. 3 *Open Items about WTC, CSC and FE*

(1)	Describe some situations in Zoom TLS settings which encouraged you to speak.
(2)	From your experience, what are the main differences between Zoom and face-to-face settings in TLS1 and TLS2 with regard to your willingness to speak English?
(3)	Consider your communication with your peers in TLS 1 and TLS2. How and to what extent do the two delivery modes Zoom and face-to-face impact your communication with your peers?
(4)	Consider your communication with your lecturer in TLS 1 and TLS2. How and to what extent do the two delivery modes Zoom and face-to-face impact your communication with your lecturer?
(5)	In how far does your anxiety about communication vary in Zoom and face-to-face TLS course settings?
(6)	In how far does your self-perceived communicative competence vary in Zoom and face-to-face TLS course settings?
(7)	Compare the frequency of your communication in Zoom and face-to-face TLS course settings. In which of the two settings do you tend to speak more English? Why?

The data corpus used for content analysis consisted of 7306 words with all student responses to these seven questions.

3.2.3 Data Coding and Analysis

Qualitative data were submitted to content analysis (Mayring, 2022) based on concept-driven (deductive) and data-driven (inductive) category formation (Kuckartz & Rädiker, 2022). This analysis followed a cyclical procedure of developing categories, coding, discussing, and recoding the data, using MAXQDA 2022 analysis software (VERBI Software, 2021). The concept-driven categories were based on the seven open questions in the questionnaire (see Table A2 in the appendix). The corresponding data-driven categories were developed through individual trial-coding (2 coders), and, after revisiting personal interpretations of category definitions in a research meeting, were revised and used for individual coding of the entire data material. After merging the two coders' coding results, the intercoder agreement data analysis yielded rather low Kappa values (Brennan & Prediger, 1981), with $\kappa = 0.25$ at 90% segment level and $\kappa = 0.46$ at 10% segment

level. In a final coding session, some code assignments were deleted, added, and changed until full agreement was reached. All concept-driven and data-driven categories as well as the number of codes after the first and second coding session can be found at OSF at https://osf.io/tyx3b/

3.2.4 Results

The research focus of the qualitative study was the interactional context (FTF, SCMC) and in how far it affected the three interaction components (WTC, CSC, FE). Here, we looked into the quantity of and reasons for potential differences the students perceived in both settings. In addition, we explored student-peer and student-lecturer communication in the two delivery modes, and we also traced factors, which encouraged students to communicate in SCMC settings. In the following, the main results are presented with reference to the research questions. The student teachers' statements (S01–S37), selected to corroborate or illustrate these findings, are listed in Table A3 at https://osf.io/tyx3

The wording, spelling and punctuation of verbatim quotations in these statements correspond to the original transcription, a common practice in qualitative content analysis (cf. Kuckartz & Rädiker, 2022).

WTC. With three 'More WTC in SCMC', 22 'Less' and 7 'No difference' code assignments, the students' responses concerning differences in perceived WTC in FTF versus SCMC settings suggest a tendency towards higher self-perceived WTC in FTF delivery mode settings. The students' responses indicating possible reasons for these self-perceived differences can be subsumed under four types.

Interactional reasons. One of the key interactional factors which positively impacts the students' willingness to speak in FTF settings seems to be the higher degree of mutual familiarity with fellow course participants (S01). Additionally, the students reported interactional differences in terms of turn-taking (S02) and due to absence (or lack of) physical features in Zoom settings, such as facial expressions and gestures (S03). This perceived lack of physical features, in turn, may be one of the reasons for further interaction-related issues in Zoom settings, such as students' anticipation difficulties (S04), and their personal impression of not feeling addressed (S05).

Pedagogy-related reasons. The students' responses suggest negative effects of online settings on student participation in discussions. Apparently, in face-to-face classes it is easier to take part in discussions and it is also considered more natural (S06).

Personal-affective reasons. The students' responses for self-perceived WTC differences in FTF versus SCMC settings yielded rather ambiguous results. On the one hand, quite a few students reported that they experienced FTF settings as more personal, more fun and more motivating, they also reported about the distance Zoom creates, and the artificiality, weirdness and awkwardness of talking to

a computer screen. On the other hand, there were also responses that suggested more willingness to answer questions online, either because the students felt less anxious than in live contexts (S07) or wanted to overcome the silence-induced awkwardness in a Zoom session (S08).

Technology-related reasons. Most interactional factors impacting the students' willingness to speak in the two different settings are closely linked to technology-related reasons. In addition to internet connection problems, key issues the students listed were webcam- and audio-related, such as feeling watched, the fear of being audio-/video-recorded or photographed by peers or others (S09), and minor technical hassles such as unmuting oneself before speaking. Finally, many student responses suggested distraction as another technology-related (in a broader sense) reason that played a crucial role in online course delivery (S10).

CSC. To trace potential differences the students would list with regard to CSC, we asked several questions that focused on the two key concepts of the construct separately. Thus, the presentation of the results will also be divided into those related to communicative competence, on the one hand, and language anxiety, on the other.

Communicative competence. In terms of self-perceived communicative competence, the students did not report major differences, although there are more responses suggesting less in Zoom settings. Apart from *contextual* factors, such as more opportunities for successful communication in FTF settings, it appears as if *linguistic* factors affected the students' communicative competence. Some students, for instance, emphasised the benefits of face-to-face sessions for improving speaking skills through developing pronunciation, fluency and accuracy (S11). There also seem to be different ways of dealing with language errors and mistakes in the two settings (S12). Yet, while Zoom is considered less spontaneous (S13), overall, perceptions differ.

Language anxiety. In terms of students' self-perceived language anxiety, the analysis yielded mixed results. With 15 code assignments suggesting more anxiety, 15 less anxiety, and 13 no difference in SCMC settings, the two delivery modes seem to have both increasing and reducing effects on the students' language anxiety levels. Among the anxiety-inducing factors that are at play in Zoom sessions are *interactional* and *personal-affective* factors, such as the anonymous setting of the online course (S14), the fear of interrupting others (S15), prolonged periods of silence (S16), and feelings of inhibition. Some students, on the other hand, also reported reduced anxiety levels in Zoom settings, where they experienced less exposure, less stress (S17), and for *pedagogy-related* reasons, such as the limited number of conversation partners in breakout rooms (S18).

In sum, with regard to CSC, the qualitative data did not show major differences between FTF and SCMC-settings.

FE. As for potential differences in frequency of English use in FTF versus SCMC settings, the students' responses revealed a clear tendency towards higher frequencies in FTF course settings (more speaking in SCMC: 4, less speaking in SCMC: 28, no difference: 8). And yet, there are inconclusive results concerning the frequency of L1 and L2 use in the two settings. While some code assignments suggested hardly any or no differences, some students reported more L1 speaking in FTF sessions and frequent code switching from L2 to L1, particularly in breakout rooms (cf. statements S19, S20, and S21). As the data shows, the reasons for the differences in FE are often *pedagogy-related* (e.g. better monitoring through the lecturer in FTF settings) and *personal-affective*, including, e.g., less pressure to speak in Zoom sessions, and feeling better and more comfortable in FTF classes). Additionally, Zoom courses were often reported to offer fewer speaking opportunities, not least due to *technology-based* reasons (S22).

Student-peer communication. In terms of impact on student-peer communication, the number of code assignments for positive responses for FTF settings by far exceeded those for Zoom settings (35/6), while for negative responses the latter received 32 responses (face-to-face: 1). Some key factors, repeatedly put forward by respondents, were interactional in nature. They included better opportunities for getting to know each other FTF (S23), time for informal conversation before and after class (S24), and higher level of interaction with peers in general (S25). In addition, the responses show that a number of pedagogy-related factors, such as the importance of playing games for getting to know each other better, as well as personal-affective factors (e.g. the importance of feeling comfortable) and technology-related factors (e.g., appropriate sound volume in breakout rooms) have an impact on student-peer communication.

Student-lecturer communication. The results for student-lecturer communication showed a similar picture. Again, the number of responses suggesting a positive impact through FTF was considerably higher than in the online setting (24/1), while the responses for negative impact revealed an inverse ratio (1/16). As the code assignments suggest, a key factor impacting student-lecturer communication, at interactional level, appears to be student perceptions of the lecturer's reduced approachability (cf., e.g., S26). Similar to student-peer communication, the code assignments also included personal-affective factors, such as student perceptions of good atmosphere and feelings of (dis)comfort. Interestingly, personal-affective factors can positively impact student-lecturer communication in Zoom settings, too (S27).

Factors encouraging communication in SCMC. When asked to describe situations in Zoom contexts which encouraged them to speak, the students mentioned affective, content-related, lecturer-related, task-related, and technology-related factors. Affective factors included the significance of knowing course members, avoiding long periods of silence and the importance of making online

sessions more personal (S28). The importance of personalising Zoom sessions in order to encourage speaking is reflected in the content-related factors the students mentioned, which also included sharing personal experiences (S29). Additionally, interesting topics and activities (S30), giving feedback, asking (or being asked) meaningful questions (S31) were also reported to encourage students to speak. As far as lecturer-related factors are concerned, unsurprisingly, the lecturer's friendliness, was mentioned as one of the encouraging factors (S32). Task-related factors included, for instance, working in breakout rooms with peers (S33), but also working through new material and sharing personal experiences about topics in the plenary (S34). Finally, there are also technology-related factors which students consider conducive to speaking. In addition to open access to Zoom rooms prior to and after online meetings (S35), for instance, different ways of using webcams also seemed to matter. Here, code assignments for preferences, however, were ambiguous, ranging from advocating camera use (S36) to avoiding it (S37).

In sum, the qualitative data illustrate that WTC tends to be higher in FTF settings, with a high number of personal-affective and interactional reasons. With regard to CSC, no major differences could be determined in the data, FE was higher in FTF settings, mainly due to pedagogy-related reasons. Both student peer and student lecturer interaction were perceived as generally better in FTF settings.

4. Triangulation[4] and Discussion

While the quantitative results revealed comparable mean and median scores across instructional modes for CSC and FE, the scores for WTC decreased significantly in the SCMC setting. This suggests a negative impact on students' confidence in engaging with the language in online settings. Similarly, a substantial effect of delivery mode on WTC was also revealed in cumulative link mixed models, with online classes significantly reducing the students' WTC. However, despite minor negative trends in FE and communicative competence in SCMC delivery mode, there were no statistically significant differences.

Qualitative findings provided in-depth insights into students' perceptions and highlighted factors influencing WTC discrepancies. Interactional factors, including familiarity with peers and turn-taking, proved to be key contributors to higher WTC in face-to-face settings. Pedagogy-related issues, such as ease of participation and naturalness of discussions, were also influencing factors. Additionally, technology-related issues, such as webcam problems and distractions, also had negative effects on WTC in online settings. Among the factors potentially impacting the students' communicative competence were linguistic factors and lan-

4 The following two paragraphs provide a summary of both results sections supported by ChatGPT (OpenAI. (2024). ChatGPT (Version 3.5) [Large Language Model]. https://chat.openai.com/).

guage anxiety, while FE usage showed a clear preference for face-to-face settings. Student-peer and student-lecturer communication were found to be more positive in face-to-face settings, indicating the importance of personal interaction.

As can be seen in Table 4, the complex interplay between delivery mode and the students' learning experience becomes evident after triangulation of the quantitative and qualitative results.

TAB. 4 *Integrated Results Matrix for Student Teachers' WTC, CSC and FE in FTF versus SCMC Settings*

Quantitative results	**Qualitative results**	**Example quotes**
With regard to WTC, a significant and strong positive effect could be observed for FTF class delivery.	With regard to WTC, the students' responses suggest a tendency towards higher self-perceived WTC in FTF delivery mode settings.	"I generally am more willing to talk in a face-to-face setting than over Zoom."
With regard to CSC, only a mild positive trend could be observed for FTF class delivery.	With regard to CSC, no major differences were described by the participants in FTF versus SCMC settings.	"I know what I am capable of so it does not matter for me if communcation takes place online or in person."
With regard to FE, only a mild positive trend could be observed for FTF class delivery.	With regard to FE, students reported that they had more opportunity to speak in FTF than SCMC settings.	"You in a classroom setting encourages you more to talk to the peers in class, even if it's not course-related".

The importance of interactional reasons such as familiarity with the interlocutor(s) is also stressed by other researchers who reported higher WTC in FTF settings (e.g., Lee & Hsieh, 2019; Zarrinabadi, 2021). Similarly, pedagogy-related reasons such as instructional support and the creation of a good and safe learning environment have been determined as being conducive to WTC (Lee & Lee, 2019). Seyyedrezaei & Ziafar (2014) describe a high correlation between learners' WTC and their belief that their feelings can be expressed in a computer-mediated setting. However, the quality of this expression could depend on factors such as the visibility of facial expression and body language, which might be limited in the SCMC setting. Another potentially relevant factor has been reported in Topalov et al. (2022). They argue that willingness was highest in situations which could be regarded as particularly important for their participants in their role as students, e.g., asking for clarification at the end of the lesson. Moreover, as self-confidence has been found to be an important factor affecting WTC (Cao & Philp, 2006), one might argue that because of the fact that the participants in our study had reached a high level of language proficiency already, they were more likely to be self-confident in using the language.

Our data suggest that neither perceived CSC nor FE usage were actually hampered in a computer mediated delivery mode when compared with the FTF settings. Both the level of CSC and FE were only mildly decreased in SCMC. This is somewhat striking, as WTC is considered to be a strong predictor of the frequency with which learners communicate in the target language (Darasawang & Reinders, 2021). Apparently, although students felt profoundly impeded in their willingness to communicate, they nevertheless contributed on a level comparable to traditional live teaching; and while their willingness to communicate decreased, their self-confidence in communication did not. Thus, it seems that students somehow managed to overcome initial concerns about SCMC communication, eventually contributed as they would have done in FTF mode, and, retrospectively, turned out to be self-confident about their communicative contributions. Thus, while there are more or less pronounced detrimental SCMC mode effects on students' L2 behaviour, those effects might be reducible and compensable.

There are several limitations to our study that need to be acknowledged. First, the study only investigated a small sample of students over a limited time period of two semesters. Second, a number of influencing factors such as a certain degree of familiarity with the overall course design, instructors and fellow students were not taken into account. Third, the definition of the construct FE, as used in the present study, needs further revision, as does CSC and related scales. Fourth, a within-subjects comparison of delivery mode, with students serving as their own control group, could diminish random influencing factors commonly associated with between-group designs. And finally, in order to trace in how far students actually engage in foreign language interaction and communication, additional target measurements, such as foreign language classroom engagement (FLCE, Guo et al., 2023), foreign language enjoyment (FLE, Dewaele et al., 2023) or standardised holistic procedures (van Batenburg et al., 2018) would help grasp the construct of interaction more precisely and comprehensively.

5. Conclusion and Implications for Teacher Education

The implications of our findings for the training of English teachers are three-fold. First, WTC can be increased by providing ample opportunity for informal and spontaneous conversations, by fostering familiarity among peers and lecturer, and preventing situations which trigger a feeling of awkwardness and allow for distraction, which was shown to be more successfully accomplished in FTF settings than in SCMC. Second, for proficient language users, communicative self-confidence does not seem to be severely affected by the teaching mode but rather robust, whereas language anxiety is sometimes also described as being lower in SCMC settings. As being able to communicate confidently in FTF settings is of particular importance for future language teachers, one might argue that especially for stu-

dent teachers whose anxiety seems to be higher in FTF settings, it is of great importance to be given ample opportunity to overcome such fear in many forms of social interaction. Third, the frequency of use of the target language, it seems that even advanced language users tend to use their L1 in situations where they are not observed by the instructor, which was more often the case in SCMC. Hence, FTF settings might provide more opportunity for further language development. In FTF, however, the L1, was also used to get to know each other better, which, in return, probably had a positive effect on WTC.

Considering the key findings and given that there are current trends to offer teacher education programmes in synchronous and asynchronous delivery modes rather than FTF, the study seems to suggest that in the context of EFL teacher education reinforcing this trend might come with a risk.

References

American Psychological Association (2017). *Ethical principles of psychologists and code of conduct* (2002, amended effective June 1, 2010, and January 1, 2017). https://www.apa.org/ethics/code/

Aubrey, S., & Philpott, A. (2023). Second language task engagement in face-to-face and synchronous video-based computer-mediated communication modes. Performances and perceptions. *System, 115*(6), 103069. https://doi.org/10.1016/j.system.2023.103069

Aoyama, T., & Takahashi, T., (2020). International Students' Willingness to Communicate in English as a Second Language: The Effects of L2 Self-Confidence, Acculturation, and Motivational Types. *Journal of International Students, 10*(3), 703–723. https://doi.org/10.32674/jis.v10i3.730

Ayers-Glassey, S. & MacIntyre, P. D. (2019). *Measuring Willingness to Communicate in First and Additional Languages: An Annotated Bibliography*. (Unpublished Manuscript). Cape Breton University.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Badertscher, H., & Bieri, T. (2009). *Wissenserwerb im Content and Language Integrated Learning*. Haupt.

Berglund, T. (2009). Multimodal student interaction online: An ecological perspective. *ReCALL, 21*(2), 186–205. https://doi:10.1017/S0958344009000184

British Association for Applied Linguistics. (2021). Recommendations on good practice in applied linguistics. 4th edition, available at https://www.baal.org.uk/wp-content/uploads/2021/03/BAAL-Good-Practice-Guidelines-2021.pdf

Cao, Y., Philp, J. (2006). Interactional context and willingness to communicate: A comparison of behavior in whole class, group, and dyadic interaction. *System*, 34, 480–493. https://doi.org/10.1016/j.system.2006.05.002

Carifio, J., & Perla, R. J. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences*, *3*(3), 106–116. https://doi: 10.3844/jssp.2007.106.116

Champely, S. (2020). pwr. Basic functions for power analysis. R package version 1.3-0. [Computer software]. Available at https://CRAN.R-project.org/package=pwr

Christensen, R. H. B. (2019). *Ordinal*. Regression models for ordinal data. (Version 2019.12-10) [Computer software]. Available at https://CRAN.R-project.org/package=ordinal.

Clément, R., & Kruidenier, B. (1985). Aptitude, attitude and motivation in second language proficiency: A test of Clément's model. *Journal of Language and Social Psychology*, *4*(1), 21–37. https://doi.org/10.1177/0261927X8500400102

Clément, R., Dörnyei, Z., & Noels, K. A. (1994). Motivation, self-confidence, and group cohesion in the foreign language classroom. *Language Learning, 44*(3), 417–448. https://doi.org/10.1111/j.1467-1770.1994.tb01113.x

Clément, R., Baker, S. C., & MacIntyre, P. D. (2003). Willingness to communicate in a second language: The effects of context, norms, and vitality. *Journal of Language and Social Psychology*, *22*(2), 190–209. https://doi.org/10.1177/0261927X03022002003

Core Team. (2022). R: A language and environment for statistical computing (Version 4.2.2 "Innocent and Trusting") [Computer software]. Vienna, Austria: R-Foundation for Statistical Computing. Available at https://www.R-project.org/

Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL. Content and Language Integrated Learning*. Cambridge University Press.

Creswell, J. W., & Plano Clark, V. L. (2006). *Designing and conducting mixed methods research* (1st ed.). Sage.

Creswell, J. W., Plano Clark, V. L., Gutmann, M., & Hanson, W. (2002). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 209–240). Sage.

Cunnings, I., & Finlayson, I. (2015). Mixed effects modeling and longitudinal data analysis. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 159–181). Routledge.

Dalton-Puffer, C., Hüttner, J., & Llinares Garcia, A. (2022). CLIL in the 21st century: Retrospective and prospective challenges and opportunities. *Journal of Immersion and Content-Based Language Education, 10*(2), 182–206. doi: dhttps://doi.org/10.1075/jicb.21021.dal

Dao, P., Duong, P. T., & Nguyen, M. X. N. C. (2021). Effects of SCMC mode and learner familiarity on peer feedback in L2 interaction. *Computer Assisted Language Learning, 36*(7), 1206-1235. doi: https://doi.org/10.1080/09588221.2021.1976212

Darasawang, P. & Reinders, H. (2021). Willingness to communicate and second language proficiency: A correlational study. *education sciences, 11*(517). https://doi.org/10.3390/ educsci11090517

DeKeyser, R. M. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125–151). Cambridge University Press.

DeKeyser, R. M. (2017). Knowledge and skill in ISLA. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 15–32). Routledge.

Dewaele J. M., Botes, E., & Meftah, R. A. (2023). A three-body problem: The effects of foreign language anxiety, enjoyment, and boredom on academic achievement. *Annual Review of Applied Linguistics,* 43, 7–22. https://doi.org/10.1017/S0267190523000016

Dörnyei, Z., & Kormos, J. (2000). The role of individual and social variables in oral task performance. *Language Teaching Research, 4*(3), 275–300. https://doi.org/10.1177/136216880000400305

Ducker, N. T. (2022). Bridging the gap between willingness to communicate and learner talk. *The Modern Language Journal, 106*(1), 216–244. https://doi.org/10.1111/modl.12764

Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. Routledge.

Fallah, N. (2014). Willingness to communicate in English, communication self-confidence, motivation, shyness and teacher immediacy among Iranian English-major undergraduates: A structural equation modeling approach. *Learning and Individual Differences, 30,* 140–147. https://doi.org/10.1016/j.lindif.2013.12.006

Fatima, I., Malik Mohamed Ismail, S. A., Pathan, Z. H., & Memon, U. (2020). The power of openness to experience, extraversion, L2 self-confidence, classroom environment in predicting L2 willingness to communicate. *International Journal of Instruction, 13*(3), 909–924. https://doi.org/10.29333/iji.2020.13360a

Flores, M. A. (2020). Learning to teach: knowledge, competences and support in initial teacher education and in the early years of teaching. *European Journal of Teacher Education, 43*(2), 127–130. https://doi: 10.1080/02619768.2020.1733828

Fox J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Sage.

Guo, Y., Xu, J. & Chen, C. (2023). Measurement of engagement in the foreign language classroom and its effect on language achievement: the case of Chinese college EFL students. *International Review of Applied Linguistics in Language Teaching, 61*(3), 1225–1270. https://doi.org/10.1515/iral-2021-0118

Hall, J. (2010). Interaction as method and result of language learning. *Language Teaching, 43*(2), 202–215. https://doi.org/10.1017/S0261444809005722

Hung Y.-W., & Higgins S. (2016). Learners' use of communication strategies in text-based and video-based synchronous computer-mediated communication environments: Opportunities for language learning. *Computer Assisted Language Learning, 29*(5), 901–924. https://doi.org/10.1080/09588221.2015.1074589

Huth, T. (2021). *Interaction, language use, and second language teaching. Routledge advances in second language studies.* Routledge.

Janda, L. A., & Endresen, A. (2017). Five statistical models for Likert-type experimental data on acceptability judgments. *Journal of Research Design and Statistics in Linguistics and Communication Science, 3*(2), 217–250. https://doi.org/10.1558/jrds.30822

Jung, I., & Lee, Y. (2015). YouTube acceptance by university educators and students: a cross-cultural perspective. *Innovations in Education and Teaching International, 52*(3), 243–253. https://doi.org/10.1080/14703297.2013.805986

Junn, H. (2021). L2 communicative competence analysis via synchronous computer-mediated communication (SCMC) as an alternative to formal classrooms. *Innovation in Language Learning and Teaching 17*(1), 1–17. https://doi.org/10.1080/17501229.2021.1895802

Keplinger, G., Spann, H., & Wagner, T. (2021). Interaction types in distance learning – experiences and perspectives of Austrian EFL student teachers. *Pädagogische Horizonte, 5*(2), 181–202.

Khajavy, G. H., Ghonsooly, B., Hosseini Fatemi, A., & Choi, C.W. (2016). Willingness to communicate in English: A microsystem model in the Iranian EFL classroom context. *TESOL Quarterly, 50*(1), 154–180. https://doi.org/10.1002/tesq.204

Kim, H. Y. (2014). Learning opportunities in synchronous computer-mediated communication and face-to-face interaction. *Computer Assisted Language Learning, 27*(1), 26–43. https://doi.org/10.1080/09588221.2012.692386

Kohl, M. (2023). *MKpower:* Power analysis and sample size calculation. (Version 0.7). [Computer software]. Available at https://github.com/stamats/MKpower.

Krashen, S. (1981). *Second language acquisition and second language learning.* Oxford University Press.

Kreitchmann, R. S., Abad, F. J., Ponsoda, V., Nieto, M. D., & Morillo, D. (2019). Controlling for response biases in self-report scales: Forced-choice vs. psychometric modeling of Likert items. *Frontiers in Psychology,* 10, 2309. https://doi.org/10.3389/fpsyg.2019.02309

Kuckartz, U. & Rädiker, S. (2022). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computunterstützung (*5th ed.) Beltz Juventa.

Lantolf, J. P., & Thorne, S. L. (2006): *Sociocultural theory and the genesis of second language development.* Oxford University Press.

Lantolf, J. P., & Thorne, S. L. (2007). Sociocultural theory and second language learning. In B. van Patten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 201–224). Lawrence Erlbaum.

Lantz, B. (2013). Equidistance of Likert-type scales and validation of inferential methods using experiments and simulations. *Electronic Journal of Business Research Methods, 11*(1), 16–28.

Lasagabaster, D., & Doiz, A. (2016) (Eds.). *CLIL experiences in secondary and tertiary education. In search of good practices*. Peter Lang.

Lee, J. S., & Hsieh, J. C. (2019). Affective variables and willingness to communicate of EFL learners in in-class, out-of-class, and digital contexts. *System*, 82, 63–73. https://doi.org/10.1016/j.system.2019.03.002

Lee, J. S., & Lee, K. (2019). Affective factors, virtual intercultural experiences, and L2 willingness to communicate in in-class, out-of-class, and digital settings. *Language Teaching Research, 24*(6), 813–833. https://doi.org/10.1177/1362168819831408Liu

Léger, D. S., Storch, N. (2009). Learners' perceptions and attitudes: Implications for willingness to communicate in an L2 classroom. *System, 37*(2), 269–285. https://doi.org/10.1016/j.system.2009.01.001

Leiner, D. J. (2019). *SoSciSurvey*, version 3.1.06. [Computer software]. Available at https://www.soscisurvey.de

Lenkaitis, C. A. (2020). Teacher candidate reflection: Benefits of using a synchronous computer-mediated communication-based virtual exchange. *Teaching and Teacher Education*, 92, 10304. https://doi.org/10.1016/j.tate.2020.103041

Liddell, T. M., & Kruschke, J. K. (2018). Analysing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology, 79*, 328-348. https://doi.org/10.1016/j.jesp.2018.08.009

Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning, 65*(S1), 185–207. https://doi.org/10.1111/lang.12117

Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching, 51*(3), 285–329. https://doi.org/10.1017/S0261444818000125

Lyster, R., & Sato, M. (2013). Skill acquisition theory and the role of practice in L2 development. In M. García Mayo, J. Gutierrez-Mangado & M. Martínez Adrían (Eds.), *Contemporary approaches to second language acquisition* (pp. 71–92). John Benjamins.

MacIntyre, P. D. (2007). Willingness to communicate in the second language: Understanding the decision to speak as a volitional process. *The Modern Language Journal, 91*(4), 564–576. https://doi.org/10.1111/j.1540-4781.2007.00623.x

MacIntyre, P. D., Noels, K. A., & Clément, R. (1997). Biases in self-ratings of second language proficiency: The role of language anxiety. *Language Learning, 47*(2), 265–287. https://doi.org/10.1111/0023-8333.81997008

MacIntyre, P. D., Clément, R., Dörnyei, Z., Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: a situational model of L2 confidence and affiliation. *The Modern Language Journal, 82*(4), 545–562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x

MacIntyre, P. D., Baker, S. C., Clément, R., & Conrod, S. (2001). Willingness to communicate, social support, and language learning. Orientations of immersion students. *Studies in Second. Language Acquisition, 23*(3), 369–388. https://doi.org/https://doi.org/10.1017/S0272263101003035

MacIntyre, P. D., Baker, S. C., Clément, R., & Donovan, L. A., (2003). Talking in order to learn: willingness to communicate and intensive language programs. *Canadian Modern Language Review, 59*(4), 589–607. https://doi.org/10.3138/cmlr.59.4.589

Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition, 21*(4), 557–587. https://doi.org/10.1017/S0272263199004027

Mayring, P. (2022). *Qualitative Inhaltsanalyse. Grundlagen und Techniken (*13th rev. ed.). Beltz.

Moratinos-Johnston, S., Juan-Garau, M., & Salazar-Noguera, J. (2019). The effects of English-medium instruction in higher education on students' perceived level and self-confidence in EFL. In C. Pérez Vidal, S. López-Serrano, J. Ament & D. J. Thomas-Wilhelm (Eds.), *Learning context effects: Study abroad, formal instruction and international immersion classrooms* (pp. 75–100). Language Science Press.

Mulyono, H., & Saskia, R. (2021). Affective variables contributing to Indonesian EFL students' willingness to communicate within face-to-face and digital environments. *Cogent Education*, 8, 1911282. https://doi.org/10.1080/2331186X.2021.1911282

Noels, K. A., Pon, G., & Clement, R. (1996). Language, identity, and adjustment: The role of linguistic self-confidence in the acculturation process. *Journal of Language and Social Psychology, 15*(3), 246–264. https://doi.org/10.1177/0261927X960153003

Olsson, U. (2022). Power properties of ordinal regression models for Likert type data. *Practical Assessment, Research, and Evaluation, 27*(6). https://doi.org/10.7275/er28-5992. Available at: https://scholarworks.umass.edu/pare/vol27/iss1/6.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. Methodological review article. *Language Learning, 64*(4), 878–912. doi: https://doi.org/10.1111/lang.12079

Qui, X., & Bui, G. (2022). Pre-task planning effects on learner engagement in face-to-face and synchronous computer-mediated communication. *Language Teaching Research, 0*(0), 1–20. https://doi.org/10.1177/13621688221135280

Rädiker, S., & Kuckartz, U. (2019). Analyse qualitativer Daten mit MAXQDA: Text, Audio und Video. Springer VS.

Riazi, A. M., & Candlin, C. N. (2014). Mixed-methods research in language teaching and learning: Opportunities, issues, and challenges. *Language Teaching, 47*(2), 135–173. https://doi.org/10.1017/S0261444813000505

Ryan, R. M., & Deci, E. L. (2009). Promoting self-determined school engagement: Motivation, learning, and well-being. In K. R. Wenzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 171–195). Routledge/Taylor & Francis Group.

Sanchez-Meca, J., Chacon-Moscoso, S., & Marin-Martinez, F. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods, 8*(4), 448–467. https://doi.org/10.1037/1082-989X.8.4.448

Seyyedrezaei, S. H., & Ziafar, M. (2014). The relationship between computer mediated communication (CMC) and willingness to communicate (WTC). *Journal of Language Sciences & Linguistics, 2*(1), 21–26. https://www.researchgate.net/publication/277274263_The_Relationship_between_Computer_Mediated_Communication_CMC_and_Willingness_to_Communicate_WTC

Shih, Y. C. (2014). Communication strategies in a multimodal virtual communication context. *System*, 42, 34–47. https://doi.org/10.1016/j.system.2013.10.016

Skidmore, M. (2023). Effects of participation in an online intercultural exchange on drivers of L2 learning motivation. *Language Teaching Research, 0*(0), 1–19. https://doi.org/10.1177/13621688231153622

Smit, N., van Dijk, M., de Bot, K., & Lowie, W. (2021). The complex dynamics of adaptive teaching: observing teacher-student interaction in the language classroom. *International Review of Applied Linguistics in Language Teaching, 60*(1), 23–40. https://doi.org/10.1515/iral-2021-0023

Sonderen, E. v., Sanderman, R., & Coyne, J. C. (2013). Ineffectiveness of reverse wording of questionnaire items: Let's learn from cows in the rain. *PLoS One, 8*(7). e68967. https://doi.org/10.1371/journal.pone.0068967

Spratt, M. (2017). CLIL teachers and their language. *Research Papers in Language teaching and Learning, 8*(1), 44–61.

Taylor J. E., Rousselet, G. A., Scheepers, C., & Sereno, S. C. (2023). Rating norms should be calculated from cumulative link mixed effects models. *Behavior Research Methods.* Epub ahead of print. https://doi.org/10.3758/s13428-022-01814-7

Topalov, J., Knezevic, L., & Halupka-Resetar, S. (2022). What it takes to communicate: comparing EFL students' willingness to communicate in traditional and online classroom. Annual Review of the Faculty of Philosophy XLVII(2), 41–57. https://doi.org/10.19090/gff.2022.2.41-57

van Batenburg, E. S. L., Oostdam, R. J., van Gelderen, A. J. S., & de Jong, N. H. (2018). Measuring L2 speakers' interactional ability using interactive speech tasks. *Language Testing, 35*(1), 75–100. https://doi.org/10.1177/0265532216679452

VERBI Software. (2021). MAXQDA 2022 [computer software]. Available from maxqda.com.

Weinert, F. E. (Ed.). (2001). *Leistungsmessungen in Schulen*. Beltz Pädagogik.

Wigham, C. R., & Chanier, T. (2015). Interactions between text chat and audio modalities for L2 communication and feedback in the synthetic world second life. *Computer Assisted Language Learning, 28*(3), 260–283. https://doi.org/10.1080/09588221.2013.851702

Yanguas, I. (2010). Oral computer-mediated interaction between L2 learners: It's about time! *Language Learning & Technology, 14*(3), 72–93. https://doi.org/10125/44227

Yanguas, I. (2012). Task-based oral computer-mediated communication and L2 vocabulary acquisition. *CALICO Journal, 29*(3), 507–531. https://doi.org/10.11139/cj.29.3.507-531

York, J., Shibata, K., Tokutake, H., & Nakayama, H. (2020). Effect of SCMC on foreign language anxiety and learning experience: A comparison of voice, video, and VR-based oral interaction. *ReCALL, 33*(1), 49–70. https://doi.org/10.1017/S0958344020000154

Yu, L. T. (2022). The effect of videoconferencing on second language learning: A meta-analysis. *Behavioral Sciences, 12*(169), 1–13. https://doi.org/10.3390/bs12060169

Zarrinabadi, N., Lou, N. M., & Shirzad, M. (2021). Autonomy support predicts language mindsets: Implications for developing communicative competence and willingness to communicate in EFL classrooms. *Learning and Individual Differences, 86*(6). https://doi.org/10.1016/j.lindif.2021.101981

Ziegler, N. (2016). Synchronous computer-mediated communication and interaction. *Studies in Second Language Acquisition, 38*(3), 553–586. doi:10.1017/S027226311500025X

Zoom Video Communications Inc. (2022). *Zoom* [Computer software]. Available at https://zoom.us/

Zou, C., Li, P., & Jin, L. (2021). Online college English education in Wuhan against the COVID-19 pandemic: Student and teacher readiness, challenges, and implications. *PLoS ONE, 1*(16), e0258137. doi: 10.1371/journal.pone.0258137