5-12-2024

Investigating Language Usage in the Language Classroom: A Comparative Analysis of Operationalization Methods and Implications for Educational Research

Manal Bani Humayyim
Syracuse University

Abstract

Multilingual approaches in second language (SL) classrooms, where learners use first language (L1) or any other known language beside the target language (TL), is an issue of ongoing research. While various research methods like word analysis, time analysis, and classroom observation exist, a standardized approach to measuring language use is lacking. Examining the same data set, this study analyzed the frequency of L1 English and TL Arabic use in two language classes through three methods: word count, time analysis, and impressionistic judgments from live observations and made comparisons of results generated by three methods as well as the feasibility. Focusing only on whole-class discussions, video recordings were transcribed, then words were counted in each language. Time was stamped every 5 seconds using Mangold INTERACT software, and observation data from Brown (2023) were analyzed to calculate majority/minority language use. The labor time spent using each method was also recorded for a feasibility analysis. Results showed that each method captured different levels of language use, but all indicated Arabic as the predominant language and that the method of counting words was the most time-consuming method. This study provides a valuable tool for researchers, encouraging their consideration of different language measurement methods for robust comparisons and meta-analyses such that the results of their research have optimal outcomes for future pedagogy.

Investigating Language Usage in the Language Classroom: A Comparative Analysis of Operationalization Methods and Implications for Educational Research

by

Manal Bani Humayyim
B.A., Najran University, 2011

Thesis

Submitted in partial fulfillment of the requirements for the degree of
Master of Arts in Linguistic Studies

Syracuse University
May 2024

Acknowledgment

I am especially grateful for Dr. Amanda Brown for allowing me to explore the depths of her data. Her unwavering support, insightful guidance, and the immense amount of time she has generously given have been invaluable throughout this journey. No single language, past or present, could ever fully express my appreciation.

My gratitude extends to the thesis committee members: Dr. Rania Habib, Dr. M. Emma Ticio Quesada, and Dr. Adam Singerman for their support, patience, and insightful feedback.

I would also like to express my gratitude to the Ministry of Education in Saudi Arabia for their full support and the resources that empowered me to achieve my goals.

Finally, my heartfelt appreciation goes to my family. Despite the distance, my parents, Saman and Fatimah, provided unwavering support, love and wisdom. Mohammed was a beacon of light during challenging times. My thanks also extend to my wonderful friends and colleagues for their constant encouragement and support.

Thank you all, without you this thesis would not have been possible.

Table of Contents

CHAPTER 1: INTRODUCTION 1

CHAPTER 2: LITERATURE REVIEW 4

2.1. A HISTORICAL PERSPECTIVE ON THE ROLE OF THE NON-TARGET LANGUAGE IN THE CLASSROOM 4

2.2. THE MONOLINGUAL AND THE MULTILINGUAL TEACHING APPROACHES 7

2.3. OPERATIONALIZING MEASUREMENTS OF LANGUAGE USAGE IN LANGUAGE CLASSROOMS 9

2.3.1. Word Frequency Method 9

2.3.2. Time-Analysis Method 15

2.3.3. Observational and Impressionistic Judgement Analysis Method 18

CHAPTER 3: CURRENT STUDY 22

CHAPTER 4: METHODOLOGY 25

4.1. PEDAGOGICAL CONTEXT 25

4.2. PARTICIPANTS 29

4.3. FREQUENCY ANALYSIS OF LANGUAGE USE IN THE CLASSROOM 29

4.3.1. Words Analysis 30

4.3.2. Time-Analysis 34

4.3.3. Classroom Observations and Impressionistic Judgement Analysis 38

4.3.4. Intra-coder Reliability 40

4.4. FEASIBILITY ANALYSIS OF MEASUREMENTS OF CLASSROOM LANGUAGE 41

CHAPTER 5: RESULTS 42

5.1. FREQUENCY OF LANGUAGE USE IN THE CLASSROOM 42

5.1.1. Word Counts 42

5.1.2. Time-Stamping 45

5.1.3. Classroom Observations and Impressionistic Judgement Analysis. 52

5.2. FEASIBILITY OF MEASUREMENTS OF CLASSROOM LANGUAGE USE 53

5.2.1. Feasibility of Word Count Measurement 53

5.2.2. Feasibility of Time-Count Measurements 54

5.2.3. Feasibility of Classroom Observations and Impressionistic Judgement Measurements 55

CHAPTER 6: DISCUSSION 56

6.1. Research Question 1: Frequencies of Classroom Language Use as Captured by Three Methods 56

6.2. Research Question 2: Feasibility of Three Methods of Capturing Frequencies of Classroom Language Use 61

6.3. The Impact of Language Teaching Approaches on the Methods of Measuring Language Use 62

6.4. Additional Considerations in the Application of the Three Methods of Measuring Language Use in L2 Classrooms 63

6.5. Limitations and Recommendations for Future Research 68

6.6. Conclusion 71

REFERENCE 73

VITA 81

List of Tables

Table 4.1 Topics Covered in the Arabic Course 28
Table 4.2 An Overview of Language Categorization..... 33
Table 4.3 An Overview of Codes for Analyzing Teacher's Talk..... 36
Table 4.4 Examples of Short and Long Pauses in Learners' Talk 37
Table 4.5 An Overview of Codes for Analyzing Learners' Talk 38
Table 4.6 Results of ICC Calculation in SPSS Using 2-Way Random-Effects Model for Word Counts 40
Table 4.7 Results of ICC Calculation in SPSS Using 2-Way Random-Effects Model for Time-Analysis..... 41
Table 5.1 Duration of Speaking Time in Weeks 7 and 8 45
Table 5.2 Temporal Analysis: Language Distribution by Teacher in Week 7..... 46
Table 5.3 Temporal Analysis: Language Distribution by Teacher in Week 8..... 47
Table 5.4 Temporal Analysis: Language Distribution by Learners in Week 7..... 48
Table 5.5 Temporal Analysis: Language Distribution by Learners in Week 8..... 48
Table 5.6 Impressionistic Judgments of Teacher and Learners' Language Use in Weeks 7 & 8.. 52
Table 6.1 Summary of the Frequency of Teacher Language Usage Resulting from Three Methodologies..... 57
Table 6.2 Summary of the Frequency of Learners' Language Usage Resulting from Three Methodologies..... 58
Table 6.3 Summary of the Feasibility for the Three Methodologies. 62
Table 6.4 Excerpt Illustrating Lengthening Sounds from Classroom Discussions..... 64
Table 6.5 Excerpts for Verbal Fillers and Short Pauses from Classroom Discussions..... 65
Table 6.6 Excerpt for a Long Pause from Classroom Discussions. 66
Table 6.7 Excerpt for Arabic Morphemes in Classroom Discussions 67

List of Figures

Figure 5.1 Frequency Analysis: Distribution of Language Use by the Teacher and Learners in Week 7 44
Figure 5.2 Frequency Analysis: Distribution of Language Use by the Teacher and Learners in Week 8 44
Figure 5.3 Analysis of Teacher's Brief Language Use in Week 7 50
Figure 5.4 Analysis of Teacher's Brief Language Use in Week 8 50
Figure 5.5 Analysis of Learners' Brief Language Use in Week 7 51
Figure 5.6 Analysis of Learners' Brief Language Use in Week 8 51

Chapter 1: Introduction

A growing body of research has investigated the origins of teaching a second language (L2) with solely the target language (TL) and has begun to reconsider the role of the first language (L1) in L2 classrooms (Hall & Cook, 2012). Following a period of discouraged L1 use and being ignored in Communicative Language Teaching (CLT) for the last two decades, the L1 has been shown to be useful in L2 classrooms, to have valuable functions, and to be utilized as a pedagogical resource (Ellis & Shintani, 2014). Thus, the question of L1 use in the L2 classroom has re-emerged as a critical issue for research on SL pedagogy.

From the perspective of professional practice, supporters of a monolingual approach emphasize following the recommendation by the American Council on the Teaching of Foreign Languages (ACTFL) advocating for maximal TL use, at 90% or above, in foreign language classrooms where the TL is not the language of the wider community. Supporters also stress on the importance of developing monolingual teaching practices and modifying TL input when necessary (LeLoup et al., 2013). On the other hand, from a theoretical perspective, Cook (2016a) emphasizes the essential role of the L1 in students' learning process, asserting that second language learners are multi-competent speakers of the L1 and L2, not unsuccessful native speakers of the L2. In addition, he draws a distinction between monolingual and bilingual perspectives in the context of language learning. One perspective evaluates L2 learning by comparing it solely to monolingual speaker proficiency, whereas the other perspective, the bilingual perspective, assesses L2 learners as learners who speak two or more languages, all of which make up their total multi-competent language system. Supporting the bilingual perspective, advocates of the multilingual L2+L1 approach recommend changing the reality of

the classroom from one dominated by native monolingual speakers' mindset to one that embraces bilingual practices, such as codeswitching, as the classroom is an environment for emerging bilinguals (Cook, 2001; Levine, 2011).

In order to engage with the above debate, several studies have measured the amount of L1 used in L2 or FL classrooms by audio/video-recording classes and counting both L1 and L2 words or utterances (e.g., Dicamilla & Anton, 2012; Storch & Wigglesworth, 2003; Thompson & Harrison, 2014; Yussof & Sun, 2020; Zhang, 2021). Other studies have analyzed the time spent in each language (e.g., Brevik & Rindal, 2020; Edstrom, 2006; Macaro, 2001). Yet other studies have estimated language use in language classes through observing teachers' instruction and classroom interactions (e.g., Brown, 2023; Kantzou & Vasileiadi, 2021; Tammenga-Helmantel et al., 2022). These methods have been applied in two settings: foreign language (FL) environment, which, according to Klein (1986), refer to learners studying a target language not commonly spoken in their everyday lives, or in a second language[1] (SL) environment, which involve learners who use the target language daily as their primary means of communication within their communities.

By employing the three methods of measuring classroom language use, researchers have sought to gather data that sheds light on important research questions. These question include proportion of L1 and TL use in L2 classrooms (e.g., Duff & Polio, 1990; Edstrom, 2006; Izquierdo et al., 2016; Zainil & Arsyad, 2021), the existence and amount of pedagogical codeswitching (e.g., Brevik & Rindal, 2020; Levine 2014; Ma, 2019; Vold & Brkan, 2020), roles

[1] Note the dual use of "second language" here, in some contexts referring to the learning of a new language (L2) and in some contexts emphasizing the wider environment of learning (SL).

that languages play in classroom interactions (e.g., Chavez, 2016), and the languages used within form-focused utterances (e.g., Nakatsukasa & Loewen, 2015).

While it is important to propose recommendations for improved teaching practices grounded in empirical evidence, the empirical evidence must be weighed on a common foundation. Thus, a critical initial step is to establish a methodological framework for capturing proportions of language use in the classroom (Macaro, 2022; Vold & Brkan, 2020). Moreover, motivated by the hypothesis proposed by Macaro (2022) regarding possible differences in percentages of language use when employing different methods, this study aims to contribute methodologically to the existing research base by being the first study to apply three different methodologies – word analysis, time analysis and classroom observations with impressionistic judgements – to the same dataset and compare the results of measurements of language use by both teachers and learners. In addition to assessing the methods in terms of the measurement results obtained, we will also assess the feasibility of each method so as to provide researchers with essential information on accuracy and practicability in the crucial area of research methodology for language use in L2 classrooms and language pedagogy.

Chapter 2: Literature Review

2.1. A Historical Perspective on the Role of the Non-target Language in the Classroom

Language teaching has undergone significant transformation over time. Changes and improvements have largely been from two perspectives: the purpose of learning a language and how the first language is viewed within each method (Celce-Murcia, 2014). Views of the TL i.e., the language the learner is attempting to learn, versus the non-target language—which is often the native or L1—and how to use it in the classroom has also been inconsistent.

By the end of the nineteenth century, an alternative to the classic Grammar-Translation method, which focused on analyzing and understanding the grammar of the target language rather than using it to communicate orally, gained popularity in America and later in Europe. The Direct Method by Maximilian Berlitz, though not of his original invention, promoted a conversational learning experience within educational institutions. In this method, only native (and monolingual) speakers of the target language would teach students, eliminating any chance of resorting to their native tongue. In addition, any common language, if present, was strictly prohibited for instructional purposes (Howatt, 2004).

The theory behind the Direct Method was that learning a second language is similar to that of a first language; therefore, a conversational approach and absence of translation are key to acquire a second language. Hence, the Direct Method had only one objective: conveying meaning directly in TL, as its name implies. Since the L1 was not allowed, the teacher was instructed to use visual aids and gestures to help convey the meaning in the TL (Larsen-Freeman & Anderson, 2011).

During the 1940s, another teaching method with a different theory and goals emerged. The Audiolingual Method became popular after World War II, especially in the 1960s. It was based on the belief that language consists of patterns and structures, and learning a language rested on 'habit-formation'. Thus, the assumption was that learners would acquire the language through repetition drills and exercises to develop correct habits in the TL. Nonetheless, it shared a similar perspective on the L1 with the Direct Method in that any incorporation of the L1 was prohibited by both teachers and learners within the classroom setting (Cook, 2016b).

However, Selinker (1972) questioned the view of second language learning as a process of habit formation, noting that after experiencing Audiolingualism, learners were able to use the language successfully inside the classroom but failed to communicate effectively outside the classroom. He stated that language is more than mere habits, and mistakes are not indicative of learners failing to acquire a habit; instead, they are evidence of the learner's interlanguage (IL). According to Selinker, IL is a linguistic system that develops as learners of a second language naturally convey meaning using the language they are currently acquiring. The IL is shaped by L1, strategies of second-language learning, strategies of second-language communication, and overgeneralization of TL linguistic materials.

By the late twentieth century, a shift to a new teaching approach – Communicative Language Teaching (CLT)—emerged as a response to perceived limitations in the former language teaching methods as well as learners' need for functional language skills targeted towards social interactions (Duff, 2014). CLT was based on the concept of communicative competence by Hymes (1972), and the theory behind CLT was that the primary function of language use is communication, and the goal of the method was to develop communicative competence. Canale and Swain (1980) outlined three kinds of competence:

1. Grammatical competence: which means the ability to utter and interpret sentences of a language accurately.
2. Sociolinguistic competence: the ability to produce suitable expressions within a social context.
3. Strategic competence: the ability to use strategies and repair breakdowns in a conversation.

These three competencies were argued to represent the learner's ability to use the language effectively, whether in an educational context or not. Canale and Swain argued that there is a lack of empirical justification to support the belief that having a strong grasp of grammar is more or less sufficient for effective communication compared to having sociolinguistic understanding or strategic skills. Later, Canale (1983) developed the theory of components of communicative competence and added a fourth competency: discourse competence, which reflects the ability to integrate meanings and grammatical structures to create cohesion in a spoken or written communication.

CLT, the currently prevailing method in teaching a second or foreign language, at least in North America, has developed a number of misconceptions, among which concerns the use of L1 in the classroom (Spada, 2007). A large body of literature on CLT often overlooks or gives minimal attention to the role of the L1. It also implies that the teacher will primarily employ the L2 and that students' use of L1 is problematic within the CLT framework. With no overt statement about L1 use, educators have understandably interpreted that L1 has little pedagogical function and maximal L2 instruction is more effective (Ellis & Shintani, 2014).

It is important to note that the literature related to CLT has mainly been focused on English as a second or foreign language. Thompson and Harrison (2014) state that while

language teaching practices are generalizable, there is also a need to examine how practitioners implement CLT when teaching Languages Other Than English (LOTE).

2.2. The Monolingual and the Multilingual Teaching Approaches

We move beyond the history of language teaching methods over the past century explicitly to deal with the use of languages in the SL or FL classrooms. Cook (2016b) defines monolingual teaching as an approach that prohibits second or foreign language learners and teachers from employing their L1 or other proficient language skills to enhance their learning of L2. There are, as Hall and Cook (2012) state, other terms for the monolingual pedagogy such as the immersive or L2-only approach. Monolingual Teaching promotes using the TL entirely while teaching learners, and that a successful language learning-process involves extensive use of the TL.

Language teaching pedagogy has often neglected or suppressed the use of multilingual methods, advocating for a primarily monolingual approach that prioritizes exclusive or near-exclusive use of the TL for better learning outcomes (Inbar-Lourie, 2010). The belief in the benefits of immersive TL teaching is pervasive throughout the profession. For example, The American Council on the Teaching of Foreign Languages (ACTFL) advise that teachers instruct, engage and provide feedback in the TL to cultivate students' linguistic and cultural competence: "ACTFL therefore recommends that language educators and their students use the target language as exclusively as possible (90% plus) at all levels of instruction during instructional time and, when feasible, beyond the classroom" (American Council on the Teaching of Foreign Languages [ACTFL], 2023, para. 1).

In contrast to the monolingual approach, other educators advocate for a more inclusive approach. Levine (2011) supported the multilingual approach, which integrates students' existing linguistic knowledge to enhance their learning of the TL. This approach also involves classroom code-switching, defined by Lin (2017) as the use of two or more languages in the classroom setting, and it involves both the inter-sentential (switching between languages within sentences) and the intra-sentential (switching languages within a single sentence), a common practice in educational research. Cook (2001) was among the pioneers in acknowledging the pedagogical value of L1 within language instruction. Cook emphasized the distinct processes of acquiring L1 and learning TL, and the integral role L1 plays in a student's learning process. He also discussed that learners of second languages are multi-competent speakers, not failed learners for not meeting the level of native speakers. Therefore, there is a distinction between being a native speaker of a specific language and becoming a successful L2 user.

A modest amount of research has identified reasons behind using L1 or a non-target language in the classroom through observing and analyzing classroom interactions in the last decades. The empirical studies examining the functions of L1 in L2 classrooms revealed a general consensus regarding its functions in language classrooms (e.g, de la Campa & Nassaji, 2009; DiCamilla & Antón, 2012; Edstrom, 2006; Ma, 2019; Polio & Duff, 1994; Storch & Wigglesworth, 2003; Temesgen & Hailu, 2022; Thompson & Harrison, 2014; Tsagari & Giannikas, 2020; Üstünel & Seedhouse, 2005; Zhang, 2021). Macaro (2022) stated that the roles of L1 in L2 classrooms exhibit global similarities across various languages, some of L1 roles are: contrasting forms between the languages, translating, and explaining some concepts in L1 and management.

Given recent empirical studies yielding some mixed findings in this area, with some finding a positive impact of using L1 (e.g., De la Fuente & Goldenberg, 2020; Ma, 2019; Zhang, 2021) or not negatively or positively impacting learning outcomes (Brown & Lally, 2019), the debate within language pedagogy regarding whether to adopt a monolingual or a multilingual approach remains vibrant and ongoing. However, an underlying methodological issue of critical foundational importance has received very little explicit discussion thus far: namely, the existence and efficacy of different systems for measuring and assessing language use in the second and foreign language classroom.

2.3. Operationalizing Measurements of Language Usage in Language Classrooms

Several research studies have examined the nature of language classrooms to capture the actual practices of language usage. To quantify proportions of languages being used, researchers usually apply one or more of the three following methodologies: word frequency analysis, time analysis, or classroom observations and analysis of impressionistic judgements.

2.3.1. Word Frequency Method

The method of analysing word frequency has included counting each word or counting utterances in the L1 and TL. As part of the analyses, researchers have calculated words per minute, provided descriptive statistics on proportions of L1 versus TL use and proportions of codeswitching, and investigated the languages used within form-focused utterances, defined below.

A study conducted by Nakatsukasa and Loewen (2015) involved a total of 12 hours of recorded Spanish classes, to an approximately low-intermediate proficiency level, in order to

investigate the use of L1 English and TL Spanish within Focus on Form Episodes (FFEs). FFEs are defined as classroom episodes that direct learners' attention to linguistic aspects of the language in vocabulary, semantics, or grammatical structures during communicative-based activities. The study included 23 participants in their second year, second semester at a US university, with intensive classes conducted for 120 minutes per class four times a week over a six-week period in the summer. Recording from weeks 2 to 4 were transcribed and only the teacher's talk was coded. The coding system involved segmenting the teacher's talk into sentences and then coding each sentence as entirely in L1, entirely in TL, an equal mix of both languages, or unknown for unintelligible utterances. If an utterance was in L1 English but included some TL Spanish words, it was coded as entirely L1, and vice versa. Also, the teacher's incomplete utterances were included in counting. Then, an FFE was identified in the transcriptions whenever a specific language feature was emphasized in the conversation, and the FFE ended when the discussion shifted back to a general communicative topic. After that, all identified FFEs were sorted based on four categories—grammar, semantics, vocabulary and other if it was about pronunciation or pragmatics— and then the percentages of language use in each category were calculated. The results revealed that the total number of utterances of teacher talk was N=1707 in which TL Spanish comprised 47.6% of utterances, L1 English comprised 39.7%, while codeswitching accounted for 11% of utterances and a further 1.8% of utterances were coded as unknown language use. In addition, the teacher produced N=457 FFEs, with 47.7% of them concerning grammar, followed by 36.1% for vocabulary. Semantics accounted for 11.5% of FFEs while 5.7% addressed other functions such as pronunciation and pragmatics. Regarding teacher's language use in FFEs, 45% of the grammar FFEs occurred in L1 English and 44% were in TL Spanish, showing almost equal usage between the languages. However,

when FFEs targeted vocabulary, 60% of FFE were in TL and 26% were in L1. These descriptive statistics indicated considerable L1 use for FFEs relating to grammar and considerable TL use for FFEs relating to vocabulary. The chi-square analysis revealed a statistically significant association between the category of FFE and teacher's choice of language with a p-value less than 0.001 ($p < .001$) and Cramer's V effect size calculated at 0.1225, which indicate a moderate relationship between the teacher's choice of language and the linguistic category of the FFEs.

Similarly, Izquierdo et al. (2016) conducted a longitudinal design study that examined features of English instruction by nine Mexican teachers at five distinct secondary schools in Southeast Mexico. By recording 50-minute classes every second week—which resulted in five classes per teacher— in the third term, the study collected 45 hours of instructional data. These recordings were transcribed, and then teachers' utterances were segmented into independent semantic clauses, each of which was subsequently coded as either: 1) exclusively in L2 or L1, 2) involving code-switching, or 3) involving translation, which is different than category 2 as the same semantic meaning was expressed first in one language then immediately in the other. Then, the teachers' utterances were analyzed and categorized for their purpose according to one of the three main categories as follows: 1) Management, which included two subcategories: procedural or disciplinary utterances within a lesson, 2) Language, which comprised four subcategories: structure, function, discourse, and sociolinguistic aspects of language, 3) Other topics, which included two subcategories: narrow topics that focused on engaging learners in discussions about their immediate environment, or broad topics that were about broader issues beyond their immediate context. The findings revealed that overall, teachers produced a total of N=15768 utterances in which N = 7522 were in L1 Spanish, and that surpassed the utilization of L2 English (N = 6427). This was followed by translation (N = 1112) and finally codeswitched

utterances (N= 707). These results indicate that all teachers provided immediate translations from L1 to L2 at levels that surpassed their engagement in codeswitching practices. There was also variability in L1 usage among the teachers in which seven of them heavily relied on L1 Spanish use in their utterances, with percentages ranging from 27% to 71% of utterances, while only two teachers used L1 at 16% and 18% of their talk. The data showed that Spanish was frequently used as the primary language for conveying meaning in teachers' discourse, rather than codeswitching or using L2 only, especially in discussions on narrow topics and delivering class procedures, with 55.3%[2] and 56.4% of utterances being in L1 respectively. However, when the goal was to draw students' attention on form of the TL (i.e., English), teachers used English at 64.9% of utterances which was nearly four times more than Spanish (16%). The authors did not provide details about excluded utterances.

An exploratory study by Chavez (2016) conducted with three fluent, CLT-trained and experienced teaching assistants (TAs) teaching German as foreign language to 61 sophomores in a US university aimed to measure language use by TAs and learners during whole class interaction and in peer+group activities. The TAs in this case shared the same language as the students, and the use of L1 was not prohibited; instead, TAs and learners were encouraged to use it if it facilitated a more effective or advanced use of L2. In total, 21 classes were videotaped over a 14-week period, but only nine classes were selected, with three classes selected per TA. The nine selected classes were then transcribed in order to count occurrences of words in each language by TAs and students. Non-verbal fillers and incomplete words were not counted as well as TA contributions during peer-work. Also, some English words were produced by learners

[2] Raw numbers rather than percentages were included in the original article; thus, the percentages noted here were calculated from the published raw numbers.

with German-like pronunciation, such words were categorized as L1, except for cognates, which were classified based on the language they were intended to represent. Findings showed that the TAs varied in their L1 use during teacher led-talk, with TA 1 employing L1 at 52.23% of total words, which was more than her students who used L1 at 45.44% of students' total words. TA 2 and her students used L1 almost in similar percentages, with the former at 21.63% of total words and the latter at 22.39%. TA 3 had an L1 usage at 8.45% of total words while his students accounted for 17.90%. In addition, an analysis of language use revealed that learners in all three classes spoke more L2 German during peer activities than in whole-class discussions.

Zainil and Arsyad (2021) conducted a comparative analysis between teachers' perceptions of codeswitching in teaching English language and their actual implementation of codeswitching in the classroom by analyzing a video corpus and using stimulated recall interviews. The study involved five teachers across four different junior high schools in Indonesia, and the number of students was between 25 to 30 in each class. Each class lasted 80 minutes, totaling 160 minutes per week and recorded classes included 20 English classes throughout the second semester, resulting in 25hrs20mins of recorded material. Teachers' talk was transcribed, and words were coded as either L1 Bahasa Indonesia or L2 English. Teachers were also asked to estimate in percentages the amount of TL they used. The results indicated that teachers frequently employed at least two languages—specifically Bahasa Indonesia and English—with occasional incorporation of the local language Bahasa Minang. While all teachers engaged in codeswitching practices during instruction, only one teacher consistently used the TL over L1 Bahasa Indonesia, using it in nearly 60% of the words he spoke, whereas another teacher used Bahasa Indonesia and TL English almost equally. The remaining three teachers utilized more Bahasa Indonesia than TL with rates of TL at 36.07% or lower of the words they spoke. Following the

recordings, stimulated recall interviews were conducted to allow teachers to describe their own practices, perceptions, and reasons for codeswitching in their classrooms. It was interesting that teachers' beliefs did not align consistently with their actual practices, leading some to their surprise upon discovering the extent of their own L2 English language usage.

Additional studies aimed at uncovering language usage in educational settings have adopted similar methodologies of counting words from whole-class transcriptions (e.g., Aull, 2021; de la Campa & Nassaji, 2009; Kawafha & Al Masaeed, 2023; Miri et al., 2017; Thompson & Harrison, 2014; Zhou & Li, 2022), including in pair activities (e.g., Dicamilla & Anton, 2012; Storch & Aldosari, 2010; Zhang, 2021), or the coding of sentences and utterances in place of words (e.g., Storch & Wigglesworth, 2003; Yussof & Sun, 2020).

Despite the different approaches in counting words or utterances, prior studies lack details on what was counted or what was excluded in the counting process. It is highly likely that mispronounced or repeated words, proper names, borrowed terms, and cognates will occur in language classes. However, studies have not provided descriptions on how to handle the inclusion or exclusion of these instances during counting (e.g. Aull, 2021, Dicamilla & Anton, 2012; de la Campa & Nassaji, 2009; Kawafha & Al Masaeed, 2023; Miri et al., 2017; Storch & Aldosari, 2010; Thompson & Harrison, 2014; Zhou & Li, 2022). In addition, coding words might be challenging with character-based languages such as Mandarin Chinese. A study by Zhou and Li (2022) offered percentages for word counts in Chinese and English, while Zhang (2021) offered the average use of Chinese and English, since it was challenging to compare words in English and Chinese due to the differences in the linguistic system of the two languages.

2.3.2. Time-Analysis Method

Among the first studies to raise teachers' awareness of how languages were being used inside the classroom, Duff and Polio (1990) used an analysis of time to observe and document language use in 13 FL classes in a range of languages, although the languages were not identified. These classes were taught by 13 individuals who were both native speakers of the L2 and fluent or bilingual in English. The language learners were in their second semester of L2 learning. Two 50-minute classes for each language were recorded at the University of California, Los Angeles. For a fine-grained analysis, every fifteen seconds, teacher utterances were coded into one of five categories: entirely in L1 or TL, in L1 but with a word in TL or vice versa or codeswitched in equal amount of L1 and TL. The results showed that percentages of L2 in teacher talk varied from 100% to 10% of teacher talking time, with only six classes following the recommended maximal L2 use (90% and above). However, more than half the classes were described as utilizing TL from 10% to 79% and some teachers showed inconsistent use of TL in the two classes.

Several other researchers have also investigated the extent of L2 usage in instructed classrooms for adult beginners. In a longitudinal assessment, Edstrom (2006) aimed to compare her actual teaching practices with her teaching beliefs as an instructor in university-level Spanish classes. She recorded 24 of her own classes throughout the entire semester of teaching Spanish 101 and wrote one entry after each class in a reflective journal where she documented her observations of her own language use after each class. Using the recordings to tally the time spent in each language, Edstrom's findings indicated that her use of L1 (English) ranged from 6% to 33% of class time across the course—except in the last two classes before the finals,

where she utilized L1 at 54% and 71% of class time respectively—which exceeded her own estimation of L1 use between 5 to 10% of class time.

In the above studies, researchers have focused only on teacher's talk without investigating whether learners used only L2, some L1, or switched codes. Vold and Brkan (2020) examined both language use by the teachers and students in the context of Third Language (L3) instruction. The Norwegian education system mandates learners to learn English as the first FL from 1st grade until upper secondary school, and other additional FLs such as French, German or Spanish are optional and introduced in lower secondary schools. The researchers documented 45 classes to investigate the utilization of L1 Norwegian and TL French by six French teachers teaching six classes with a total of 85 beginner students in grades 9 and 10, across six different lower secondary schools. Following the national policy, the teachers adopted CLT—which emphasizes interaction as both the means and the ultimate goal of learning a language—for overall 2-3 hours classes per week. Using a quantitative analysis of time, three codes were used: L1 for Norwegian, TL for French, and L1-TL when mixing the two languages. Codes were activated whenever speech commenced and ceased when speech concluded. Also, the authors differentiated language use according to both speaking time and class time which led to varied results. In regard to speaking time, which comprised 32 hours, the results showed that approximately 25 hours were in L1 (79%), 357 minutes, approximately 6 hours, in TL (19%) and 50 minutes in both (2%). However, the proportions of use of languages in relation to total classes time (39h 15mins) resulted in L1 use at 66%, TL at 16%, codeswitching at 2% of class time, along with a further 16% of class time that was silent and therefore not coded. In addition to analyses of overall language usage across all classes, individual analyses revealed that for each class, the usage of L1 surpassed that of TL, varying between 70% to 90% of speaking time by

both the teachers and learners, except for one grade 9 class, where the TL French and L1 Norwegian were used at approximately equal percentages. Furthermore, the study highlighted that TL usage was not different among grades 9 and 10 despite the fact that grade 10 might be exposed to more French, and that teachers' proficiency level did not appear to have a correlation with the use of the TL. The authors reported that frequent pauses that lasted three seconds or less were captured by the code L1-TL which is for mixing languages, but pauses that occurred within the other categories (L1 or TL) were not mentioned. The authors did not code the interaction between learners due to audio quality issues.

A similar study was conducted by Brevik and Rindal (2020) investigated the practices of EFL teachers teaching 13-15-year-old students at the CEFR-B1 intermediate English proficiency level across seven schools. Although Norwegian was the L1 for the majority of students, classes also included multilingual students who had different L1s. The data included video recordings of 60 lessons—between four to six lessons in each school—and the researchers employed a five-second sampling method to quantify language use into the following: L1 Norwegian, L2 English, a combination of both, or other (in case other languages were used) by the teachers and learners. Due to inadequate clarity of the recording equipment, some pairs' interaction was not coded. In addition, any language use lasting fewer than 3 seconds was excluded from coding. The findings indicate that classes in two schools utilized L2 at comparatively low percentages, between 40% to 60% of class time or using both languages almost equally. Classes in other five schools showed variation and different language practices, utilizing more L2 at 70% or above of class time. The study did not indicate how much time in learners' interactions was excluded as well as details about language use that was excluded (details such as total duration and language choice).

Some studies have applied the time-analysis method such as Macaro (2001) and Wolthuis et al. (2020), but perhaps due to the labor intensity of the method or the lack of supporting technologies, relatively few studies were found to have done so. In addition, comparatively few studies operationalized their time-analysis method in detail. Further, some studies only focused on L1/TL without analyzing codeswitching (e.g. Edstrom, 2006; Wolthuis et al., 2020), or did not give details about the nature and duration of excluded language use (e.g. Brevik & Rindal, 2020).

2.3.3. Observational and Impressionistic Judgement Analysis Method

Finally, the third methodology used to examine language practices in the language classroom has involved employing observations completed by trained individuals, or sometimes collected by the teachers themselves, in order to compile impressionistic judgements of classroom language uses.

A case study conducted by Tammenga-Helmantel et al., (2022) used multiple methods to assess the use of the TL by three Dutch EFL student teachers. The methods included four surveys, three live classroom observations, and a written teacher's reflection on TL use. The surveys and observations were conducted at different times during the teachers' final year and early career to provide insights into their TL use when teaching EFL classes in secondary schools, as well as the factors that influenced their use of the TL. Classroom observations—carried out by a trained observer—quantified language use in percentages based on estimations. The findings indicated that, in general, student teachers tended to use the TL at 70% or higher of speaking time in senior classes in secondary schools, whereas in junior classes, TL usage ranged between 30% and 40% of speaking time. However, the observations, which were conducted a

year after the student teachers had officially become EFL teachers in educational institutions, revealed inconsistent TL usage, where one teacher reduced TL use, another increased it, and one remained stable.

In the context of CLT, Aoyama (2020) conducted a study examining the utilization of both L1 Japanese and L2 English among advanced-level senior high school students in Japan during English language activities. The researcher observed 10 classes, each lasting 50 minutes, involving a total of 39 students who belonged to different groups. The study focused on peer and small group discussions and the observations focused on categorizing L1 use into speech functions rather than language proportions. The findings indicated that all Japanese high school students utilized their L1 and varied in L1 use during the tasks. Specifically, students commonly employed L1 as speech fillers and when they encountered difficulties in expressing unfamiliar English words or phrases immediately. Also, students expressed concepts or ideas more accurately by incorporating Japanese terms that lacked direct English equivalents, thereby enhancing their communication proficiency during English tasks. The researcher did not document language use during whole-class discussions.

Another Study that implemented class observations was by Brown (2023), where 53 learners participated in courses teaching French or Modern Standard Arabic to beginners through a CLT approach. The courses were taught over 10 weeks by a native speaker each of French or Arabic, with one 2.5-hour class taught weekly with a 15-minute break. For each language, one class was conducted with the inclusion of L1, by teachers and learners, while the other class was conducted following a L2-only policy. An observer—who was physically present in the classroom —documented language use in real time by the teacher and learners every week using a structured observation protocol. Using evidence from weekly quizzes in listening, writing, and

vocabulary learning, the results of the study showed that the non-immersion groups statistically significantly outperformed the immersion groups in both French and Arabic.

A longitudinal study by Brown et al. (2022) employed impressionistic self-report on language usage by the language instructors as well as the instructors' observations of student interactions. Two researcher-practitioners instructed 50 international undergraduate students who were all at the CEFR-B2 proficiency level in English for Academic Purposes classes over a four-month period. The teaching approach employed CLT, and both researcher-practitioners each taught two groups, with all groups experiencing both monolingual and multilingual teaching conditions. First, each group was exposed to one treatment, then reversed to the other treatment halfway through the study, with treatment order counterbalanced across groups. Scores were collected through students' writing assignments as well as mid and final grades. The assignments were assessed based on the students' demonstrated mastery of five rhetorical modes as well as grammatical and lexical accuracy. Descriptive analyses of learning outcomes showed an advantage of the multilingual approach, but inferential analyses showed no statistically significant differences in student performance between the two teaching conditions. Reflective journals by the researcher-practitioners were completed after each lesson over the four-month period and documented self-reflections on teachers' instructions and students' language use. Journals revealed instances of encouraging students to use L1, occasional L1 usage in monolingual conditions, and frequent codeswitching in the multilingual group.

While classroom observations have been used for a long time, the field has a lack of standardized tools and still relies primarily on informal and non-standardized designs (Klette, 2023). The following studies have employed classroom observations, primarily utilizing non-

standardized observation instruments created on-site with personal notes or custom categories: Kantzou and Vasileiadi (2021), Kawafha and Al Masaeed (2023), and Tekin and Garton (2020).

Chapter 3: Current Study

Investigating language use in L2 teaching practices is an issue that is critically important for applied linguistics theory and practice. Previous and current research has examined the functions and impact of monolingual, immersive target language-only instruction versus multilingual, non-immersive language instruction that involves the target language and other languages in which students have proficiency such as the L1. These studies have been operationalized using one or more of the following three methods: (1) recording classroom discourse, transcribing recordings and then counting words or utterances in each language; (2) recording classroom discourse and time-stamping intervals based on language use: typically TL only, L1 only, or an equal or unequal mix of both; (3) or employing live observation protocols on which impressionistic judgments about proportions of language use are based. In the case of studies employing method (1), a number have not included the precise details of the method followed in counting words, for example whether they included mispronounced or repeated words, and whether they included proper names, borrowed terms, and cognates (e.g. Aull, 2021, Dicamilla & Anton, 2012; de la Campa & Nassaji, 2009; Kawafha & Al Masaeed, 2023; Miri et al., 2017; Storch & Aldosari, 2010; Thompson & Harrison, 2014; Zhou & Li, 2022). Comparatively fewer studies have employed method (2), operationalizing time-analysis, and of those, some only focused on L1/TL without analyzing codeswitching (e.g. Edstrom, 2006; Wolthuis et al., 2020), or did not give details about the nature and duration of excluded language use (e.g. Brevik & Rindal, 2020). Finally, regarding method (3), it is crucial to approach estimates of classroom language use with caution since they rely on impressionistic judgements during synchronous observations or rely on memory when paired with later journaling.

Indeed, as Vold and Brkan (2020) state, this area of critical research in applied linguistics is in need of more precise descriptions of research methodology. Furthermore, Macaro (2022) hypothesized possible differences in the implementation and results of a word count method and a time-stamping method, however, stopped short of testing his hypotheses. He has also argued for necessary consideration of methods (1) and (2) above; however, he did not include method (3). Therefore, there are basic and currently open questions surrounding the efficacy and feasibility of methods (1)-(3) above. In addition, regarding characterizations of the interactions themselves, some of the previous studies have focused only on teacher talk (e.g. Duff & Polio 1990; Ed 2006), or total language use by teachers and learners (e.g. Brevik & Rindal, 2020; Vold & Brkan, 2020) without distinguishing teacher's use versus learners' use of the respective languages. Furthermore, proportions of language use have been found to vary based on factors such as the teacher's proficiency, the learner's proficiency, educational context, and teaching methods, etc. However, a significant aspect influencing these proportions may be the methodology adopted in studies to measure language use by both teachers and learners.

The aim of this study, then, is to make a methodological contribution to the existing research base. To the best of my knowledge, this will be the first study to apply three different methods of analysis of language usage in the L2 classroom to the same dataset, comparing them in terms of frequencies of language use. Furthermore, I offer insights into the feasibility of each method. Thus, the following research questions were targeted in this study:

> How do measurements of language usage by teachers and students in the L2 classroom as operationalized by (a) word count, (b) time count, (3) impressionistic judgements based on live classroom observations compare in terms of :
>
> 1. results on occurrences of language usage?

2. feasibility of method for research?

The first research question aims to examine frequency of language usage captured by three different methods: word count, time-analysis, and impressionistic judgements based on live classroom observations. In this context, frequency refers to the extent or how often each language was used as measured by each of the three methods. This analysis aims to compare the quantitative findings obtained from each of the three methods.

The second question shifts focus to the feasibility of these diverse methods of research operationalization for measuring language usage in the language classroom. It addresses logistical considerations, and explores potential challenges or advantages associated with each method and how their feasibility may relate to reliability and comprehensiveness of language use analysis.

In sum, previous studies have shed light into how L1 is used in language classrooms. This methodological study aims to provide a detailed framework for doing word and time analyses as well as a quantitative comparison of the three methods to help enhance methodological rigor in the field.

Chapter 4: Methodology

4.1. Pedagogical Context

The current study utilizes data that were collected by another researcher. This data is considered secondary data. According to Tight (2019) "Secondary data are data that have been collected by someone else (or, perhaps, by yourself some time ago, and probably for a different purpose)" (p. 164). Tight argues that secondary data has strong potential for contributing new perspectives and results, which is contrary to the prevailing notion in academic domains that researchers ought to gather or collaborate in generating their own empirical data as a fundamental aspect of the research journey. Similarly, Lombardi et al. (2023) outline several reasons and scenarios for using secondary data in new research. One of the different scenarios is when the original study yields new questions, prompting a need to delve deeper in order to address these inquiries.

This study uses recordings of previously collected classroom data for Brown's (2023) longitudinal intervention research, which aimed to compare the learning outcomes achieved from implementing two distinct pedagogies—monolingual and multilingual— in teaching the commonly and less-commonly taught languages of French and Arabic at beginning levels of proficiency in the United States. For quantifying language use, Brown (2023) utilized the methodology of observation and impressionistic judgements, and the data for language frequency were reanalyzed to serve as a comparison to the additional analyses of words and time conducted as part of this study (see below).

The pedagogical context of the courses in Brown (2023) was the teaching of French or Modern Standard Arabic to beginners in a community-based educational setting through a CLT

approach. The courses were advertised to the local community and recruited diverse members, and were taught over 10 weeks, with one 2.5-hour class per week with a 15-minute break, finishing with progress quizzes in most lessons. For the current study, we focus only on the video recordings of one group of learners in Arabic, documenting the frequency of teacher's and learners' use of Arabic (TL) versus English (L1[3]) by comparing results from the three different measures of classroom language usage. The Arabic course provided authentic listening and speaking activities while reading and writing content was modified. Due to time constraints, the nature of the community-derived program, and the need to keep the French and Arabic courses somewhat comparable for the original study, the Arabic alphabet was transliterated into the Roman alphabet.

Feasibility constraints for the current study permitted transcribing and systematically analyzing two lessons, since each class lasted 2.5-hours, comprising a total of five hours of complex multi-party interactions in a classroom context. The lessons in weeks seven and eight of the ten-week course were targeted for several reasons. First, classes in these two weeks included a total of six activities comprising transcribable whole-class discussions, while other lessons had more pair and small group work which was not transcribable. Weeks five and ten were also atypical with mid-course and final quizzes, and week six included a review of the mid-term quiz and provided feedback to learners. Further, the community-based courses in Arabic suffered considerable attrition and weeks seven and eight contained a consistent number of students. Thus, the extent of whole class discussions with a stable number of students in weeks seven and

[3] Some learners were not native English speakers and had different L1s such as Italian, Somali, Spanish and Swedish, but all were proficient speakers of English.

eight offered an opportunity to capture more of the language use as part of classroom interactions to submit to the different methods of analyzing the frequency of language use.

Second, as shown in Table 3 in Brown (2023, p.1645) about data from observation and impressionistic judgement of classroom language use, weeks seven and eight were characterized by variation in language use that would be interesting for the current study. As part of the original study, which utilized the methodology of observation and impressionistic judgements of classroom language use, learners were observed to use more L1 English than TL Arabic in week seven, while the opposite was observed for week eight. Further, in week seven the teacher was observed to instruct only in the target language, Arabic while in the remaining weeks, he employed some English in his instruction.

The Arabic language course for the beginner-level covered different topics. All class materials, designed by the teacher, eliminated the need of a textbook. Achievement quizzes were conducted at the end of every class. Table 4.1 shows the topics covered during the course:

Table 4.1 Topics Covered in the Arabic Course

Weeks	Topics
Week 1	Greeting, nationalities and occupations + quiz.
Week 2	Ordinal and cardinal numbers, days of the week, months, time, date of birth + quiz.
Week 3	Question words + quiz.
Week 4	Body parts, clothing items, and colors + quiz.
Week 5	General review (exam).
Week 6	Possessions + quiz.
Week 7	Prepositions, cardinal directions, asking about places and giving directions + quiz.
Week 8	Daily routines + quiz.
Week 9	Talking about past events and future plans + quiz.
Week 10	Course revision (grammar error game).

The students in the classroom were arranged in a U-shaped seating. All lessons were recorded simultaneously using two cameras, one focused on the teacher and the other on the learners. These cameras recorded whole-class interactions but not pair activities. An observer was present in all classes and sat at the back of the class to document language use in each segment of the class using a structured observation protocol.

The quality of the cameras and recorders was very good and captured all whole class conversations, which could be transcribed and coded to examine languages use i.e., English and Arabic. Primary reliance was on video recordings that focus on the teacher; when certain

conversations were not understood, it was important to check the student-focused video recording.

4.2. Participants

The number of learners in the target course was 19, aged between 20-58 years old, although due to participate attrition, only 12 and 11 learners were present in weeks seven and eight, respectively. Participants were recruited through community organizations and advertisements. Although some learners were not native English speakers and had different L1s, they were all proficient speakers of English; thus, English was used as a contrast to the target language, Arabic. Prior to the beginning of the course, all learners completed a self-assessment grid that contained statements based on the 2012 version of the Common European Framework of Reference (CEFR), and their level was determined to be pre-CEFR A1 (lowest level). The instructor was a native Moroccan-Arabic speaker. He was also a proficient speaker of English and French, with credentials and experience in teaching languages.

4.3. Frequency Analysis of Language Use in the Classroom

Data analysis comprised analysis of word count, time spent, and language use based on live observations from Brown's (2023) observation protocol in each language. The goal is to compare the results of these analyses. The methods of analysis are outlined below.

Before analyzing the data, the recordings were viewed multiple times to become familiar with the data. Also, the structured observation protocol followed by the observer in Brown (2023) was thoroughly examined to grasp details about the learners, seating arrangements, lesson components, and language use. The total duration of the two video recordings was 4hr49mins.

The coding for both word and time analyses began when the teacher initiated the class and stopped whenever students quietly began answering the quiz. Each lesson incorporated both pair activities and whole-class discussions, but coding could only be applied to the latter due to all pair work not being captured on the class video recorders. Nevertheless, there were instances in which the teacher intervened in pair activities, calling on learners to pay attention for follow-up explanations, resulting in learners engaging in whole-class discussions. Therefore, coding started whenever the teacher called everyone and stopped when the teacher confirmed they understood a certain point and then learners continued to do the activity in pairs.

4.3.1. Words Analysis

In order to examine the frequency of words in English and Arabic, class recordings were first transcribed following Seedhouse's (2004) transcription conventions with slight modifications. The conventions included symbols such as T for teacher speech and L for learner speech, with numerals like L1 (learner 1) and L2 (learner 2) assigned to individual learners based on their respective seating positions within the class. These conventions also included symbols indicating short and long pauses, lengthening sounds, speaker's emphasis on certain words, and when overlaps occurred and stopped. Additionally, notations were used for mispronounced words, unfinished words, and when the transcriber was unsure of a word.

Speech in Arabic was italicized, and a corresponding English translation was provided within angle brackets right after the statement. All teacher-learner interactions were transcribed, including pauses and any errors in pronunciation. Pair-work activities were excluded because the video recorders positioned at the front and back of the classroom only captured discussions involving the two closest pairs. However, when the teacher clapped to gain everyone's attention to provide further explanation on the task, learners participated as a whole class. Therefore,

transcribing these instances began when the teacher addressed the entire group and concluded when learners resumed the activity. This aligns with the segmentation of the lesson made by the observer in Brown's (2023) observation protocol, and only segments for whole class discussions were coded.

After transcription, spoken words were categorized based on language and speaker into T-Arabic, T-English, and T-Other in case other languages were used given the multilingual status of the teacher. Spoken words by the learners were categorized into Ls-Arabic, Ls-English and Ls-other since some students were multilinguals. Excel was used to categorize words spoken by the teacher and learners. Each line from the transcriptions underwent systematic classification based on the speaker and language. Subsequently, words within each line were examined in depth. If a line spoken by the teacher was solely in Arabic or English, it was added to the corresponding columns (T-Arabic or T-English). However, if a line contained words from both languages or any other language, Arabic words were extracted and placed in the T-Arabic column, English words were extracted and placed in the T-English column, and words in any other language placed in the T-Other column. After that, words in each line were tallied and the number was recorded next to the line. Finally, a sum formula was employed to calculate the total words spoken by the teacher in each language. The same process was adopted for the learners' speech.

The selection of what to count from the transcriptions and the reasons behind it posed a significant challenge. Proper names e.g., names of learners, were excluded as well as names of places and streets if they had the same pronunciation in both languages such as New York, New Jersey and Florida. However, some proper names spoken in Arabic such as *al'urdun* (Jordan) and *amriika* (America) were counted as Arabic because the pronunciation differs in both

languages. Additionally, some borrowed words such as *yoga* and *internet* were not counted since it was difficult to classify such words as either English or Arabic. Repeated words that occurred in the same utterance were counted because they were part of the teacher-learner interaction or the teacher's instructional talk.[4] Speakers' errors were also categorized since they represent attempts of production, even if there's a mispronunciation of a sound, stress on the wrong syllable, or difficulty with challenging Arabic sounds such as the pharyngeal fricative /ʕ/ or the emphatic consonants such as /tˤ/, /dˤ/ and /sˤ/. Contracted words such as "haven't" were counted as one. Finally, unfinished words, speech fillers (e.g. 'uhh', 'aha') and expressions (e.g., 'yikes', 'oh wow!') were transcribed but not counted as completed or countable words.

Table 4.2 below summarizes the criteria followed in the analysis of word counts. All the examples are from the current data set; Arabic is in italics followed by a translation in angle brackets:

[4] Repetitions were counted from all participants because they may signify linguistic challenges, cognitive processing strategies, or language input in the classroom.

Table 4.2 An Overview of Language Categorization

Category	Description
T-Arabic	Words in Arabic used by the teacher, including proper names in Arabic such as *almaghrib* <Morocco>, *al'urdun* <Jordan> and repeated words.
T-English	Words in English used by the teacher.
T-Other	Words in languages other than English and Arabic.
Ls-Arabic	Words in Arabic used by the learners, examples: ▪ Proper names in Arabic *yoluo*[5] <July>, *al'urdun* <Jordan> ▪ Pronunciation errors **nifS* < *half>, *'idhabii *kharb* < go west >, **bada daliik* < *after that> ▪ Repeated words *mataa tasta- tastayqiDiin: mina nawm?* <when do you wak- wake up from sleep?>, *qariib, qariib min hunaa* < close, close from here >
Ls-English	Words in English used by the learners, including proper names with pronunciation differing pronunciation from Arabic, e.g., Morocco, Jordan.
Ls-Other	Words in languages other than English and Arabic
Excluded Words	Words by the teacher or learners that were not counted such as: ▪ Learners' names ▪ Places and streets: I-81, Hall of Languages ▪ Proper names with similar pronunciation in both languages: New York, Florida ▪ Borrowed words from English with similar pronunciation: Internet, Yoga

[5] Although this word is borrowed from Romance languages, it was counted as Arabic as the teacher did not use a possible equivalent term in Arabic which is Tammuz, used in some Arabic speaking countries.

4.3.2. Time-Analysis

In order to examine the class time spent using the target language, Arabic, versus the first or other proficient language, English, a time analysis was conducted by applying a language code at five-second intervals, a methodology adopted from Macaro (2001), Brevil and Rindal (2020), and Vold and Brkan (2020), using the video-analysis software Mangold INTERACT (2022). Although Macaro did not discuss pauses during interactions or briefer utterances within 5-second intervals, the studies by Brevil and Rindal and Vold and Brkan adopt the same methodology for 5-second analysis but decided for the exclusion of any language use less than 3 seconds. In the current study, we continued the five-second analysis as it is both feasible and efficient for analysis compared to other timeframes (see Duff and Polio, 1990). However, we further refined the time analysis by differentiating between continuous speech and briefer utterances as it will enable a more comprehensive exploration of measuring language in whole-class discussions. In addition, given the critical focus in this study on both teacher and students talk, the time analysis was conducted twice, with the first pass focusing solely on the teacher's talk, while the second pass focused on learners' talk during whole-class instruction, aiming to calculate the time spent in English and Arabic. As noted above, each lesson incorporated both pair activities and whole-class discussions, but coding could only be applied to the latter due to all pair work not being captured on the class video recorders, which aligned with lesson segments made by the observer in the observation protocol. Nevertheless, the instances in which the teacher intervened in pair activities were coded when the class interacted as a whole. Therefore, coding started whenever the teacher called everyone and stopped when the teacher confirmed they understood a certain point and then learners continued to do the activity in pairs.

To attain precise measurements of language use in the classroom, four codes were employed for teacher-talk and four other codes for learner-talk. Codes for the teacher included the following: TA for teacher talk in Arabic, TE for teacher talk in English, TCS for the teacher switching between English and Arabic within the 5-second interval, and TB for brief language use by the teacher.

First, TA for teacher talk in Arabic, denoted instances where the entire utterance was in Arabic and the teacher continuously spoke during the 5-second interval. Moreover, some extended utterances spanned two five-second intervals or more, with the first part of the utterance falling fully in one interval but the last part of the utterance—which may have contained one word or two—falling into to the second five-second interval. One or two words would usually take one second or less. In such a scenario, the second interval, given its continuity, was also coded as TA. Second, the code TE was used for teacher talk fully in English during the 5-second interval. Usually, English sentences fell in one 5-second interval and did not extend to more than one. Third, the code TCS indicated teacher switching between English and Arabic within the 5-second interval. Finally, in cases where a five-second interval contained only a short language use that was not continued from a prior five-second interval, the code TB was assigned to indicate brief language use, signifying instances where the teacher spoke only one word in Arabic in the form of recast or echoic repetition during sharing answers with the whole class activities, or one word in English when a learner asked the teacher directly for the meaning of an Arabic word in English. These single words uttered by the teacher between learners' utterances that corresponded to a 5-second interval were coded as TB, regardless of the language used. Later, all the brief language use was examined again and classified as either English or Arabic.

Table 4.3 presents an overview of the codes for the teacher's talk. Speaker codes are also provided in the table to show the identity of the speaker (T for teacher). The targeted examples occurred in a single interval and are shown between slashes, except the first example in TA which occurred in three intervals; the last word *matHaf* <museum> occurred in a five-second interval and it was coded as TA since it's part of the same utterance.

Table 4.3 An Overview of Codes for Analyzing Teacher's Talk

Code	Description
TA	Utterances fully in Arabic spoken by the teacher, for example: ▪ T: /*'inahu bijaanib* (.)/ *'inahu, maktab albariid, bijaanib /matHaf* (.)/ < it's next to, it's, the post office, next to the museum >
TE	Utterances fully in English spoken by the teacher, for example: ▪ T: /uh-uh, it's not the same. /
TCS	The teacher switches between English and Arabic, for example: ▪ T: /masculine uhhm. *mudakar wa mu'anath*/ < masculine. masculine and feminine >
TB	Brief language use by the teacher, typically one word in the form of recast or echoic repetition, for example: ▪ T: /*bilHaliib/* < with milk>

Regarding analysis for learners' talk, four similar codes were also employed, although coding learner' talk was challenging due to a slower pace and frequent short pauses. The codes comprised the following: LA for learners' talk in Arabic, LE for learners' talk in English, LCS

for learners switching between English and Arabic within the 5-second interval, and LB for brief language use by the learners.

First, the code LA denoted an entire utterance by a learner or learners in Arabic. Short pauses, lasting five seconds or less within an utterance, were tallied in the coding. However, coding stopped whenever a longer pause, exceeding five seconds, occurred. This exclusion aimed to maintain the focus on an active language interaction to enhance the accuracy of language usage analysis within L2 classrooms. Therefore, coding started whenever the learner resumed speaking. Examples of pausing in LA talk are presented below in Table 4.4:

Table 4.4 Examples of Short and Long Pauses in Learners' LA Talk

Code	Example
Short Pause	L11: *'ajjrii: um* **(2.90)** *talatiin saa^{c}ah?* <I run for thirty, hours?>
Long Pause	L5: *um* **(.)** *thumma* **(.)** *um* **(07.09)** <then> T: *uhhm* L5: *ba^{c}da dhaalik?* <after that?>

Second, the code LE was used for a learner or learners using English fully in one interval. Third, the code LCS was employed when a learner or learners switched codes between English and Arabic. Fourth, the code LB denoted brief language use by the learners regardless of the language. Brief language use meant only one word from a learner in a five-second interval, usually in the form of short answers to the teacher's questions or guessing the meaning of a word in Arabic. Later, an analysis of learners' brief language use was conducted to determine which words were in English and which were in Arabic.

Table 4.5 includes a description of the codes used in time-stamping with examples from the data translated between angle brackets. Speaker codes are also provided to show the identity of the speaker (T for teacher and L+numeral for a learner). The targeted examples are in bold, and all the examples given with each code occurred in one interval, except for the first example in LA which occurred in two intervals.

Table 4.5 An Overview of Codes for Analyzing Learners' Talk

Code	Description
LA	Utterances fully in Arabic spoken by a learner or learners, for example: ▪ L6: /*athhab* (.) *illa lmutanazah/* < I go to the park > */um m:aᶜa kalbi/* < with my dog >
LE	Utterances fully in English spoken a learner or learners, for example: ▪ L9: /to the right, not turn right./
LCS	A learner or learners switching between English and Arabic, for example: ▪ L9: /is *maqha* coffeeshop? / < does cafe mean coffeeshop? >
LB	Brief language use, typically one word in the form of short answers or immediate translations, example: ▪ L3: = /newspapers/

4.3.3. Classroom Observations and Impressionistic Judgement Analysis

As noted above, the original study for which the data were collected (Brown, 2023) examined learning outcomes after the application of different instructional approaches: monolingual versus multilingual. The original study employed live classroom observations using a structured observation protocol. The protocol required the content of each lesson to be

described in lesson segments and for each segment to be classified as containing a majority of teacher language use in Arabic or English and a majority of learner language use in Arabic or English. The real time descriptions yielded an impressionistic judgement of overall majority-minority language use in each entire lesson by teacher versus learners, also in real time. The reliability of this method of reporting classroom language use was validated through a secondary analysis of word counts based on a transcription of one of the lesson segments. The impressionistic judgements of majority-minority language usage for teachers and learners in each lesson were reported in Table 3 in Brown (2023, p. 1645) along with the validating secondary analysis.

The document containing observations and impressionistic judgments from Brown (2023) is secondarily used here. It was re-examined with a specific focus on teacher-learner language use and recalculated to analyze the majority-minority language use during whole-class discussions. Pair activities and language use in student-student interaction are excluded from this analysis. Therefore, all the segments for the whole class discussions are the same segments in the three analyses which creates consistent examination.

A key difference between this study and the results of impressionistic judgements published in Brown (2023) lies in the observers' focus. Brown's study included observations of majority-minority language use during both whole class discussions and pair work. In contrast, this study focused solely on whole class discussions–since the analyses of word counts and time conducted here only focused on whole class discussions–to ensure valid comparisons of the three methods analyzing the same data set.

4.3.4. Intra-coder Reliability

Setting a clear framework for the two methodological operationalizations, words analysis and time-stamping, that were implemented as part of the current study helped to ensure reliable analysis by the same rater over different points in time. As Mackey and Gass (2005) suggest, when a coder makes judgments in two different periods and obtains a high level of agreement – intra-coder reliability testing – it can reasonably indicate a consistent approach to coding. Therefore, the coding of words and time were conducted at two different points of time. Then, the Intraclass Correlation Coefficient (ICC) was used to test absolute agreement of coding between the first session and second session. The agreement included the whole data of both word and time analyses.

The results of two-way mixed effects model showed an ICC value 0.999, which indicated excellent reliability in the coding and counting of words. Table 4.6 below shows the reliability analysis from SPSS.

Table 4.6 Results of ICC Calculation in SPSS Using 2-Way Random-Effects Model for Word Counts

Intraclass Correlation Coefficient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0		
		Lower Bound	Upper Bound	Value	df1	df2
Single Measures	.999[a]	.997	1.000	2966.637	7	7
Average Measures	1.000[c]	.999	1.000	2966.637	7	7

Similarly, results of absolute agreement for time-analysis showed a value of 0.933 indicating excellent reliability in the coding at two different times, which is shown in Table 4.7

Table 4.7 Results of ICC Calculation in SPSS Using 2-Way Random-Effects Model for Time-Analysis

Intraclass Correlation Coefficient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0		
		Lower Bound	Upper Bound	Value	df1	df2
Single Measures	.933[a]	.881	.960	761.610	15	15
Average Measures	.990[c]	.990	1.000	761.610	15	15

4.4. Feasibility Analysis of Measurements of Classroom Language

To assess the feasibility of the three methodological operationalizations of measuring classroom language use, the amount of labor spent on word counts and time stamping was measured to reveal distinct advantages and challenges. The results for observations yielding impressionist judgements of classroom language use were reanalyzed from Brown's (2023) observation protocol, but given that observations were conducted in real time, the length of the lessons could be used as a measure of the feasibility of this latter method.

Chapter 5: Results

Results in terms of language frequency of the applications of each of the three measurements to analyze language use are presented below. These are followed by results assessing the feasibility of the applications of each of the three different measurements of analyzing classroom language use. A comparison of each of the methods in terms of language frequency and feasibility of methods follow in Chapter 6, Discussion.

5.1. Frequency of Language Use in the Classroom

The first research question aimed to provide the outcomes of language usage measurements in the language classroom, exploring three distinct operationalizations: word counts, time stamping, and impressionistic observations.

5.1.1. Word Counts

Overall, the total number of words uttered by the teacher in both classes was 4653 words. In week 7, the teacher spoke 2273 words, while in week 8, the number increased to 2381 words. Similarly, learners spoke a total of 3279 words, with 1544 words in week 7 and 1735 words in week 8. The number of excluded words was 618 in both classes.

In week 7, the teacher predominantly used Arabic, with 2246 Arabic words (99%), while English accounted for only 27 words (1%). Among the learners, Arabic comprised 913 words (59%) and English made up 631 words (41%). No use of any other language was observed by both the teacher and learners.

While in week 8, the teacher's use of Arabic accounted for 2330 words (98%), while English comprised 49 words (2%), and only two French words were utilized, amounting to less

than 1%. Additionally, learners utilized Arabic with 978 words (56%) compared to 757 English words (44%).

Figures 5.1 and 5.2 below show the distribution of language use by the teacher and learners in weeks 7 and 8. The figures utilize a double circle design to represent two sets of data for each week. The inner circle shows language use in words by the learners each week, while the outer circle shows the language use in words by the teacher during the same week.

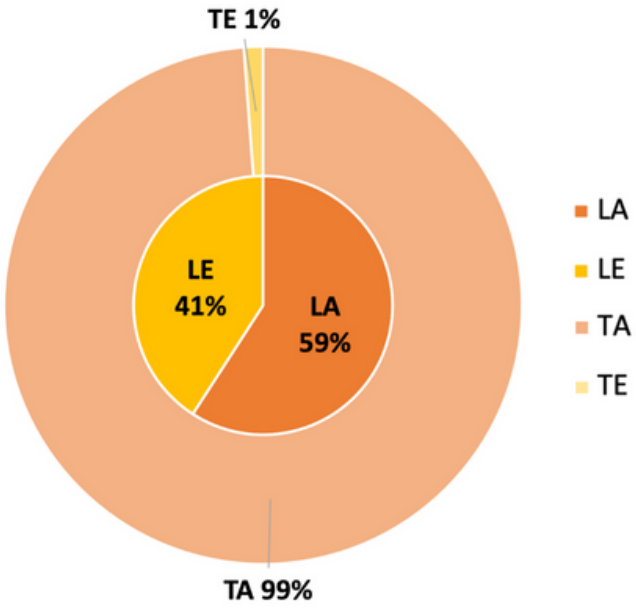

Figure 5.1 Frequency Analysis: Distribution of Language Use in Words by the Teacher and Learners in Week 7

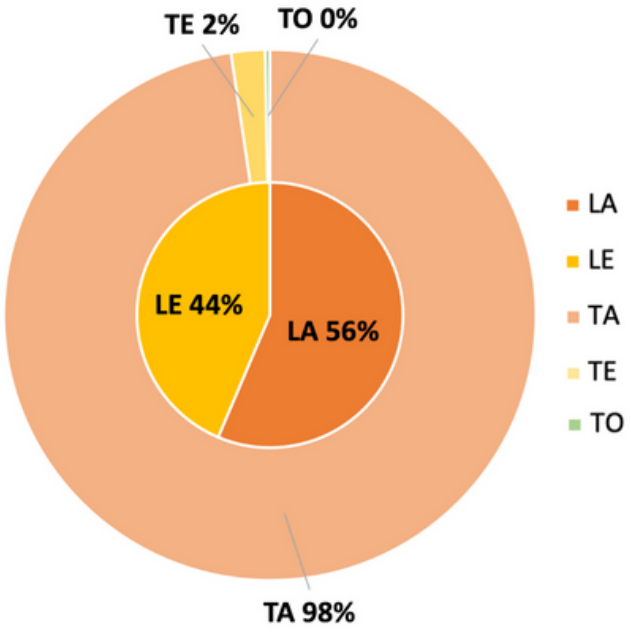

Figure 5.2 Frequency Analysis: Distribution of Language Use in Words by the Teacher and Learners in Week 8

5.1.2. Time-Stamping

Each class lasted for 2hrs30mins with 15-minute break in the middle and 15-minute quiz at the end of the class. The methodology of examining time to measure the duration of time spent by the teacher and learners in each language simultaneously revealed the actual speaking time during whole class discussions over the two classes. Since time was measured in seconds, the results showed that in week 7, the teacher spent 3789 seconds (approximately 63 minutes) instructing and interacting with learners in whole class discussions. While in week 8, he spent 3992.03 seconds (approximately 67 minutes). Learners on the other hand spent 2548.69 seconds (approximately 42 minutes) speaking and discussing in whole class activities and discussions. Similarly, duration of learners' output in week 8 was 2588.57 seconds (approximately 43 minutes). Table 5.1 below shows the duration of speaking time by the teacher and learners in seconds and minutes during whole class activities in weeks 7 and 8.

Table 5.1 Duration of Speaking Time in Whole Class Activities in Weeks 7 and 8

Week	Speaker	Speaking Time (in seconds)	Speaking Time (in minutes)	Class Time
7	Teacher	3789	63.15	2hrs30mins
7	Learners	2548.69	42.47	2hrs30mins
8	Teacher	3992.03	66.53	2hrs30mins
8	Learners	2588.57	43.14	2hrs30mins

The language analysis conducted every 5 seconds revealed a dynamic mix of Arabic, English, codeswitching, and brief instances of language use. During weeks 7 and 8, Arabic was the predominant language for teacher instruction in whole-class activities, accounting for 87.9% and 85.6% of speaking time, respectively. The teacher's brief language usage represented the second-highest percentage in both weeks, occupying a maximum of 11.1% of speaking time in both weeks. The use of English and codeswitching in both weeks fell at 2.5% or below of teacher speaking time in whole-class activities.

Table 5.2 shows the distribution of language use during the teacher's speaking time (approximately 63mins) in whole-class activities in week 7. It also presents the language distribution for the entire class time (2hrs30mins). However, the analysis of teacher's speaking time excludes pair work, while the entire class time includes pair work, a 15-minute break and a 15-minute quiz. This explains the seemingly large decrease in the teacher's use of Arabic (to 43%) when considering total class time.[6]

Table 5.2 Temporal Analysis: Language Distribution by the Teacher in Week 7

Week	Code	% over Whole Class Teacher Speaking Time	Total Duration [sec]	% over Total Class Time
7	TA	87.9%	3336.18	43%
7	TCS	2.2%	80.87	1%
7	TE	0.13%	4.95	0.1%
7	TB	9.8%	367.01	5%

[6] Software limitations prevented exclusion of pair work and breaks.

Similarly, Table 5.3 below presents distribution of language use during teacher's speaking time (about 67mins) during whole-class activities in week 8 as well as distribution of language use during whole class time, including pair work, break, and quiz.

Table 5.3 Temporal Analysis: Language Distribution by Teacher in Week 8

Week	Code	% over Whole Class Teacher Speaking Time	Total Duration [sec]	% over Total Class Time
8	TA	85.6%	3414.42	44%
8	TCS	2.5%	102.92	1%
8	TE	0.8%	35.36	0.5%
8	TB	11.1%	439.33	6%

Learners on the other hand showed varied percentages of language use. In week 7, learners spoke for a total of 43 minutes in whole-class activities in which Arabic constituted the majority of their speaking time (57%). Codeswitching followed closely, accounting for 17.4%. Learners' use of English reached 14.4% of their speaking time. Table 5.4 details the distribution of languages in both learners' total speaking time during whole-class activities in week 7, and language distribution during the entire class time, including pair work and two breaks.

Table 5.4 Temporal Analysis: Language Distribution by Learners in Week 7

Week	Code	% over Whole Class Learner Speaking Time	Total Duration [sec]	% over Total Class Time
7	LA	57%	1448.35	19%
7	LCS	17.4%	442.89	6%
7	LE	14.4%	369.91	5%
7	LB	11.2%	287.54	4%

In week 8, learners' use of Arabic in whole-class activities decreased to 52% compared to week 7. However, speaking in English and codeswitching increased, reaching 18.3% and 19.7% respectively. Table 5.5 below shows the distribution of language use in both learners' total speaking time during whole-class activities (approximately 43 minutes) as well as language use in the entire class time, including pair work and two breaks.

Table 5.5 Temporal Analysis: Language Distribution by Learners in Week 8

Week	Code	% over Whole Class Learner Speaking Time	Total Duration [sec]	% over Total Class Time
8	LA	52.5%	1359.60	18%
8	LCS	19.7%	507.93	7%
8	LE	18.3%	471.89	6%
8	LB	9.6%	249.16	3%

Regarding brief language use, which mainly contained one word in a 5-second interval and was not a part of a stretch of language that occupied five seconds or more, the findings showed a majority of Arabic use by the teacher. In week 7, 70 Arabic words (95.9%) were used compared to just three English words (4.1%) in the form of recast, echoic repetition or one-word questions. Similarly in week 8, the total number of words comprising brief language use spoken by the teacher was 88, with 86 in Arabic (97.7%) and only two words in English (2.3%). The following figures (5.3 and 5.4) illustrate the distribution of languages within brief language use by the teacher.

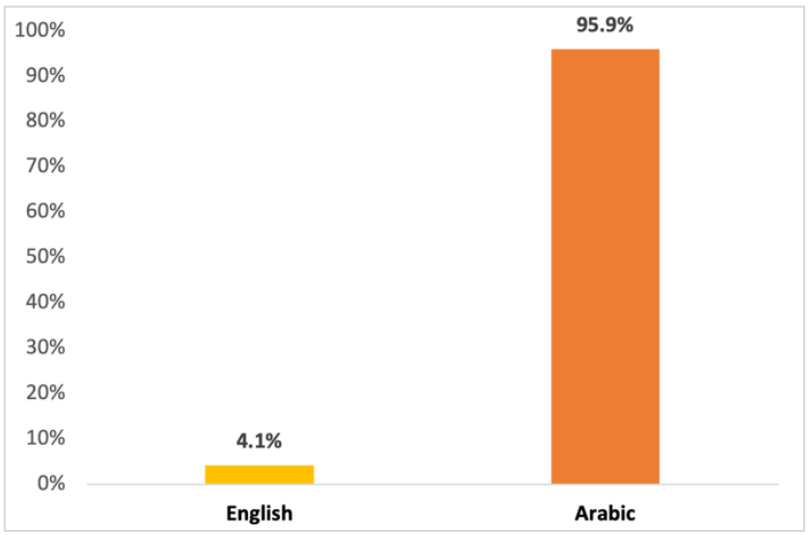

Figure 5.3 Analysis of Teacher's Brief Language Use (TB) in Week 7

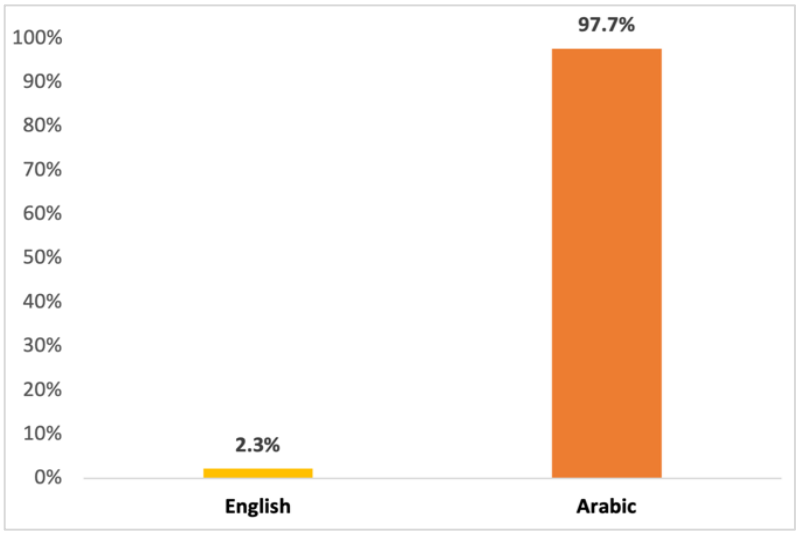

Figure 5.4 Analysis of Teacher's Brief Language Use (TB) in Week 8

Analysis of learners' brief language use showed that learners in week 7 used 16 words in English (28.6%) and 40 words in Arabic (71.4%) in the form of short answers, providing translations and sometimes asking one-word questions, totalling in 56 words of brief language use. While in week 8, learners uttered 49 words of brief language use: 33 were in Arabic (67.3%) and 16 words in English (32.7%), as shown in figures 5.5 and 5.6 below:

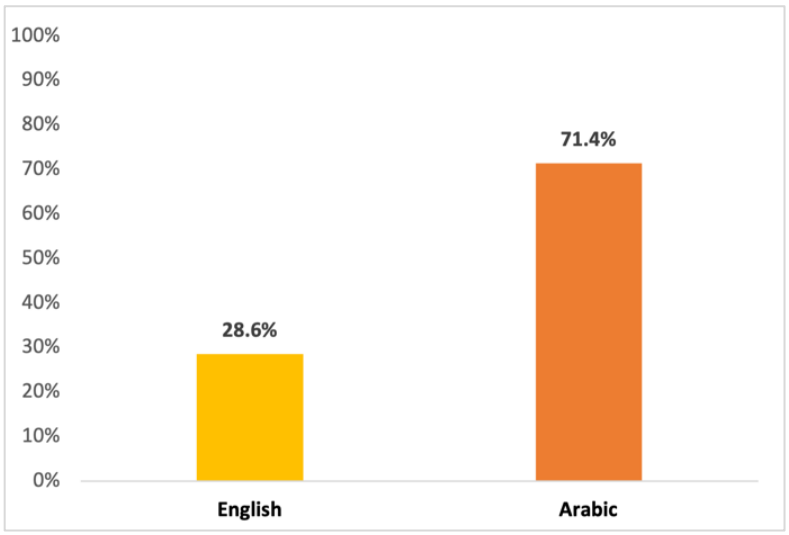

Figure 5.5 Analysis of Learners' Brief Language Use (LB) in Week 7

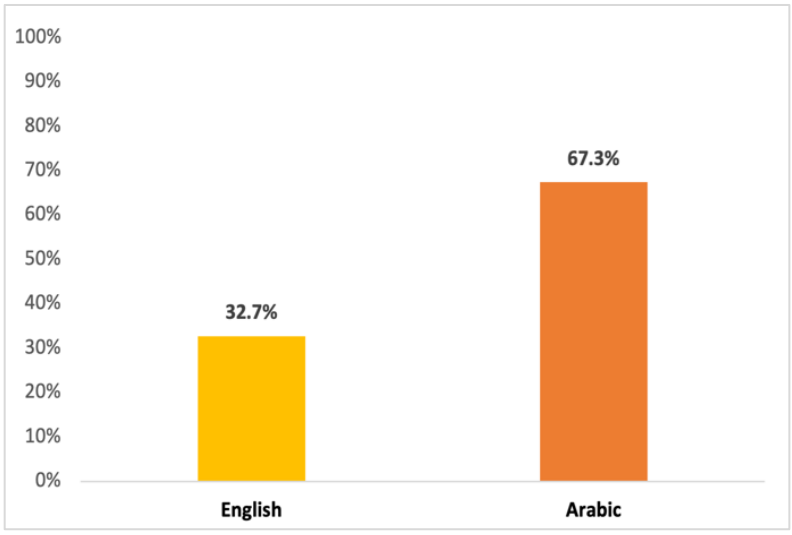

Figure 5.6 Analysis of Learners' Brief Language Use (LB) in Week 8

5.1.3. Classroom Observations and Impressionistic Judgement Analysis.

In Brown (2023), the observer documented language use by the teacher and learners during whole-class discussions and pair-activities using a structured observation protocol across the 10 weeks. The observer divided each class into lesson segments, with each class including 13 to 18 segments. Language use by the teacher and learners in each segment was described, and from judgements of language use in individual segments, an overall majority-minority language use for each lesson was determined. Since the focus here is on weeks 7 and 8, and only on whole-class discussion with exclusion of pair work, the observation document has been re-examined and reanalyzed for the results below.

After excluding pair work and language use between students, the reanalysis of impressionistic judgments resulted in that the teacher employed more Arabic than English during instruction and interaction in both weeks. Similarly, the analysis for learners' language use showed that Arabic use surpassed English during discussions and whole-class activities in both weeks. Table 5.6 illustrates the results of the impressionistic judgments made by the observer after excluding pair work and student-student interactions. A comparison of each of the methods in terms of language frequency follows in the Discussion.

Table 5.6 Impressionistic Judgments of Teacher and Learners' Majority-Minority Language Use in Weeks 7 & 8

Week	Teacher	Learners
7	Majority Arabic - Minority English	Majority Arabic - Minority English
8	Majority Arabic - Minority English	Majority Arabic - Minority English

5.2. Feasibility of Measurements of Classroom Language Use

The second research question concerned assessing the feasibility of the three different methods of measuring language use in L2 classrooms: word count, time stamping, and impressionistic observations. Feasibility was operationalized in amount of labor needed for each method, which is presented below.

5.2.1. Feasibility of Word Count Measurement

The recorded video of week 7 was 2hrs30mins long, but not all of it was transcribed. The 15-minute break, pair-activities where the teacher did not interrupt, and 15-minute quiz at the end of class lowered the task of transcribing. This resulted in transcribing 1.33hr of instruction and whole-class discussions that involved both the teacher and leaners.

Using Seedhouse's (2004) transcription conventions and occasionally revisiting some utterances for clarity, transcribing one 15-minute classroom segment where all speakers used English and Arabic in a CLT context required 2hrs30mins. After that, translating the Arabic utterances into English required 1 hour. Finally, the process of categorizing and counting words in either Arabic or English and by the teacher and learners during the 15-minute classroom segment needed 50 minutes of labor.

Extrapolating from the analysis of labor employed in analyzing a 15-minute segment of classroom discourse in week 7 indicates that transcribing, translating and coding the full 1.33hr (93 minutes) of classroom discourse took approximately 26hrs52mins. Similarly, excluding the time for the break, quiz and uninterrupted pair-activities in week 8 resulted in transcribing 1hr42mins out of the whole 2hr30mins video recording. Using the above extrapolation, we can

estimate that transcribing, translating, and coding words in 1hr42mins (102mins) required approximately 29hrs33mins.

5.2.2. Feasibility of Time-Count Measurements

Time-analysis required analyzing the recordings two times, one for examining language use by the teacher and one for language use by the learners. Analyzing language proportions during a segment of 15 minutes by stamping every five seconds as either Arabic, English, codeswitch, or brief language use twice resulted in the following: coding of teacher talk took approximately 47 minutes, while coding the same segment to examine learners' use of language required 42 minutes.

Although coding was repeated, once for the teacher and once for the learners, the entire recording of the week 7, which lasted 2hr30mins, was not coded. The time for the break, quiz, uninterrupted pair-activities were excluded. Thus, extrapolating from the feasibility analysis of time-stamping 15 minutes of teacher talk, we can estimate that the coding of 63 total minutes of teacher talk in week 7 required approximately 3hrs17mins. Similarly, time-stamping the total of 43 minutes of learners' talk in week 7 required approximately 1hr58mins. In total, the labor required for coding week 7 twice—combining the teacher and the learners—was approximately 5hrs15mins.

Regarding week 8, teacher talk comprised 67 minutes, which required approximately 3hrs30mins of coding. Meanwhile, learners' talk comprised nearly 44 minutes and required approximately 2hrs3mins. Thus, the total time spent in time stamping class 8 combining teacher and learner talk was approximately 5hrs33mins.

5.2.3. Feasibility of Classroom Observations and Impressionistic Judgement Measurements

Reporting language use using the method of classroom observation and impressionistic judgement within a 15-minute segment of classroom interaction required 15 minutes for the observer to take notes and use those notes to make impressionistic judgements of classroom language use by teacher and learners. Extrapolating from a 15-minute segment analysis, conducting a two-hour classroom observation and making impressionistic judgments of language use would require approximately 1hr33mins for week 7 and 1hr27mins for week 8. This excludes pair activity time, which lasted 27 minutes in week 7 and 31 minutes in week 8.

A comparison of each of the methods in terms of feasibility follows in the Discussion.

Chapter 6: Discussion

The aim of this methodological research was to investigate the applications of three different methods for capturing the proportions of language use (target and non-target) in the language classroom: word count, time-analysis, and live observations with impressionistic judgements. The study aimed to test Macaro's (2022) hypothesis that different methods would generate different language use frequencies by comparing the results generated from these methods of the frequency of language use as well as the feasibility of implementation of each method.

6.1. Research Question 1: Frequencies of Classroom Language Use as Captured by Three Methods

To capture language practices by the teacher and learners in two Arabic language classes, the first operationalization was counting words from two classroom transcriptions in which the teacher and learners used English, TL Arabic and a seldom use of French in whole classroom discussions. The second operationalization was time-stamping every 5 seconds in four codes: Arabic, English, codeswitching, or brief language use. Timed analysis allowed to present two additional categories that were not present in counting words: codeswitching and brief language use, beside English and Arabic. Time-stamping was done two times per class to first show the language use by the teacher, then the language use by the learners. Finally, the third operationalization was a reanalysis of Brown's (2023) observation protocol data, excluding pair work and student-student interaction. The data comprised classroom observations made by the

observer to provide impressionistic judgements about language proportions used by the teacher and learners during whole-class discussions.

Tables 6.1 and 6.2 below summarize the findings from the application of each of the three methods regarding the frequencies of the languages used by teacher and learners, respectively, focusing only on whole class discussions and excluding pair work.

Table 6.1 Summary of the Frequency of Teacher Language Usage Resulting from Three Methodologies.

Week	Language Use	Word Analysis of Total Teacher Talk	Time-Analysis of Total Teacher Talk	Live Observation and Impressionistic Judgement Analysis of Teacher Talk
7	Arabic	99%	87.9%	Majority
	English	1%	0.13%	Minority
	Other	0%	N/A	0
	Codeswitching	N/A	2.2%	N/A
	Brief Use	N/A	9.8%	N/A
8	Arabic	98%	85.6%	Majority
	English	2%	0.8%	Minority
	Other	<1% French	N/A	0
	Codeswitching	N/A	2.5%	N/A
	Brief Use	N/A	11.1%	N/A

Table 6.2 Summary of the Frequency of Learners' Language Usage Resulting from Three Methodologies.

Week	Language Use	Word Analysis of Total Learner' Talk	Time-Analysis of Total Learner' Talk	Live Observation and Impressionistic Judgement Analysis of Learner' Talk
7	Arabic	59%	57%	Majority
	English	41%	14.4%	Minority
	Other	0%	N/A	0
	Codeswitching	N/A	17.4%	N/A
	Brief Use	N/A	11.2%	N/A
8	Arabic	56%	52.5%	Majority
	English	44%	18.3%	Minority
	Other	0%	N/A	0
	Codeswitching	N/A	19.7%	N/A
	Brief Use	N/A	9.6%	N/A

In a comparison of the three methodologies, the method of word counts showed high percentages of TL Arabic use by the teacher, at 99% and 98% in weeks 7 and 8, respectively, but the results from the method time-stamping yielded decreased percentages of target language Arabic use by the teacher, at 87.9% and 85.6% in both weeks. This difference can in part be explained by the existence of more coding categories for the method of time stamping. Thus,

while less present in the category of TA (Teacher's use of Arabic), the teacher's use of Arabic in time stamping could also be represented both in the category of codeswitching (TC) and in the category of brief language use (TB), especially since Figures 5.3 and 5.4 in Chapter 5 indicate that the vast majority of brief language uses by the teacher were conducted in Arabic.

When examining use of English (the non-target language) by the teacher, we see a similar pattern to that described across methods for teacher use of Arabic, namely that the method of time stamping classified English use at 0.13% while the method of word counting noted a higher use of English at 1% in week 7, and at 0.8% and 2%, respectively, in week 8. Again, the difference can be explained in part by the fact that time stamping included more coding categories, and the use of English could also be represented in teacher codeswitching (TC) and brief language use by the teacher (TB).

Switching to language use by learners, we see that the percentages of TL Arabic use by learners was comparable as measured by both time and word counts in both week 7 (57% and 59%, respectively) and in week 8 (52.5% and 56%, respectively). However, results for the use of L1, English, were considerably different across measurement methods, with learners reportedly using English at 14.4% according to the analysis of learner speaking time and 41% according to the analysis of learner words produced in week 7 (similarly 18.3% and 44% respectively, in week 8). At 17.4% of learner speaking time in week 7 and 19.7% in week 8, learners were observed to engage in much more codeswitching, according to the analysis of time, than was the teacher (2.2% in week 7 and 2.5% in week 8).

In addition, the difference of percentages in language use as measured by time stamping versus word counts is generally in line with Macaro's (2022) observation:

> "it is highly likely that a five-second sampling technique will produce comparatively lower levels of … L1 [non-target language English] use than word analysis, because the rate of speech for the L1 [non-target language English] tends to be faster than that for L2 [target language Arabic], the latter being often characterized by repetitions, slower articulation and longer pauses." (p. 210)

The observation fits the learners' data of language use, showing a decline in English use from 41% in words count to 14.4% in time-stamping during week 7, and from 44% of English usage in words count to 18.3% in time-stamping. Though the teacher did not share the L1 with the learners, his use of the non-target language English decreased from 1% in words count to 0.13% in time-stamping during week 7, and similarly from 2% to 0.8% during week 8.

Regarding the method of classroom observations and impressionistic judgements, results obtained for teacher talk aligned with the findings obtained through application of the analyses of time and words, namely that Arabic was the predominant language used by the teacher in both weeks. Similarly, the method of classroom observations and impressionistic judgements of learners talk during whole-class discussions yielded similar results to the findings borne out by both time and word count analyses. In weeks 7 and 8, learners were observed to use more TL Arabic than English in classroom discussions.

Overall, when evaluating the three different ways to measure language use in the L2 classroom, we see that live observation with impressionistic judgements of language yields a broad picture, lacking in detail about codeswitching. Regarding the remaining two methods, an analysis of time may under-report use of the L1 relative to the TL in comparison to an analysis of word count, likely due to the differences in speaking speed for the L1 and L2. However, the analysis of time has the capacity to include more coding categories – codeswitching and small

language use (including full words and morphemes) - and thus may give researchers a more fine-grained picture of the distribution of different types of language interactions in the SL classroom.

6.2. Research Question 2: Feasibility of Three Methods of Capturing Frequencies of Classroom Language Use

The second research question discussed the feasibility of the three methods used to document language use in two Arabic classes. The first method, word counts, included different phases: transcribing, translating, and then classifying and counting words in each language. After extrapolating from measuring labor time in the processing of one 15-minute sample of classroom interaction, the process for transcribing, translating lines, classifying words by speakers then counting words in each language was estimated to take approximately 56hrs25mins for both classes.

The second method, time-analysis, was conducted two times; one time each for the teacher and learners separately in week 7 and one time each for the teacher and learners separately in week 8. After extrapolating from measuring labor time in the processing of one 15-minute sample of classroom interaction, these four passes of coding were estimated to take a total of approximately 10hrs 48mins of work.

Finally, the estimated time to conduct live observations and make impressionistic estimates of language use in both classes was approximately 3hrs2mins. This excludes break time (15 minutes per class), quiz time (15 minutes per class), and pair activities time (27 minutes in week 7 and 31 minutes in week 8) for a total observation time of 5 hours across both classes. Table 6.3 below summarizes the time spent for each method.

Table 6.3 Summary of the Feasibility for the Three Methodologies.

Method	Duration of Class Sample	Labor Time Required	Total Classes Duration	Approximate Labor Time Required
Time Analysis for Teacher talk	15mins	47mins	2hrs10mins	6hrs47mins
Time Analysis for Learners talk	15mins	43mins	1hr27mins	4hrs1min
Total Time Analysis			3hrs37mins	10hrs48mins
Word Analysis	15mins	4hrs20mins	3hrs15mins	56hrs25mins
Live Observation + Impressionistic Judgements	15mins	15mins	4hrs	3hrs2mins

Among the three methods, counting words was the most time-consuming method compared to the other two methods (approximately five times more labor time than the total time-analysis, and 14 times more than live observation with impressionistic judgements) since it compromised three steps: transcribing, translating, and counting. Alternatively, time-stamping was the method involving the second largest commitment of labor time. Finally, the live observation and impressionistic judgement approach required the least amount of labor time among the three methods, since documenting language use occurred within actual class time.

6.3. The Impact of Language Teaching Approaches on the Methods of Measuring Language Use

During the process of application of different methods of counting language use in the SL classroom, the potential impact of the teaching approach became apparent. Although the teacher followed a CLT approach in teaching the TL Arabic most of the time, there were a few activities

in both classes where he incorporated some audiolingual drills in which the students read aloud short conversations or repeated a list of words.

In a comparison of the three methodologies, the first method—counting word—included transcribing the discourse first before counting. Transcribing audiolingual activities was markedly easier because that included transcription of frequent repetition of the same language items. This contrasted with transcribing CLT activities, since these were characterized by unique, non-repeated discussions and multiple speakers. In addition, translating, analyzing discourse, and classifying words from utterances during audiolingual activities was also easier than applying the same methodological research procedures for CLT activities.

Similarly, coding audiolingual activities using the time-analysis method required less time than coding CLT activities as utterances in the former were formulaic, with a focus on fluent pronunciation and therefore fewer short pauses and almost no long pauses of five seconds or more.

6.4. Additional Considerations in the Application of the Three Methods of Measuring Language Use in L2 Classrooms

Counting words from discourse transcriptions showed only proportions of English and Arabic, with no instances of intra-word codeswitching in the data. Therefore, classifying words according to languages was fairly straightforward. Also, classifying words did not include concerns about learners lengthening a specific sound or measuring pauses within utterances, which were frequently produced since the classes were designed for beginners. On the other hand, the analysis of time had the capacity to include more coding categories – codeswitching and brief language use - which may give researchers a more fine-grained picture of the

distribution of different types of language interactions in the L2 classroom. Also, choosing the interval of five seconds over other numbers of seconds was because analyzing five seconds was in line with previous research, inclusive of language use with condensation of short and long pauses, and most importantly feasible. When using the method of time-stamping, tallying pauses that occurred between learner's answers were included, only if the pause was less than five seconds, as well as lengthening sounds which usually took one to two seconds. These features affect the analyses of time, lengthening the representation of the target language Arabic (see Macaro, 2022). Table (6.4) illustrates instances of learner lengthening a sound as symbolized by [**:**].

Table 6.4 Excerpt Illustrating Lengthening Sounds from Classroom Discussions.

No.	Speaker	Line
754	L1	um (.) *'astayqiDu, 'as:tay:qi:Du:* < I wake up, I wake up >
755	T	uhhm
756	L1	*fii:* um (.) *ataa:ss:i^{c}ah SabaaHan* <at nine o'clock in the morning>
757	T	uhhm
758	L1	um *'umaarisu* (.) *a:jjrii* < I go *run (verb) > ((grammatical error))

Verbal fillers were also included the time-stamping when they occurred in the same utterance when the teacher or learners spoke. In addition, proper names and borrowed words with similar pronunciation were included in the time-analysis as they occurred within the teacher and learners' speech. This was in contrast with the first method, counting words, as verbal fillers, proper names and borrowed words were transcribed but excluded from word counts. The following Excerpts in Table 6.5 show two examples with many verbal fillers and short pauses,

indicated in bold font, which were included in the timed analysis but were excluded in the word count.

Table 6.5 Excerpts for Verbal Fillers and Short Pauses from Classroom Discussions.

Excerpt	No.	Speaker	Line
1	979	L11	**iii** *'ajjrii:* **um (3.17)** *thalathiin* **(.)** *saa^{c}ah?* < I run for thirty, hours? >
	980	T	*thalathiin saa^{c}ah?* < for thirty hours? >
	981	L11	((laughs)) *laa?* < no? >
2	1021	T	**uhhm**, *jayyid* < good >
	1022	L4	**um** *'inahu* **(1.44) uhh uhh (2.19)** *fii* Hall of Languages **uhh** *bijanib almaktab* **(.)** so, next to the office. < it's in Hall of Languages, next to the office. So, next to the office >

However, there were instances where the pauses extended beyond five seconds within the same utterance. Such pauses were excluded in the timed analysis; the coding stopped then resumed once the learner started again their answer. The decision was made to exclude long pauses because including silent intervals within the analysis could potentially misrepresent the language use. Also, pauses did not have a designed code. While in counting words, pauses did not present a challenge in measuring language use in the classroom. Examples are given in Table 6.6.

Table 6.6 Excerpt for a Long Pause from Classroom Discussions.

No.	Speaker	Line
949	T	*thumma* (.) *baᶜda thalik* < then, after that>
950	L5	[oh
951	T	[*nafsu alshay'*] <similar>
952	L5	So **(08:04)** ((turning the pages))

One of the merits of using time-analysis is that this method can capture language use in the small linguistic forms in any language. Arabic, for example, is rich in inflectional morphology and there were many instances of language usage where the teacher and the learners sometimes uttered only individual morphemes such as the inflectional morpheme for first person singular. These morphemes were transcribed but were not counted in the analysis of countable words, since the focus was on full words only. The two excerpts in Table 6.7 below show examples for some Arabic morphemes that were included in the time-analysis but not in word counts. The excerpts are all from transcribed data and the morphemes are in bold font:

Table 6.7 Excerpt for Arabic Morphemes in Classroom Discussions

Excerpt	No.	Speaker	Line
1	504	T	*waraa'* < behind >
	505	L2	*waraa'* (.) *waraa'* < behind, behind >
	506	T	***ii*** < 1st person singular possessive suffix >
	507	L2	***ii*** < 1st person singular possessive suffix >
	508	T	*jayyid* < good >
2	1694	T	*'anaa 'a^{c}mal, 'at- atasawaq 'ashtarii 'adhab 'amshii aakul* '*ashrab* ***'a 'a 'a 'a 'a*** *'anaa* < I work, I shop, I buy, I walk, I eat, I drink, 1st person singular prefix for verbs (4 times), I >
	1695	L2	***'a*** < 1st person singular prefix for verbs >
	1696	L11	***'a*** < 1st person singular prefix for verbs >

The above examples support the notion reported in Macaro (2022) and Vold & Brkan (2020) that the time-analysis is particularly useful for analyzing languages that have a high number of compound words (e.g., German) or rely on characters (e.g., Chinese) (see Zhang, 2021). This aligns with observations in this study of languages like Arabic, where frequent morphological constructions in speech by teachers and learners can benefit from time-analysis.

However, it is important to note that the word analysis does not differentiate between TL and proficient language use since it deals with counting words. Time analysis, on the other hand, might be biased towards TL use. This bias might show higher percentages of TL use because it included pauses and TL produced at a slower rate. While learners' speech in native or proficient language is characterized by faster rate and shorter pauses.

6.5. Limitations and Recommendations for Future Research

To gain a deeper understanding of language use in language classrooms, this study compared three methods for analyzing teacher and learner interactions during whole-class discussions. While the three methods provided valuable insights, some limitations emerged.

As this study created a framework for coding, at least for two methods, the categories for each method made the comparison of the results challenging. While the method of word analysis resulted in three categories: Arabic, English and French, the time analysis showed four categories: Arabic, English, codeswitching and brief language use, and the observations resulted in two categories: majority and minority language use. Standardization of categories is important, and this can be improved by analyzing complete utterances, rather than isolated words, because complete utterances will likely include a codeswitching category and that would be particularly beneficial for a comparison with the method of time-analysis, and it would allow for a more comprehensive understanding of code-switching patterns across different methodologies. In addition, counting utterances might show different percentages of language use by the teacher and learners. Since the method of word counts showed higher percentages of English use and TL Arabic use as compared to the percentages reported by the time analysis, counting utterances might also show lower percentages of English-only and Arabic-only, more comparable to the time analysis. For the method of observations, the manuals for observations could include a category of codeswitching along with L1 and L2. While this might not be possible or difficult in live observations, researchers can use video recordings and conduct asynchronous observations that will allow for a more reliable analysis.

Another limitation in this study is the exclusion of pairs' interactions in the three methods. The exclusion of these discussions, which accounted for approximately a third of class

time, in the analyses likely impacted the reflection of actual language practices in the classroom. Dyadic discussions are as important as whole-class discussions and reflect critical language use in the classroom. While including pair activities means adding more language use by the learners, since the teacher will mostly check pairs and won't interrupt them to enable them to use TL, it can be very useful to compare codeswitching between pair activities and whole class discussions. In addition, the results shown by learners did not track any individual language use, but showed the total language use by all learners. Therefore, it is not possible to say that all learners used TL Arabic at 57% of total words or so.

Finally, the method of word analysis required a significant amount of time: 56hrs25mins for both classes. This is because the method used in this study did not incorporate transcription services. The labor time for word analysis can definitely vary depending on the TL and teaching method, but it could be reduced by using transcription services that incorporate AI speech recognition. While such services might not support all languages, researchers can still consider using AI speech recognition to measure the time required for audio transcriptions and determine if human intervention is needed or not. Also, all the work was done by a single coder to assess feasibility, so it is important to remember that individual work times vary.

Researchers are advised on the choice of analyses based on their objectives and resources. Ideally, if time and technology allow, conducting a combined analysis of word and time in language classroom is recommended, especially considering ACTFL's recommendation of 90% TL, but unspecified as words or time. However, if constraints exist, prioritizing word analysis is advisable when resources permit both time and budget. But if researchers face limitations in time, opting for a time-analysis proves to be a good solution because it offers insights with multiple categories of language use. The time-analysis also seems to be a balanced

option, offering considerable feasibility while still yielding reasonably accurate results, thus making it a favorable choice for many research projects, especially with software programs such as Mangold INTERACT or BORIS. Although calculating class time will include some silent time and that might not be relevant, calculating speaking time only may miss teachers' non-verbal signals that are not spoken. Classroom observations are recommended if researchers face time constraints or lack access to coding software technologies, or when video/audio recordings are not permitted due to potential consent issues from all students, which may cause IRB concerns. Using standardized observation protocol instead of informal observations is advised to capture language practices in the classroom (see Klette, 2023). Online language classes offer a valuable feature for researchers with the built-in recording and transcription capabilities within programs such as Zoom and Microsoft Teams. This allows researchers to easily revisit classes and engage in an asynchronous observation.

Future research might consider establishing a more precise time threshold for the category 'brief language use' in the method of time analysis. Since one word took less than a second, the 5-second interval likely captured much silent time.

Also, since this course was designed for beginners, the teacher made extensive use of nonverbal gestures and mimics. It will be beneficial to explore the role of nonverbal communication alongside language use in language classroom.

Finally, the results of the current study reported different language proportions by the teacher and learners according to the methods used: word analysis, time analysis, and live observations and impressionistic judgments. It is critical for future research comparing studies of language usage in the L2 classroom, particularly those involved in meta-analysis, to take the methods of language measurement into consideration alongside the levels of proficiency and

educational context, since results obtained from each method of language measurement can vary greatly. Moreover, this study does not have direct implications for changes in language teaching practices, but it contributes valuable insights for researchers seeking to improve language pedagogy and investigate language use in classroom settings.

6.6. Conclusion

This study investigated three measurements of language usage in the language classroom, operationalized through analysis of word counts, time, and impressionistic judgements based on (live) class observations and compared and contrasted them in terms of frequency of language use and feasibility of implementation. Also, this study provided a detailed framework for implementation of word and time analyses in an attempt to help improve research methodological practices. The findings suggest that each operationalization method presents different advantages and challenges.

Regarding the proportions of language use in the classroom, the method of word analysis provided a quantitative measure of TL Arabic and English use, allowing for precision in analysis; however, it may not capture the nuances of language practices. The method of analyzing time offered insights into the aspects of language use and showed codeswitching practices by the teacher and learners but was skewed by speaking speed and hesitations. The impressionistic observations, while subjective, enabled a holistic understanding of language usage patterns but may lack the thoroughness and accuracy of words and time analyses. In terms of feasibility, word analysis proved to be the most time-consuming method, followed by time-analysis—which was moderate—, then live observations with estimates of language use which required the least amount of time.

In both classes, the teacher demonstrated maximal L1 use and maintained higher usage of TL Arabic. With current calls for the optimal use of L1 and reconsidering its role in the classroom, and the concept of multi-competence and making use of learners' linguistic repertoires, it is important to capture actual practices of language usage accurately with the proper methodology. Future research should analyze utterances, to show language use as well as codeswitching practices in the classroom. In addition, including pair-activities in the analysis will give valuable insights about language practices in educational contexts.

Reference

Aoyama, R. (2020). Exploring Japanese High School Students' L1 Use in Translanguaging in the Communicative EFL Classroom. *TESL-EJ (Berkeley, Calif.)*, *23*(4), 1.

Aull, B. (2021). Pre-primary Teacher Talk: L1 Use across Different Activity References. *ELT Journal*, *75*(3), 278–289. https://doi.org/10.1093/elt/ccab013

Brevik, L. M., & Rindal, U. (2020). Language Use in the Classroom: Balancing Target Language Exposure with the Need for Other Languages. *TESOL Quarterly*, *54*(4), 925–953. https://doi.org/10.1002/tesq.564

Brown, A. (2023). Monolingual versus Multilingual Foreign Language Teaching: French and Arabic at Beginning Levels. *Language Teaching Research: LTR*, *27*(6), 1634–1659. https://doi.org/10.1177/1362168821990347

Brown, A., & Lally, R. (2019). Immersive Versus Non-immersive Approaches to TESOL: A Classroom-Based Intervention Study. *TESOL Quarterly*, *53*(3), 603–629. https://doi.org/10.1002/tesq.499

Brown, A., Lally, R. & Lisnyczyj, L. (2022) Multilingual versus Monolingual Classroom Practices in English for Academic Purposes: Learning Outcomes, Student Attitudes, and Instructor Observations. In C. Bardel, C. Hedman, K. Rejman, & E. Zetterholm (Eds.), *Exploring Language Education. Global and Local Perspectives* (pp. 393-402). Stockholm University Press. https://doi.org/10.16993/bbz.d

Canale, M. (1983). From Communicative Competence to Communicative Language Pedagogy 1. In J. C. Richard, & R. W. Schmidt (Eds.), *Language and Communication* (1st ed., pp. 2–27). Routledge. https://doi.org/10.4324/9781315836027

Canale, M., & Swain, M. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics*, *I*(1), 1–47. https://doi.org/10.1093/applin/I.1.1

Celce-Murcia, M. (2014). An Overview of Language Teaching Methods and Approaches. In M. Celce-Murcia, D. Brinton, & M.A. Snow (Eds.), *Teaching English as a Second or Foreign Language* (4th ed., pp. 2-14). National Geographic Learning.

Chavez, M. (2016). The First Language in the Foreign Language Classroom: Teacher Model and Student Language Use—An Exploratory Study. *Classroom Discourse*, *7*(2), 131–163. https://doi.org/10.1080/19463014.2016.1149499

Cook, V. (2001). Using the First Language in the Classroom. *Canadian Modern Language Review*, *57*(3), 402–423. https://doi.org/10.3138/cmlr.57.3.402

Cook, V. (2016a). Premises of multi-competence. In V. Cook & L. Wei (Eds.), *The Cambridge Handbook of Linguistic Multi-Competence* (pp. 1–25). Cambridge University Press https://doi.org/10.1017/CBO9781107425965.001

Cook, V. (2016b). *Second Language Learning and Language Teaching* (5th ed.). Routledge. https://doi.org/10.4324/9781315883113

De La Campa, J. C., & Nassaji, H. (2009). The Amount, Purpose, and Reasons for Using L1 in L2 Classrooms. *Foreign Language Annals*, *42*(4), 742–759. https://doi.org/10.1111/j.1944-9720.2009.01052.x

de la Fuente, M. J., & Goldenberg, C. (2022). Understanding the Role of the First Language (L1) in Instructed Second Language Acquisition (ISLA): Effects of Using a Principled Approach to L1 in the Beginner Foreign Language Classroom. *Language Teaching Research*, *26*(5), 943–962. https://doi.org/10.1177/1362168820921882

DiCamilla, F. J., & Antón, M. (2012). Functions of L1 in the Collaborative Interaction of Beginning and Advanced Second Language Learners. *International Journal of Applied Linguistics*, *22*(2), 160–188. https://doi.org/10.1111/j.1473-4192.2011.00302.x

Duff, P. (2014). Communicative Language Teaching. In M. Celce-Murcia, D. Brinton, & M.A. Snow (Eds.), *Teaching English as a Second or Foreign Language* (4th ed., pp. 2-14). National Geographic Learning.

Duff, P. A., & Polio, C. G. (1990). How Much Foreign Language Is There in the Foreign Language Classroom? *The Modern Language Journal*, *74*(2), 154–166. https://doi.org/10.2307/328119

Edstrom, A. (2006). L1 Use in the L2 Classroom: One Teacher's Self-Evaluation. *Canadian Modern Language Review*, *63*(2), 275–292. https://doi.org/10.3138/cmlr.63.2.275

Ellis, R., & Shintani, N. (2014). *Exploring Language Pedagogy through Second Language Acquisition Research* (1st ed.). Routledge. https://doi.org/10.4324/9780203796580

ACTFL. (2024). *Facilitate Target language Use*. https://www.actfl.org/educator-resources/guiding-principles-for-language-learning/facilitate-target-language-use

Hall, G., & Cook, G. (2012). Own-language Use in Language Teaching and Learning. *Language Teaching*, *45*(3), 271–308. https://doi.org/10.1017/S0261444812000067

Howatt, A. P. R. (2006). Language Teaching: History. In K. Brown (Ed.), *Encyclopedia of Language & Linguistics* (2nd ed., pp. 634–646). Elsevier. https://doi.org/10.1016/B0-08-044854-2/01325-0

Hymes, D. H. (1972), On Communicative Competence, In J.B. Pride and J. Holmes (Eds), *Sociolinguistics. Selected Readings* (pp. 269–293). Harmondsworth: Penguin.

Inbar-Lourie, O. (2010). English only? The Linguistic Choices of Teachers of Young EFL Learners. *International Journal of Bilingualism*, *14*(3), 351–367. https://doi.org/10.1177/1367006910367849

Izquierdo, J., García Martínez, V., Garza Pulido, M. G., & Aquino Zúñiga, S. P. (2016). First and Target Language Use in Public Language Education for Young Learners: Longitudinal Evidence from Mexican Secondary-School Classrooms. *System*, *61*, 20–30. https://doi.org/10.1016/j.system.2016.07.006

Kantzou, V., & Vasileiadi, D. M. (2021). On Using Languages Other Than the Target One in L2 Adult Language Education: Teachers' Views and Practices in Modern Greek Classrooms. *Journal of Lanugage and Education*, *7*(2), 155–174. https://doi.org/10.17323/jle.2021.11250

Kawafha, H., & Al Masaeed, K. (2023). Multidialectal and multilingual translanguaging in L2 Arabic classrooms: Teachers' beliefs vs. Actual practices. *Frontiers in Education*, 8. https://doi.org/10.3389/feduc.2023.1060196

Klein, W. (1986). *Second Language Acquisition*. Cambridge University Press. https://doi.org/10.1017/CBO9780511815058

Klette, K. (2023). Classroom observation as a means of understanding teaching quality: Towards a shared language of teaching? *Journal of Curriculum Studies*, *55*(1), 49–62. https://doi.org/10.1080/00220272.2023.2172360

Larsen-Freeman, D., & Anderson, M. (2011). *Techniques and Principles in Language Teaching* (3rd ed). Oxford University Press.

LeLoup, J. W., Ponterio, R., & Warford, M. K. (2013). Overcoming Resistance to 90% Target Language Use: Rationale, Challenges, and Suggestions. *NECTFL Review*, *72*, 45–60.

Levine, G. S. (2011). *Code Choice in the Language Classroom*. Multilingual Matters. https://doi.org/10.21832/9781847693341

Levine, G. S. (2014). Principles for Code Choice in the Foreign Language Classroom: A focus on grammaring. *Language Teaching*, *47*(3), 332–348. https://doi.org/10.1017/S0261444811000498

Lin, A. M. Y. (2017). Code-Switching in the Classroom: Research Paradigms and Approaches. In K. A. King, Y.-J. Lai, & S. May (Eds.), *Research Methods in Language and Education* (3rd ed., pp. 487–501). Springer, Cham. https://doi.org/10.1007/978-3-319-02249-9_34

Lombardi, A. R., Rifenbark, G. G., & Taconet, A. (2023). Quality Indicators of Secondary Data Analyses in Special Education Research: A Preregistration Guide. *Exceptional Children*, *89*(4), 397–411. https://doi.org/10.1177/00144029221141029

Ma, L. P. F. (2019). Examining the Functions of L1 Use through Teacher and Student Interactions in an Adult Migrant English Classroom. *International Journal of Bilingual Education and Bilingualism*, *22*(4), 386–401. https://doi.org/10.1080/13670050.2016.1257562

Macaro, E. (2001). Analysing Student Teachers' Codeswitching in Foreign Language Classrooms: Theories and Decision Making. *The Modern Language Journal*, *85*(4), 531–548. https://doi.org/10.1111/0026-7902.00124

Macaro, E. (2022). The L2-only versus the Multilingual Debate. In E. Macaro, & R. Woore (Eds.) *Debates in Second Language education* (1st ed., pp. 201–221). Routledge. https://doi.org/10.4324/9781003008361

Mangold, P. (2018). Discover the Invisible Through Tool-Supported Scientific Observation. (Translation of: Das Unsichtbare entdecken durch werkzeuggestützte wissenschaftliche

Beobachtung). In Böttger, H., Jensen, K., Jensen T. - Mindful Evolution. Conference Proceedings. Bad Heilbrunn: Klinkhardt.

Miri, M., Alibakhshi, G., & Mostafaei-Alaei, M. (2017). Reshaping Teacher Cognition About L1 Use Through Critical ELT Teacher Education. *Critical Inquiry in Language Studies*, *14*(1), 58–98. https://doi.org/10.1080/15427587.2016.1238286

Nakatsukasa, K., & Loewen, S. (2015). A Teacher's First Language Use in Form-focused Episodes in Spanish as a Foreign Language Classroom. *Language Teaching Research, 19*(2), 133–149. https://doi.org/10.1177/1362168814541737

Polio, C. G., & Duff, P. A. (1994). Teachers' Language Use in University Foreign Language Classrooms: A Qualitative Analysis of English and Target Language Alternation. *The Modern Language Journal*, *78*(3), 313–326. https://doi.org/10.2307/330110

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, *10*(1–4), 209–232. https://doi.org/10.1515/iral.1972.10.1-4.209

Spada, N. (2007). Communicative Language Teaching. In J. Cummins & C. Davison (Eds.), *International Handbook of English Language Teaching* (pp. 271–288). Springer US. https://doi.org/10.1007/978-0-387-46301-8_20

Storch, N., & Aldosari, A. (2010). Learners' Use of First Language (Arabic) in Pair Work in an EFL Class. *Language Teaching Research*, *14*(4), 355–375. https://doi.org/10.1177/1362168810375362

Storch, N., & Wigglesworth, G. (2003). Is There a Role for the Use of the L1 in an L2 Setting? *TESOL Quarterly*, *37*(4), 760–770. https://doi.org/10.2307/3588224

Tammenga-Helmantel, M., Mossing Holsteijn, L., & Bloemert, J. (2022). Target language Use of Dutch EFL Student Teachers: Three Longitudinal Case Studies. *Language Teaching Research*, *26*(4), 726–754. https://doi.org/10.1177/1362168820911195

Tekin, S., & Garton, S. (2020). L1 in the Primary English Classroom: How Much, When, How and Why? *Iranian Journal of Language Teaching Research*, *8*(3), 77–97.

Temesgen, A., & Hailu, E. (2022). Teachers' Codeswitching in EFL Classrooms: Functions and Motivations. *Cogent Education*, *9*(1), 2124039. https://doi.org/10.1080/2331186X.2022.2124039

Thompson, G. L., & Harrison, K. (2014). Language Use in the Foreign Language Classroom. *Foreign Language Annals*, *47*(2), 321–337. https://doi.org/10.1111/flan.12079

Tight, M. (2019). *Documentary Research in the Social Sciences*. SAGE Publications Ltd. https://doi.org/10.4135/9781529716559

Tsagari, D., & Giannikas, C. N. (2020). Re-evaluating the Use of the L1 in the L2 Classroom: Students vs. Teachers. *Applied Linguistics Review*, *11*(1), 151–181. https://doi.org/10.1515/applirev-2017-0104

Üstünel, E., & Seedhouse, P. (2005). Why that, in that Language, Right Now? Code-switching and Pedagogical Focus. *International Journal of Applied Linguistics*, *15*(3), 302–325. https://doi.org/10.1111/j.1473-4192.2005.00093.x

Vold, E. T., & Brkan, A. (2020). Classroom discourse in lower secondary French-as-a-foreign-language classes in Norway: Amounts and contexts of first and target language use. *System*, *93*, 102309. https://doi.org/10.1016/j.system.2020.102309

Wolthuis, F., Bloemert, J., Tammenga-Helmantel, M., & Paran, A. (2020). A Curriculum in Transition: TL/L1 Use in Dutch EFL Literature Lessons. *Language, Culture and Curriculum*, *33*(4), 335–350. https://doi.org/10.1080/07908318.2019.1697279

Yussof, N. T., & Sun, H. (2020). Mismatches between Teacher Beliefs, Practices and Reasons for English Use in Preschool Malay Language Classrooms. *Language and Education*, *34*(4), 363–382. https://doi.org/10.1080/09500782.2020.1720230

Zainil, Y., & Arsyad, S. (2021). Teachers' Perception of Their Code-Switching Practices in English as a Foreign Language Classes: The Results of Stimulated Recall Interview and Conversation Analysis. *SAGE Open*, *11*(2), 215824402110138. https://doi.org/10.1177/21582440211013802

Zhang, M. (2021). Understanding L1 and L2 Interaction in Collaborative writing: A lexico-grammatical analysis. *Language Teaching Research*, *25*(3), 338–359. https://doi.org/10.1177/1362168819859911

Zhou, W., & Li, G. (2022). Chinese Dual-language Bilingual Education Teachers' Pedagogical and Languaging Practices in American Immersion Schools. *Journal of Multilingual and Multicultural Development*, *43*(6), 569–583. https://doi.org/10.1080/01434632.2021.1996582

Vita

Manal Bani Humayyim, from Najran City, Kingdom of Saudi Arabia. She graduated in 2011 with a Bachelor of Arts in Education in English Language Studies from Najran University. In 2018, she completed a training program at the Ontario Institute for Studies in Education at the University of Toronto.

She was accepted in the Master of Arts program in Linguistic Studies with a concentration in Teaching Language (TESOL/TLOTE) at Syracuse University in the Spring of 2022. She presented a poster at the Language Educators of Central New York (LECNY) Conference in 2023 and two oral presentations at the TESOL, Applied Linguistics, Foreign Languages (TALFL) Conference at West Chester University in 2023.

She is a current EFL practitioner for the Ministry of Education in Saudi Arabia, her research interests lie in second language teaching and learning and pedagogy.