



UNIVERSITAT DE
BARCELONA

Department of Modern Languages and Literatures and
English Studies

M.A. Thesis

**Longitudinal Classroom-Based Research: A Synthesis of L2
studies**

Christina Bridgewater

Supervisor: Dr. Elsa Tragant Mestres

Applied Linguistics and Language Acquisition in Multilingual Contexts

Academic year: 2019-2020

**Màster Oficial en Lingüística Aplicada
i Adquisició de Llengües en Contextos Multilingües
LAALCM**

.....Elsa Tragant Mestres..... com a supervisor/a del treball (Tesina
de
(nom i cognoms)

Màster) presentat com a requeriment per a l'avaluació de l'assignatura **Projecte de**

Recerca en Lingüística Aplicada

presentat per l'alumne/a:**Christina Bridgewater**.....
(nom i cognoms)

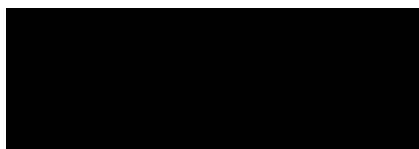
amb el títol de:

..... Longitudinal Classroom-Based Research: A Synthesis of L2 Studies.....

certifico que he llegit el treball i l'aprovo perquè pugui ser presentat per a la seva defensa pública.

I perquè consti i tingui els efectes oportuns signo aquest certificat en

Barcelona, a8..... desetembre..... de 2020.....



.....

Dr/a.

**Official MA programme in
Applied Linguistics and Language Acquisition in Multilingual Contexts (LAALCM)
Universitat de Barcelona**

Non-Plagiarism Statement

This form must be completed, dated and signed and must be included at the beginning of every copy of the MA Thesis you submit for assessment.

<i>Name and surnames:</i>	Christina Bridgewater
<i>MA Thesis title:</i>	Longitudinal Classroom-Based Research: A Synthesis of L2 Studies
<i>Supervisor:</i>	Elsa Tragant Mestres

I HEREBY DECLARE THAT:

- This MA Thesis that I am submitting for assessment is entirely my own work and I have written it completely by myself.
- I have not previously submitted this work or any version of it for assessment in any other programme or institution.
- I have not used any other sources or resources than the ones mentioned.
- I have identified and included the source of all facts, ideas, opinions and viewpoints of others through in-text referencing and the relevant sources are all included in the list of references at the end of my work. Direct quotations from books, journal articles, internet sources or any other source whatsoever are acknowledged and the sources cited are identified in the list of references.

I understand that plagiarism and copying are serious offences. In case of proof that this MA Thesis fails to comply with this declaration, either as negligence or as a deliberate act, I understand that the examiner has the right to exclude me from the assessment act and consequently all research activities conducted for this course will be declared null and the MA Thesis will not be presented for public defense, thus obtaining the lowest qualification.

Date:
8/9/20

Signature:



Abstract

Classroom spaces are common contexts second language acquisition. In addition, they provide congenial and manageable contexts for study of the second language acquisition process.

Because that process is inherently extended through time, classroom-based longitudinal studies should provide credible and useful information about the conditions, and trajectories over time, of second language learning. This thesis Is a research synthesis of such studies included in nine journals from 2017 through 2020. Using Ortega and Iberri Shea's (2005) chapter on longitudinal research in second language acquisition as a reference, results of analysis of the 27 qualifying studies are presented : with a focus on five dimensions: design, duration, focus change over time, methodological innovations and role of teachers and researchers. Further examination of these studies, categorized by their quantitative, qualitative, or mixed methodologies, is conducted in order to determine whether, within the context of classroom-based research, trends noted by Ortega and Iberri- Shea have remained constant, or have evolved. Noteworthy studies within each methodology are highlighted and innovation of design or methodology are discussed. We close with a bottom-up definition of longitudinal research and summary of some trends observed as well as possible implications for future research.

Acknowledgements

This thesis is dedicated to the memory of Sha Smith. You are sorely missed. Things are difficult without the third tripod leg, but I'm so grateful for the time we did have and our special moments. Thank you for always being a steadfast support (in studies and outside life), my best friend, and for always making me belly laugh in the midst of stress. This academic odyssey would have been much more of a struggle without your support.

I'd also like to thank my Advisor Dr. Elsa Tragant for taking this thesis journey with me. Thank you for the many round table brainstorms which eventually morphed into quarantined skype sessions. I'm so appreciative of the support and encouragement you've offered me at all points of this process, both as a mentor as well as a fellow human. Your patience is saintly and I'm glad to have worked together with you.

It would be remiss not to express how amazing my parents have been during the madness that is completing a thesis in the midst of a global pandemic. Thank for encouraging me in moments of self-doubt, your advice, your prayers, and dog photos.

To my classmates from my time in this program, I'm forever grateful for the friendships that have been built. Thank you for making this city 100 times better because of the people in it. My fondest memories will be the ones we made together.

I have such gratitude for my friends. Thank you for believing in me when I felt overwhelmed and for being there when I needed to vent or brainstorm out loud. Onur, you are a champion for dealing with my thesis in confinement craziness. I am so appreciative of your excellent company and thoughtful reminders to stay grounded and positive. Jordan you are truly such a gift. Thank you for answering my panicked messages and speaking truth to me. I couldn't ask for a better tripod leg. Andy, bless you for your suggestions on my rather slipshod first draft and for not passing too harsh of judgement.

Finally thank you to the coffee shop owners that allowed me to stay and continue working after they were closed because they saw I was in the flow of things (Morning Glory & Brew Coffee). You are sanctuaries for the tired and otherwise distracted at home. Thank you for providing me with a productive space when my go-to libraries were closed.

Table of Contents

1. Introduction.....	1
1.1 Longitudinal Research	2
1.1.1 Ortega & Iberri Shea's Previous Trends Within Longitudinal Research	4
1.2 Classroom-Based Research.....	5
1.2.1 The Language Classroom	5
1.2.2 Classroom-Based Research.....	5
1.2.3 Action and Collaborative Action Research.....	6
2. Methodology	7
2.1 Selection:.....	7
2.2. Analysis:	9
3. Results of the Review	9
3.1 Overview	9
3.2 Quantitative Longitudinal Classroom-Based Studies	11
3.2.1 Design	11
3.2.2 Study Duration	12
3.2.3 Role of Researcher and Teacher	13
3.2.4 Focus on Change Over Time	13
3.2.6 Innovation Within Quantitative Methodology	14
3.3 Qualitative Longitudinal Classroom-Based Studies	15
3.3.1 Design	15
3.3.2 Study Duration	17
3.3.3 Role of Researcher and Teacher	18
3.3.4 Focus on Change Over Time	19
3.3.5 Innovation Within Qualitative Methodology	19
3.4 Mixed Methods Longitudinal Classroom-Based Studies	20
3.4.1 Design	20
3.4.2 Study Duration	21
3.4.3 Role of Researcher and Teacher	22
3.4.4 Focus on Change Over Time	22
3.4.5 Innovation Within Mixed Methodology	23
3.4.6 General Observations.....	24
4. Discussion	24

5. Concluding Remarks.....	27
References.....	29
Appendices.....	34
Appendix B: Journal Totals	34
Appendix B: Longitudinal Classroom-Based Studies Included for Synthesis	35

1. Introduction

This study was prepared for a symposium at the 19th AILA World Congress with the title: *Language teaching and learning in the classroom: challenges and opportunities for longitudinal research*. The synthesis will be presented at the symposium with the hope to document patterns and trends within recent research in addition to identifying strengths and limitations found in the literature, that ought to be given further attention in forthcoming research.

This work has been inspired by a previous synthesis in Ortega and Iberri-Shea's (2005) “Longitudinal Research in Second Language Acquisition: Recent Trends and Future Directions,” which investigated longitudinal research trends in Second Language Acquisition (SLA). The article examined 38 studies, published between 2002-2004, and divided them into different groups to better summarize trends within each. Beginning with the trends they outlined, the present study adopts their synthesis as a framework, but narrows the scope to longitudinal research in classroom-based scenarios. Specifically, such studies published from 2017- May 2020 have been considered. The aim of this research is to create a synthesis, similar to Ortega and Iberri-Shea's, that will be a beneficial reference for current and future researchers planning to conduct longitudinal classroom-based research.

Within applied linguistics, the bridge between research and classroom implementation of findings is not always clear. This is unfortunate, because increased communication and collaboration would greatly benefit researchers, teachers, and students. Research conducted within the context of a classroom setting provides valuable insight and assures a higher ecological validity of the data collected and the implications derived therefrom. (Loewen & Plonsky, 2016). It allows us the opportunity to examine learning within an authentic setting and recognize beneficial directions in which to advance research. Along with the benefits of classroom-based research, adding a longitudinal component is advantageous in allowing us to document the development of language in the setting over time (Dörnyei, 2007) This research is, of course, accompanied by challenges relative to maintaining participant numbers among others. Nevertheless, research that takes place in a classroom setting and uses longitudinal data or opts for a longitudinal design can provides a more comprehensive understanding of learning and teaching.

In the following sections (1.1, and 1.2) terms relevant to the longitudinal classroom-based research domain will be outlined. These definitions inform the methodology implemented for

selection of the studies in this review which is discussed in section 2. Subsequent observations from analysis of studies, categorized by qualitative, quantitative, and mixed methodology, will be explicated and connections to past trends seen in Ortega and Iberri-Shea will be explored (see section 3). The hope is that these observations will identify patterns and give insight into the attributes of studies recently published. Concluding remarks and recommendations are provided in section 4.

1.1 Longitudinal Research

The definitions used for “longitudinal” research can be at times ambiguous. There seems to be no universal consensus of what classifies a study as longitudinal for all frameworks. While Loewen and Plonsky, (2016) posit that if a study takes place over one to two weeks or longer it can be considered longitudinal, most definitions omit temporal specification. Perhaps this can be attributed to the lack of a fixed temporal framework within the field or the bureaucracy involved with securing funding and negotiating longer spans of observation and data collection.

Paltridge and Phakiti (2015) adopt a brief definition of longitudinal research in their glossary as “a study in which data is collected from the same participants at more than one point in time. It is carried out over a relatively long period of time, often with a small number of participants” (p. 567) which is not necessarily in agreement with studies observed in past synthesis (Ortega & Iberri-Shea, 2005).

Looking at a denser conception, Dörnyei (2007) characterizes longitudinal research as “a family of methods that share one thing in common: information is gathered about the target of research (which can include a wide range of units such as people, household institutions, nations, or conceptual issues) during a series of points in time” (p.79) In addition, Dörnyei also considers retrospective studies that gather information on change and development as longitudinal research. Dörnyei later cites Menard (2002) as establishing the following parameters for longitudinal research: “a) data collected for 2+ distinct time periods; b) the subjects or cases analyzed are the same or comparable (drawn from the same population) from one period to the next; c) the analysis involves some comparison of data between periods. ...(analysis) would permit the measurement of differences or change in a variable from one period to the next (p. 2).”

Jackson (2017) outlines some factors that should be considered within longitudinal studies: appropriate length, frequency of data collection, attrition of participants, and informed,

strategic selection of analysis (e.g., qualitative, quantitative, or mixed methods). These factors prompt the following lines of inquiry:

How long is enough?

The length of the study should be appropriate for the measurement of the selected phenomenon or variables. There should be regard for the possibility that a study with a longer timeframe of observation and data collection may reveal claims and findings contrary to those of a study with more limited length.

How often is necessary?

In order to draw conclusions about change over time, the frequency of data collection needs to be considered. Too infrequent collection, or neglect of important milestone moments in learning will compromise the reliability of final analysis (Barkaoui, 2014). Singer and Willet (2009) assert that multiple data collection waves (3+) in longitudinal studies allow for more complex statistical models and, as such, offer a richer portrait of the trajectory of change. However, if data collection is too frequent and intrusive, or goes over an extended period of time, consequences such as fatigue and attrition may imperil the validity of the data (Barkaoui, 2014; Jackson, 2017; Ortega & Iberri Shea, 2005).

How feasible is it to sustain long-term participant involvement?

A major challenge within longitudinal research is maintaining participant involvement over the course of the study. While longer studies allow for more accurate analysis of trends in learning, they generally struggle with attrition among their pool of subjects as time passes. For this reason, while not always the most theoretically adequate time frame, many studies are limited to institutional time units (semesters, quarters, academic years) to conduct research (Barkaoui, 2014; Ortega & Byrnes, 2008).

What research approach should be taken?

Given the wide spectrum of analyses and forms of data collection that can be used within longitudinal research, it is important that researchers carefully consider the best approach for their focus (Jackson, 2017).

Longitudinal research is of particular value in SLA and foreign language acquisition (FLA) research because it allows us the opportunity to observe development (or sustainment) of a variety of aspects of language over time (Jackson, 2017), and within specific individuals

(Loewen, & Plonsky, 2016). Compared to conclusions drawn from cross-sectional data collections, the longitudinal design may allow for more confidence in the quality and consistency of data collected; thereby causing the analysis, and the conclusions drawn, to have broader application. Longitudinal research lessens the uncertainty that accompanies results from cross-sectional, one-shot, design studies by allowing an opportunity to confirm sustained results and continued development (Jackson, 2017; Ortega & Iberri Shea, 2005).

1.1.1 Ortega & Iberri Shea's Previous Trends Within Longitudinal Research

Ortega and Iberri Shea (2005) outline four broad trends within longitudinal studies published at the time of their analysis: 1) descriptive-quantitative longitudinal studies of L2 development, 2) longitudinal research on L2 program outcomes, 3) longitudinal investigation of L2 instructional effectiveness, and 4) qualitative longitudinal SLA research (i.e., sociocultural SLA longitudinal studies, and ethnographies of L2 learners). Each of these trends has distinct features that attest to the versatility of a longitudinal design. They found the studies which focused on L2 development (trend 1) ranged from 4 months to 4 years, and tracked quantifiable variables from a small pool of participants multiple times during the timeframe. The strength of this type of study, is that they accommodate for the effect of time on given variables as well as the cause and effect relationships present between particular variables in different contexts. In contrast the second trend of studies that looked at outcomes of L2 programs and curriculum designs tended to draw from larger samples of participants, span over longer time periods (four to six years of institutional time) and have less frequent data collection. Such studies are of benefit because they give evidence of the success or failure of curriculum design trends that are being implemented. Because these studies are oftentimes focused on institutionally implemented programs (e.g., CLIL in Spain, or Canadian French immersion programs), they offer an enriched perspective on language development and the effectivity of these programs. The third trend observes studies largely focused on instructional effectiveness, involving an intervention which on average lasted eight weeks. These studies included the then growing trend of delayed post-tests one to three months following the initial post-test. On the qualitative end of studies (trend 4), ethnographies and sociocultural perspectives were the most prevalent; normally ranging in length from 15 weeks to 3 years and lacking a homogeneity in focus. An issue observed within some qualitative studies taking an ethnographic approach was that while they may have involved longer

timeframes, and lacked consideration of the role of change, making them incongruent with general definitions of longitudinal research resulting in dense but stagnant analysis.

1.2 Classroom-Based Research

1.2.1 The Language Classroom

For purposes of defining the concept of a classroom, Dörnyei (2007) understandably proposes a broad definition of “any physical space in which scheduled teaching takes place” (p. 176). Collins and Munoz (2016) elaborate this definition postulating that language classrooms have three main features: “It is a physical space that serves as learning environment and is bounded in time; it is managed by a facilitator who normally has expertise in the FL and in FL pedagogy; and it is populated by groups of people who share the common purpose of learning/using the target L2 (although their individual learning goals may differ)” (p.134). This definition omits individual or one-on-one learning scenarios (tutoring), as well as non-traditional classroom settings (writing centers, daycare centers, online courses). Understanding that the concept of a classroom is constantly evolving due to innovation in learning, the current study nonetheless adopts Dörnyei (2007) and Collins and Muñoz’s (2016) material interpretations of a classroom, for the sake of standardization and systematic synthesis.

1.2.2 Classroom-Based Research

Taking into consideration the just-discussed definition of “classroom,” a “classroom-based” study suggests a formal investigation conducted with the classroom as the main research site and focuses on how teaching and learning take place within this context (Dörnyei, 2007). While Williams (2013) adopts a similar definition of classroom research, she limits the meaning of “classroom-based” by specifying that classroom-based research examines the role of the teacher and of inter-learner dynamics. However, other definitions that are broader in scope allow for more variation of research focus. They look at classroom-based research as an investigation by external researchers and/or teachers that takes place within the classroom space and examines what occurs within the classroom when learners and teachers come together (Allwright, 1983; Gass & Mackey, 2007; Loewen & Plonsky, 2016).

Contemporary research in the field looks toward “understanding of learning, documenting, and analyzing the dynamic interplay of various classroom processes and

conditions that contribute to variation in learning outcomes” (Dörnyei 2007 p 178) This investigation into the classroom may range from a qualitative, holistic approach where the classroom environment’s naturalistic occurrences are of interest (e.g., ethnography, case study), to a more quantitative and quasi-experimental approach noting the measurable changes that occur as a result of intentional intervention or curriculum design (Loewen & Plonsky, 2016; Mackey, 2017).

Classroom-based research is important because it creates a bridge and open dialog between the research of academics and the practicalities encountered by classroom language instructors that leads to more relevant and applicable investigation and findings. Thoughtfully designed language classroom research help us not only examine theoretical assumptions but also what transpires in true educative practice within a language learning classroom (Harbon & Shen, 2015; Loewen & Sato, 2017). When classroom-based research is conducted in cooperation with or solely by the teacher, there is the additional benefit of learner comfort and familiarity which facilitates the integration of the treatment or data collection into the normal classroom domain (Loewen & Sato, 2017)

1.2.3 Action and Collaborative Action Research

The action research model can be envisioned as a helix of repeated processes of inquiry where the teacher is the main conductor (Stringer, 2008). It begins with the careful observation, and systematic collection, of data concerning the everyday classroom. These observations can then be analyzed to identify features and elements of concern or interest. Considering this new information, solutions can be devised and decisions about future practice can be made. The process is intended to be ongoing and repeated to foster a classroom of ongoing innovation (Wallace, 1998). Action research is that which aims to gain a better understanding of the learning environment and recognize the ways in which it can be improved. It is particularly valuable because it provides valuable insight and draws applicable connections that can be of instant benefit to the successful running of a classroom.

Two main approaches exist within the realm of action research. Traditional action research is a means of professional learning for language teachers. Taking a socio-constructivist approach, teachers are viewed as having the role of actors and investigators within their personal contexts of instruction and learning (Burns, 2019). By releasing the investigative reins to the

teacher, they are able to enhance their teaching and integrate inquiry into the classroom process with the aim of improving students' learning experience (Kemmis, 2014; Stringer, 2008). As self-conducted research within a classroom can at times be seen as laborious and time consuming for a teacher, a contemporary, alternative approach to action research, known as "collaborative action research," is often seen as a more realistic mode of teacher research. It involves cooperation between colleagues or an outside researcher to design, implement, and analyze research following a practice to theory trajectory rather than theory to practice (Burns, 1999; Wallace, 1998). In longitudinal classroom-based research, the relationship between those in the roles of researcher and teacher is of particular importance, due to the potential complexities that accompany extended collection of data within the classroom. Thus, action research with teacher as researcher or collaborative action research lends itself to the methodology.

2. Methodology

This study will conduct a systematic review of recent longitudinal classroom-based research. The review is intended to recognize thematic patterns within the methodologies used and areas of interest in recent research in order to illuminate directions and techniques for future research (Li & Wang, 2018). The selection process for the review was approached following the steps specified in a book chapter titled Research Synthesis by Ortega (2015) in *Research Methods in Applied Linguistics*). First, the scope was defined by examining different definitions of the terms "Classroom-based Research" and "Longitudinal Research." These definitions were used to create a standard by which to examine the literature and decide which studies would be included, and why. The following section describes this initial process in the present research review.

2.1 Selection:

The literature search was limited to online and published journal articles from nine major applied linguistics journals: *Applied Linguistics*, *Language Learning*, *Language Teaching Research*, *Second Language Research*, *TESOL Quarterly*, *The Modern Language Journal*, *Annual Review of Applied Linguistics*, *Journal of Second Language Writing*, and *Studies in Second Language Acquisition*. The journals were chosen for their high impact factors, ranging from 1.75- 4.2 (see table 1) . Masters theses, books, and unpublished materials were excluded from the review. Thus, the following synthesis should be read with this caveat in mind, and with the understanding that

the information that follows applies more specifically as a synthesis for research published in journals.

The review began by searching the online platforms of the journals for the term “longitudinal,” limiting our timeframe from 2017 to the present (as of May, 2020). We included studies which had been published in volumes as well as those available only online at the time of writing this synthesis. Similar to the methodology of Ortega and Iberri-Shea (2005) and Marsden et al (2018), only studies where the authors self-labelled their study as longitudinal and used the term in reference to the design were included. Studies that might qualify as longitudinal but made no statements using the term were excluded to create a more standardized selection process. Out of the roughly 640 articles published in the years under review (9 journals x 3,5 years x 4 issues x 5 articles on average), 97 were self-labelled as longitudinal. Sixty-seven additional studies did not have a longitudinal design, but included the term ‘longitudinal’ in the main body of the text, and were noted to have expressed the need for future longitudinal research. The 97 studies with longitudinal designs were then read more closely to identify if they were classroom-based and fulfilled a set of additional criteria: First, each study was examined to ensure that it focused on second or foreign language learners, excluding naturalistic or uninstructed foreign language learning. The studies were further filtered to clarify that the studies’ participants were teachers and/or students learning in a classroom setting (this would include studies concerning study abroad programs which collected data from the participant’s classrooms to analyze). All studies included in the review occurred within a traditional classroom setting, thereby omitting studies conducted within virtual foreign language classrooms, studies where the students were removed from the class during data collection and only tested in an isolated or lab setting. Studies were also excluded from analysis that concerned experience (such as study abroad) rather than data collected within or pertaining to the foreign language classroom. Similar to DeKeyser & Prieto Botana (2019), the studies selected did not include “...studies that were carried out with classroom learners (and teachers) but where nothing of what happened in the classroom was a variable in the study” (p. 2). In the studies selected, the experiments and intervention should be integrated into the content of the classroom or curriculum or the study focus should look at elements of regular classroom curriculum processes.

Resulting from the above selection process, a total of 27 articles that were longitudinal and class-based were identified. *Tesol Quarterly* was found to have published the most longitudinal classroom-based research, with seven studies being published in the 2017-May 2020 time frame. In comparison, the two journals, *Second Language Research* and *Studies in Second Language Research* were not found to have published any. Furthermore, these two journals also only published four studies of longitudinal design. This disparity suggests that some journals (i.e., these two journals in particular) may prefer to publish research with cross-sectional designs taking place in controlled contexts (see Appendix A). The remaining 6 journals were found to have published between one and six articles within the 2017-2020 time frame.

2.2. Analysis:

The 27 studies which met criteria for inclusion were then analyzed more closely for identification of particular characteristics including context, design, and certain methodological attributes. Looking at these characteristics, the studies were divided into three subgroups: mixed methods, qualitative, and quantitative design. This categorization resulted in 11 studies of qualitative design, 10 of quantitative, and 6 of mixed (see Appendix A). Using these subgroups, patterns were noted and particular attention was given to study design, duration, roles of the researcher and teacher, focus on change over time, and innovation noted within each group of studies.

3. Results of the Review

3.1 Overview

Appendix B offers an overview of the 27 studies under review including such aspects as educational context, target languages, countries where research was conducted, and areas of research represented in recent publications. of The majority of the selected studies, 23 in total, were conducted in courses for learning English, however Chinese (Lee & Kinginger, 2018), Swedish (Cekaite, 2017), Spanish (Menke & Strawbridge, 2019), Arabic and Hebrew (both in Schwartz & Gorbatt, 2017) were also represented as target languages. The analysis found that 10 of the studies occurred within the context of a Second Language (SL) classroom, compared to 16 in a Foreign Language (FL) context. Teng and Zheng (2020) used the terms ESL and EFL interchangeably, making it difficult to categorize their study. Four studies had particularly unique

contexts. Kibler (2017), and Kibler and Hardigree (2017) during their eight-year long studies, followed writing development of a single native Spanish speaker who began in ESL classes but eventually joined mainstream classes. Zheng (2017) took an ethnographic case study perspective that followed the experiences of non-native English teaching assistants instructing college composition classes to native speakers. Nguyen (2018) investigated the interactional practices of a non-native English speaker studying pharmacology including role plays in a communication course and recorded consultations during a clerkship at a community pharmacy.

There was a clear predominance of studies (thirteen in number) solely conducted at a university level. This saturation of studies within a university context represents a trend that is echoed in SLA research in general (DeKeyser & Prieto Botana, 2019). This may be explained by the fact that most researchers have positions in higher education making these settings and participant groups ones of convenience. In addition, university instructors tend to have more flexibility and authority with regard to the curriculum of their classes so. Thus, it can be easier at a university level to adapt plans to accommodate outside researchers or the instructors' own research endeavors than in lower levels of education (Polio & Lee, 2019, p. 148).

Two studies had designs that included data collected in both university and secondary school settings (Kibler, 2017; Kibler & Hardigree, 2017). One study involved data collected in university class and an outside clerkship (Nguyen, 2018). The second largest group of studies, numbering seven, were carried out in secondary school. Language schools, with two studies, and primary schools and preschools, with one study each, were the least represented contexts.

The largest number of studies, twelve, were carried out in the United States. Even so, a variety of other countries were represented in the literature used for this study including: Japan (3), Korea (2), China (2), Sweden (2), Netherlands (2) and Poland, Taiwan, Canada, and Israel, which each had one study. This array of international contribution to the field is exciting as it shows a diversification of arenas for conducting longitudinal classroom-based research.

Only five studies (Sato et al., 2019; Schwartz & Gorbatt, 2017; Tammenga-Helmantel et al., 2020; Worden, 2018; Zheng, 2017) focused primarily on teachers in contrast with 22 which focus primarily on the student as subjects.

Regarding areas of research there is a predominance of studies (12) which involve some aspect of analyzing writing and/ or take place in the context of a writing course. In many institutions students are already asked to write essays regularly, thus making these sources more

readily available or convenient for data collection. However, this observed trend may not be of merit as five of these studies were extracted from a writing focused journal.

3.2 Quantitative Longitudinal Classroom-Based Studies

In order to further categorize the studies which met the criterion for longitudinal classroom-based studies, the following characteristics were adopted to analyze research specifically of a quantitative design

- Descriptive quantitative studies of L2 development (Menke & Strawbridge, 2019; Vercellotti, 2017; Yoon & Polio, 2017)
- Intervention-based studies (involving two or more conditions or an intervention) (Aubrey, 2017; Kim & Emeliyanova, 2019; Lin & Lee, 2019; Otwinowska et al., 2020; Rahimi, 2019; Teng & Zhang, 2020a)
- Correlational studies (Lowie and Verspoor)

A total of ten studies exhibited these characteristics.

3.2.1 Design

There were six intervention-based studies. Three of these studies involved teacher-led interventions such as three groups of differing percentages of data-driven learning (DDL) and traditional deductive approach (TDA)(Lin & Lee, 2019), writing instruction using Self-Regulated Strategy Development model to instruct the targeted SRL strategies (Teng & Zhang, 2020), and training of participants in noticing cognates and searching for cross-linguistic similarity (Otwinowska et al., 2020). The other three involved performing tasks under different conditions including: task based interactions of students in intercultural conditions with their Japanese peers versus intracultural conditions where students were partnered with an international student (Aubrey, 2017), the impact of focused versus comprehensive corrective feedback and revision on ESL learners' writing (Rahimi, 2019), and comparing collaborative or individual revision of WCF (Kim & Emeliyanova, 2019). Only four of the above six studies used control groups in their studies nonetheless, they are also quasi-experimental, given the use of established intact classes over the ideal experimental randomization of individuals into groups.

There were three descriptive-quantitative studies focusing on individual development (Menke & Strawbridge, 2019; Vercellotti, 2017; Yoon & Polio, 2017). In Yoon and Polio and

Vercelotti participants (n=37 and n=66 respectively) were analyzed as a group while in Menke and Strawbridge participants (n= 3) were analyzed individually. Of these studies, only Yoon and Polio implemented the use of a control group (n=46 native speakers).

As previously discussed, one issue with longitudinal research design is the increased likelihood of attrition. Several studies reported the loss of some participants over the duration of the study. Otwinowska et al. (2020), stated that the 44 participants did not include 4 students which were excluded due to absences. Kim and Emeliyanova's (2019) 36 participants had initially numbered 45 but 9 were ultimately excluded due to attrition. Other studies included no explicit information about attrition.

Nevertheless, it should be noted that none of the studies had particularly low sample sizes, except for Menke and Strawbridge (2019), which quantitatively analyzed the texts of three Spanish studies undergraduate students for development of syntactic complexity. Viewing Menke as exceptional in its small sample size, the next smallest number of participants was 36 (Kim & Emeliyanova, 2019) and the largest 80 (Teng & Zhang, 2020). Overall, the mean sample size (excluding Menke) was 56.14 participants

3.2.2 Study Duration

The length of data collection for all the studies of quantitative methodology ranged from 6 weeks to 10 months with a mean length of 15 weeks. Nevertheless, several of these studies acknowledged time limitations. For example, Kim & Emeliyanova (2019), which was 8 weeks in duration and involved three error correction sessions, found their number of treatment sessions to be a significant limitation of the study. Even Vercelotti (2017), the longest study spanning three academic semesters, noted that, with regard to future research, a more extended study might find different trajectories.

If we compare the duration of the three descriptive-quantitative longitudinal studies of L2 development in this corpus with those in Ortega and Iberri-Shea we find some consistency in regards to duration. Their survey of studies found those of this trend ranged from four months to four years, the studies identified herein collected data according to the institutional measures of their educational contexts (semesters), consequently ranging in duration from one to six semesters—roughly three months to 3 years.

Lowie and Verspoor (2019), whose innovative design separated it as a correlational study, collected data for seven months; the course of an academic year.

The six studies which looked at interventions conducted in the classroom spanned timeframes of five weeks to five months, with an average of 10 weeks. This was a slight increase from, Ortega and Iberri Shea's observed trend of longitudinal investigation of L2 instructional effectiveness which reported interventions of eight weeks duration. All of these studies compared intact groups under different conditions or undergoing different interventions.

3.2.3 Role of Researcher and Teacher

Only one of the 10 quantitative studies could be considered action research given that the researcher explicitly acknowledged being one of the four teachers involved (Rahimi, 2019). Five of the studies involved collaboration with the teacher to administer interventions (Lin & Lee, 2019; Otwinowska et al., 2020; Teng & Zhang, 2020), or implement different conditions of learning into their curriculum (Aubrey, 2017; Y. J. Kim & Emeliyanova, 2019). Only one study seemed to have the researcher observing or analyzing texts apart from collaboration with teachers or the students (Yoon & Polio, 2017). The remaining three studies (Lowie & Verspoor, 2019; Menke & Strawbridge, 2019; Vercellotti, 2017) involved analysis of data generated in the class room, but gave no indication of relationship between researcher and teacher.

3.2.4 Focus on Change Over Time

The three descriptive-quantitative studies (Menke & Strawbridge, 2019; Vercellotti, 2017; Yoon & Polio, 2017) focus on change over time as they all aim to describe L2 development over time. Within these studies, measurement of this development is done by the collection of several pieces of data (writing samples or recorded oral tasks) from each student. over the duration of the study. Both Vercellotti (by means of 3-7 speaking samples per subject) and Yoon and Polio (analyzing 6 essays per subject) investigate how measures of complexity, accuracy, lexical complexity, and fluency develop for individuals over the course of their prospective studies. Menke & Strawbridge followed the development of syntactic complexity for their subjects by collecting texts throughout the course of their studies.

Among the intervention-based studies, Aubrey (2017), Rahimi (2019), and Teng and Zhang (2020) have a focus on change over time as one of their main aims. Lin and Lee (2019) have a secondary focus on change over time and report student' shifting opinions toward their prospective intervention groups (TDA or two different blends of data driven learning (DDL) with TDA). Otwinowska et al. (2020) and Kim and Emeliyanova (2019) instead look at the

differences between two intervention conditions and use time to have contextual acumen rather than to document change.

While several studies incorporated analysis that could also be used with cross-sectional data, several stood out for their analysis fitted to and exclusively appropriate for longitudinally collected data. Ortega and Iberri-Shea (2005) noted the promise of using time series design to explore longitudinal effects of instruction, but at the time of their review, could only find a singular example (Ishida, 2004). For the present analysis, if we interpret the term ‘time-series design’ as studies where one or more collections of data occur between the start of the intervention and its end, then it can be said that three studies implemented this design (Aubrey, 2017; Rahimi, 2019; Yoon & Polio, 2017).

Interestingly, in the majority of these studies productions were analyzed but process data, which would document change was not examined. Only in Kim and Emeliyanova (2019) were students recorded while revising in one of the conditions (pairs). However, these recordings don’t appear to have been analyzed in the study. Otwinowska et al. (2020) and Teng and Zhang (2020) additionally would have been appropriate studies to investigate process, because they involved teacher led interventions, however, the intervention was not the focus.

3.2.6 Innovation Within Quantitative Methodology

Presenting as a correlation study, Lowie and Verspoor (2019) explored the role of individual differences (ID) by means of both a group study and 22 longitudinal case studies. This design was implemented with the intention of investigating to what extent results from group studies on ID could be generalized on an individual level as well as contrariwise. Regression analysis was run between holistic and analytic scores given to texts written by students over the course of an academic year, as well as between the results of a survey measuring ID (aptitude and motivation.) and calculated proficiency gains. Overall, ID were not shown to be significant predictors of final ratings received on texts. However, it is revealing that there was little congruence between individual and group trajectory analysis. The incorporation of both group and case study analysis, provides findings that would otherwise be incomplete and perhaps misleading without the complement of the other. Group studies provide insight into the broad scope influence of ID on L2 development, but should be viewed with consideration of the fact that findings may not be representative of longer trajectories, or trajectories of individuals.

Conversely, results of case studies should not be generalized to larger populations because they are unique to the individual. This understanding contributes to the larger realm of longitudinal classroom-based research in that it suggests both group and individual case studies are needed when exploring L2 development

Although attrition is a common issue within longitudinal studies, on close examination, attrition appeared to be less problematic in Vercellotti because the study made use of hierarchical linear and non-linear modeling in analysis (HLM). The use of this analysis is original because, it analyzes and details individual trajectories over time rather than averages of groups at single points. HLM was particularly appropriate for this study since not all students remained for the three semesters the study lasted. Their approach is particularly innovative in this sense because it eliminated the need to exclude students from the sample.

3.3 Qualitative Longitudinal Classroom-Based Studies

Definition parameters for the studies included in this grouping were adapted from De Costa, Li, and Rawal, (2019) who extracted their points from Benson (2012) and Friedman (2012).

- Analytical, descriptive in nature, and aims to explain phenomena which occurs in the classroom. As it relies heavily on observation data, coding and analysis usually emerge from the data collected rather than being previously established. Usually the variables are not manipulated as the aim is to collect more naturalistic data.
- Multiple perspectives and data collection methods are used in order to validate findings (i.e., class observations to validate both student and teacher experience as revealed by questionnaire).
- Report findings that are context specific. They may range from case studies of individual student development within a particular classroom to studies spanning several years and learning contexts. They look at what the data reveals on a micro individual scale, while also considering possible macro level influences, be they cultural (ethnographic observation), or institutional circumstances, attitudes, and policy (examining norms and regulations of the educational context).

Eleven of the studies in our corpus fall into this group.

3.3.1 Design

The qualitative studies in this review all favor a case study approach and some of them have adopted a range of analytical frameworks such as investment (Jiang, Yang, & Yu, 2020), activity theory (Lee & Kinginger, 2018), or systemic functional approach to multimodal discourse analysis (Shin et al., 2020). Two studies in particular adopt approaches worth highlighting given their specificity towards longitudinal data: longitudinal histories approach in Kibler (2017), and learning behavior tracking in Kunitz and Marian (2017).

The number of participants in the studies within this group was significantly smaller than those that were in the quantitative studies. Studies ranged from one to three individuals. Seven focused on a single subject, two tracked and compared two subjects (Cekaite, 2017; Zheng, 2017), and two looked at three subjects in the same (Kunitz & Marian, 2017) or similar (Tammenga-Helmantel, Holsteijn, & Bloemert, 2020) contexts. Of particular note were the studies that reported more in depth about the instructors, recognizing their role as secondary subjects as well as valuable resources for analysis (Jiang et al., 2020; Lee & Kinginger, 2018; Shin et al., 2020).

Lee and Kinginger (2018) stood out among these studies because they also conducted and provided analysis based on secondary sources of interviews with the subject's teacher, as well as his eight classmates. These secondary source interviews allowed for a multifaceted portrayal of both the classroom context and how the study subject was situated role-wise in relation to his peers throughout his process of adjusting and reintegrating into the Chinese language classroom environment at his university following a study abroad experience in China.

Six studies note that they are parts of larger projects, or state that only a few samples were selected for inclusion in the analysis out of larger samples of data (Cekaite, 2017; Jiang et al., 2020; Kibler, 2017; Kibler & Hardigree, 2017; Shin et al., 2020; Worden, 2018). Cekaite (2017) states that out of the ninety sessions that were recorded they reported on only a selection of the interactions which included interactional participation of the two girls of focus (p.48). Similar selective samples were noted by Ortega and Iberri- Shea's review and may be due to the extensive involvement that a qualitative approach to data collection and analysis entails. (However, it may also be due, as Ortega and Iberri-Shea (2005) surmise, to inevitable pressure within academia to publish at a rate not conducive to extended length, collaborative projects.)

The lack of homogeneity of these qualitative data can be seen by looking at the sources employed. As summarized in Table 2, the primary data sources for four of the studies in this

group were recordings which were later transcribed, coded, and carefully analyzed. Three relied heavily on writing samples and conducted interviews for further data and insight. The other four studies made use of multiple sources of data with observations often used for triangulation of data and to serve as a backdrop of contextualization for data gathered from other sources.

Table 2: Data Sources of Qualitative Longitudinal Classroom-Based Studies

	Primarily based on classroom recordings.	Primarily based on writing samples (+interviews)	Various Sources
Cekaite, 2017	x		
Jiang et al., 2020			Primary interviews, conversations Triangulated with observations and reflections
Kibler & Hardigree, 2017		x	
Kibler, 2017		x	
Kunitz & Marian, 2017	x		
Lee & Kinginger, 2018	x		
Marjon Tammenga-Helmantel et al., 2020			Observations paired with surveys
Nguyen, 2018	x (and clerkship recordings)		
Shin et al., 2020		X- digital	
Worden, 2019			Interviews, stimulated recalls, class observations, concept maps, instructional artifacts
Zheng, 2017			Interviews, observations, survey, course documents

3.3.2 Study Duration

While all of these studies self-identified as longitudinal, data collection length for these studies varied from as short as six days (Kunitz & Marian, 2017) to up to 8 years (Kibler & Hardigree, 2017). Three studies (Cekaite, 2017; Lee & Kinginger, 2018; Nguyen, 2018) reported duration in terms of lengths of institutional time (e.g., semester, academic year). Assuming the institutional periods were of standard lengths, the general trend was for most of the studies having a duration of between 10 weeks and one year.

Viewing Kunitz & Marian's (2017) six day timeframe as exceptional, the next shortest study was ten weeks (Zheng, 2017), a timeframe more within more traditional understanding of “longitudinal”. The study looked at the experiences of international teaching assistants teaching a college composition class. The data gathered is presented thematically and provides rich

descriptions of each individual's experiences in their role. The two subjects' experiences are sometimes compared and a cultural commentary is proposed to explain their respective differences. The duration of the study was used as a tool to gather interview and observational data for analysis and for recognition of the strengths and struggles of each individual as a teacher, rather than as a variable by which to report changes in their perspectives and approaches as teachers, or changes that occurred in their classrooms.

A limitation seen in Nguyen (2018), Kibler (2017), and Kibler and Hardigree (2017), were the gaps during the duration of the studies with regard to data collection. Both Kibler studies collected data during only six out of the full eight years. Nguyen collected recorded data for one semester and then again after a year for two months during a clerkship. While these gaps may be due to convenience, they do create space for questions about the validity of the trajectories observed and what undocumented change may have transpired in these pauses of data collection. They also create gaps in the tracking of the process of subjects in context, creating holes in the full picture of the subject's development.

3.3.3 Role of Researcher and Teacher

Two studies made statements of collaboration (collaborative action research) with the teachers (Jiang et al., 2020; Shin et al., 2020). This collaboration included teacher involvement with the larger research project from where the article stemmed (Jiang et al., 2020), and teacher involvement with the design, implementation of a multimodal task, and generation of feedback exchanges which were one of the main domains of data analyzed (Shin et al., 2020). The remainder of the studies commented on the researcher conducting the study as an observer or outside analyst. This was surprising given that qualitative research would be more credible if conducted by the teacher-researcher as an insider within the classroom context.

In De Costa et al.'s chapter on Qualitative Classroom research (2019), and TESOL quarterly's guidelines for qualitative research (Chapelle, & Duff, 2003), there is a concerted call for reflection on behalf of researchers about their role in the studies they conduct. This role should be explicitly accounted for and as equally analyzed as that of the participants involved. Nonetheless, only five studies (Jiang et al., 2020; Kibler, 2017; Kibler & Hardigree, 2017; Shin et al., 2020; Zheng, 2017) made any comments about the researcher's role or contribution and of those studies only three (Kibler, 2017; Kibler & Hardigree, 2017; Zheng, 2017) contained

detailed accounts of the researcher's contributions as players in the broader context of their investigations.

3.3.4 Focus on Change Over Time

Ortega and Iberri-Shea (2005) found most qualitative studies at their time of review exhibited “little attention to illuminating change over time” (p.36) and were instead focused on obtaining a deeper understanding of the issues at play in the research. The current analysis found the opposite to be true. In this group of longitudinal studies, nine reported on change over the course of the study (Cekaite, 2017; Jiang et al., 2020; Kibler, 2017; Kibler & Hardigree, 2017; Kunitz & Marian, 2017; Lee & Kinginger, 2018; Nguyen, 2018; Shin et al., 2020; Tammenga-Helmantel et al., 2020) and only two (Worden, 2018; Zheng, 2017) aimed to give a more detailed account, and deeper insight, concerning the study's focus.

One of the studies which looks at change over time is Cekaite (2017). This study is also unique in that it looks at young learners. Learning trajectories of two seven year old girls in a Swedish as a second language classroom for refugee and immigrant children are documented over 90 hours of video recordings. Through analysis of the girls' interactional routines, they were able to chronicle the growth and change of these girls' interactional repertoires from limited to more diverse over the course of a school year.

An exceptional study included within this group was Kunitz and Marian (2017), which followed three seventh grade boys in an EFL class in Sweden working as a group on a presentation project over the course of 10 classes in 6 days. During the study researchers utilized a longitudinal conversation-analytic methodology called learning behavior tracking. This limited duration of observation, while subject to dense and detailed analysis, gives cause for consideration of our understandings of “longitudinal” design. Moreover, this limited timeframe begs the question of whether the patterns of change observed are truly translatable to more extended timeframes.

In the next section, we will discuss several studies considered in the present synthesis that took a particularly innovative approach to change over time by appearing to accept Ortega and Iberri-Shea's invitation to examine turning points and transitions.

3.3.5 Innovation Within Qualitative Methodology

At the time of Ortega and Iberri-Shea's (2005) review of studies they suggested that "whether biological or institutional scales are chosen for longitudinal SLA research, they can be better motivated when key events and turning points in the social or institutional context investigated are considered" (p. 38). Six out of the eleven studies examined here appear to have adopted that innovative attention to transitions: Kibler (2017) (feminist identity development leading to decision to study gender and women's studies as university), Kibler and Hardigree (2017) (writing development from ESL classes to participation in mainstream classes), Lee and Kinginger (2018) (transition from study abroad to a domestic classroom), Nguyen (2018) (development from communication course role-playing to real-world consultations during a clerkship), Tammenga-Helmantel et al., (2020) (new teachers use of target language during teacher education to one year after graduation), and Worden (2018) (first time instructor's transition from a flexible to more fixed definition of genre).

Of these, Nguyen (2018) is a particularly interesting example. The study design is explained as taking an emic microanalysis of the interactional practices of a pharmacy student. The study looks at the student's development from role plays in a communication course to real consultations during a clerkship 1 year later. The design stands out because it documents change over time but also across settings. It originates in the classroom and continues into naturalistic interaction creating a narrative which shows the progression of the subject in her role as a pharmacist and validates the relevance of the initial practiced task by connecting it to application beyond the classroom.

3.4 Mixed Methods Longitudinal Classroom-Based Studies

The six studies within this group involved collection and/or analysis of data (concurrently or sequentially) using both qualitative and quantitative approaches or methods. They were also categorized as mixed method if the integration of both methodologies aimed to draw more credible and insightful conclusions and deepen understanding of the focus by establishing corroboration, seeking paradox across results, expanding the range of inquiry, or entwining data for more thorough or narrative reporting (Creswell, 2018; Johnson, Onwuegbuzie, & Turner, 2007). The rationale for the use of this methodology extends further to reaching audiences which might not otherwise be responsive to one of the approaches when applied solely (Dörnyei, 2007).

3.4.1 Design

The analysis of studies with a mixed methods design found little homogeneity in this group with respect to design. Four of the them could be described as descriptive mixed-methods studies. They adopted designs of case study (Sato et al., 2019; Schwartz & Gorbatt, 2017), group study (Han & Hiver, 2018), or a combination of both group analysis and case study(Kim et al., 2017).

Brown and Lally (2019) is the only mixed methods study comparing an intervention and control group, making it the most experimental in design. In the study, they looked at the effects of an immersive versus non-immersive language classroom condition for both low and high intermediate learners. Thus, the study tracked change over time for purposes of determining any difference in the effectiveness of the two classroom approaches. Through qualitative and quantitative analysis of class grades, tasks, course evaluations and researchers' journals, the study found neither immersion nor non-immersion instruction had a measurable advantage over the other with respect to a benefit to language learning.

Leeming's (2019) correlational study suggests that emergent leadership within a group could promote more active participation in group conversation in a classroom. The study conducted a regression analysis of students' individual performance paired with analysis of group engagement in a conversation test. Given that the primary focus was to explore the role of leadership, the longitudinal design allowed collection of enough corroborative data to report deeply, but focused little on change throughout the duration of the study

As noted in the next section, mixed method studies seemed to fall into two distinct duration subgroups. Interestingly, two studies within the subgrouping of longer duration studies both focused on teachers as subjects rather than on students. Sato et al.'s (2019) study looked at five secondary school level EFL teachers in a master's degree program in Japan, and aimed to document their personal learning, professional development and engagement with continuous collaborative action research. This particular study was included within mixed methodology grouping because it repeatedly self-labeled as mixed methods. However, there is little indication of the quantitative methods utilized in this study; instead, the primary analysis reports on information garnered from thematic analysis of essays, excerpts of a final project, classroom observation, and interview. As the study is actually more qualitative in nature, it may be the case that a more mixed methods approach was applied to the larger research project that included the study.

3.4.2 Study Duration

The six mixed methods studies identified in this synthesis appeared to have two groupings in regards to length of data collection and number of subjects involved.

- 1) Two studies between one and four years in duration, which adopted case study design and tended to focus on smaller numbers of subjects ranging from two to five individuals. (Sato et al., 2019; Schwartz & Gorbatt, 2017) These studies usually incorporated mixed methods analysis
- 2) Four studies of shorter duration spanning from nine to sixteen weeks, which looked at larger samples; ranging from 27 to 174 subjects (Brown & Lally, 2019; Han & Hiver, 2018; Kim et al., 2017; Leeming, 2019). These studies tended to involve mixed methodology in data collection.

3.4.3 Role of Researcher and Teacher

An exciting trend emerged within a little over half (four) of the mixed methodology studies regarding the role of researcher as a participant in action (Brown & Lally, 2019; Y. Kim et al., 2017; Leeming, 2019), or collaborative action research (Han & Hiver, 2018). This may suggest that being in an instructor role, or working hand in hand with practitioners, facilitates a wide range of methods of data collection (both qualitative and quantitative). It also increases the likelihood of more naturalistic data uninfluenced by possible effects resulting from the presence of an outside researcher. This is a promising development that hopefully contributes to bridging research and practice. Given the limited number of studies in this group, each was looked at individually to understand the respective approach to longitudinal design, how the classroom was situated in their research, and the various characteristics, strengths and limitations of their design.

As previously discussed, Sato et al. (2019) provided a fairly ethnographic examination of five teachers. Another study that focused primarily on teachers was Schwartz and Gorbatt (2017). The research presented a case study looking at the ways in which two teachers (one Hebrew model teacher and one Arabic model teacher at a preschool in Israel) encouraged their students, with Arabic or Hebrew as an L1, to use their L2. Although this study relied heavily on data collected within the classroom context, the researcher stated an intent to have the researcher and assistant involved as little as possible in the teacher-student interactions.

3.4.4 Focus on Change Over Time

The studies in this group had a range of subject matters which vacillated between focusing on change over time and viewing time as a vehicle for more thorough study. Those studies which documented change looked at professional development of EFL teachers over the course of four years (Sato et al., 2019), emergent leadership (Leeming, 2019), evolving attitudes towards task based learning (Kim, Jung, & Tracy-Ventura, 2017), changes in writing specific psychological factors (Han & Hiver, 2018), and measuring whether immersive or non-immersive conditions are more conducive to the acquisition of English (Brown & Lally, 2019). Schwartz & Gorbatt, (2017) however stood apart as they used their academic year time frame to get a deeper picture from teacher–child conversations of how teachers were encouraging children to use their L2.

The study by Kim et al., (2017) sought to measure perceptions (e.g., emotions, task performance confidence, task enjoyment, task motivation, and learning beliefs) of TBLT among South Korean university students over the course of a semester. They did so by having students complete a perception at the close of each unit of study (each involving four tasks) over the course of a 16 week semester. Subsequently they also looked at an individual student's portfolio to more descriptively discuss the change which occurred for her attitudes towards the tasks from unit to unit. The analysis concluded that in there was a general positive shift in students' attitudes toward TBLT improved over time.

Han and Hiver (2018), by means of time-series analysis, was able to track patterns of change in motivational profiles among 174 middle school language learners over the course of nine weeks in a genre-based writing class. Their study revealed that over the semester, students developed stronger self-regulation and self-efficacy, and had an overall elevation of writing anxiety. Furthermore, this elevated anxiety, when paired with satisfactory levels of self-efficacy and self-regulation could potentially be constructive.

3.4.5 Innovation Within Mixed Methodology

Leeming (2019) is similar to Lowie and Verspoor (2019) in that it could be considered a correlation study. Students' individual scores for leadership, extroversion, and proficiency were measured to explore which was a greater predictor of individual performance and participation as a part of a group. The study found all three to contribute to varying extents within different individuals, yet leadership to be the only measure of significance. Analysis of recordings of conversation tests revealed that groups with stronger leaders showed more participation in these

tests. Inclusion of individual performance further allowed the researcher to recognize and attempt to account for individual variance from the general trend.

The implications of this study are relevant to instructors beyond Leeming. It points to the advantage of strategic organization of conversation groups. By recognizing students of higher leadership, and dispersing them through different groups, more optimal group dynamics can foster participation and increased speaking. With regards to limitations, the study acknowledges that a larger sample would have been of benefit given that only three groups within each category of low, medium, and high were studied. Furthermore, the researcher was the only individual involved with coding; this raises questions about whether the results were affected by conscious or unconscious bias due to the dual role as instructor.

3.4.6 General Observations

In sum, a mixed methods approach can be efficacious in longitudinal classroom-based study by allowing for greater exploration and a multi-dimensional analysis of the focus of study. An unfortunate consequence of this design, however, is that it gives such a context specific analysis that it makes findings difficult to generalize to broader scopes and within different environments. Given the growing diversification of epistemological approaches within longitudinal and classroom-based research, these six studies respond to Ortega and Iberri-Shea's hope that this diversity would "also encourage longitudinal research that capitalizes on the strengths of mixed methods designs" (p.37). Studies with this methodology should continue to be improved in the future and should be seen rather as food for practitioner and researcher thought and inspiration for further investigation. They represent a new trend of studies containing amalgams of the qualitative and quantitative trends previously reported by Ortega and Iberri Shea (2005) that triangulate data and offer more sources for analysis in order to elicit comprehensive results.

4. Discussion

In this study Ortega and Iberri Shea (2005) was used as a platform from which to begin and as a model to follow when applicable. It has been used as a baseline from which to measure shifts and growth within linguistic longitudinal SLA research. With the focus further narrowed to studies that were classroom based, the present study documents several recent trends that have evolved or emerged since the observations reported in 2005. For example, from the 2003-2005 time period, the Ortega and Iberri-Shea examination of 20 journals only elicited 38 studies of L2

learning with a longitudinal design. By comparison, the present examination of only nine journals from 2017 to the present found 97 examples of this type. This growth suggests a significant, and promising, shift in present day linguistics research toward more common utilization of longitudinal frameworks.

This synthesis only identified 27 classroom-based studies out of the total of 97 longitudinal studies. This small proportion could be explained by the difficulty of conducting research in classrooms. McKinley (2019) acknowledges this saying ‘These days many researchers are apprehensive of conducting highly contextualized classroom-based research due to perceived methodological messiness’ (p. 880).

Our review showed revealed some trends (like the dominance of studies on writing) but also high levels of heterogeneity in the studies under review (with some studies focusing on teachers and others on students; qualitative, quantitative and mixed methods approaches; a very wide range of topics both linguistic and non-linguistic and analytical frameworks, sample sizes ranging from 1 to 147 participants). This shows how versatile longitudinal research is that takes place in the context of the classroom.

This versatility, however, poses a challenge when it comes to defining longitudinal research. At the beginning of this thesis when reviewing definitions, we noted that most of them referred to the duration of the studies but were quite open in their references to time (‘over one or two weeks or longer’ (Loewen & Plonsky, 2016); ‘a relatively long period of time’ (Paltridge & Phakiti, 2015)). In our review of longitudinal studies, we have come to realize that longitudinal research cannot be described only in reference to study duration (the shortest study lasted 6 days and the longest 8 years) but also in reference to attention to change over time. Thus, we would like to propose a new definition of longitudinal research as: “a quantitative, qualitative or mixed methods study that (1) is either carried out over an extended period of time OR (2) has a focus on change over time OR (3) both. In our corpus Zheng’s (2017) ethnographic narrative of teaching assistants experience over 10 weeks would exemplify case 1 in the definition. Kunitz and Marian’s (2017) tracking of the spelling of a word over 10 hours of task-based work is a good illustration of a study with a short duration (6 days) that has a primary focus on change over time. Finally, Aubrey (2017) fits case 3 in the definition since it explores how flow changes in strength over a series of five tasks (11 weeks).

Many of the studies reviewed could be categorized using the four types noted in Ortega and Iberri Shea. However, it should be noted that although Ortega and Iberri-Shea (2005) identified a trend of longitudinal research on L2 Program Outcomes trend. At the time of their synthesis, Ortega and Iberri-Shea found that this particular longitudinal design was “infrequently employed in SLA” (p. 30). Within the corpus of this synthesis, this infrequent employ of this design persists, and is in fact unrepresented (as we identified none) in recent classroom-based longitudinal research despite their recommendation for increased implementation in future research.

Of the three methodological groups of studies considered here, the quantitative studies generally seem to evidence the least amount of evolution from the comparable studies examined in Ortega and Iberri-Shea (2005). There are notable exceptions, however. In addition to Rahimi’s (2019) adoption of a time-series design, the use in Vercellotti (2017) of the innovative statistical analysis (HLM), which allowed for analysis of individuals with different totals of data sources and promises one way to deal with the problem of attrition that often afflicts longitudinal research. In addition, Lowie and Verspoor (2019), emerges as an innovative correlational study which possessed a design which looked at the prediction and assessment of relationships.

The present analysis did, however, find potentially significant difference in the character of studies classified as qualitative. Analysis of qualitative studies found a trend of studies considering turning points and transitions, consistent with Ortega and Iberri-Shea's (2005) recommendation for such inquiry (p. 38). Ortega and Iberri-Shea found many qualitative, ethnographies to be focused on obtaining a deeper understanding of researched issues, rather than documenting change over time. In contrast, a majority of the qualitative studies examined herein did report change over time.

The mixed method studies in the present review served as an expansion of research methodology that had been hoped for at the time of Ortega & Iberri-Shea’s (2005) synthesis (p.37). This may be taken as evidence of an increasing role of this methodological approach in longitudinal research. It also reflects the increasing importance of mixed-methods studies have in SLA research (Mirhosseini, 2018). This is a promising development as the use of mixed qualitative and quantitative methodologies achieves in generating intriguing insights.

Another trend in regard to researcher role in the studies, was an emergence of studies which positioned the researcher as the instructor (action research) or working in collaboration

with instructors (collaborative action research) in order to implement an intervention, or facilitate in data collection and analysis. The literature also revealed a need for researchers to consider and be more explicit about their role within the research in order to consider the possible impact they themselves have on data which is generated during their time of observation. By fostering collaborative investigation on longitudinal and cross-sectional scales, we undoubtedly encourage a fruitful exchange of ideas and perspectives between researcher and practitioner.

5. Concluding Remarks

Although there is wide variation in the studies considered here, it is recognized that this thesis is by no means exhaustive or capable of capturing and synthesizing all that is occurring within the field of longitudinal classroom-based research. Furthermore, the number of journals under review may not be representative of larger trends. Given that the focus was on methodologies rather than a particular topic, it was at times challenging to draw meaningful connections and recognize shared themes across the studies. Nonetheless, this review attempted to document interesting details of recent literature and find connections and comparisons with past emergent trends.

As circumstances change on a global level, there will be a need for continued conversation about what warrants definition of a “longitudinal” or “classroom based” design. Scientific desire for uniformity and control tempts us to standardize parameters for these terms, but it is clear from the present review that understandings of these designs are evolving. Platforms and programs for learning are becoming more varied and the integration of technology for target exercises or blended and fully online courses has caused us to reexamine and redefine our perceptions of what qualifies as a classroom (Collins & Muñoz, 2016; Nunan, 2005). This shift in learning environment and experience is one that will be of continued importance and will have an effect on how future research will interpret the classroom context and conduct research within. Moreover, with this extension of the classroom beyond a fixed space, or with limited group interactions in times of global crisis, research will be need to examine whether these adapted learning spaces can be considered classrooms as well.

Cross-sectional approaches to research are informative in their own sense and can produce data that may be less cumbersome to manage. However, it is clear that adopting a longitudinal design allows for much more immersion into an issue in context. In addition, by

placing this investigation within a classroom environment rather than a laboratory we authenticate and give instant significance to the analysis we conduct on data originating within the natural site of learning. Thus, kept in context, the data has clearer implications and can inform program choices and reform. The application and future directions are less hypothetical, because the context where they will be of greatest value is the one which revealed the information in the first place.

As researchers we must question the intentionality behind the research we conduct and recognize the value of both controlled cross-sectional investigation and studies having prolonged and invested interest in specific contexts.

Our review revealed that longitudinal classroom-based research is not very popular but highly versatile. Through our synthesis we have been able to identify exciting methodological innovations regarding design, testing and methodological approaches in studies following quantitative, qualitative and mixed-methods approaches. This has allowed us to propose a bottom up definition of longitudinal research that is more comprehensive than existing definitions in the field of SLA.

Word count:10,458

References

*Note: *denotes inclusion in the synthesis*

- Allwright, D. (1983). Classroom-Centered Research on Language Teaching and Learning: A Brief Historical Overview. *TESOL Quarterly*, 17(2), 191. <https://doi.org/10.2307/3586649>
- *Aubrey, S. (2017). Measuring Flow in the EFL Classroom: Learners' Perceptions of Inter- and Intra-Cultural Task-Based Interactions. *TESOL Quarterly*, 51(3), 661–692.
- Barkaoui, K. (2014). *Quantitative approaches for analyzing longitudinal data in second language research. Annual Review of Applied Linguistics* (Vol. 34).
- Benson, P. (2012). Qualitative Methods: Overview. In *The Encyclopedia of Applied Linguistics*. Oxford, UK: Blackwell Publishing Ltd.
- Brian, P., & Phakiti, A. (Eds.). (2015). *Research methods in applied linguistics : a practical resource*. London : Bloomsbury Academic.
- *Brown, A., & Lally, R. (2019). Immersive Versus Nonimmersive Approaches to TESOL: A Classroom-Based Intervention Study. *TESOL Quarterly*, 53(3), 603–629.
- Burns, A. (2019). Action Research: Developments, Characteristics, and Future Directions. In *The Cambridge Handbook of Language Learning* (pp. 166–185). Cambridge University Press.
- Burns, A. 1945-. (1999). *Collaborative action research for English language teachers / Anne Burns*. Cambridge [etc.] : Cambridge University Press.
- *Cekaite, A. (2017). What Makes a Child a Good Language Learner? Interactional Competence, Identity, and Immersion in a Swedish Classroom. *Annual Review of Applied Linguistics*, 37(2017), 45–61.
- Chapelle, Carol A. Duff, P. a. (2003). Some Guidelines for Conducting Quantitative and Qualitative Research in TESOL. *TESOL Quarterly*, 37(1), 157–178.
- Collins, L., & Muñoz, C. (2016). The Foreign Language Classroom: Current Perspectives and Future Considerations. *Modern Language Journal*, 100, 133–147.
- Creswell, J. W. . (2018). *Designing and conducting mixed methods research / John W. Creswell (Department of Family Medicine, University of Michigan), Vicki L. Plano Clark (School of Education, University of Cincinnati)*. Thousand Oaks, California : SAGE Publications Ltd.,.
- De Costa, P. I., Li, W., & Rawal, H. (2019). Qualitative Classroom Methods. In *The Cambridge Handbook of Language Learning* (pp. 111–136).

- DeKeyser, R. M., & Prieto Botana, G. (2019). Chapter 1. Current research on instructed second language learning. In *Soing SLA Research with Implications for the Classroom* (pp. 1-7)
- Dörnyei, Z. (2007). Longitudinal versus cross-sectional research. In *Research Methods in Applied Linguistics* (pp. 78–91). Oxford University Press.
- Dörnyei, Z. (2007). *Research methods in applied linguistics : quantitative, qualitative, and mixed methodologies*. Oxford [etc.] : Oxford University Press.
- Friedman, D. A. (2012). How to Collect and Analyze Qualitative Data. In *Research Methods in Second Language Acquisition* (pp. 180–200). Chichester, UK: John Wiley & Sons, Ltd.
- Gass, S., & Mackey, A. (2007). *Data elicitation for second and foreign language research. Data Elicitation for Second and Foreign Language Research*. Wiley.
- *Han, J., & Hiver, P. (2018). Genre-based L2 writing instruction and writing-specific psychological factors: The dynamics of change. *Journal of Second Language Writing*, 40(November 2017), 44–59.
- Harbon, L., & Shen, H. (2015). Researching Language Classrooms. In B. Paltridge & A. Phakiti (Eds.), *Research Methods in Applied Linguistics: A Practical Resource* (pp. 457–470). London, England: Bloomsbury Publishing.
- Jackson, D. (2017). Longitudinal Research. In S. Loewen & M. Sato (Eds.), *The Routledge Encyclopedia of SLA* (pp. 404–405). Routledge.
- *Jiang, L., Yang, M., & Yu, S. (2020). Chinese Ethnic Minority Students’ Investment in English Learning Empowered by Digital Multimodal Composing. *TESOL Quarterly*, 0(0), 1–26.
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research*, 1(2), 112.
- Kemmis, S. (2014). *The action research planner : doing critical participatory action research / Stephen Kemmis, Robin McTaggart, Rhonda Nixon*. Singapore : Springer.
- *Kibler, A. K. (2017). Becoming a “Mexican feminist”: A minoritized bilingual’s development of disciplinary identities through writing. *Journal of Second Language Writing*, 38(October), 26–41.
- *Kibler, A. K., & Hardigree, C. (2017). Using Evidence in L2 Argumentative Writing: A Longitudinal Case Study Across High School and University. *Language Learning*, 67(1), 75–109.
- *Kim, Y. J., & Emeliyanova, L. (2019). The effects of written corrective feedback on the

- accuracy of L2 writing: Comparing collaborative and individual revision behavior. *Language Teaching Research*.
- *Kim, Y., Jung, Y., & Tracy-Ventura, N. (2017). Implementation of a Localized Task-Based Course in an EFL Context: A Study of Students' Evolving Perceptions. *TESOL Quarterly*, 51(3), 632–660.
- *Kunitz, S., & Marian, K. S. (2017). Tracking Immanent Language Learning Behavior Over Time in Task-Based Classroom Work. *TESOL Quarterly*, 51(3), 507–535.
- *Lee, S. H., & Kinginger, C. (2018). Narrative Remembering of Intercultural Encounters: A Case Study of Language Program Reintegration After Study Abroad. *Modern Language Journal*, 102(3), 578–593.
- *Leeming, P. (2019). Emergent Leadership and Group Interaction in the Task-Based Language Classroom. *TESOL Quarterly*, 53(3), 768–793.
- *Li, S., & Wang, H. (2018). Traditional Literature Review and Research Synthesis. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *Palgrave Handbook of Applied Linguistics Research Methodology* (pp. 123–144). Palgrave Macmillan.
- *Lin, M. H., & Lee, J. Y. (2019). Pedagogical suitability of data-driven learning in EFL grammar classes: A case study of Taiwanese students. *Language Teaching Research*, 23(5), 541–561.
- Loewen, S., & Plonsky, L. (2016). *An A–Z of Applied Linguistics Research Methods*. Palgrave.
- Loewen, S., & Sato, M. (2017). *The Routledge Handbook of Instructed Second Language Acquisition*. *The Routledge Handbook of Instructed Second Language Acquisition*.
- *Lowie, W. M., & Verspoor, M. H. (2019). Individual Differences and the Ergodicity Problem. *Language Learning*, 69(March), 184–206.
- Mackey, A. (2017). Classroom-based research. In *The Routledge Handbook of Instructed Second Language Acquisition* (pp. 541–561).
- Marsden, E., Morgan-Short, K., Thompson, S., & Abugaber, D. (2018). Replication in Second Language Research: Narrative and Systematic Reviews and Recommendations for the Field. *Language Learning*.
- Mckinley, J. (2019). Evolving the TESOL Teaching–Research Nexus. *TESOL Quarterly*, 53(3), 875–884.
- Menard, S. (2002). *Longitudinal Research*. Thousand Oaks, California: Sage.

- *Menke, M. R., & Strawbridge, T. (2019). The writing of Spanish majors: A longitudinal analysis of syntactic complexity. *Journal of Second Language Writing*, 46(September), 100665. <https://doi.org/10.1016/j.jslw.2019.100665>
- Mirhosseini, S. A. (2018). Mixed Methods Research in TESOL: Procedures Combined or Epistemology Confused? *TESOL Quarterly*, 52(2), 468–478.
- *Nguyen, H. T. (2018). Interactional Practices across Settings: From Classroom Role-plays to Workplace Patient Consultations. *Applied Linguistics*, 39(2), 213–235.
- Nunan, D. (2005). Classroom Research. In *Handbook of Research in Second Language Teaching and Learning* (pp. 225–240). Taylor and Francis.
- Ortega, L. (2015). Research Synthesis. In B. Paltridge & A. Phakiti (Eds.), *Research Methods in Applied Linguistics* (pp. 225–244). Bloomsbury Academic.
- Ortega, L., & Byrnes, H. (2008). *The longitudinal study of advanced L2 capacities: An introduction. The Longitudinal Study of Advanced L2 Capacities.*
- Ortega, L., & Iberri-Shea, G. (2005). Longitudinal research in second language acquisition: Recent trends and future directions. *Annual Review of Applied Linguistics*, 25(2005), 26–45.
- *Otwinska, A., Forys-Nogala, M., Kobosko, W., & Szewczyk, J. (2020). Learning Orthographic Cognates and Non-Cognates in the Classroom: Does Awareness of Cross-Linguistic Similarity Matter? *Language Learning*, 70(3), 685–731.
- Polio, C., & Lee, J. (2019). Experimental Studies in L2 Classrooms. In *The Cambridge Handbook of Language Learning* (pp. 137–165).
- *Rahimi, M. (2019). A comparative study of the impact of focused vs. comprehensive corrective feedback and revision on ESL learners' writing accuracy and quality. *Language Teaching Research*. <https://doi.org/10.1177/1362168819879182>
- *Sato, K., Mutoh, N., & Kleinsasser, R. C. (2019). Longitudinal research on EFL teacher professional development in (Japanese) contexts: Collaborative action research projects. *Language Teaching Research*. <https://doi.org/10.1177/1362168819893814>
- *Schwartz, M., & Gorbatt, N. (2017). “There is No Need for Translation: She Understands”: Teachers' Mediation Strategies in a Bilingual Preschool Classroom. *Modern Language Journal*, 101(1), 143–162.
- *Shin, D. shin, Cimasko, T., & Yi, Y. (2020). Development of metalanguage for multimodal composing: A case study of an L2 writer's design of multimedia texts. *Journal of Second*

- Language Writing*, 47(February), 100714. <https://doi.org/10.1016/j.jslw.2020.100714>
- Singer, J. D., & Willett, J. B. (2009). *Applied Longitudinal Data Analysis*. *Applied Longitudinal Data Analysis*.
- Stringer, E. T. (2008). *Action research in education*. Upper Saddle River, NJ : Pearson Prentice Hall,.
- *Tammenga-Helmantel, M., Mossing Holsteijn, L., & Bloemert, J. (2020). Target language use of Dutch EFL student teachers: Three longitudinal case studies. *Language Teaching Research*.
- *Teng, L. S., & Zhang, L. J. (2020). Empowering learners in the second/foreign language classroom: Can self-regulated learning strategies-based writing instruction make a difference? *Journal of Second Language Writing*, 48(November), 100701.
- *Vercellotti, M. Lou. (2017). The development of complexity, accuracy, and fluency in second language performance: A longitudinal study. *Applied Linguistics*, 38(1), 90–111.
- Wallace, M. J. (1998). *Action Research for Language Teachers*. 134976. Cambridge, England: Cambridge University Press.
- Williams, J. (2013). Classroom research. In *The Routledge Handbook of Second Language Acquisition* (pp. 541–554).
- *Worden, D. (2018). Balancing stability and flexibility in genre-based writing instruction: A case study of a novice L2 writing teacher. *Journal of Second Language Writing*, 42(May), 44–57.
- *Yoon, H. J., & Polio, C. (2017). The Linguistic Development of Students of English as a Second Language in Two Written Genres. *TESOL Quarterly*, 51(2), 275–301.
- *Zheng, X. (2017). Translingual Identity as Pedagogy: International Teaching Assistants of English in College Composition Classrooms. *Modern Language Journal*, 101, 29–44.

Appendices

Appendix B: Journal Totals

Journal	Impact Factor	Longitudinal Studies	Longitudinal + Classroom-based
<i>Tesol Quarterly</i>	2.056	14	7
<i>Applied Linguistics</i>	3.593	17	2
<i>Language Learning</i>	2	19	3
<i>Language Teaching Research</i>	2.086	11	5
<i>Second Language Research</i>	1.75	4	0
<i>The Modern Language Journal</i>	2.789	14	3
<i>Annual Review of Applied Linguistics</i>	2.481	3	1
<i>Journal of Second Language Writing</i>	4.2	11	6
<i>Studies in Second Language Acquisition</i>	2.6	4	0
TOTALS		97	27

Appendix B: Longitudinal Classroom-Based Studies Included for Synthesis

Author, Date of Publication	Journal	Education Level	Context	Language	Country	Topic	Method
Aubrey, 2017	TQ	University	FL	English	Japan	Flow, Perceptions of Inter/Intra cultural task-based interactions	Quant
Brown & Lally, 2019	TQ	University	SL	English	USA	Immersive/nonimmersive tTESOL effects on language acquisition	MM
Cekaite, 2017	ARAL	Primary*	SL	Swedish	Sweden	Interactional Competence, Identity, Immersion, Classroom discourse analysis	Qual
Han & Hiver, 2018	SLW	Secondary	SL	English	Korea	Genre-based writing instruction, processes of motivational change	MM
Jiang et al., 2020	TQ	University	FL	English	China	Digital Multimodal Composing, minority student's investment/experience	Qual
Kibler, 2017	SLW	Secondary-University	other	English	USA	development of disciplinary identities through writing	Qual
Kibler & Hardigree, 2017	LL	Secondary-University	other	English	USA	Use of evidence in argumentative writing	Qual
Kim & Emeliyanova, 2020	LTR	IEP	FL	English	USA	WCF processing in groups vs individually, accuracy	Quant
Kim et al., 2017	TQ	University	FL	English	Korea	TBLT, student perceptions	MM
Kunitz & Marian, 2017	TQ	Secondary	FL	English	Sweden	Language learning behavior tracking	Qual
Lee & Kinginger, 2018	MLJ	University	FL	Chinese	USA	transition from study abroad to domestic classroom	Qual
Leeming, 2019	TQ	University	FL	English	Japan	Emergent Leadership, behavioral engagement, TBL	MM
Lin & Lee, 2019	LTR	University	FL	English	Taiwan	Grammar development in data driven learning conditions	Quant
Lowie & Verspoor, 2019	LL	Secondary	FL	English	Netherlands	Individual differences, CAF Writing	Quant
Menke & Strawbridge, 2019	SLW	University	FL	Spanish	USA	Development of syntactic complexity in writing	Quant
Nguyen, 2018	AL	University-Clerkship	other	English	USA	Interactional practices across settings	Qual
Otwinowska et al., 2020	LL	Secondary	FL	English	Poland	Vocabulary development (cognates/noncognates), Awareness effects	Quant
Rahimi, 2020	LTR	University	SL	English	Canada	focused/unfocused written corrective feedback (WCF), accuracy and quality of writing	Quant
Sato et al., 2019	LTR	Secondary (T)	FL	English	Japan	EFL Teacher development, collaborative action research	MM
Schwartz & Gorbatt, 2017	MLJ	Preschool (T)	SL	Hebrew & Arabic	Israel	L2 mediation strategies, teacher-child conversations	MM
Shin et al., 2020	SLW	Secondary**	SL	English	USA	development of metalanguage for multimodal composing	Qual
Tammenga-Helmantel et al., 2020	LTR	Secondary (T)	FL	English	Netherlands	Target Language use	Qual
Teng & Zhang, 2020	SLW	University	SL/FL	English	China	Self-regulated learning strategies-based writing instruction	Quant

Vercellotti, 2017	AL	IEP	FL	English	USA	Speaking, CALF	Quant
Worden, 2019	SLW	University (T)	SL	English	USA	genre-based writing instruction, Teacher Pedagogical Content Knowledge of Genre	Qual
Yoon & Polio, 2017	TQ	University	SL	English	USA	CALF development, Genre Writing	Quant
Zheng, 2017	MLJ	University (T)	other	English	USA	translingual identities, identity-as-pedagogy	Qual

Note:

Journals: AL- *Applied Linguistics*, ARAL- *Annual Review of Applied Linguistics*, LL- *Language Learning*, LTR- *Language Teaching Research*, MLJ- *The Modern Language Journal*, SLW- *Second Language Writing*, TQ- *Tesol Quarterly*,

Education Level: (T)- Teacher focused study, IEP- Intensive English Program,

Context: FL- Foreign Language Classroom, SL- Second Language classroom, other- mainstream classroom or exceptional context

Method: MM- Mixed Methods

* class was a "mottagningsklass" for refugee children aged 7-10,

** 6th grade in USA can be either primary or secondary