Metacognitive Instruction in L2 Listening: an intervention study.

by

Jesús Leopoldo Toapanta Túarez

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Spanish and Latin American Studies

Department of Modern Languages and Cultural Studies
University of Alberta

Abstract

This study compared three intact classes of Spanish learners who received three different types of listening instruction: a metacognitive pedagogical cycle following Vandergrift and Tafaghodtari (2010), an awareness-raising approach which exposed L2 listeners to the factors associated with successful L2 listening, following the Metacognitive Awareness Listening Questionnaire (Vandergrift et al., 2006), and an approach that incorporated conventional practices such as pre-listening and post-listening activities in which key vocabulary and guiding questions were used to construct a better understanding of the audio text (Field, 2012). The intervention consisted of eight listening lessons delivered over the course of a semester to these three groups of intermediate language learners. Listening performance was measured before the intervention at pre-test, and after the intervention at post-test and at delayed post-test. The results of this study corroborate previous findings regarding metacognitive instruction in L2 listening (e.g., Cross, 2011; Vandergrift & Tafaghodtari, 2010), and provide support to the notion that metacognitive awareness impacts the development of L2 listening positively (e.g., Goh, 1997). In this respect, participants exposed to the metacognitive pedagogical cycle showed statistically significant improvement from pre-test to post-test. Also, after statistically controlling for initial differences in listening ability, the Metacognitive Pedagogical Sequence group outperformed the Conventional Approach group at post-test. Moreover, participants in the Awareness-Raising group showed significant improvement from pre-test to post-test. The results also show evidence of long-term effects which can be attributed to metacognitive instruction in L2 listening. This

study addresses concerns such as the need for intervention studies in L2 listening that identify what works best (Berne, 2004; Macaro et al., 2007) as well as long-term effects (Plonsky, 2011).

Preface

This thesis is an original work by Jesus Toapanta. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name: "Metacognitive Strategy Instruction in L2 listening", No. Pro00066085, July 20th, 2016.

Acknowledgements

I want to thank the University of Alberta and the Department of Modern Language and Cultural Studies for giving me the support I needed during my studies. I would also like to thank my examiners, my supervisory committee and my supervisor for reading, commenting, and providing feedback on this dissertation.

I want to thank Dr. Leila Ranta, Dr. Greg Thomas, and Dr. Suzanne Graham for all their comments and suggestions on previous drafts of this dissertation. I want to thank Dr. Yvonne Lam and Dr. Claudia Kost for their discerning comments and suggestions in early versions of this paper. I want to thank my supervisor, Dr. Xavier Gutiérrez, for not leaving me alone during this process and walking with me from the beginning of my studies. Thank you Xavier for sharing your experience and knowledge with me. I would not have been able to go through all the different stages of the program without your encouragement and support. Your tactful, meticulous, and thoughtful feedback made this journey a wonderful learning experience for me.

I would also like to thank the instructors and students that participated in this study. Without their consent and participation, this investigation would not have been possible. Also, I would like to thank my family who accompanied me during this adventure. Thank you Elie for your patience, enthusiasm and understanding.

Thank you all,

Jesús

Table of Contents

List of Tables

List of Figures

1 Introduction

The study presented in this dissertation compares three instructional approaches in second language (L2) listening: a metacognitive pedagogical cycle that guided learners through stages such as prediction, verification, problem-solving, and reflection (Vandergrift, 2004); an awareness-raising approach which exposed L2 listeners to the factors associated with successful L2 listening, following the Metacognitive Awareness Listening Questionnaire (MALQ; Vandergrift, Goh, Mareschal & Tafaghodtari, 2006); and a conventional approach in which vocabulary was introduced as a pre-listening activity and comprehension questions were used to build a better understanding of the listening text. This study addresses the paucity of experimental studies about L2 listening instruction in the field of Second Language Acquisition (SLA).

1.1 L2 Listening in Language Teaching

During the first half of the twentieth century, L2 listening did not receive much attention. Although the Direct Method emphasized oral communication skills such as speaking and listening (Richards & Rogers, 2001), the influence of the Grammar Translation Method, which emphasized long and elaborate explanations of grammar with little or no attention to pronunciation (Prator & Celce-Murcia, 1979), dominated the field of language teaching and learning. During this time, language teaching was more concerned with the development of reading knowledge leaving L2 listening completely unattended. Even with the emergence of the Audiolingual Method at the beginning of the second half of the twentieth century, L2 listening in

language teaching was mainly limited to pronunciation drills and imitation of dialogues (Richards & Rodgers, 2001).

In the 1970's, the language teaching profession experienced the emergence of methods and approaches such as Community Language Learning (Curran, 1972), Silent Way (Gattegno, 1972), Total Physical Response (Asher, 1977), and the Natural Approach (Krashen & Terrell, 1983). These innovative approaches brought insightful perspectives such as discovery learning, development of student autonomy, non-threatening classroom environments, comprehension preceding production, and active engagement of learners. These perspectives have contributed to the current teaching approaches that highlight the communicative properties of language. During the last quarter of the twentieth century until the present, L2 listening has been recognized as an essential skill in second language learning and teaching. The emergence of communicative competence (Hymes, 1972) in the early 1970's and the emphasis on approaches that develop communicative ability clearly identify L2 listening as an important language skill.

However, in practice, L2 listening in language teaching continues to be underrepresented. In fact, L2 listening is often regarded as the Cinderella of the language skills (e.g., Nunan, 2002; Vandergrift, 1997), and of the four language skills, L2 listening is the least-researched skill (Cross, 2015). The amount of time devoted to developing this skill is often minimal when compared to the amount of time devoted to other class activities. In this regard, Feyton (1991) commented that L2 listening should not be conceived as a mere listening activity in the language classroom but should rightfully be treated as a language skill.

1.2 L2 Listening

Instruction in L2 listening has mostly relied on the comprehension approach (Graham, Santos & Francis-Brophy, 2014; Seigel, 2014). This approach brings learners benefits such as exposure to different voices and oral interactions as well as practice in different types of listening (Field, 2012). At the same time, this approach allows teachers to verify comprehension due to its focus on the product of L2 listening (i.e., correct or incorrect answers). However, an exclusive focus on the right answer "often creates a high level of anxiety, especially since an interest in the correct answer is often associated with evaluation" (Vandergrift, 2011: 464). Additionally, while this approach informs instructors on how to carry out L2 listening in the classroom, when looking at L2 listening from the learners' perspective, it may not necessarily help listeners address L2 listening more tangibly. That is, while conventional practices such as pre-listening activities, intensive and extensive listening, and post listening activities (Field, 2012) provide guidance to instructors on how to develop their listening lessons, these practices might not necessarily provide learners with guidance on how to address L2 listening more strategically.

Recent research in the field of L2 listening instruction has found that skilled listeners display a wide range of strategies and employ metacognitive strategies (e.g., Goh, 2000) in an efficient and orchestrated manner (e.g., Vandergrift, 2003). These findings also indicate that teaching individual and isolated strategies does not impact L2 listening as positively and substantially as instruction that focuses on clusters of cognitive and metacognitive strategies (e.g., Graham & Macaro, 2008). Currently, L2 listening instruction is experimentally exploring the benefits of metacognitive instruction; that is, an instructional approach to L2 listening that aims at developing "learners' person knowledge, task knowledge, and strategy knowledge and

their ability to self-manage their listening through a range of process-based instructional activities which stimulate metacognitive experiences" (Cross, 2015: 886).

Metacognitive instruction in L2 listening is a relatively new pedagogical approach. It enhances listening skills substantially by guiding learners through a learning process that develops metacognitive knowledge and regulatory skills that allow listeners to address L2 listening more strategically (e.g., Vandergrift & Goh, 2012). In fact, empirical evidence from an intervention study that implemented a metacognitive pedagogical cycle that guided learners through stages such as prediction, verification, monitoring, problem-solving and reflection (Vandergrift & Tafaghodtari, 2010) indicates that metacognitive instruction in L2 listening impacts L2 listening significantly. Nevertheless, the lack of intervention studies, especially studies that compare different approaches (e.g., Berne, 2004), has somehow hindered its applicability and popularity. The field of L2 listening instruction is in need of intervention studies that compare the effects of different types of instruction to determine which type of listening approach is more beneficial. As Macaro, Graham and Vanderplank (2007) noted, the field is still in its infancy and "urgently need[s] more intervention studies that identify more clearly what kind of strategy instruction works with what kinds of learner" (p. 185).

Therefore, while it is important to experimentally verify the robustness of this metacognitive pedagogical cycle (Vandergrift & Tafaghodtari, 2010), it is also important to compare the effects of this approach with current conventional pedagogical practices to determine the appropriateness of these instructional approaches. Thus, this intervention study compares three intact classes of Spanish learners who received three different types of listening instruction: a metacognitive pedagogical cycle following Vandergrift (2004); an awareness-raising approach in which learners discussed strategies associated with problem solving,

planning and evaluation, mental translation, person knowledge, and direct attention (MALQ factors) and were encouraged to reflect on their listening experience; and an approach that incorporated conventional practices such as pre-listening and post-listening activities in which key vocabulary and guiding questions were used to construct a better understanding of the audio text (Field, 2012). This study provides support to the notion that metacognitive awareness impacts the development of L2 listening positively (e.g., Goh, 1997) and significantly (e.g., Vandergrift & Tafaghodtari, 2010). The results of this study indicate that teaching learners how to listen in the second language is possible and beneficial when L2 listening is carried out as a learning process in which learners are encouraged to take an active role in the listening activities.

Including this introductory chapter, this dissertation contains six chapters: 1) Introduction, 2) Literature Review, 3) Methodology, 4) Results, 5) Discussion, and 6) Conclusion. Chapter 2 provides a review of the literature on language learning strategies, metacognition and metacognitive instruction in second language listening, and intervention studies in L2 listening. The Literature Review chapter concludes with the research questions that guided the present study. The Methodology chapter describes the details associated with participants, interventions, procedures of data collection, and data analysis. The Results chapter presents the findings of the study in relation to the research questions; that is, short- and long-term effects associated with the intervention, and the participants' perceptions about the instructional approaches implemented. In chapter 5, I provide a discussion of the findings in relation to previous research and their pedagogical implications. I also discuss the limitations of this study and the directions for future studies. Finally, the Conclusion chapter summarizes the findings of the present study.

2 Literature Review

This chapter provides a review of the literature regarding L2 listening. It starts with a review of the linguistic and semantic processing involved in listening and L2 listening. Then, after providing an overview of the field of language learning strategies, the chapter discusses aspects associated with metacognition, metacognition in L2 listening, and metacognitive instruction in L2 listening. The chapter continues with a review of intervention studies dealing with L2 listening strategy instruction and L2 listening metacognitive instruction. Finally, the chapter concludes with the most relevant findings pertaining L2 listening instruction as well as with the gap in knowledge, the relevance and purpose of the present study, and the research questions that guided the study.

2.1 Listening

Although listening can simply be described as the act of decoding and determining what the speaker means or intends to express, this description does not fully capture what understanding spoken language actually entails. Listening involves complex cognitive processes that operate at the phonological, syntactic, lexical, semantic, and pragmatic levels. Thus, in addition to perceiving sounds and matching the speech signal with the mental linguistic knowledge, listeners also construct meaning based on knowledge of the language system, prior knowledge, situation, and context (Vandergrift & Goh, 2009). In this sense, the complexity of listening is not only due to linguistic, perceptual, and cognitive factors (e.g., reduced forms, hesitations, rate of delivery, accent, processing ability), but it is also based on pragmatic (e.g.,

interlocutors' intentions) and sociolinguistic (e.g., socially and/or culturally encoded language) factors. Therefore, to understand and interpret spoken language, listeners process and triangulate all available information in the phonology, prosody, syntax, semantics, pragmatics, as well as visual information if available (Guellai, Langus & Nespor, 2014).

An informative and interesting perspective on how listeners construct meaning was articulated by Anderson (1995), who conceived listening within three interconnected phases: 1) perceptual processing, 2) parsing, and 3) utilization. During the perceptual processing phase, listeners create a phonetic representation by attending to the speech signal, discriminating and recognizing phonemes, and grouping sound categories according to the phonological system of the language. Then, in the second phase, listeners parse the phonetic representation and activate potential word candidates in their mental lexicon. This creates a meaning-based representation in working memory. After that, during the utilization phase, listeners create a conceptual framework by associating the meaning-based units with information stored in long-term memory. During this phase, listeners triangulate all available information in their schemata and context to interpret the linguistic input.

This model provides an insightful perspective on how listeners construct meaning; however, it is important to note that these processes operate simultaneously at the different levels of organization—phonological, syntactic, lexical, and pragmatic. With this in mind, Rost (2005) described listening as consisting of three basic processing phases: decoding, comprehension, and interpretation. Thus, listeners basically engage in a simultaneous process that involves: a) word recognition and grammatical parsing (i.e., Decoding Phase), b) activating prior knowledge, constructing a conceptual framework, and drawing logical inference (i.e., Comprehension

Phase), and c) comparing meanings with expectations and checking for congruency at the discourse level (i.e., Interpretation Phase).

While in essence these two descriptions portray a similar picture of how spoken language is processed, it is important to highlight that listening involves simultaneous cognitive processes characterized by their receptive, constructive, and interpretative nature. Listening involves cognitive processes such as discriminating sounds, identifying word boundaries, segmenting utterances into linguistic units (i.e., linguistic processing) as well as cognitive processes that utilize prior knowledge and expectations (i.e., semantic processing).

2.1.1 Linguistic Processing

Linguistic processing refers to the act of decoding language from smaller to larger units. This decoding process, which is often referred to as bottom-up, goes from phonetic features (e.g., aspiration, nasalization, +/- voiced), to phonemes, to syllables, to words, to utterances (i.e., sentences). Therefore, when the speech signal reaches the ear, this speech stream is broken down into segments and linguistic categories are assigned (Best, 1995). Thus, in addition to a phonological representation, listeners also need a syntactic representation to decode the linguistic input. That is, listeners need to have knowledge of grammar, sound system, and vocabulary since parsing of speech occurs simultaneously at different levels of linguistic organization (e.g., phonological, syntactic, lexical) (Rost, 2016).

In order to process linguistic input, listeners engage in a cognitive task that requires the identification of sounds and the activation of knowledge related to these sounds. For example, when listeners hear the utterance ‘I read a book’, listeners need to identify phonological and

syntactic elements that can assist them in understanding what the utterance means. In this case, listeners need to be able to perceive whether the verb sounds as [ri:d] or [red] to assign the corresponding linguistic category to the utterance. If listeners hear [red], they know that it is not the color 'red', but the past form of the verb 'read'. Additionally, this utterance does not occur without intonation, so listeners have to process intonation units as well. Thus, listeners need to simultaneously segment the speech stream and activate knowledge associated with the linguistic units in a parallel fashion across the phonology, prosody, grammar, and lexicon. All of this processing is accomplished in real time while attending to the new incoming input.

2.1.2 Semantic Processing

Unlike linguistic processing, which originates in the speech signal, semantic processing originates in the mind of the listener. Semantic processing refers to reasoning, inferring, and constructing meaning based on background knowledge and contextual information. This type of processing, which is often referred to as top-down, involves the activation of modules of knowledge in the brain and allows "listeners to *intuit meanings* through their connection to the speakers, the events, and their inner world, and through their intention to understand" (Rost, 2016: 64). These modules of knowledge or schemata are connections and associations of concepts that aid listeners in constructing an idea of what is meant, implied, or intended to mean.

The activation of these modules of knowledge allows listeners to generate and develop a general idea of the listening input. Thus, if for whatever reason listeners are not able to identify words in the speech signal, they can still interpret the message by drawing inferences based on their world knowledge and contextual information. For instance, emotional tones (e.g., anger,

happiness, despair, enthusiasm, sadness) can actually clue listeners in constructing an idea regarding the listening input. That is, listeners can draw inferences about states, intentions, and perspectives based on emotional tones and the activation of knowledge related to these tones (Barsties & De Bodt, 2015). The activation of these modules of knowledge automatically triggers the activation of other related portions of knowledge in the network of associations. Thus, for example, if the listener hears the utterance 'summer vacation', interrelated associations such as beach, sun, travel, seafood, and others are activated. This simultaneous activation of schemata allows listeners to generate a general understanding of the listening input.

2.1.3 Linguistic and Semantic Processing in L2 listening

A fundamental and perhaps the most important aspect in understanding language is recognizing words and lexical phrases in fluent speech. Recognizing words in the speech stream entails identification of the word and activation of lexical knowledge associated with that particular word. In this process, listeners need to identify word boundaries in a continuous string of acoustic speech where, unlike readers, the spaces that separate words are not present. While recognizing words and identifying word boundaries in fluent speech are not problematic for first language listeners, who do this automatically, these key aspects in the processing of oral input, are certainly major challenges for second language learners (Cutler, 2001; Field, 1998). Particularly for adult second language learners, identifying words can be very difficult because two distinct phonemes in the second language can sound as if they are the same sound to the second language learner (Kuhl, 2000). For example, the words 'bag' and 'bug' can sound the same to a Spanish learner of English because these two vowel sounds do not exist in the

phonological system of Spanish. The closest phoneme in Spanish to both of these English vowels is the central open vowel /a/ as Spanish has a five vowel system. Moreover, identifying word boundaries and recognizing lexical items can become even more complicated in fluent speech due to factors such as cluster reductions, assimilation of sounds, weakening and/or omission of vowels, intonation patterns, among others. In this sense, the pronunciation of isolated words can change when embedded in utterances. Thus, even if second language listeners can identify individual words, they might not be able to identify these same words in connected speech (Broersma & Cutler, 2008).

Additionally, listeners have to recognize words from a speech signal which is characterized by dialectal variation. For instance, Caribbean Spanish in Latin America (e.g., Puerto Rico, Dominican Republic, Cuba) is characterized by the deletion of the alveolar fricative sound /s/ at the end of syllables. That is, the utterance 'Las costas de las islas son preciosas.' is pronounced as [la kota de la ila son presiosa] (Zentella, 1997: 43). This linguistic variation not only occurs at the phonological and lexical levels, but also at the syntactic level. Questions in Caribbean Spanish, for example, do not exhibit the usual verb-subject order as in ¿Cómo estás tú? In this variety, this question is formulated as ¿Cómo tú estás? and pronounced as [kómo tú etá] (Lamboy, 2008: 159). In this sense, second language listeners indeed face great difficulties in recognizing words and identifying word boundaries (Goh, 2000).

Nevertheless, although spoken language can become quite incomprehensible to second language listeners due to the difficulty to segment the stream of speech and recognize words in the linguistic input (i.e., approaching L2 listening from a bottom-up perspective), second language listeners can still make inferences about the meaning of the utterances based on the context, common sense, and background knowledge (i.e., approaching L2 listening from a top-

down perspective). Listeners do not know precisely what the listening input will be like or what the intended meaning will be, so they need to make inferences to interpret the message (Basnakova, Weber, Petersson, Van Berkum & Hagoort, 2013). Similarly, second language listeners do not exactly know how the listening input will unfold. Thus, they need to infer meaning by drawing logical associations and filling in gaps based on their world knowledge and available contextual information. In this sense, L2 listening becomes a problem-solving activity in which listeners triangulate all available information with their stored world knowledge to construct an idea of what the listening input means and/or implies. It is this inferring, reasoning, and triangulation of information that makes it possible for listeners to interpret and generate a global understanding of the listening input.

Contextual information, background knowledge, and activation of these modules of knowledge are essential aspects in building a general understanding of the listening input. However, this type of processing can also lead listeners to misinterpreting the input if the activation of these conceptual connections is not corroborated with more information (Macaro, Vanderplank & Graham, 2005). Thus, listeners need to continuously monitor and update their understanding of the listening input to be able to profit from this type of processing. Likewise, in extreme conditions of stress, linguistic ambiguity, and background noise, second language listeners will tend to focus and rely on lexical items to decode the speech stream (McGowan, 2015). That is, because second language listeners might not be able to utilize their background knowledge and contextual information due to anxiety, background noise, and/or linguistic ambiguity, they have to rely on linguistic processing to decode the speech stream. However, this speech stream can be undecipherable to second language listeners due to difficulties in identifying word boundaries. In other words, although semantic processing is quite powerful and

essential in L2 listening because it triangulates information from different sources, semantic processing alone is not enough to guarantee an appropriate interpretation of the listening input. In fact, the activation of background knowledge is often triggered by word recognition (Rost, 2005), and hence, to a certain degree, semantic processing depends on linguistic processing.

Semantic processing and linguistic processing operate in a complementary fashion and they rarely operate independently. Linguistic processing (bottom-up processing) and semantic processing (top-down processing) interact together to create an understanding of the listening input. However, before being able to successfully utilize both types of processing, second language listeners have to overcome difficulties such as inability to segment the speech stream and inadequate activation of schemata as discussed above. Also, second language listeners have to overcome difficulties associated with anxiety and insufficient vocabulary. For example, anxiety can deliberately affect the outcome of listening because it is interrelated with other factors such as motivation and confidence. In this sense, confident second language listeners are more likely to be more motivated and to exhibit lower levels of anxiety. In fact, in a study with French learners, Mills, Pajares and Herron (2006) found a significant negative correlation between listening ability and anxiety; that is, higher scores in listening ability were associated with lower levels of reported anxiety. Higher levels of anxiety can cause second language listeners to avoid listening situations and therefore limiting their exposure to the language. Anxiety in L2 listening is associated with the difficulty of processing the linguistic input, which is often perceived as too fast and incomprehensible (Vogely, 1999). That is, since second language listeners are not able to identify words in the speech stream due to aspects such as reduced forms, rate of delivery, accents, and others, their level of anxiety increases causing L2 listening to suffer.

Additionally, L2 listening is often associated with evaluation in the language classroom (Mendelsohn, 1994), which naturally causes a lot of anxiety and frustration in second language listeners. Although they allocate a lot of effort to the task of listening, they often perceive second language listening as the most difficult skill (Graham, 2006) because they do not know how to address L2 listening more tangibly and systematically. Therefore, while it is important to reduce the levels of L2 listening anxiety as well as to work on the learners' attitudes and beliefs about second language listening (Arnold, 2000), it is also important to provide second language listeners with guided practice that actually helps them approach the L2 listening task more systematically so that they can take control of the task of listening (i.e., help them self-regulate listening processes).

Also, among the linguistic and cognitive factors that affect L2 listening, vocabulary has been known to play a central role in L2 listening. In fact, vocabulary knowledge is a significant predictor for L2 listening ability. Mecartty (2000), for example, found that 14% of the variance in L2 listening ability can be explained by L2 vocabulary knowledge. Also, Staehr (2009) found that 51% of the variance in listening comprehension can be explained by L2 vocabulary knowledge. In other words, L2 listening is facilitated by the vocabulary size. Additionally, "the activation of background knowledge (content schemata and cultural schemata) that is needed for comprehension of speech are linked to and launched by word recognition" (Rost, 2005: 508). In this respect, vocabulary is an essential aspect in L2 listening because it contributes to both types of processing. That is, while vocabulary establishes a baseline to segment the speech stream (i.e., linguistic processing), it also contributes to the activation of schemata (i.e., semantic processing).

In this regard, while vocabulary knowledge is indeed a key aspect in L2 listening, Vandergrift and Baker (2015) argued that second language listeners with good metacognitive

abilities can manage unknown words. This is attributed to metacognitive abilities that help learners regulate L2 listening processes. That is, although vocabulary is clearly important in L2 listening, second language listeners with good metacognitive abilities may actually be able to handle unknown vocabulary. Vandergrift and Baker (2015) argued that metacognitive instruction sensitizes learners to the process of listening by helping them bridge gaps in unknown vocabulary as they reconstruct the message. Conversely, Wang & Treffers-Daller (2017) found that vocabulary knowledge was more important than metacognitive awareness in L2 listening. They argued that "to improve learners' listening comprehension, in L2 teaching, teachers should focus on enhancing learners' vocabulary knowledge in particular" (p. 148).

Successful second language listeners do not only apply knowledge associated with linguistic and semantic processing, but also metacognitive knowledge about the listening processes. Metacognitive knowledge helps learners devise strategies to overcome deficiencies in their linguistic and semantic processing. Regarding these cognitive processes in L2 listening as well as the ability to control these processes, Vandergrift and Goh (2012) have proposed the need to develop metacognitive knowledge in L2 listening in order to assist listeners in regulating and coordinating these cognitive processes. This is discussed in the section below.

2.2 Metacognition

Before discussing aspects associated with metacognition, instruction, and L2 listening, it is important to provide a background of the field of language learning strategies since research on such strategies preceded the interest in metacognitive instruction and it is the field that has

been researched more extensively. In fact, metacognitive instruction in L2 listening is a relatively new field that has come out of the field of language learning strategies.

2.2.1 Language Learning Strategies

Initial studies in the field of language learning strategies started in the mid 1970's. These studies focused on 'the good language learner' and aimed at identifying the characteristics of successful language learners (Naiman, Frohlich, Stern & Todesco, 1978; Rubin, 1975; Stern, 1975). The quest was motivated by the assumption that if the characteristics and/or the strategic behavior of successful language learners are identified, then these characteristics, in the form of strategies, could be transferred to less successful learners in an attempt to enhance their language learning. In L2 listening, 'the good language learner' studies revealed that more proficient listeners monitor their attention, focus better on the task, associate the information to their own experiences, judge how coherent the information is, and attend to larger chunks of information (O'Malley, Chamot & Küpper, 1989). These studies also showed that more successful listeners use metacognitive strategies more often and do not usually translate (Vandergrift, 1997). However, it is still not clear whether it is the strategies that lead the successful learners to attain higher levels of proficiency, or it is the proficiency that leads these learners to display such strategies. In fact, Kojic-Sabo and Lightbown (1999) argued that "we must, however, be careful not to consider all strategies as universally valid or useful to all learners, or to assume falsely that strategies used by successful students will undoubtedly be helpful to less successful ones" (p. 190). Learning a language involves many factors that usually interact in particular ways depending on the context, situation, purpose, and individuals. Thus, a strategy that is effective

and useful in a specific situation for a particular purpose cannot be assumed to be effective for all individuals because learners are different in so many ways.

In the early 1980's, with the development of communicative competence (Hymes, 1972) and the emergence of the communicative language teaching approach (e.g. Littlewood, 1981; Spada, 2007), the importance of language learning strategies in language teaching and learning was reaffirmed since strategic competence was incorporated as a key component in models of communicative competence (e.g., Bachman, 1990; Bachman and Palmer, 1996; Canale & Swain, 1980; Canale, 1983; Celce-Murcia, Dörnyei and Thurrell, 1995).

The field of language learning strategies has grown enormously since its appearance in the 1970's. It has attracted the attention of both teachers and researchers. The field of language learning strategies has generated a lot of discussion and its active development can be observed in the several definitions generated around the field. For example, Rubin (1975) defined language learning strategies as "techniques or devices which a learner may use to acquire knowledge" (p. 43). Oxford (1990) defined them as "specific actions taken by the learner to make learning easier, faster, more enjoyable, more self-directed, more effective, and more transferable to new situations" (p. 8). Anderson (2005) said that "Strategies are the conscious actions that learners take to improve their language learning" (p. 757). And more recently, in an attempt to unify the different perspectives and to provide a more encompassing definition, Griffiths (2013) defined language learning strategies as "activities consciously chosen by learners for the purpose of regulating their own learning" (p. 36). This last definition contains the core essence of language learning strategies. It describes strategies as activities which are conscious and serve the purpose of helping learners self-regulate.

Also, several classifications of language learning strategies (e.g. O'Malley et al., 1985; Rubin, 1987; Stern, 1992) have been generated. Rebecca Oxford, for instance, recently re-classified her former classification of language learning strategies into cognitive, affective, sociocultural-interactive, and metastrategies (Griffiths & Oxford, 2014). While all of these classifications include a category for metacognitive strategies, Rubin's (1987) classification grouped metacognitive and cognitive strategies together. This is particularly important because Graham and Macaro (2008) attributed the positive results of their study to L2 listening instruction that focused on clusters of cognitive and metacognitive strategies. They argued that "individual cognitive strategies cannot be considered, and therefore taught, in isolation" (p. 770). Thus, it may seem that classifying cognitive and metacognitive strategies separately is not appropriate if the aim is to teach listening skills.

The field of language learning strategies has indeed been very prolific, and although the different definitions and classifications have generated criticism, such as that the field has developed chaotically with a lot of ambiguity (Dörnyei & Skehan, 2003), language learning strategies have had a great impact on language learning and teaching. Nowadays, language learning strategies are included in influential documents such as the Common European Framework of Reference for Languages (CEFR) (Oxford, Rubin, Chamot, Schramm, Lavine, Gunning & Nel, 2014).

The essential purpose of discovering how successful language learners learn was conceived within a pedagogical perspective. The idea that instructors also bear the responsibility of helping learners learn on their own was the main stance in 'the good language learner' studies (Rubin, 1975). However, the concern of whether strategies should or should not be taught has generated some discussion as class time might be more productively employed in teaching

language than in teaching strategies (Rees-Miller, 1993). While some researchers argue that teaching language learning strategies can indeed make language learning more effective (Griffiths & Oxford, 2014), other researchers argue that the teaching of language learning strategies (i.e., strategy instruction) has not been successful mainly due to uncontrolled variables such as age, educational and cultural backgrounds, beliefs about how to learn a language, and cognitive styles (Rees-Miller, 1993). In fact, in a meta-analysis of studies about second language strategy instruction, Plonsky (2011) argued that the effectiveness of strategy instruction has been obscured due to methodological flaws such as the complexity of variables involved, the absence of reliable and valid instruments, issues associated with cost-benefit ratios, and whether or not there are long-term effects.

More specifically, regarding strategy instruction in L2 listening, Graham and Macaro (2008) noted that the teaching of second language learning strategies cannot be attained by teaching one or two strategies in isolation; instead, the teaching of second language learning strategies should be addressed by teaching clusters of cognitive and metacognitive strategies. In this respect, Cross (2011) argued that strategy instruction in L2 listening has a narrow focus and "does not really go far enough in providing learners with adequate knowledge about the nature of L2 listening, associated challenges, and the cognitive and emotional factors involved" (p. 408). Instead, Cross advocated for metacognitive instruction in L2 listening; that is, a relatively new approach that aims at fostering metacognitive knowledge and regulatory skills. In this sense, although metacognitive instruction in L2 listening is much younger than the field of language learning strategies, metacognitive instruction in L2 listening has subsumed language learning strategies.

Strategy instruction in L2 listening has usually been addressed by demonstrating how specific strategies can be useful and then providing learners with opportunities to practice such strategies. This approach, however, has focused on one or two strategies at a time for a short period of time (Vandergrift & Tafaghodtari, 2010). On the other hand, metacognitive instruction in L2 listening aims at teaching learners how to listen by developing critical aspects of L2 listening such as person knowledge, task knowledge, strategy knowledge and strategy use.

2.2.2 What is metacognition?

Metacognition refers to the knowledge, awareness, and control of a person's thinking and learning strategies (Thomas, 2003). It can be described as consciously thinking about thinking. While cognitive skills are necessary to perform a task, metacognition aids in understanding how a task is performed. That is, if an individual can perform a task, this individual has the necessary cognitive skills. In addition, if this same individual can understand how he/she performs the task, this individual also has metacognitive knowledge. Schraw and Dennison (1994) laid it out as the ability to "reflect upon, understand, and control one's learning" (p. 460).

Metacognition distinguishes between two major components: knowledge of cognition and regulation of cognition. On the one hand, knowledge of cognition, which refers to what individuals know about their own cognition or cognition in general, includes 1) declarative knowledge, 2) procedural knowledge, and 3) conditional knowledge. Schraw (1998) noted that declarative knowledge refers to "knowing about things", procedural knowledge refers to "knowing how to do things", and conditional knowledge refers to "knowing the why and when aspects of cognition" (p. 114). That is, declarative knowledge refers to knowledge about the

learner himself/herself and the factors that influence his/her performance; procedural knowledge refers to how individuals perform tasks more automatically by using and orchestrating strategies; and, conditional knowledge refers to when and why to use declarative knowledge and procedural knowledge. Conditional knowledge also assists learners in adjusting to the different situations and demands of the learning tasks.

On the other hand, regulation of cognition, which refers to activities that facilitate the management and control of learning, essentially consists of three regulatory skills: planning, monitoring, and evaluation. Planning refers to the selection of appropriate strategies to perform a task, monitoring refers to awareness of task performance, and evaluation refers to assessment of task performance and learning (Schraw, 1998). For example, when engaged in a second language listening comprehension activity, planning may involve making predictions and/or allocating one's attention to aspects such as what strategies and information one is going to need. During the listening activity, monitoring may involve a regular checking on aspects such as whether the task makes sense or not and whether the interpretation or understanding of the audio text needs to be modified or not. Once the listening activity is completed, evaluation may involve an assessment of what worked or did not work as well as what could be done differently the next time one engages in a listening comprehension activity.

Metacognition is the ability to think about our own thinking. It is the knowledge about our own cognitive processes and our active monitoring, regulation, and orchestration of these processes (Flavell, 1976: 232). Metacognition is our ability "to step back, as it were, from what occupies our mind at a particular moment in time to analyze and evaluate what we are thinking" (Vandergrift & Goh, 2012: 84). Metacognition is the ability that allows individuals to visualize and think of how they perform as they actively engage in analyzing, monitoring, evaluating, and

regulating their cognition. It allows individuals to construct an understanding of themselves as well as to control and to monitor their thoughts and behaviours. In this sense, as noted above, metacognition comprises knowledge of cognition and regulation of cognition.

Metacognition is essential in learning, and its development is an important goal in education since metacognition enables individuals to use their cognitive skills more efficiently. It contributes positively to "problem solving and the development of higher-order thinking skills" (Thomas & Mee, 2005: 221). Metacognitive knowledge is very important because, as a specific type of knowledge, it can be constructed in the classroom by means of organized instruction and effective use of learning strategies. Similarly, metacognitive regulation contributes to a better allocation of attentional resources, improves the use of existing strategies, and creates awareness of comprehension breakdowns. That is, when metacognitive regulatory skills are part of classroom instruction, significant improvements in learning occur (Schraw, 1998). In fact, Schraw and Dennison (1994) indicated that "metacognitive awareness allows individuals to plan, sequence, and monitor their learning in a way that directly improves performance" (p. 460).

2.2.3 Metacognition in L2 Listening

The construct of metacognition was first brought to second language learning by Wenden (1987) who mentioned that metacognition contributes to learner autonomy. Wenden (1998) argued that guiding and fostering metacognitive behaviours is beneficial to second language learners because metacognition contributes to the development of knowledge and awareness about how to "learn more efficiently and how best to improve their learning outcomes" (p. 531).

In this sense, Wenden (1987, 1998) argued that metacognitively aware second language learners are self-directed, reflective, and can take control of their own learning.

Although this perspective is quite congruent with approaches that emphasize learner-centeredness and learner autonomy (e.g., the Communicative Language Teaching approach), metacognition in language learning has not been sufficiently explored. In second language listening, for example, metacognition has relatively recently been empirically explored (e.g., Vandergrift & Tafaghodtari, 2010). In this regard, Vandergrift and Goh (2012) proposed a framework to address listening in the second language classroom. This framework, which places metacognition as an overarching category that subsumes language learning strategies, aims at developing knowledge about cognition (metacognitive knowledge) and regulation of cognition.

The proposed framework serves two main functions in language learning: 1) self-management or control of cognition, and 2) self-appraisal or knowledge about cognition. Vandergrift and Goh (2012) argued that self-management aids in orchestrating and regulating cognition, and self-appraisal "occurs through personal reflections about one's ability and means to meet the demands of a cognitive goal" (p. 85). They addressed these two functions on the basis of three components: metacognitive experience, metacognitive knowledge, and strategy use.

Regarding the first component of the framework, Vandergrift and Goh (2012) noted that metacognitive experience "is a thought or feeling that occurs to a person during and about the main thought" (p. 86). A metacognitive experience happens, for example, when a second language learner does not understand a word, a phrase, or a sound in the listening input, but remembers a similar situation in which the learner managed to solve a comprehension problem and applies the strategy she/he used before to solve the current comprehension problem.

Metacognitive experiences can have a lasting impact as it is the case when a learner recurrently associates and refers back to similar previous experiences to solve current listening comprehension problems in the second language. On the other hand, a metacognitive experience can have no impact at all and just disappear without trace. This happens when a second language learner perceives an unfamiliar word, phrase, or sound in the listening input but immediately forgets it. This makes the metacognitive experience not productive.

Therefore, for a metacognitive experience to be useful and productive, it needs to have a lasting impact on the listener. As Vandergrift and Goh (2012) noted, a metacognitive experience is only useful "if it leads to some productive application of strategies or further understanding about the task, themselves, and/or the world around them" (p. 86). Metacognitive experience is at the center of this framework and can influence either of the other two components (i.e., metacognitive knowledge and strategy use).

The second component of the framework is metacognitive knowledge. This knowledge refers to "an individual's understanding of the ways different factors act and interact to affect the course and outcome of learning" (Vandergrift & Goh, 2009: 401). This knowledge, which is "similar in structure and function to other kinds of knowledge in long-term memory" (Borkowski, 1996: 392), is concerned with person knowledge, task knowledge, and strategy knowledge.

Person knowledge refers to personal factors that support or hinder listening comprehension (e.g., anxiety and/ or negative feelings). It is concerned with cognitive and affective factors that support listening comprehension and listening development. It is knowledge about "ourselves as learners and the beliefs we have about what leads to success or failure in learning" (Vandergrift & Goh, 2012: 86). Person knowledge has to do with how we learn and the

factors that affect our learning. For instance, a second language learner who experiences problems in second language listening may develop the belief that listening comprehension is impossible. This in turn may cause negative feelings and make the learner avoid such situations.

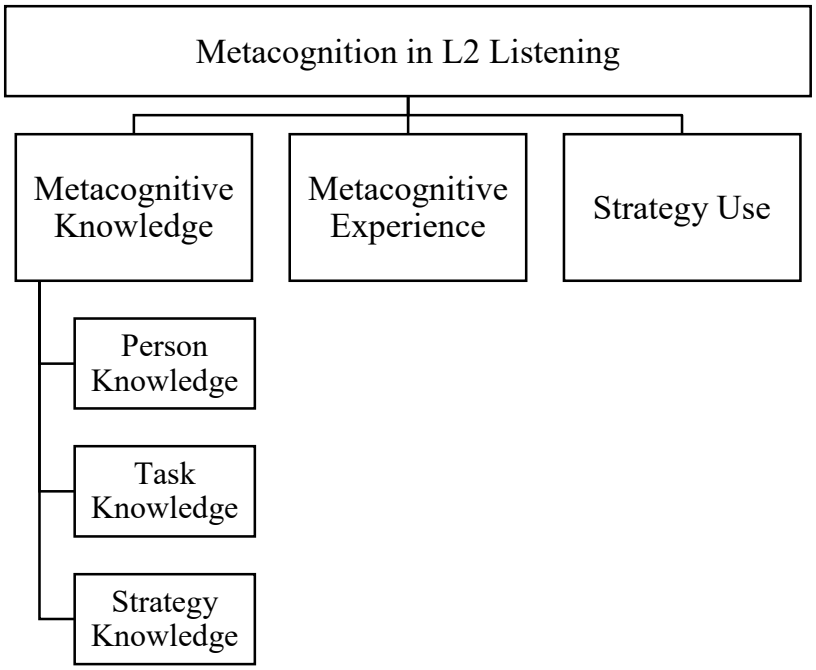
Task knowledge refers to the organization, demands, purpose and nature of the listening task. Task knowledge includes knowing about phonetic and phonological features, grammatical forms, and discourse structures. This knowledge is also concerned with "how to approach and complete a real-life listening task" (Vandergrift & Goh, 2012: 86). For example, in a real-life interaction that involves a telephone conversation, the listener needs to be aware of different accents as well as the discursive organization and conversational rules of the interaction (e.g., negotiation, clarification, turn-taking, maintenance and termination).

The last element of metacognitive knowledge is strategy knowledge. This knowledge refers to strategies that enhance listening comprehension. It is concerned with "knowing which strategies can be used to accomplish a specific goal, be it achieving comprehension in a specific communicative context or improving one's listening ability" (Vandergrift & Goh, 2012: 87). Strategy knowledge is different from strategy use in that strategy knowledge is only concerned with knowing about strategies. Also, in the proposed framework, strategy knowledge is one of the three elements of metacognitive knowledge (i.e. person knowledge, task knowledge, and strategy knowledge), but strategy use is one of the three major components of metacognition (i.e. metacognitive experience, metacognitive knowledge, and strategy use).

The third major component of the framework is strategy use. Strategy use builds on strategy knowledge and refers to our ability to select and combine appropriate strategies to successfully meet the demands of the listening task. Vandergrift and Goh (2012) defined strategy use as "the deployment of specific procedures or actions to make learning easier, faster, more

enjoyable, more self-regulated, more effective, or more transferable to new situations" (p. 89). Interestingly, this definition of strategy use is very similar to Oxford's (1990) definition of language learning strategies: "specific actions taken by the learner to make learning easier, faster, more enjoyable, more self-directed, more effective, and more transferable to new situations" (p. 8). The similarity of these two definitions shows how the framework of metacognition for second language listening has subsumed language learning strategies under the component of strategy use. In fact, strategy use corresponds to language learning strategies, which is a field that has been explored more extensively and which compiles most of the literature in second language listening. Metacognitive strategies (e.g., planning, monitoring, and evaluating) have always been part of the field of language learning strategies; however, these strategies only constitute one aspect of metacognition. Figure 1 below summarizes the different components of the framework proposed by Vandergrift and Goh (2012).

Figure 1: Graphic Representation of the Metacognitive Framework in L2 Listening.



Metacognition in second language listening is critical in the process of developing listening skills. The cognitive, social, and affective factors that metacognition brings to second language listening aid learners in the process of 'learning how to listen'. Aspects such as person knowledge, task knowledge, strategy knowledge, and strategy use enable learners to take control over their learning because processes that are usually unseen become tangible. This allows learners to self-regulate and address listening more strategically. Learners develop a positive concept of themselves as they become capable of tackling listening comprehension more strategically. In fact, research shows a positive and significant relationship between metacognition and listening performance. For example, Goh and Hu (2014) found that metacognitive ability accounted for 22% of the variance in L2 listening performance.

As metacognition in second language listening becomes central in the process of learning how to listen, metacognitive instruction emerges as a means to address second language listening more tangibly and more systematically. Research in the systematic development of metacognition in second language listening is a relatively recent development. Metacognitive instruction aims at triggering metacognitive experiences and fostering metacognitive knowledge and strategy use in an effort to raise listeners' awareness of second language listening processes as well as the ability to self-regulate these processes.

2.2.4 Metacognitive Instruction in L2 Listening

Metacognitive instruction in second language listening is a pedagogical approach that aims at developing learners' knowledge about themselves and about the demands of second language listening as they explore, discover, and synchronize appropriate ways to manage their listening comprehension and listening development. It is a process-based approach that enhances the act of listening by helping learners plan, focus their attention, self-monitor their progress, employ and modify multiple strategies, and self-evaluate their goals and progress. Cross (2015) laid out metacognitive instruction as a pedagogical approach to listening that aims at developing "learners' person knowledge, task knowledge, and strategy knowledge and their ability to self-manage their listening through a range of process-based instructional activities which stimulate metacognitive experiences" (p. 886).

In this respect, metacognitive instruction in second language listening is concerned with developing metacognitive knowledge and strategy use. It aims at fostering a deeper understanding of ourselves as second language listeners as well as an understanding of the nature

and demands of second language listening. It also promotes knowledge and use of cognitive, metacognitive, and socio-affective strategies for self-management and listening development. In this sense, metacognitive instruction underscores the act of 'learning to listen' as an individual cognitive activity and as a constructive social development. As Vandergrift and Goh (2012) put it:

> Metacognitive instruction refers to pedagogical procedures that enable learners to increase awareness of the listening process by developing richer metacognitive knowledge about themselves as listeners, the nature and demands of listening, and strategies for listening. At the same time, learners also learn to plan, monitor and evaluate their comprehension efforts and the progress of their overall listening development. (p. 97)

Metacognitive instruction in second language listening provides learners with opportunities to engage in self-appraisal and self-management as they experience listening comprehension activities and plan for overall listening development. It guides learners through a systematic and principled pedagogical cycle that engages learners in developing person knowledge, task knowledge, strategy knowledge, and effective strategy use. Some objectives associated with metacognitive instruction in second language listening include: 1) to examine personal beliefs about can-do attitudes and self-concepts, 2) to identify listening problems and possible solutions, 3) to be aware of the factors that influence L2 listening performance, 4) to identify and demonstrate use of appropriate strategies for specific situations and different types of listening such as listening for details and/or listening for global understanding, and 5) to be able to plan, monitor, and evaluate listening comprehension and overall listening development (Vandergrift & Goh, 2012).

As such, metacognitive instruction in second language listening is beneficial in several respects. It empowers learners as they self-regulate their L2 listening and take control of the act

of listening. It helps learners be more confident, less anxious, and more motivated. In this respect, Graham and Macaro (2008) highlighted the importance of learners' beliefs about their own abilities and competences. They noted that those learners who possess control over their learning display higher levels of motivation and 'can-do' attitudes. In their words, "learners who attribute the level of their achievement on academic tasks to factors within their control (e.g., to effort expended or to strategies employed) are likely to have higher levels of self-efficacy and to be motivated to attempt similar learning tasks again" (p. 755).

Metacognitive instruction turns second language learners into self-regulated listeners capable of setting goals, monitoring, regulating, and controlling their thinking, motivation, and behaviour (Pintrich, 2000). It boosts learners' listening comprehension ability by increasing their second language listening capacity as they actively engage in a process of learning to listen that allows them to monitor, regulate, and control their thinking and strategic behaviour. Furthermore, metacognitive instruction not only boosts learners' listening ability in classroom activities, but it also has the potential to increase learners' listening ability in situations outside the classroom. As Field (2012) argued, "metacognitive training has potential benefits beyond the classroom. Enabling learners to handle listening exercises strategically boosts their confidence when they come to real-life listening encounters" (p. 211).

Overall, metacognitive instruction impacts the act of listening positively by helping learners self-regulate their abilities and better manage their listening comprehension and listening development. This pedagogical approach, which fosters reflection and self-regulation, employs diaries, questionnaires, discussions, and tasks. Diaries, for example, are useful teaching tools to promote discussions about strategies, beliefs, and attitudes towards listening. This helps learners evaluate their strategies and try out new ones (Goh, 1997). Also, tasks that engage

learners in predicting, monitoring, problem solving, and evaluation provide learners with the practice they need to tackle second language listening more effectively. This kind of task, which Vandergrift (2004) termed "metacognitive cycle", aids in the construction of metacognitive knowledge and the development of self-regulated listening (pp. 10-13).

While reflection helps learners develop knowledge about themselves and about what second language listening entails, tasks in the form of metacognitive cycles help learners with guided listening practice to tackle second language listening more systematically and strategically. In fact, after examining data from several sources (i.e., questionnaire, simulated-recall protocols, diaries, think-alouds and an open-ended questionnaire), Mareschal (2007) observed how French language learners better regulated their listening comprehension as a result of this metacognitive cycle.

Metacognitive instruction to second language listening is not only about 'thinking about thinking'. It is about critically reflecting on one's "knowledge about learning, before, during, or after a particular listening experience or task" (Vandergrift & Goh, 2012: 93). It is about constructing knowledge and involving other agents in the environment such as classmates and realia to facilitate learning. In this particular sense, listeners need to do more than just thinking about their cognition. They need to act on their thoughts, knowledge, experiences, and strategic behaviours. They need to critically analyze their actions, reflect, and take action upon the feedback they generate or receive.

As noted earlier, metacognition in second language listening is an overarching term that has subsumed language learning strategies. In addition, although the field of language learning strategies has been explored more extensively, the metacognitive framework described here provides a more comprehensive ground to handle second language listening in the classroom. As

Cross (2015) noted, “metacognitive instruction is a holistic approach to L2 listening instruction which aims to enhance each of the three knowledge factors [i.e., Metacognitive Experience, Metacognitive Knowledge, and Strategy Use] to be inculcated in L2 listeners and does not just focus narrowly on promoting strategy knowledge (and use)” (p. 886). In this sense, metacognitive instruction in second language listening is a broader construct that encompasses language learning strategies under strategy use (see Figure 1 above).

2.3 Intervention Studies in L2 Listening

This section describes intervention studies in L2 listening in detail. Because the field of language learning strategies preceded metacognitive instruction in L2 listening and it is the field which has been researched more extensively, the first part of this section focuses on intervention studies in L2 listening strategy instruction, whereas the second part of this section deals with intervention studies in L2 listening metacognitive instruction.

2.3.1 Intervention Studies in L2 Listening Strategy Instruction

Rubin (1990) and her colleagues conducted an experiment about teaching listening strategies. The study looked at the effects of teaching listening strategies using videos on listening comprehension. Their study was conducted in seven high schools in a Californian school district with 394 participants. The study included three treatment groups and two control groups, and implemented a pre/post-test design.

The study was conducted over eight weeks during which twelve interventions were implemented. The interventions consisted of "training in the use of strategies [that] would improve listening comprehension" (Rubin, 1990: 313). These consisted of 40-45 minute lessons that included two five-minute videos each. The strategy training used was a 'story line strategy' which required participants to figure out the main plot of the video. Then, after examining the data and running statistical analyses, Rubin and her colleagues argued that when the video material was easy, strategy instruction had no effect on listening comprehension. However, when the video material was difficult, strategy instruction did have an effect on listening comprehension: "When the text or task is just hard enough, strategy training can improve the performance of students" (Rubin, 1990: 313). They concluded that combining videos that provide sufficient clues for information processing and strategy instruction with effective listening strategies helped learners cope with more challenging listening material and also helped improve their motivation.

It must be noted, however, that in this study, it seems that the probability of committing a type I error is higher than the set alpha level; this is particularly the case when several t-tests are run separately in one single analysis. A more stringent procedure would have perhaps combined a between-groups and within-subjects design. Additionally, although the videos were carefully selected, this introduces biases in the study because it is difficult to know objectively when the videos are hard enough for the students.

In another intervention study, Thompson and Rubin (1996) examined the effects of cognitive and metacognitive strategy instruction on listening comprehension performance of university learners of Russian over two academic years. In the first academic year, participants were assigned to two separate groups by random assignment. In the second academic year, only

an experimental group was formed with an intact class. The authors used the same instrument as pre-test and post-test, which consisted of a video comprehension test and an audio comprehension test. The video comprehension test was created for this particular study and the audio comprehension test was the listening portion of a Russian proficiency test designed by the Educational Testing Service. Participants in the experimental and control groups used the same course material and followed the same syllabus. However, participants in the experimental group focused on developing listening strategies and participants in the control group focused on using the content of the videos as a basis for speaking and writing activities.

The results of this study showed that participants receiving strategy instruction scored significantly higher on the video listening test, but not on the audio test. Thompson and Rubin (1996) argued that the 15 hours of instruction were not sufficient to reflect a larger gain in listening comprehension. They also noted that listening strategy instruction should be tailored to different proficiency levels since learners with low listening comprehension skills may not benefit much from aural input without visual support. The researchers concluded that "…systematic instruction in the use of cognitive and metacognitive strategies will result in improvement of listening comprehension…" (p. 336).

The fact that the participants receiving strategy instruction scored significantly higher on the video listening test but not on the audio listening test might be a concern because the latter test was the only standardized measure used in the study. Thus, the significant results may be due to the similarity of the tasks in the interventions and the video listening test used as post-test, and not necessarily due to listening strategy instruction.

In a study that included only a post-test with random assignment of participants to groups, Kohler (2002) looked at how metacognitive strategies affected language knowledge,

performance, and strategy use of adult second language learners of Spanish. In this study, a language learning program that combined computer-assisted and classroom activities was implemented in the experimental group for eight weeks. The computer-assisted component of strategy instruction consisted of an online tutorial and tasks dealing with the importance, meaning, and elements of metacognitive language learning strategies. The classroom activities component was a hands-on training that gave participants an opportunity to actually use the strategies as opposed to simply attending a lecture about strategies. Additionally, participants in the experimental group engaged in discussions about metacognitive strategies. The amount and frequency of the interventions decreased from once a day to a couple of times a week towards the end of week eight. Participants were measured on grammar knowledge, listening comprehension, task performance, vocabulary knowledge, phrase knowledge, spontaneous strategy use, and metacognitive strategy use.

After examining the data and running multivariate analyses of covariance, Kohler's (2002) study found the following: a) low-achieving learners trained with metacognitive strategies improved in vocabulary and phrase knowledge, b) significant improvement in L2 listening performance was also found, and c) regarding language learning strategy use, significant differences in speaking, listening and vocabulary strategy use were found. Kohler argued that her findings supported O'Malley and Chamot's (1990) findings regarding the effectiveness of metacognitive strategies over cognitive strategies in listening (p. 67). Also, participants who had strategy instruction on metacognitive strategies did significantly better in listening, vocabulary, and phrase knowledge.

Among the limitations, Kohler (2002) mentioned that generalizing the results of this study to other students, languages, ages, cultures, or training systems should be cautiously

interpreted. Even for the population from which the sample was taken, generalizations of these results may not be appropriate due to the small size of the sample (70 English-speaking missionaries) in relation to the population of interest. Also, participants were not measured at pre-test and the data was largely self-reported reflecting what participants thought they did as opposed to what they actually did.

In a study investigating the effectiveness of strategy instruction on listening proficiency, Graham and Macaro (2008) found that strategy instruction focusing on clusters of cognitive and metacognitive strategies had a positive impact on listening comprehension. The participants in this study (N=107) were students of French as a foreign language and their ages ranged from 16 to 17. The samples were taken from 15 schools in the south of England. This study incorporated two intervention groups and one control group. Three tests were used to measure listening ability at pre-test, post-test and delayed post-test. These tests consisted of different audio-recordings that were based on the same topic. The level of difficulty of these tests was calibrated by the ratings of an additional group of participants who were not involved in the study. Participants listened to these audios twice and had to write down what they thought they understood. Also, after the pre-test and at post-test, participants were asked about their perception regarding their listening ability and how confident they felt about their listening skills.

After statistically controlling for initial differences in listening ability (i.e., pre-test scores were used as a covariate), the results showed that the intervention groups (n=29 & n=39) scored significantly higher than the control group (n=39) at post-test. Similarly, at delayed post-test, the intervention groups (n=20 & n=11) scored significantly higher than the control group (n=28). Graham and Macaro (2008) attributed the positive results of their study to the high level of focus on clusters of cognitive and metacognitive strategies such as prediction, direct attention,

phonemic segmentation, inference, verification, monitoring, and evaluating. In their words, "clustering of strategies was the first theoretical underpinning of our intervention program. We would argue that individual cognitive strategies cannot be considered, and therefore taught, in isolation" (p. 770). An interesting feature of this study has to do with the instruments of data collection and the intervention. The study implemented a listening measure that was not practiced during the intervention. This is particularly important because it ensures that it is the intervention, and not the practice, what actually causes the effect. In their words, "we ensured that the testing procedure did not bias the intervention students by implementing a listening-test type that was not practiced during the strategy instruction program. By doing so, we believe that we also provide evidence that students were able to transfer their strategic behaviour from the tasks they engaged in during the instruction to a different task in the listening tests" (p. 772).

To summarize, the very few intervention studies in listening strategy instruction that are currently available show the following: strategy instruction which combines videos and strategy instruction can help learners better deal with more challenging listening material (Rubin, 1990). Also, strategy instruction on metacognitive strategies appears to be more effective with listening comprehension and vocabulary (Kohler, 2002). Additionally, strategies should not be taught in isolation but as clusters of cognitive and metacognitive strategies such as prediction, direct attention, phonemic segmentation, inference, verification, monitoring, and evaluating (Graham & Macaro, 2008). Also, systematic instruction in the use of strategies can improve listening comprehension (Thompson & Rubin, 1996). Future studies need to consider methodological concerns such as a) inflated alpha levels due to running multiple statistical tests instead of running one single test, b) the similarity of the tasks in the interventions and the outcome measure can actually give the treatment group an advantage over the control group, c) self-

reported data can reflect what participants think they do as opposed to what they actually do. In fact, the validity of strategy instruction has been questioned due to methodological flaws, the complexity of variables involved, the lack of reliable instruments to measure outcome variables, and the absence of evidence of long term effects (Plonsky, 2011). These intervention studies have, however, provided evidence congruent with the notion that instruction of second language listening should focus on teaching clusters of cognitive and metacognitive strategies as well as with the notion that listening instruction should be addressed more systematically in the language classroom. Also, regarding long-term effects, Graham and Macaro (2008) provided preliminary evidence of such effects.

2.3.2 Intervention Studies in L2 Listening Metacognitive Instruction

In a relatively recent study which marks the beginning of metacognitive instruction in L2 listening, Vandergrift and Tafaghodtari (2010) conducted an intervention study that investigated the effects of a metacognitive pedagogical cycle. This pedagogical cycle incorporated stages of listening such as planning, predicting, verification, and reflection. In addition, embedded within these stages, strategies such as planning, direct attention, monitoring, problem solving and evaluation guided the process of listening (see section 2.2.4). This metacognitive pedagogical cycle aims at developing overall strategic second language listeners by teaching learners how to listen in the second language. The participants in this study were university students learning French as a second language. This study used a pre-/post-test design with control and experimental groups. The six intact classes in this study were randomly assigned to either the control group (n=47) or the experimental group (n=59). Then, based on the pre-test scores,

participants in each cohort were identified as less skilled listeners or more skilled listeners. The pre-test and post-test were a version of a placement test used at this university and its internal consistency was .94. The tests took approximately 20 minutes to complete. Since random assignment was not feasible, group differences on the post-test were analysed after statistically controlling for initial differences. The development of metacognitive awareness over time was also assessed using a questionnaire—the Metacognitive Awareness Listening Questionnaire (MALQ).

The results showed that less skilled listeners in the experimental cohort outperformed the less skilled listeners in the control cohort; however, more skilled listeners in the experimental cohort did not significantly differ from the more skilled listeners in the control cohort. Additionally, the less skilled participants in the experimental group showed greater improvement than the more skilled participants in the experimental group. Also, the less skilled listeners in the experimental group reported greater growth in two of the five factors associated with metacognitive knowledge; that is, problem solving and mental translation. This intervention study showed that metacognitive instruction in L2 listening can improve listening performance, and it is the less skilled listeners who actually benefit more. This pedagogical cycle appears to be effective, especially when the pedagogical intervention is conceived as a process-based approach. This study provides evidence of the effectiveness of teaching learners 'how to listen'. However, although a common practice, statistically controlling for initial differences (i.e., ANCOVA) without random assignment of participants to groups might be a limitation.

In a small-scale study investigating how less-skilled listeners benefited from metacognitive instruction, Cross (2011) tailored five listening lessons using the pedagogical cycle proposed by Vandergrift (2004). Each of the ninety-minute listening lessons incorporated

features such as predicting information, confirming initial predictions, talking about strategies, discussing and reconstructing the audio text, matching words in the audio transcript with the audio sounds, and reflection and evaluation of listening performance and strategic behaviour. One group of participants composed of 20 EFL learners participated in this study. After comparing the percentages of the pre-test and post-test, Cross (2011) found that less-skilled listeners improved their scores on the post-test. This study provides support to the notion that this pedagogical cycle affects second language listening positively. However, the lack of a comparison group places a limitation on the study. Furthermore, even though the study was not set up to apply statistics, a t-test could have been run to assess whether the differences at post-test were statistically significant.

In a study that apparently applied the same metacognitive cycle as Vandergrift (2004), Rahimirad (2014) found statistically significant differences between an experimental group and a control group. The study implemented eight ninety-minute listening lessons during five weeks. The study involved 50 intermediate EFL students at an Iranian University who were randomly assigned to the two groups. Their listening performance was measured before and after the intervention with a pre-test and a post-test, which were modified versions of a TOEFL practice test. The tests had 31 questions and the allocated time for the administration was 45 minutes.

The first five sessions in the intervention mainly consisted of lectures and examples about planning, prediction, monitoring, reflection and evaluation. The remaining three sessions seemed to be similar to the pedagogical cycle used in Vandergrift and Tafaghodtari (2010). Unfortunately, there is no specification of how the listening lessons were implemented, especially the last three sessions. It appears that during the first five listening lessons the

researcher engaged in teaching metacognitive strategies, and during the last three remaining sessions, the metacognitive cycle proposed by Vandergrift (2004) was applied.

This study exhibits a lack of theoretical consistency, mixing terminology such as 'metacognitive strategies' and 'metacognitive instruction'. Also, although the author mentioned that the only difference between the two groups was 'strategy instruction', there are no specifics about how the control group was handled. This is important because if the control group did not engage in any listening activity, it may seem obvious that this group will not show as much improvement as the group receiving instruction. That is, the significant results of the study could be attributed to the extra practice the experimental group experienced during the intervention. While this study applied random assignment of participants to groups, which seemed to work well since the pre-test scores of the two groups were very similar, alpha levels might have been inflated due to running multiple statistical tests separately. A more appropriate procedure is perhaps running one factorial ANOVA instead of running three t-test separately.

In a similar study also with Iranian EFL learners, Fahim and Fakhri (2014) examined a group of 30 participants who were exposed to metacognitive instruction in L2 listening. The pedagogical cycle was very similar to Vandergrift (2004). The study examined the effects of metacognitive instruction on L2 listening performance and metacognitive awareness of L2 listening, but it did not incorporate a control group. The 30 participants in this study completed a standardized pre-test and post-test, as well as a questionnaire (i.e., Metacognitive Awareness Listening Questionnaire—MALQ). The intervention was carried out over a period of 10 weeks. However, it is not mentioned how long each intervention lasted and how many intervention sessions were conducted. Also, to analyze the results of listening performance, the researchers used a non-parametric test (i.e., Wilcoxon signed-rank), whereas to analyze the results of the

questionnaire, the researchers used a pair-samples t-test. Although the results turned out to be significant for both listening performance and metacognitive awareness, the insufficient information provided on how the treatment was conducted and the lack of specifics regarding the statistical tests used to analyse the data undermine the validity of this study.

Wang (2016) designed a study to investigate whether or not metacognitive instruction in L2 listening had an effect on listening performance and metacognitive knowledge. This study involved two classes at a university in China and the cohort consisted of 100 non-English major students who were learning English as a foreign language. The participants were randomly assigned to either the experimental group (i.e., 45 participants) or the control group (i.e., 55 participations), and the intervention was conducted once a week for an hour during 10 weeks. The intervention used the same metacognitive pedagogical cycle as Vandergrift (2004). The experimental group engaged in reading and watching short videos, predicting vocabulary and information, sharing and comparing information in small groups as well as a whole class in an attempt to reconstruct the audio text, matching words in the audio transcript with the audio sounds, and evaluating and reflecting on their listening performance. As for the control group, they listened to the same material as the experimental group, but they followed the pre-established listening activities in their textbook.

Listening comprehension was measured at pre-test and post-test with a standardized test designed and used by a Chinese national proficiency testing agency. These results were analysed with a paired-samples t-test which turned out significant for both groups. That is, both control and experimental groups made significant gains. Also, the groups were compared at post-test using an analysis of covariance (ANCOVA). The results of this analysis were not significant, indicating that both groups performed at a similar level. The researcher attributed the non-

significant results of the post-test analysis to biases in the testing. That is, the post-test favoured the control group because the exercises and questions that were practiced during class time were very similar to those in the post-test. Thus, to a certain degree, the instructional approach used in the control group replicated the testing situation at post-test. Also, although Wang (2016) found growth in metacognitive knowledge in the experimental group (i.e., person knowledge, task knowledge, strategy knowledge), this growth could not be compared with the control group because this type of data was not collected from the control group. An interesting feature in this study is the random assignment of participant to the groups. This is an important aspect in any study because random assignment controls for extraneous variables that threaten the internal validity of a study. In this study, two separate paired-samples t-tests were used to analyse the data. This may have inflated the alpha levels and increased the probability of Type I error. Although the gains in listening ability in this study were large, a repeated measures ANOVA would have been a more appropriate model for the analysis of the data.

To summarize, the metacognitive pedagogical cycle proposed by Vandergrift (2004) seems to significantly improve listening performance (e.g., Vandergrift & Tafaghodtari, 2010; Wang, 2016). Also, this metacognitive pedagogical cycle impacts listening positively in that less-skilled listeners are the ones who benefit the most (e.g., Cross, 2011). However, the very few intervention studies in the field of metacognitive instruction in L2 listening limit the feasibility of generating stronger conclusions regarding the effectiveness of this metacognitive pedagogical cycle. Moreover, methodological flaws such as inappropriate statistical procedures, lack of a comparison group, and incomplete information regarding the intervention (e.g., Fahim & Fakhri, 2014) raise concerns about the validity of certain studies. Also, the lack of understanding of the construct of metacognitive instruction in L2 listening (e.g., Rahimirad,

2014) limits the validity of some studies. Although these studies provide information regarding this metacognitive cycle in different contexts and for different learners, the methodological flaws in these studies make it difficult to assess the effectiveness of this metacognitive cycle across studies.

2.3.3 Final Remarks regarding Intervention Studies in Strategy Instruction and Metacognitive Instruction in L2 Listening

Although these intervention studies about strategy instruction in L2 listening and metacognitive instruction in L2 listening have used treatments, control groups, pre-/post-tests, and/or statistical procedures, there are methodological and theoretical concerns that need to be considered carefully and addressed more rigorously. For example, methodological flaws associated with the design and statistical procedures such as inflated alpha levels and lack of specifics regarding the interventions raise concerns about the validity of some studies. Also, misinterpretations of the construct of metacognitive instruction in L2 listening result in an inappropriate articulation of perspectives that end up mixing 'metacognitive strategies' and 'metacognitive instruction'. This generates confusion and misunderstanding among researchers and scholars in the field. Additionally, it is important to carefully evaluate the situation of comparing a treatment group with a control group. This is particularly important because it is not unreasonable that a group receiving an intervention will somehow do better than a group not receiving such intervention. That is, improvement might be attributed to the extra practice and not necessarily to the intervention. Therefore, while addressing these methodological and theoretical concerns, it is also necessary to actually contrast different approaches in L2 listening

as opposed to comparing a group that receives a pedagogical intervention and a group that does not receive any instruction.

2.4 The Present Study

The literature review in this chapter showed that segmenting and identifying words pose serious difficulties to second language learners, and that L2 listening causes anxiety that often impacts second language learners negatively. This review also showed that although L2 listening is an important skill in language teaching and learning, this skill is often taken for granted. In fact, L2 listening is often referred to as the Cinderella of the language skills, and it is the least-researched skill (Cross, 2015). The small number of intervention studies indicates that teaching isolated strategies does not impact L2 listening positively. Instead, the teaching of L2 listening is more effective and meaningful when clusters of cognitive and metacognitive strategies are considered together (Graham & Macaro, 2008). Also, listening instruction needs to be implemented systematically in the language classroom since "systematic instruction in the use of cognitive and metacognitive strategies will result in improvement of listening comprehension" (Thompson & Rubin, 1996: 336). Research also indicates that instruction in L2 listening should be viewed as a learning process in which learners are given opportunities to experience the challenges and demands of L2 listening as well as to employ their strategies and try out new ones (Vandergrift & Tafaghodtari, 2010).

However, these important findings have not been empirically verified due to the very limited number of intervention studies. As Macaro et al. (2007) noted: the field "urgently need[s] more intervention studies that identify more clearly what kind of strategy instruction works with

what kinds of learner" (p. 185). In this sense, the current study contributes to this limited number of intervention studies and fills the gap regarding the lack of studies that compare the effects of different types of L2 listening instruction (Berne, 2004). It also addresses concerns associated with long-term effects (Plonsky, 2011). Therefore, the current study contributes to the advancement of L2 listening instruction by examining the effects of three approaches to L2 listening: 1) a metacognitive pedagogical cycle in which learners predicted information, confirmed or disconfirmed information, reconstructed the audio text, and evaluated and reflected on their performance (Vandergrift & Tafaghodtari, 2010); 2) an awareness-raising approach in which learners discussed L2 listening strategies associated with successful L2 listening, and evaluated and reflected on their performance; and, 3) a comprehension approach in which vocabulary was explained and guiding questions intended to build a better understanding of the listening text were employed (e.g., Field, 2012).

This study contributes to the existing body of research in several respects because currently a) there is a very limited number of intervention studies dealing with L2 listening, b) there is no study that has assessed the effectiveness of different pedagogical approaches to L2 listening, c) there is no study that has looked at metacognitive instruction in L2 listening with learners of Spanish, and d) there is no study that has looked at long-term effects of metacognitive instruction in L2 listening. In this sense, this intervention study contributes to the body of literature by assessing the effectiveness of these instructional approaches to L2 listening in a Spanish as a foreign language context as well as by examining long-term effects. The research questions that guided this study are presented below:

RQ-1: The first research question in this study is concerned with whether or not there are significant changes in L2 listening performance among three groups of participants who were exposed to three different types of listening instruction, namely, a metacognitive pedagogical sequence, an awareness-raising approach of the factors associated with successful second language listening, and a conventional approach. More specifically,

RQ-1a: Are there significant changes in L2 listening performance within each group from pre-test to post-test?

RQ-1b: Are there significant differences in L2 listening performance between the groups at post-test?

RQ-2: Are there long-term effects associated with the different types of listening instruction?

RQ-3: What are the learners' perceptions about the L2 listening activities implemented in the three different groups?

The following chapter describes in detail the methodological decisions made in order to address these research questions. It describes aspects associated with participants, context of the study, treatment groups and intervention sessions, instruments and procedures of data collection, and the analytical tools employed to analyze the data.

3 Methodology

3.1 Overview

This study compares three groups of language learners who were exposed to three different approaches to second language listening. The pedagogical intervention comprised eight second language listening sessions which were carried out in three intact classes. In this sense, this study complies with the characteristics of a quasi-experimental design in which an experimental procedure is applied, but random assignment of participants to treatment groups is not possible (Christensen, Johnson & Turner, 2011). However, although random assignment of participants to groups was not feasible, each of these three intact classes was randomly assigned to one of the three treatment conditions.

Participants' listening ability was measured three times during the study: One week before the intervention at pre-test, two weeks after the intervention at post-test, and five months after the intervention at delayed post-test. The dependent variable was listening performance; thus, the pre-test, the post-test, and the delayed post-test were measures of second language listening. The independent variable was teaching approach to second language listening and had three treatment conditions: 1) Metacognitive Pedagogical Sequence, 2) Awareness-Raising of the Factors associated with Successful Second Language Listening, and 3) Conventional Approach (see Table 1 below).

Table 1: Schematic Representation of the Study

	Group 1	**Group 2**	**Group 3**
Week 0	Pre-Test	Pre-Test	Pre-Test
Weeks 1 through 8	Metacognitive Pedagogical Sequence	Awareness-Raising of the Factors associated with Successful L2 Listening	Conventional Approach
Week 10	Post-Test	Post-Test	Post-Test
Five months later	Delayed Post-Test	Delayed Post-Test	Delayed Post-Test

At the end of every intervention session, an in-class listening test was administered to all of the participants (see section 3.3 below). The scores of these in-class listening tests were also incorporated in this study. Examining these scores was an important aspect of the study because participants were not randomly assigned to the groups, which introduces selection biases to the study (Christensen et al., 2011). The scores of the in-class listening tests allowed to examine the influence of selection biases by verifying if there is consistency in the results yielded by the pre-test and post-test. For example, the first in-class listening test can corroborate the results obtained in the pre-test. Also, the scores of the in-class listening tests can confirm or disconfirm the results of the post-test. Thus, examining the scores of the in-class listening tests contributes to a better interpretation of the results in the absence of random assignment. Since there was a total of eight intervention sessions during the study, eight in-class listening tests were administered. This component was kept separate from the design above for two main reasons. First, the in-class

listening tests were heavily influenced by the different treatments. This caused violations of statistical assumptions associated with the in-class listening tests (see section 4.1.1). Second, the listening measures were different. That is, while the in-class listening tests were researcher-made and consisted of five questions each, the pre/post/delayed tests consisted of practice DELE (in English, 'Diplomas of Spanish as a Foreign Language') exams with 30 questions each (see section 3.4.1 below).

Additionally, this study incorporated a qualitative component to elicit data regarding the usefulness of the pedagogical approaches. That is, focus group interviews were conducted approximately 2 months after the intervention, and two open-ended questionnaires were administered on week seven during the intervention and on week 10 after the intervention.

3.2 Participants and Context of the study

For the study, participants were recruited from three low-intermediate Spanish language classes at a large University in Canada. These classes were taught by three different instructors. However, the content, the tasks, the evaluations, and the material were all the same. The students in all of these groups were exposed to task-based instruction, and all of them were required to complete the same tasks and assignments. Also, the dates for the specific task assignments and evaluations were the same for all the groups. In this sense, the groups were taught in a parallel manner under the same program, and although the three groups were taught by three different instructors, they were all using the same syllabus. The course program for this level consisted of covering the first five units of the Spanish book series "Gente Hoy 2". This is an intermediate Spanish book which consists of ten units. By the end of these ten units, students are expected to

be at the B1 level of proficiency as described in the Common European Framework of Reference for Languages.

There were 55 students in these three classes, and all of them were undergraduate students. Of these 55 Spanish learners, 32 participants were included in this study. The other participants were left out of the study due to reasons such as missing the pre-test and/or post-test, participating in fewer than seven treatment sessions, not consenting their participation, or because one of their parents was a Spanish native speaker. Therefore, the final configuration of the participants is as follows: 13 out of 21 students for the Metacognitive Pedagogical Sequence group; 11 out of 22 students for the Awareness-Raising group; and 8 out of 12 students for the Conventional Approach. Thirty-two participants for this study seemed to be an appropriate number as G*Power calculations yielded a sample size of 33 participants with an actual power of 75%, large effect size, alpha 0.05, and three groups with two measurements (Faul, Erdfelder, Lang & Buchner, 2007). All the participants were native English speakers and were majoring in different areas. Also, these participants reported a very positive attitude towards learning Spanish: None of them responded lower than five on a six-point Likert scale to the statement 'I like learning Spanish', which indicated a high level of motivation.

Additionally, only 17 participants were able to take the delayed post-test, which was administered five months after the last intervention session. The interventions were carried out during the first semester, and the delayed post-test was completed towards the end of the second semester of the next intermediate Spanish course. During this second semester, the configuration of the classes changed as the initial three course sections were reduced to two and not all of the participants who took Spanish in the first semester continued in the second semester. Thus, the final configuration of participants who took the delayed post-test was as follows: nine

participants from the Metacognitive Pedagogical Sequence group, two participants from the Awareness-Raising group, and six participants from the Conventional Approach group.

3.3 Intervention Sessions and Treatment Groups

Every week, there was one intervention session for a total of eight sessions. This number of sessions is similar to the number of interventions implemented in previous studies. Kohler's (2002) and Rubin's (1990) studies were implemented during a period of eight weeks. Similarly, using the same pedagogical approach (i.e., Metacognitive Pedagogical Cycle), Vandergrift and Tafaghodtari (2010) carried out their study in 13 weeks, Cross (2011) implemented five listening lessons in his study, and Wang (2016) implemented ten interventions in ten weeks. In this regard, Vandergrift and Tafaghodtari (2010) noted that 13 interventions can be tedious and boring for the students. Eight interventions in the present study was also logistically appropriate because after the last intervention (i.e., intervention eight at Time 8), there were only three weeks left before final exams. On average, these interventions lasted for about 25 minutes and were all delivered by the researcher. To minimize researcher bias, checklists were employed to deliver the interventions. These checklists contained the sequential order in which the interventions were to be conducted. The date, start and end times of the intervention, and a section for comments were also included in these checklists (see Appendix A).

The audio input for the interventions were taken from the textbook used in these classes. However, the listening activities were created based on the three different approaches. To avoid biases in the selection of the audios, the last two audios of every unit were systematically

selected to create the listening activities for the different groups. Consequently, the instructors were asked not to do these specific audio listening exercises with their students.

In general terms, the treatments for the three groups had the following elements in common: same audio input played twice, same in-class listening tests, and approximately the same intervention time for the three groups (i.e., on average, Metacognitive Pedagogical Sequence group 24.14 minutes; Awareness-Raising group 24.57 minutes; and, Conventional Approach 23.57 minutes). The sections below describe in more detail how the same audio input was delivered in each of the treatment groups, and how the three types of pedagogical approaches to second language listening (i.e., Metacognitive Pedagogical Sequence; Awareness-Raising of the Factors Associated with Successful Second Language Listening; and, Conventional Approach) were implemented in the groups.

3.3.1 Metacognitive Pedagogical Sequence

Metacognitive instruction in second language listening has been the subject of inquiry for some time (e.g., Goh, 1997). However, it is only recently that metacognitive instruction in second language listening has been investigated more systematically. In fact, Vandergrift and Tafaghodtari (2010) were the first ones to systematically implement a metacognitive pedagogical intervention in a quasi-experimental study. This pedagogical cycle aimed at fostering processes such as planning for the listening activity, monitoring comprehension, dealing with and solving problems in comprehension, and evaluating the outcome of listening as well as the particular strategies used. Subsequently, a few other studies have been carried out to test the effectiveness of this pedagogical cycle (e.g., Cross, 2011; Wang, 2016). This metacognitive pedagogical cycle

resembles a process-based approach that guides learners through metacognitive processes in which they explore and develop knowledge about themselves, the demands of second language listening, and the effective strategies that work for them.

Vandergrift and Goh (2012) portrayed this metacognitive approach in a single-headed arrow that goes from 1) planning and prediction, 2) to verification and planning with peers, 3) to verification and text reconstruction, 4) to final verification, and 5) to reflection and goal-setting. This cycle, which embeds metacognitive processes such as planning, monitoring, problem-solving, and evaluation, guides learners through the listening activity as they actively engage in a process of predicting and confirming information, adjusting and trying out new and old strategies, and developing a better sense of what second language listening really entails. Figure 2 below describes the different stages of this approach as well as the metacognitive processes involved in it.

Figure 2: Listening Instruction Stages and related Metacognitive Strategies

Stages of Listening Instruction	**Related Metacognitive Strategies**
Planning/predicting stage 1. Once students know the topic and text type, they predict types of information and possible words they may hear	1. Planning and directed attention
First verification stage 2. Students verify initial hypotheses, correct as required, and note additional information understood.	2. Monitoring

3. Students compare what they have written with their peers, modify as required, establish what needs resolution and decide on details that still need special attention.	3. Monitoring, planning, and selective attention
Second verification stage 4. Students verify points of disagreement, make corrections, and write down additional details understood. 5. Class discussion in which all contribute to reconstruction of the text's main points and most pertinent details, interspersed with reflections on how students arrived at the meaning of certain words or parts of the text.	4. Monitoring and problem solving 5. Monitoring and evaluation
Final verification stage 6. Students listen for information that they could not decipher earlier in the class discussion.	6. Selective attention and monitoring
Reflection stage 7. Based on discussion of strategies used to compensate for what was not understood, students write goals for next listening activity.	7. Evaluation

From Vandergrift, 2004, p. 11.

In this metacognitive pedagogical cycle, during the planning stage, for example, the instructor provides learners with information such as the topic of the listening text, the genre, and the context in which the listening text occurs. With this information, listeners get ready for the listening activity by making predictions about the possible words they might hear and/or the organization and type of the information they might find in the audio text. After this initial stage,

listeners engage in a process of verification and confirmation, reconstruction of the listening text, and reflection.

The metacognitive pedagogical cycle is the approach that constituted the treatment for the Metacognitive Pedagogical Sequence group in the current study. However, minor modifications were made so as not to take too much time from the class instructors, and also because in a testing situation learners might not have the chance of listening to the audio input more than two times. Thus, the second and third verification stages were omitted in the current study. Instead, a second listening stage followed by a set of multiple choice comprehension questions (i.e., in-class listening test) were included. As for the reflection and goal-setting stage, this was kept as an open reflection for the learners to comment on how they would address second language listening in future activities as well as the procedures they employed while doing these listening activities. The specific procedures followed in this treatment group are described below:

1. At the beginning of the listening activity, participants received a listening worksheet (see Appendix B). This worksheet was the same for the eight intervention sessions in this group. It consisted of four sections that were labeled ‘prediction’, ‘listening’, ‘peer-discussion’, and ‘reflection’. This worksheet served as a guide for the learners to follow the sequence of the different stages of the listening activity.

2. Once all participants had the listening worksheet, they were given the topic of the listening text (i.e., the topic was always written on the board and it was explained to the participants). Then, participants were asked to fill in the space under the label ‘prediction’. They were asked to think about the topic and predict the vocabulary they might hear and the possible

scenario the audio input might portray. Participants did this individually and they were free to write vocabulary, draw, and/or include any ideas associated with the topic of the listening text. This corresponds to the **Pre-listening: planning/prediction stage** in the metacognitive pedagogical cycle described in Figure 2 above.

3. After the pre-listening stage, participants were told that they could fill in the space labeled 'listening' in their worksheets either during or after the listening, and they were exposed to the audio input for the first time. They could write clues, vocabulary, ideas or any information they might have heard which could help them understand the listening input. After the first exposure to the audio input, and after having completed the 'predictions' and 'listening' sections in their worksheets, the participants were asked to pair up with a classmate or to form small groups. Then, in an attempt to reconstruct the listening text, they were asked to compare and discuss the information they had gathered with their partners in the small groups. This exchange of information with their partners gave participants a chance to verify their initial hypotheses, fill in information they might have missed, and modify their understanding as necessary. At that point, they were asked to complete the section labeled 'peer-discussion' in their worksheets. This corresponds to the **First listening: first verification stage** in the metacognitive pedagogical cycle in Figure 2 above.

4. After this exchange of information and reconstruction of the listening text, participants were given a set of five multiple choice questions (i.e., an in-class listening test). This set of questions resembled a short Spanish listening test (see Appendix C). Participants were given some time to read the questions, and they were asked to check with the researcher if they did

not understand any of them (i.e., since the questions were not in the participants' native language, the questions were clarified if participants had trouble understanding them). Although participants were asked to complete these questions individually, they were allowed to use their notes from the listening worksheets. At that point, participants were not asked to complete the questions on the in-class listening test. The purpose of reading the questions and asking for clarification before listening to the audio text for the second time was to guide the participants and to help them focus their attention.

5. After clarifying issues associated with the meaning of the questions in the in-class listening tests, the audio input was played for the second time. At that time, participants were told that they could answer the questions while they were listening to the audio text, or they could answer the questions once the audio was finished. That is, participants were given time to answer the questions after the audio stopped.

6. Finally, after participants finished completing the in-class listening test, they were asked to go back to their listening worksheets and fill in the space provided for 'reflection'. They were asked to reflect on how successful their listening was, or to simply answer the question "*Will you do anything differently the next time you listen to a text? If so, please explain what will you do? If there is nothing you will do differently, please describe what you do when listening."* This corresponds to the **Reflection and goal-setting stage** in the metacognitive pedagogical cycle described in Figure 2 above.

At the end of the listening activity, after participants finished completing the section about reflection, both the in-class listening tests and the listening worksheets were collected. These two documents were returned to the participants the following week (the researcher kept a copy of these documents); the in-class listening tests were graded. In the original version of this metacognitive pedagogical cycle, the audio text is played three times and participants engage in two types of discussions with the aim of reconstructing the main points and most pertinent details of the listening text. In the current study, the listening text is played only twice: once before participants engage in one small group discussion with the aim of reconstructing the main points and most pertinent details of the listening text; and, the second time it is played before participants answer the multiple choice questions in the in-class listening tests. Figure 3 below compares Vandergrift's (2004: 11) metacognitive pedagogical cycle and the instructional approach employed in the Metacognitive Pedagogical Sequence group.

Figure 3: A comparison of the Metacognitive Pedagogical Cycle (Vandergrift, 2004) and the instructional approach employed in the Metacognitive Pedagogical Sequence group.

Metacognitive Pedagogical Cycle	**Metacognitive Pedagogical Sequence group**
Planning/predicting stage 1. Once students know the topic and text type, they predict types of information and possible words they may hear.	**Pre-listening** After providing the topic of the audio text, learners were asked to predict possible words, situations, and/or information that might appear in the audio text.

First verification stage 2. Students verify initial hypotheses, correct as required, and note additional information understood. 3. Students compare what they have written with their peers, modify as required, establish what needs resolution and decide on details that still need special attention.	**First listening: Text Reconstruction** The audio input was played for the first time and participants were asked to write down any information they might find important. Then, they were asked to compare their notes and reconstruct the audio input in small groups.
Second verification stage 4. Students verify points of disagreement, make corrections, and write down additional details understood. 5. Class discussion in which all contribute to reconstruction of the text's main points and most pertinent details, interspersed with reflections on how students arrived at the meaning of certain words or parts of the text.	**Second listening: Task** Learners were handed out an in-class listening test. They were given some time to read the questions in the test; then, the audio text was played for the second time and learners wrote their answers.
Final verification stage 6. Students listen for information that they could not decipher earlier in the class discussion.	
Reflection stage 7. Based on discussion of strategies used to compensate for what was not understood, students write goals for next listening activity.	**Reflection** Based on the listening activity, learners were asked to write a short reflection about the listening experience they had.

3.3.2 Awareness-Raising of the Factors Associated with Successful L2 Listening

In a study aiming at developing and validating an instrument to assess learners' awareness of the L2 listening processes, Vandergrift et al. (2006) created a self-report questionnaire that addresses the perceptions of learners regarding themselves as listeners and the demands of the listening tasks as well as their awareness of the strategies they use. This questionnaire (i.e., the Metacognitive Awareness Listening Questionnaire, or MALQ for short) was primarily developed to track metacognitive development in second language listening. However, it can also be used as a "consciousness-raising tool" (Vandergrift et al., 2006: 453). Therefore, in the current study, this questionnaire was used as an awareness-raising tool for the second instructional treatment.

This questionnaire consists of 21 items that are grouped into five interconnected factors. These factors are: 1) problem-solving, 2) planning and evaluation, 3) (no) mental translation, 4) person knowledge, and 5) direct attention. These factors are associated with successful second language listening and constitute the result of a confirmatory factor analysis (Vandergrift et al., 2006). These 21 items are rated on a six-point Likert scale that goes from strongly disagree to strongly agree. In the current study, however, the items in this questionnaire were transformed into questions and used as discussion prompts to raise participants' awareness of the factors associated with successful listening. For example, item 15 reads "I don't feel nervous when I listen to English", but for the purpose of this study, this item was changed to "do you get nervous when doing listening comprehension in Spanish?".

This questionnaire was the best available option for raising participants' awareness of the listening process since it was developed under a rigorous process and validated with two large samples of participants. That is, Vandergrift et al. (2006) extensively reviewed the literature on metacognition, listening comprehension, and self-regulation. Then, they created an initial list of 88 items. These items were then subjected to expert judgements and reduced to 51 items. Subsequently, these items were piloted, and then an exploratory factor analysis was run with a sample of 996 participants. Finally, a confirmatory factor analysis was run with a different sample of 512 participants. This analysis identified five factors and the final number of items included in the questionnaire came down to 21. This rigorous process in developing and validating this questionnaire makes it a valid and appropriate instrument for raising participants' awareness of the factors associated with successful second language listening.

Therefore, in the Awareness-Raising group, with the aim of raising learners' awareness of these five factors, instruction was delivered in the following manner:

1. At the beginning of every session, the researcher presented one of these factors to the participants (i.e., problem-solving, planning and evaluation, mental translation, person knowledge, or direct attention). Then, the researcher presented the discussion questions associated with this particular factor. The factor and the discussion questions were projected on the board and were meant to provide a brief introduction to draw the participants' attention to the discussion questions and to the particular factor associated with second language listening. This took between two and three minutes and was conducted in the participants' second language (i.e., Spanish). Then, the instructor asked the participants to discuss these questions in small groups. Thus, immediately after this

brief introduction, the participants were given a handout which contained the same discussion questions in the participants' native language (i.e., English). Providing the questions in the participants' native language was considered important to guarantee that the participants fully understood the discussion questions. However, the participants were free to discuss these questions either in Spanish or in English. The handout contained three sections: a section that contained the discussion questions, a blank space for notes if needed, and a section for reflection (see Appendix D).

2. After this short introduction, participants engaged in discussing the particular factor guided by the discussion questions on their handouts. These group discussions lasted approximately 7 to 10 minutes. The purpose of these group discussions was to foster awareness of the factors associated with successful listening.

3. Then, once the group discussion was completed, the participants were given the topic of the listening text. Similar to the Metacognitive Pedagogical Sequence group, the topic was always written on the board and it was explained to the participants.

4. After the topic of the listening text was given and explained to the participants, they were exposed to the audio input for the first time. This was the same audio input as for the Metacognitive Pedagogical Sequence group. Participants could take notes and/or write down clues or any information that might be helpful.

5. After the first listening, the participants were handed out the same set of five multiple choice comprehension questions as the Metacognitive Pedagogical Sequence group (i.e., an in-class listening test). Participants were then given some time to read those questions. Similar to the Metacognitive Pedagogical Sequence group, participants were encouraged to check and clarify any issue associated with the meaning of those comprehension questions (this was considered to be important since the questions in the in-class listening tests were not in the participants' native language). Also, participants were asked to complete the questions individually, and they were told that they could use their notes to answer them. Similarly, the purpose of reading the questions and asking for clarification before listening to the audio text for the second time was to guide the participants.

6. After clarifying any issue associated with the meaning of the comprehension questions, the audio input was played for the second time. Same as the Metacognitive Pedagogical Sequence group, participants were told that they could answer the questions while they were listening to the audio, or they could answer the questions once the audio stopped.

7. Finally, after participants finished completing the questions in the in-class listening test, they were asked to go back to their handouts and complete the section for reflection. As in the Metacognitive Pedagogical Sequence group, participants in this treatment group were asked to reflect on how successful their listening was, or to simply answer the question "*Will you do anything differently the next time you listen to a text? If so, please explain what will you do? If there is nothing you will do differently, please describe what you do when listening."*

Similar to the Metacognitive Pedagogical Sequence group, after participants finished completing the 'reflection' section, both the in-class listening tests and the handouts were collected. These two documents were returned to the participants the following week (the researcher kept a copy of these documents); the in-class listening tests were graded and returned. The participants in this treatment group were exposed to the five factors that make up the metacognitive awareness listening questionnaire (MALQ). Thus, every week, one of these factors along with the discussion questions was presented to the participants. That is, in the first week, the factor 'Direct Attention' was presented; in the second week, the factor 'Mental Translation'; in the third week, the factor 'Planning and Evaluation'; in the fourth week, the factor 'Person Knowledge'; and finally, in the fifth and sixth weeks, the factor 'Problem Solving' was presented. This last factor was presented in two sessions because it had more items than the other factors (i.e., 6 items in total). As for weeks seven and eight, the five factors were summarized and presented altogether in those last two sessions.

Participants in the Awareness-Raising group were given opportunities to discuss listening strategies before being exposed to the audio input or the topic of the listening text. During this first part of the lesson, participants were prompted to get ready for the listening activity as they discussed in small groups aspects related to L2 listening (i.e., the factors included in the Metacognitive Awareness Listening Questionnaire). Then, participants were exposed to the audio input, and subsequently were asked to reflect about their L2 listening experience. In this regard, participants in the Awareness-Raising group engaged in discussing L2 listening strategies in small groups, unlike participants in the Metacognitive Pedagogical Sequence group who engaged in small group discussions to reconstruct the audio text. Also, participants in the

Awareness-Raising group, similar to participants in the Metacognitive Pedagogical Sequence group, were asked to write a reflection at the end of the listening activity (see Table 2 in section 3.3.4 below).

Although instruction in these two groups (Metacognitive Pedagogical Sequence and Awareness-Raising) was not delivered in the same way, they both constitute metacognitive instruction in L2 listening because these two approaches foster metacognitive knowledge and regulatory skills such as planning, monitoring and evaluation. That is, while participants in the Metacognitive Pedagogical Sequence group were guided through stages such as prediction, verification, reconstruction, and problem solving, participants in the Awareness-Raising group discussed aspects such as direct attention, mental translation, planning and evaluation, person knowledge, and problem solving (i.e., the factors included in the MALQ). Thus, participants in both groups were given opportunities to plan, monitor, evaluate, and reflect at different times during the lesson.

While the instructional approach implemented in the Metacognitive Pedagogical Sequence group involved a more implicit approach to teaching L2 listening, the instructional approach implemented in the Awareness-Raising group was carried out in a more explicit manner. Participants in the Awareness-Raising group were asked to explicitly discuss strategies associated with the MALQ factors at the beginning of the listening activity. Then, at the end of the listening activity, they were asked to reflect about their listening experience. Listening instruction using discussions and reflections, as it was carried out in the Awareness-Raising group, has not been subject to an intervention study. While combining guided discussions and reflections has been shown to impact L2 listening positively (Goh & Taib, 2006), the implementation of discussions and reflections has not been carried out systematically in the

context of an intervention study. In this regard, in addition to systematically implementing guided discussions and reflections about L2 listening to develop listening skills, the present study also compares implicit metacognitive instruction in L2 listening (Metacognitive Pedagogical Sequence) with a more explicit metacognitive approach in L2 listening (Awareness-Raising of the Factors Associated with Successful L2 Listening).

3.3.3 Conventional Approach

The instructional approach employed in the Conventional Approach group is what Field (2008) referred to as the comprehension approach. This approach provides learners with practice and exposure to different voices, intonations, and interactions. However, it has a heavy focus on the product of listening; that is, it focuses mainly on the outcome of the listening in the form of correct or incorrect answers to comprehension questions. Such an approach that focuses primarily on comprehension seems to be common practice in language teaching (Goh, 2008; Graham, Santos, & Francis-Brophy, 2014). Field (2012: 208) described this approach in the following manner and noted that a listening lesson conducted by an experienced teacher would typically follow this format:

> **Prelistening.** A brief (5 to 10 minute) introductory phase. Goals: to set the scene; to motivate the learners to listen; to turn learners' thinking toward the topic of the recording to be heard. It may sometimes also be necessary to preteach up to four or five critical words of vocabulary without which the recording cannot be understood.
> **Extensive listening**. First playing of the recording, followed by general questions. (*How many speakers? What are they talking about?*) Goals: to enable learners to **normalize** (adjust) to the voices of the speakers and to orient themselves in terms of where in the recording different types of information are mentioned.

Preset questions or task. Introduced ahead of the main listening phase to ensure that learners will listen in a focused way and to check that the questions / task have been fully understood.
Intensive listening. A second playing of the recording, this time to enable learners to obtain answers. The accuracy of the responses is then checked, with the teacher replaying relevant passages where comprehension levels appear to be low.
Language of the recording. One follow-up activity is for teachers to replay sentences containing unknown lexical items, asking learners to infer their meaning from context. A second is to play extracts in order to draw attention to the functional language they contain (way of threatening, offering, refusing, inviting, etc.).
Final play. Done with learners following a tapescript, it enables learners to deconstruct any sections of the recording that they have found difficult to match to words. It also provides a long-term reminder of what was heard in the lesson. (p. 208).

Thus, based on this description of the comprehension approach, second language listening in the Conventional Approach group was delivered in the following manner:

1. At the beginning of every session, the topic of the listening text and some vocabulary were presented. The topic and the vocabulary words were always written on the board, and explained to the participants. This constituted the **pre-listening phase** and was meant to activate participants' schemata as well as to provide the participants with vocabulary. The vocabulary consisted of key words that could potentially hinder comprehension and were selected based on the judgement of the researcher.

2. After introducing the topic and the vocabulary, the audio input was played for the first time. The purpose of this first listening was to enable learners to become familiar with the voices, content and different types of information in the audio input. This corresponds to the **extensive listening phase** above. The participants were allowed to take notes and/or write down clues or any information they considered important or helpful. Once the audio stopped, the researcher asked some general follow-up questions such as 'how many

speakers were interacting?', 'what were they talking about?', and so forth. These questions were directed to the whole class and any participant could answer. The purpose of asking these general questions was to guide the learners regarding where the different information in the audio text was mentioned.

3. After this short interaction between the researcher and the learners, the participants were handed out the same in-class listening test as the other two groups. Participants then were given some time to read the questions and ask for clarification if they had any problem understanding the questions. Same as the participants in the Metacognitive Pedagogical Sequence group and the Awareness-Raising group, participants in this treatment group could use their notes to answer these questions. They were asked to complete these questions individually. Similarly, the purpose of reading the questions and asking for clarification before listening to the audio text for the second time was to direct the participants' attention and to provide guidance regarding what to do. This constituted the **preset questions or task** above.

4. After clarifying any issue associated with the meaning of the questions in the in-class listening tests, the audio text was played for the second time. This constituted the **intensive listening phase** above. This time, the purpose of listening to the audio input was to obtain answers and complete the in-class listening test. Similar to participants in the other two group, participants in this treatment group were told that they could answer the questions in the in-class listening tests as the audio was running or once the audio stopped.

5. Finally, after participants finished writing their answers, the in-class listening tests were collected. As in the other two groups, these tests were graded and returned to the participants the following week (the researcher kept a copy of these tests).

This procedure was the same for all eight sessions in this treatment group. The listening lesson described by Field (2012) has two more phases: **language of the recording** and **final play**. In the current study, these two last phases were not included because it implied playing the audio input more than twice.

3.3.4 Final Remarks regarding the Intervention

Regarding the three different approaches employed in this study, there were some aspects of the entire intervention that were kept constant in the three treatment groups. These aspects were as follows: playing the recording twice, giving participants the same in-class listening tests as well as the same audios, providing participants with the topic of the audio text and delivering the listening lesson in approximately the same amount of time. Although controlling for outside classroom exposure to second language was not possible, participants' report on such exposure was on average very similar across the groups. Table 2 below shows the sequence of the listening lessons implemented in the three groups.

Table 2: Comparison of the Listening Sessions in the Groups.

Metacognitive Pedagogical Sequence	Awareness-Raising	Conventional Approach
	The researcher briefly presented L2 Listening Metacognitive Factors Learners engaged in small group discussions about these factors	
Topic was introduced Learners engaged in predicting based on the topic.	**Topic was introduced**	**Topic was introduced** Vocabulary was explained.
First audio input Learners engaged in small group discussions to compare information and reconstruct the audio text.	**First audio input**	**First audio input** The researcher asked the whole class some guiding questions intended to build a better understanding of the audio text.
In-class listening test was handed out	**In-class listening test was handed out**	**In-class listening test was handed out**
Second audio input	**Second audio input**	**Second audio input**
Learners answered the questions on the test	**Learners answered the questions on the test**	**Learners answered the questions on the test**
Reflection **End of listening activity**	Reflection **End of listening activity**	**End of listening activity**

3.4 Data Collection

Data for this study were collected before, during, and after the intervention. Two weeks before the start of any intervention, participants were informed of the study and consent was sought. Then, one week before the start of the intervention sessions, a pre-test was administered to all of the participants in the three groups. These pre-intervention data were collected from the three groups on the same day, and constituted the base-line for the comparison between the groups.

One week after the pre-test, each of the three groups received treatment in one of the approaches mentioned above (see section 3.3 above), and data were collected using an in-class listening test. The reflections written by the participants in the Metacognitive Pedagogical Sequence group and Awareness-Raising group were also collected. This pattern of intervention was used over the following seven weeks for a total of eight intervention sessions. Additionally, in week seven participants in all groups were asked to fill out an anonymous open-ended questionnaire that asked for their opinion about the listening activities carried out until that moment. 44 participants filled out this questionnaire (i.e., 18 students from the Metacognitive Pedagogical Sequence group; 19 students from the Awareness-Raising group; and, 7 students from the Conventional approach group). This questionnaire was administered at this time because participants in the three groups had just taken the course mid-term evaluation. This was an appropriate time to elicit participants' perceptions of the implemented listening activities regarding their performance during a testing situation.

Two weeks after the end of the interventions (i.e., week 10), a post-test was administered to all of the participants in the three groups. These data were collected on the same day, and

constituted post-intervention data. At this time, participants were also asked to report the amount and type of exposure to Spanish outside of class time. After participants completed this post-test, they were also asked to complete a second anonymous questionnaire. This was an open-ended questionnaire which collected data regarding the teaching approach. 40 participants filled out this questionnaire (i.e., 17 students from the Metacognitive Pedagogical Sequence group; 14 students from the Awareness-Raising group; and, 9 students from the Conventional Approach group).

Finally, focus group interviews were conducted with a subset of participants approximately two months after the last intervention session. These focus group interviews could not be scheduled earlier due to participants' unavailability. Seven participants participated in the focus groups (i.e., two from the Metacognitive Pedagogical Sequence group; three from the Awareness-Raising group; two from the Conventional Approach group). These participants volunteered to participate in the focus groups. Also, five months after the last intervention session, data were collected using another post-test. These constituted delayed post-intervention data. Data were elicited from all original participants who were still enrolled in the Spanish course. Delayed post-test data were collected from 17 participants (i.e., nine participants from the Metacognitive Pedagogical Sequence group, two participants from the Awareness-Raising group, and six participants from the Conventional Approach group).

3.4.1 Pre-test, Post-test and Delayed Post-test

The instrument used to collect data before and after the intervention was the listening section of a practice DELE test designed to measure proficiency at the B1 level. The DELE tests are standardized exams that measure language proficiency in Spanish and that reflect and

integrate the guidelines of the Common European Framework of Reference for Languages (CEFR). In that sense, these tests are used to certify competence and mastery of the Spanish language at the A1, A2, B1, B2, C1 and C2 levels of the CEFR.

The listening section of this practice DELE test contained 30 questions in total and was divided in five subsections with six questions in each subsection. The first three subsections contained multiple choice questions with three options each, and the last two subsections of the test required participants to match different statements according to what they heard in the audio. The audio input in the exam included monologues and dialogues which very much resembled real life situations. The audios were already programed to play twice during every subsection.

In this study, two listening test versions were used: Test Version A and Test Version B. Test Version A was employed to measure listening ability at pre-test and post-test, and Test Version B was employed to measure listening ability at delayed post-test. A different test version was administered at delayed post-test because the pre-test and post-test (Test Version A) were both returned to participants at the end of the intervention. Thus, administering the same test version at delayed post-test could have introduced testing biases because participants already had the tests and they had taken Test Version A twice already, at pre-test and post-test.

These two test versions were selected randomly from a set of five practice tests. These two listening practice-test versions were developed for the same proficiency level and contained the same number of questions, the same format, the same length, and the same number of tasks. However, the audio input and the comprehension questions were different. There were thirty questions in these two listening practice tests, and every correct answer was given one point for a maximum score of thirty points. There was no penalty for incorrect answers, and the final scores

were computed by adding all the correct answers. The comparability of these two test versions is discussed in more detail in section 3.5.3 below.

The procedure for the pre-test, post-test and delayed post-test involved informing the participants of the format of the tests. This included information regarding number of questions, sections, audio input, and how to answer the questions. Later, after distributing the paper copies of the tests, the volume of the audio input was adjusted. Immediately after that, participants were asked to listen to the audio input and write their answers on the tests. Participants were told that the audio input would be played twice in every section. Also, participants were told that there would be a twenty-second break separating every section in the test. Once all participants finished the pre-, post- and delayed post-test, the tests were collected. The tests were subsequently returned to the participants at the end of the study (a copy of each test was kept for the analysis) as stated in the ethics approval of the study.

3.4.2 In-class Listening Tests

Also, eight short in-class listening tests were used as instruments for data collection. These listening tests were developed by the researcher and were based on the audio files that come with the course textbook. These in-class listening tests were the same length and format; that is, every test contained five multiple choice questions. The audios for these listening tests were played twice to the participants in the three groups. These audios and listening tests were part of the intervention (see Appendix C).

The procedures were the same for the three groups. That is, right before participants were exposed to the audio input for the second time, they were handed out the in-class listening test.

Then, they were given approximately two minutes to read the questions and clarify any issue associated with these questions. Participants were encouraged to ask for clarification if there was something they did not understand in the in-class listening tests. For example, if there was a word that they did not know or if there was a question that was not clear enough, participants were given the necessary clarification. After reading and clarifying concerns related to these in-class listening tests, the audio input was played for the second time. Participants had the choice of answering the multiple choice questions in the in-class listening test while the audio was playing or after the audio finished. Once participants finished answering the questions after this second listening input, the in-class listening tests were collected.

3.4.3 Questionnaires and Focus Groups

Two questionnaires were used as instruments to collect data. These questionnaires contained open-ended questions about the instructional approach and the listening activities implemented in the different treatment groups. The first questionnaire was completed two days after participants wrote their course mid-term exam which had not been given back to them at the time. This questionnaire had only one open-ended question: 1) *Thinking about the listening section in the midterm exam, do you think the listening practice you've been receiving helped you? Please, briefly explain how it's helped you or why it has not helped you* (see Appendix E). The second questionnaire was administered two weeks after the last intervention session and had two open-ended questions: 1) *What do you think of the teaching method used during the listening activities?* and, 2) *If you could modify or change the listening activities you did in class, how and*

what would you modify/change to make them more useful and/or more effective? (see Appendix F).

The procedure was the same for the three groups and consisted of handing out the sheets of paper and asking the participants to complete the questions. Before handing out the questionnaires, participants were told that the questionnaires were anonymous and their answers were completely confidential. Anonymity in the questionnaires was important to ensure that participants commented more openly and honestly on how they perceived the listening activities. Additionally, since this type of data did not need to be matched with the pre/post-tests or the in-class listening tests, personal information was not necessary. After they completed their answers, the questionnaires were collected.

For the focus groups, a meeting with two or three participants from each treatment group was scheduled to talk about the listening activities implemented in their classes. These participants volunteered to participate in the focus group interviews. Thus, after coordinating a convenient time and date to meet, the participants and the researcher met and talked about these listening activities for about 20 minutes. At the beginning of the focus group interview, the participants were reminded that the information they provided needed to be recorded. They were assured that the information they provided was completely confidential and personal information was not being collected. After that, participants were asked to think about the listening activities they experienced during the interventions. They were asked to think in retrospection for about one minute. Then, the audio recorder was set to record and participants were prompted to talk about their experience during these listening activities.

While several questions were employed to guide the focus group interview, the researcher tried not to be intrusive so that the participants could talk freely. These questions

were: a) *How was your experience during the listening activities?* b) *Do you feel you learned something?* c) *How do you go about listening comprehension now?* d) *Is there any specific technique or strategy you employ now?* and e) *Before the start of these listening activities and after the end of these listening activities, do you feel something has changed in your approach to listening to Spanish?* Once the participants felt they had nothing else to say about the listening activities and their experience, the recorder was stopped.

Additionally, information regarding native language, exposure to Spanish outside of class time, and whether participants like Spanish or not was collected at the time of the post-test (see Appendix G). For consistency purposes, data were collected from the three groups on the same day under the same procedures. That is, the pre-test, post-test, questionnaires, and in-class listening tests were administered to the three groups on the same day.

3.5 Data Analysis

3.5.1 Pre-test and Post-test

The data collected with the pre-test and the post-test were used to run the first analysis. This first analysis addressed research question one (RQ-1) which was concerned with whether or not there are significant changes in listening performance within each group (RQ-1a), and with whether or not there are significant differences in L2 listening performance between the treatment groups (RQ-1b). Thus, to answer this question, a repeated measures ANOVA was run. This analysis combined a within-subjects factor and a between-subjects factor, and it allowed to observe changes from pre-test to post-test as well as to evaluate differences between the groups

at pre-test and at post-test. That is, this analysis permits to compare scores over time and across groups (Gamst, Meyers, & Guarino, 2008). This is the most appropriate statistical procedure because it compares the mean scores of the groups simultaneously safeguarding from inflated alpha levels caused by running multiple statistical tests, which as discussed in chapter 2, is a recurrent limitation in previous studies.

3.5.2 Delayed Post-Test

To evaluate whether or not there are long-term effects, a second repeated measures ANOVA was run. This analysis addressed research question two (RQ-2) which was concerned with whether or not there are long-term effects associated with the intervention. This second repeated measures ANOVA analysis was run separately from the repeated measures ANOVA described in section 3.5.1 above because the number of participants who were able to take the delayed post-test decreased causing the loss of one of the treatment groups.

As discussed in section 3.2 above, of the 32 participants in the study, only 17 were able to take the delayed post-test. This changed the configuration of participants in each of the groups. In the Metacognitive Pedagogical Sequence group, only nine out of the initial 13 participants took the delayed post-test; in the Awareness-Raising group, only two out of the initial 11 participants took the delayed post-test; and, in the Conventional Approach group, only six out of the initial eight participants took the delayed post-test. In relation to the initial number of participants in each group, this means 30.77% decrease in Metacognitive Pedagogical Sequence group, 81.81% decrease in Awareness-Raising group, and 25% decrease in the Conventional Approach group. This new configuration of participants in the groups made it difficult to

evaluate long-term effects in the three groups. However, considering that the number of participants in the Metacognitive Pedagogical Sequence group and the Conventional Approach group did not change drastically, the analysis was run with these two treatment groups (i.e., nine and six participants respectively). This number of participants may seem small to run this analysis. However, this is a design in which two variables are combined factorially and participants are measured repeatedly. This optimizes the use of participants, and as Aron, Aron, & Coups (2005) noted: "With a factorial design you can study both variables at once, without needing twice as many participants" (p. 313). The group with six participants might be a concern since seven participants per a cell (i.e., group) is the minimum recommended number of participants (Wilson VanVoorhis & Morgan, 2007).

Thus, to analyze this new data set, a second repeated measures ANOVA was run. This analysis examined the scores on the pre-test, post-test, and delayed post-test across the two groups to assess whether or not there are significant differences within time and between the two groups. The instruments used to measure second language listening at pre-test, post-test, and delayed post-test were not the same. As mentioned earlier (section 3.4.1), while the measure for the pre-test and post-test was the same listening test (Test Version A), the measure for the delayed post-test was a different listening test (Test Version B). The analysis of the comparability of these two test versions is discussed below in section 3.5.3.

3.5.3 Parallel Test Forms

To evaluate the comparability of Test Version A and Test Version B, both test versions were given to a different cohort of intermediate Spanish learners. The tests were administered on

two different days over the same week and the procedures were the same as those for the pre-test and post-tests. Then, the scores were calculated and a correlation between the scores of Test Version A and Test Version B was computed. This was followed by an item difficulty analysis. In this regard, Thorndike and Thorndike-Christ (2010) mentioned that parallel tests should contain the same level of difficulty across the tests, and that a correlation between the forms will provide an appropriate index of reliability (pp. 124-125). Reliability refers to consistency or stability of scores, which in this particular case refers to parallel form reliability. This type of reliability refers to "the consistency of scores obtained on two equivalent forms of a test or research instrument designed to measure the same thing. This type of reliability is measured by correlating the scores obtained by giving the two forms of the same test to a single group of people" (Christensen, Johnson, & Turner, 2011: 144).

Parallel form reliability is used to gauge the equivalence of two different versions of a test that measures the same construct. The rationale behind it is that different versions of the same test should yield very similar scores for the same individuals. In that sense, parallel test versions should be highly correlated (i.e., $r > .70$). As for item difficulty, Thorndike & Thorndike-Christ (2010) defined it as "the proportion of examinees who answered the item correctly" (p. 306). That is, the number of examinees who answered the item correctly divided by the total number of examinees. In this regard, an item that is difficult will be characterized by its low proportion (i.e., 0.3 or lower), and an item that is easy will be characterized by its high proportion (i.e., 0.8 or higher).

The parallel form reliability analysis was run with 13 participants. These participants took both Test Version A and Test Version B. Therefore, to determine the comparability of Test Version A and Test Version B, a correlation between these two test versions was run with the

scores of these 13 participants. Then, to determine the level of difficulty of these two test versions, an analysis of the test items was conducted. This item analysis was conducted with 18 participants for Test Version A, and with 16 participants for Test Version B. The number of participants in these two analyses was different because the two test versions were administered on two different days to the same group of students. That is, when Test Version A was administered, there were 18 students in class. When Test Version B was administered, there were 16 students in class. Thus, of the 18 students who took Test Version A and the 16 students who took Test Version B, only 13 students completed both test versions.

3.5.4 In-class Listening Tests

Data collected with the in-class listening tests were also examined. As mentioned earlier (section 3.1), examining these scores was an important aspect in this study because they allowed to track changes in L2 listening during the intervention. The analysis of the scores of these in-class listening tests was kept to descriptive statistics because the scores on these tests were influenced by the different treatments. For example, while participants in the Metacognitive Pedagogical Sequence group engaged in small group discussions to compare information about the audio text before receiving the in-class listening tests, participants in the Awareness-Raising group and the Conventional Approach group were not allowed to share and compare information about the audio text. Also, unlike participants in the Metacognitive Pedagogical Sequence group and the Awareness-Raising group, participants in Conventional Approach group were given and explained key vocabulary of the audio text before receiving the in-class listening tests. Thus, in a sense, participants in the Metacognitive Pedagogical Sequence group and the Conventional

Approach group would be expected to score higher on these in-class listening tests than participants in the Awareness-Raising group because participants in the Awareness-Raising group were not explained key vocabulary or allowed to compare information with their classmates. The influence of the treatments on the scores of these in-class listening tests make it difficult to meet all the ANOVA assumptions to run an inferential analysis. For this reason, this analysis was kept to descriptive statistics only (see section 4.1.1). Thus, to analyze the scores of the in-class listening tests, group mean scores were calculated and a plot was generated. Then, the results were described in terms of the mean scores and percentages.

3.5.5 Questionnaires and Focus Groups

The data collected in the focus groups was transcribed and coded. In this process, segments were identified and labeled using the third research question (RQ-3) as a guide. This research question was concerned with the perceptions of the participants regarding the L2 listening activities implemented in the groups. The data from the questionnaires were analyzed similarly, and segmenting and labeling were guided by research question three.

The process of analyzing these data was interactive in the sense of going back and forth, reading and re-reading, labelling and re-labeling, and grouping and re-grouping (i.e., open coding). In other words, these data were analysed by reading the participants' comments and answers, dividing the text into segments, labeling these segments, reducing overlapping and redundancies, and collapsing codes into themes. In this process, the comments in each group were counted and segmented on the basis of high frequency.

3.6 Summary and Conclusion

Generally, this chapter has laid out the pedagogical procedures involved in the treatment conditions, the instruments and procedures of data collection, the analytical tools employed in the analysis of the data, and all pertinent details associated with the methodological design of this intervention study. Random assignment of participants to groups was not possible. However, the three treatment conditions were randomly assigned to the groups. Also, the number of participants in this study may seem small. However, G*Power calculations yielded a similar sample size. This study displays several methodological improvements in relation to previous studies such as systematic implementation of the treatment conditions in the three groups, the incorporation of standardized measures of listening ability which were not practiced during the intervention, careful data collection procedures which were applied consistently across the three groups, monitoring of the groups during the intervention, and appropriate statistical analysis that safeguarded from inflated alpha levels. The next chapter reports the results of this intervention study in relation to the overarching research question that guided it; that is, whether or not there are significant changes in L2 listening performance among three groups of participants who were exposed to three different treatment conditions.

4 Results

This chapter lays out the results of the present study in relation to the three research questions that guided it. It starts by presenting the results of the statistical analyses concerning the effects of the intervention. This first part is divided into two subsections that address the analysis of the pre/post tests and the analysis of the in-class listening tests employed during the intervention, respectively. The second part of the chapter is concerned with the analysis of the delayed post-tests and the analysis of the two test versions that were employed as post-test (Test Version A) and delayed post-test (Test Version B) (see section 3.5.3). Finally, the third part in this chapter displays the results of the participants' comments regarding the listening activities implemented in their classes. This last section presents the results of the three treatment groups in relation to three categories that were identified during the coding process.

4.1 Pre-Test and Post-Test

This first part of the results is concerned with research question one (RQ-1). This question refers to whether or not there are significant changes in L2 listening performance among the three groups of participants who were exposed to three different types of listening instruction: a Metacognitive Pedagogical Cycle, an Awareness-Raising Approach, and a Conventional Approach (see section 3.3). This first research question was divided into research question 1a (RQ-1a) and research question 1b (RQ-1b). RQ-1a was concerned with whether or not there are significant changes in L2 listening performance within each group from pre-test to

post-test, and RQ-1b was concerned with whether or not there are significant differences in L2 listening performance between the groups at post-test (see section 2.4).

In order to answer these research questions, a repeated measures ANOVA was performed. The repeated measures ANOVA combined one within-subjects variable and one between-subjects variable. The within-subjects variable was *Time* and had two levels: Time 1 at pre-test and Time 2 at post-test. The between-subjects variable was *Teaching Approach* and had three levels: Metacognitive Pedagogical Sequence, Awareness-Raising Approach, and Conventional Approach. The dependent variable was listening performance and was measured before the intervention at pre-test and after the intervention at post-test. This mixed factorial design compares scores over time and across groups (Gamst, Meyers & Guarino, 2008). This is the most appropriate statistical analysis to address these questions because it examines the effects of the two independent variable simultaneously. Thus, it safeguards from inflated alpha levels that may occur due to running several statistical tests. In other words, running several t-tests, as previous studies did (see section 2.3), would not be appropriate because it may lead to inflated alpha levels. This design provides advantages such as more efficient use of the subjects, increased sensibility, and greater power associated with the within-subjects factor (Keppel & Wickens, 2004).

Before performing the analysis, assumptions associated with the repeated measures ANOVA were evaluated. The alpha level was set at 0.05 (i.e., $\alpha = .05$); that is, *p-values* need to be equal or lower than 0.05 to be significant. A preliminary examination of the data set showed one extreme outlier that was subsequently removed from the data set. Outlier was defined as 1.5 x IQR, where IQR is the interquartile range (Quartile 3 – Quartile 1). Thus, only values inside

the inner fences were included in the analysis. Consequently, the analysis was run with 31 cases (see section 3.2).

The assumption of normality of the sampling distribution was confirmed by the Shapiro-Wilk test. This test is used when the sample size is smaller than 50 and determines if the distribution of the dependent variable is normal. It tests the null hypothesis that the dependent variable is normally distributed. Thus, because the *p-values* in the three groups were non-significant (i.e., the values were bigger than 0.05), the null hypothesis was not rejected. As can be observed in Table 3 below, the non-significant values (i.e., 0.587, 0.155, 0.473, 0.793, 0.239, and 0.961) confirmed that the assumption of normality was not violated.

Table 3: Tests of Normality for the Repeated Measures ANOVA: Analysis of Pre-test and Post-test.

		Shapiro-Wilk		
	Teaching Approach	Statistic	Df	Sig.
Pre Test	Metacognitive Pedagogical Sequence	.947	12	.587
	Awareness-Raising	.894	11	.155
	Conventional Approach	.925	8	.473
Post Test	Metacognitive Pedagogical Sequence	.961	12	.793
	Awareness-Raising	.909	11	.239
	Conventional Approach	.980	8	.961

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

The assumption of homogeneity of variance was also confirmed. The insignificant *p-values* of the Levene's test indicated that this assumption was not violated (i.e., pre-test: $F(2, 28) = 2.664$, $p > .05$; post-test: $F(2, 28) = .825$, $p > .05$). Also, the insignificant value of Box' M test

(i.e., $p = .173$) indicated that the equality of covariance matrices assumption was not violated. The assumption of Sphericity was not evaluated because the within-subjects variable had only two levels (i.e., Time 1 at pre-test and Time 2 at post-test).

After evaluating these statistical assumptions associated with the repeated measures ANOVA, the analysis was performed. Table 4 below shows the means, standard deviations, and the number of cases in each group.

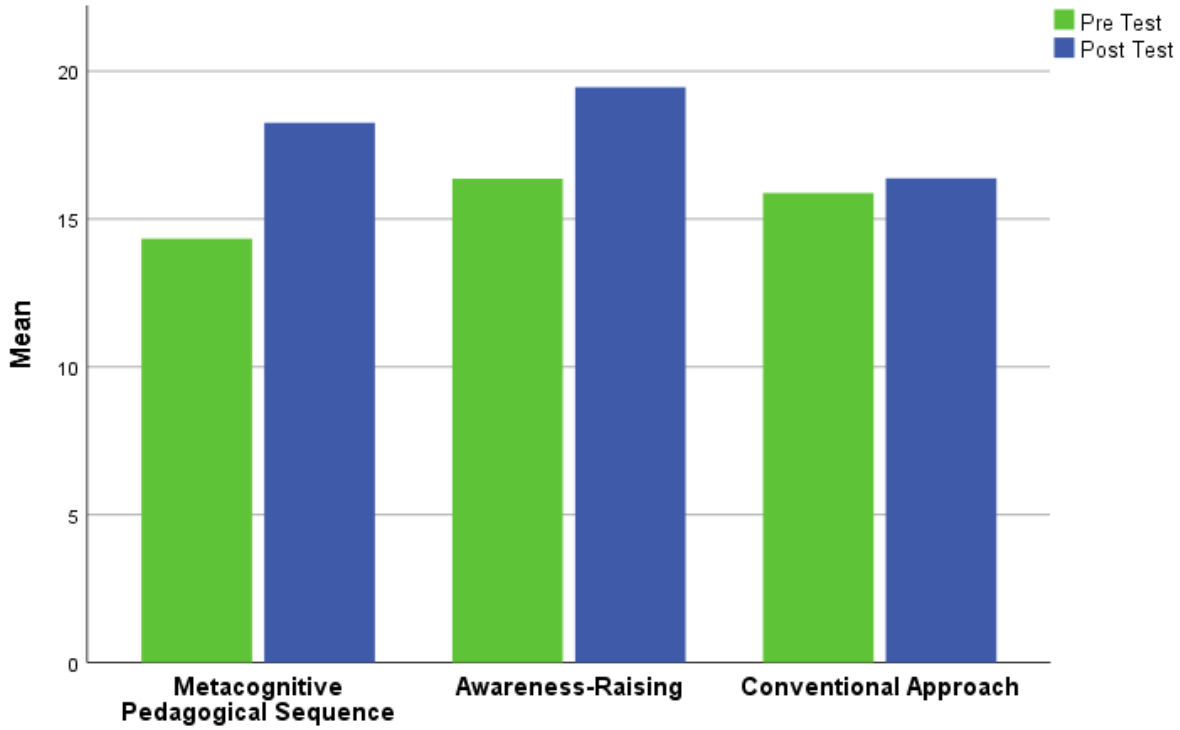
Table 4: Descriptive Statistics for the Repeated Measures ANOVA: Analysis of Pre-test and Post-test.

	Teaching Approach	Mean	Std. Deviation	N
Pre Test	Metacognitive Pedagogical Sequence	14.33	2.535	12
	Awareness-Raising	16.36	4.739	11
	Conventional Approach	15.88	3.523	8
	Total	15.45	3.686	31
Post Test	Metacognitive Pedagogical Sequence	18.25	3.545	12
	Awareness-Raising	19.45	4.886	11
	Conventional Approach	16.38	4.138	8
	Total	18.19	4.246	31

As can be observed in Table 4, the three groups experienced a gain in the mean scores from pre-test to post-test. The Metacognitive Pedagogical Sequence group experienced a gain of 3.92, the Awareness-Raising group exhibited a gain of 3.09, and the Conventional Approach

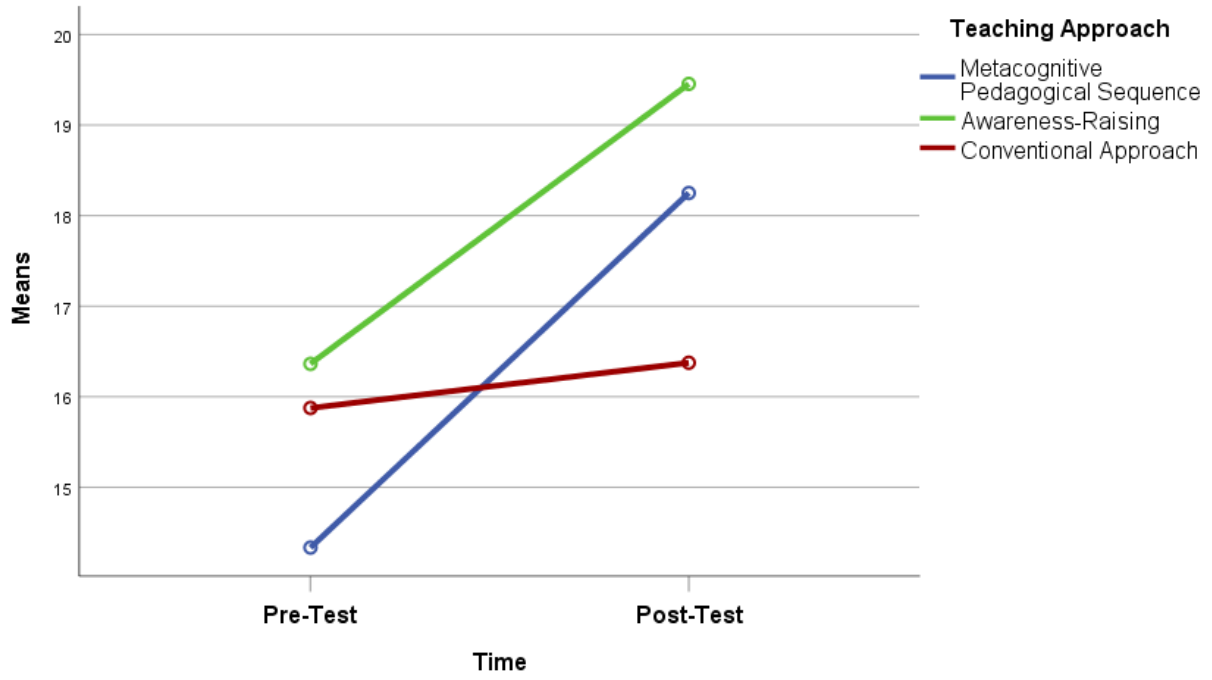
group experienced a gain of 0.50. Figure 4 below compares the mean scores of the pre-test and post-test for the three group.

Figure 4: Comparison of the mean scores at Pre-Test and Post-Test.



The omnibus statistical analysis yielded significant main effects for Time (i.e., $F(1, 28) = 28.079$, $p < .05$) with a Partial Eta Squared of .501 (i.e., the effect size for the factor Time), but not for Teaching Approach (i.e., $F(2, 28) = .723$, $p > .05$). The analysis also yielded significant effects for the interaction between Time and Teaching Approach (i.e., $F(2, 28) = 4.331$, $p < .05$). Partial Eta Squared for this interaction was .236 (i.e., the effect size for this interaction). This interaction can be observed in Figure 5 below.

Figure 5: L2 Listening performance as a function of Time and Teaching Approach across Pre-test and Post-test.



Although the main effect for the factor Time was significant with a large effect size accounting for 50.1% of the variance and associated error, this main effect was not examined further. Instead, the significant interaction effect of the factor Time and the factor Teaching Approach was examined further. In the presence of a significant interaction, the main effects are not often examined further because a significant interaction indicates that the levels of one of the factors varies depending on the levels of the other factor (Gamst, Meyers & Guarino, 2008). A significant interaction effect portrays a more complex pattern. This is examined by looking at the plot of the interaction. The effect size of this interaction indicated that 23.6 % of the variance and associated error can be accounted for by this interaction. This indicated that examining this interaction effect is relevant.

Therefore, to further examine this interaction effect, simple effect analyses were conducted. These follow-up tests indicated that the Metacognitive Pedagogical Sequence group significantly improved from pre-test to post-test: $F(1, 28) = 27.442$, $p < .05$, the size of this effect was large with $r = .70$ (Mean Difference of 3.92, Confidence Interval: 2.385—5.448). Also, the Awareness-Raising group significantly improved from pre-test to post-test: $F(1, 28) = 15.666$, $p < .05$, the size of this effect was large with $r = .59$ (Mean Difference of 3.091, Confidence Interval: 1.491—4.691). The Conventional Approach group did not show significant improvement from pre-test to post-test: $F(1, 28) = .298$, $p > .05$, the effect size was small with $r = .10$ (Mean Difference of 0.5, Confidence Interval: -1.376—2.376). These results answer RQ-1a which was concerned with whether or not there are significant changes in listening performance from pre-test to post-test. These results indicate that the Metacognitive Pedagogical Sequence group and the Awareness-Raising groups significantly improved from pre-test to post-test. Importantly, the effect sizes for the significant results of these two groups were large (i.e., .70 and .59) indicating a strong effect associated with these two instructional approaches.

Regarding RQ-1b, which was concerned with whether or not the groups differed significantly at post-test, the simple effect analyses that followed the significant interaction indicated that the groups are not significantly different from each other, neither at pre-test or at post-test. However, by looking at the interaction effect (i.e., the crossover effect) in Figure 5 above, one can observe that at pre-test, the Conventional Approach group scored higher (15.88) than the Metacognitive Pedagogical Sequence group (14.33); but, at post-test, the Metacognitive Pedagogical Sequence group scored higher (18.25) than the Conventional Approach group (16.38). This crossover effect was not taken into consideration when the groups were compared. The follow-up tests only compared if the groups were significantly different either at pre-test or

post-test without considering that the Metacognitive Pedagogical Sequence group scored lowest at pre-test, but higher at post-test. That is, the analysis compared the mean scores of the three groups at pre-test and at post-test independently from each other without considering the crossover effect. Therefore, in order to answer whether or not the groups are significantly different at post-test (RQ-1b), it is important to take into consideration that the groups did not start at the same level.

To control for these initial differences at pre-test, an analysis of covariance (ANCOVA) was performed. This analysis permits to statistically control for initial pre-existing differences (Tabachnick & Fidell, 2013). This procedure allows to statistically equate the groups, as if the groups had scored identically at pre-test, and compare the mean scores of the groups at post-test. As discussed in Best and Kahn (2006), "differences in the initial status of the groups can be removed statistically so that they can be compared as though their initial status had been equal" (p. 429). This will allow to examine whether or not the groups differed at post-test (i.e., RQ-1b) taking into consideration that the groups started at different levels. In fact, this was the statistical analysis employed by Vandergrift and Tafaghodtari (2010) to determine if the groups in their study significantly differed at post-test.

The ANCOVA procedure is a statistical analysis employed in experimental designs. In quasi-experimental designs where participants are not randomly assigned to the treatment conditions, as it is the case in this study, this procedure is questionable since "the precision of the ANCOVA estimate will be lower when the groups are non-equivalent than when they are randomly formed" (Reichardt, 1979: 157). This certainly poses a limitation of statistical control in non-equivalent group designs. In this study, however, this limitation is justified by the fact that this is a follow-up analysis of the significant interaction effect yielded by the repeated measures

ANOVA. Also, this procedure is justified by the fact that the groups do not exhibit large differences at pre-test (i.e. the mean scores at pre-test are not significantly different) because "a larger pretest difference ($X_E - X_C$) amplifies any "mistake" made in estimating β and correspondingly leads to a less precise estimate" (Reichardt, 1979: 157).

Therefore, to further examine the interaction effects of the repeated measures ANOVA, a one-factor analysis of covariance (ANCOVA) was performed. The independent variable was teaching approach and consisted of three levels: 1) Metacognitive Pedagogical Sequence, 2) Awareness-Raising of the Factors Associated with Successful L2 Listening, and 3) Conventional Approach. The scores on the pre-test were used as a covariate, and the dependent variable was listening performance at post-test.

To meaningfully interpret the univariate *F* test (ANCOVA) for the different groups**,** assumptions associated with the analysis of covariance were evaluated first. The alpha level for the statistical analysis was set at 0.05 (i.e., α = .05); that is, *p-values* need to be equal or lower than 0.05 to be significant. The non-significant values (0.793, 0.239, and 0.961) obtained in the Shapiro-Wilk Test confirmed that the dependent variable was normally distributed in the groups, as can be observed in Table 5 below.

Table 5: Tests of Normality for Analysis of Covariance (ANCOVA): Analysis of Pre-test and Post-test.

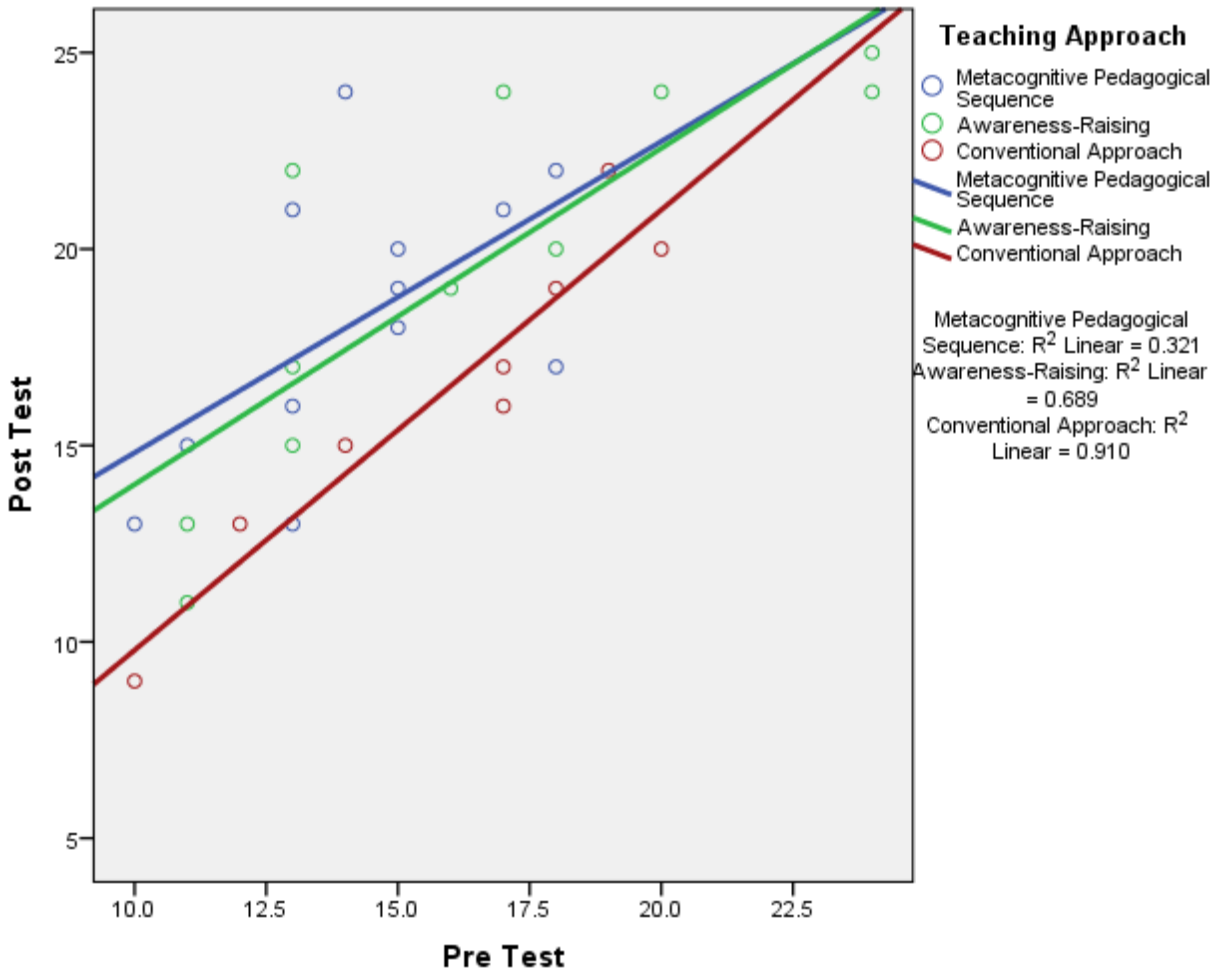
		Shapiro-Wilk		
	Teaching Approach	Statistic	Df	Sig.
Post Test	Metacognitive Pedagogical Sequence	.961	12	.793
	Awareness-Raising	.909	11	.239
	Conventional Approach	.980	8	.961

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

The assumption of homogeneity of variance was also confirmed by Levene's test (i.e., dependent variable: $F(2, 28) = .825$, $p > .05$; and, covariate: $F(2, 28) = 2.664$, $p > .05$). The non-significant *p-values* indicated that the assumption of homogeneity of variance was not violated. The assumption of linearity of regression between the dependent variable and the covariate was also met (r-squared = 0.56); that is, there is a moderate linear relationship between the dependent variable and the covariate. As for the assumption of homogeneity of regression, the graph below was created to examine if the slopes across the groups were parallel. This assumption maintains that the regression lines predicting the dependent variable from the covariate are comparable.

Figure 6: Regression Lines of the dependent variable (post-test) and the covariate (pre-test): Analysis of Covariance (ANCOVA).



The lines in Figure 6 do not exactly appear to be parallel, so the assumption of homogeneity of regression might not be met. Therefore, to evaluate and determine whether this assumption was met, a univariate analysis of variance was performed. This analysis yielded a non-significant effect (i.e., $F(2, 25) = .383$, $p > .05$) for the interaction between the independent variable and the covariate. This insignificant effect for the interaction indicated that the

regression lines do not significantly deviate from being parallel, which means that the assumption of homogeneity of regression was not violated.

Once the ANCOVA assumptions were evaluated, the one-way analysis of covariance was performed. Table 6 below shows the unadjusted mean scores of the dependent variable (post-test). As can be observed from this table, it is the Awareness-Raising group that exhibits the highest mean score (M = 19.45) followed by the Metacognitive Pedagogical Sequence group (M = 18.25), and subsequently followed by the Conventional Approach group (M = 16.38).

Table 6: Unadjusted means and standard deviations for L2 listening as a function of treatment: Analysis of Covariance (ANCOVA).

		Unadjusted	
Teaching Approach	N	M	SD
Metacognitive Pedagogical Sequence	12	18.25	3.545
Awareness-Raising	11	19.45	4.886
Conventional Approach	8	16.38	4.138
Total	31	18.19	4.246

However, since the groups started at different levels of listening ability, the scores were adjusted based on the information of the covariate (i.e., pre-test). Thus, Table 7 below shows the adjusted mean scores of the dependent variable. As can be observed in Table 7, it is now the Metacognitive Pedagogical Sequence group (M = 19.261) which actually displays the highest

mean scores. This is followed by the Awareness-Raising group (M = 18.63), and subsequently followed by the Conventional Approach group (M = 15.992).

Table 7: Adjusted means for L2 listening as a function of treatment: Analysis of Covariance (ANCOVA).

		Adjusted			
				95% Confidence Interval	
Teaching Approach	N	M	SE	Lower Bound	Upper Bound
Metacognitive Pedagogical Sequence	12	19.261a	.769	17.684	20.839
Awareness-Raising	11	18.630a	.797	16.994	20.265
Conventional Approach	8	15.992a	.926	14.093	17.891

a. Covariates appearing in the model are evaluated at the following values: Pre-Test = 15.45.

After statistically equating the groups and adjusting the score on the post-test (Dependent Variable) based on the information of the pre-test (Covariate), these scores were evaluated on the basis of significance; that is, whether or not the group mean scores are sufficiently large to be considered unlikely to occur by chance. The ANCOVA analysis yielded significant effects (i.e., independent variable, $F(2, 27) = 3.94$, $p < .05$) with a Partial Eta Squared of .226. This indicated that at least one of the group mean scores was significantly different from the others. The size of this significant effect (Partial Eta Squared = .226) indicated that it was important and relevant to conduct multiple comparisons.

Thus, to determine which group mean scores were significantly different, Bonferroni corrected multiple comparison tests were conducted. These comparisons indicated that only the Metacognitive Pedagogical Sequence group resulted in statistically significant higher listening performance in relation to the Conventional Approach group (Mean difference = 3.269, $p < .05$), the effect size for this comparison was medium with $r = .46$. The Awareness-Raising group was not found significantly different from the other two groups. However, the mean difference of the Awareness-Raising group in relation to the Conventional Approach group (Mean difference = 2.638, $p > .05$) was nearly as large as the mean difference for the Metacognitive Pedagogical Sequence group and the Conventional Approach group (3.269, $p < .05$). Interestingly, the effect size for this comparison (Awareness-Raising group and Conventional Approach group) was substantial with $r = .39$. This answers RQ-1b regarding whether or not the groups differed significantly at post-test. Thus, after statistically controlling for initial differences in listening performance, only the Metacognitive Pedagogical Sequence group was found to be significantly different from the Conventional Approach group with a medium effect size. In other words, the Metacognitive Pedagogical Sequence group outperformed the Conventional Approach group. Importantly, although the Awareness-Raising group was not found to be significantly different from the Conventional Approach group, the magnitude of the mean difference and the effect size of this comparison indicated that the Awareness-Raising group had a substantially greater impact than the Conventional Approach group.

4.1.1 In-Class Listening Tests

The scores of the in-class listening tests were also examined. The lack of random assignment of participants to the treatment groups introduces selection biases (Christensen et al., 2011), and these scores provide the "extra vigilance [needed] when using dyads, small groups, or intact groups" (Gamst, Meyers & Guarino, 2008: 52). These scores allowed to monitor how the groups responded to the intervention sessions during the study and contributed to a better interpretation of the results. Thus, consistency in the results of the in-class listening tests and the pre/post tests will strengthen the interpretation of the results.

The analysis of these scores was kept to descriptive statistics only. The scores of the in-class listening tests were influenced by the treatment conditions (see section 3.3), and the error of independence assumption was violated since participants in the groups were allowed to communicate with each other during the interventions. Thus, to analyze the scores of these tests, group mean scores for each test were calculated (see Table 8 below) and a graph of these group mean scores was generated (see Figure 7 below).

Table 8: Mean Scores and Standard Deviations of the In-class Listening Tests.

In-class Listening Tests	Teaching Approach	Mean	Std. Deviation	N
Test 1	Metacognitive Pedagogical Sequence	4.23	1.166	13
	Awareness Raising	4.36	.924	11
	Conventional Approach	4.25	.886	8
	Total	4.28	.991	32

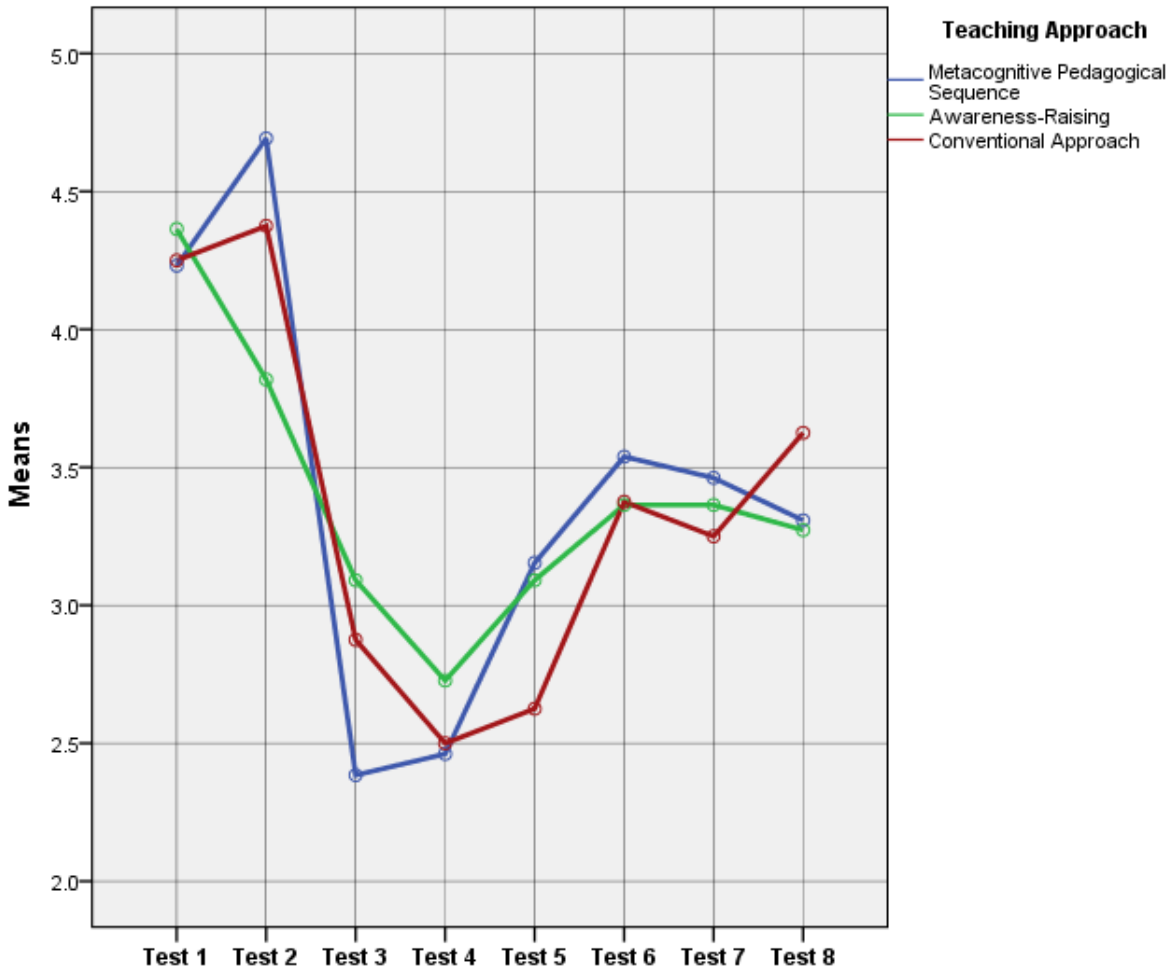
Test 2	Metacognitive Pedagogical Sequence	4.69	.630	13
	Awareness Raising	3.82	1.168	11
	Conventional Approach	4.38	.744	8
	Total	4.31	.931	32
Test 3	Metacognitive Pedagogical Sequence	2.38	1.044	13
	Awareness Raising	3.09	1.044	11
	Conventional Approach	2.88	.835	8
	Total	2.75	1.016	32
Test 4	Metacognitive Pedagogical Sequence	2.46	1.127	13
	Awareness Raising	2.73	.786	11
	Conventional Approach	2.50	1.195	8
	Total	2.56	1.014	32
Test 5	Metacognitive Pedagogical Sequence	3.15	1.068	13
	Awareness Raising	3.09	1.136	11
	Conventional Approach	2.63	1.188	8
	Total	3.00	1.107	32
Test 6	Metacognitive Pedagogical Sequence	3.54	.877	13
	Awareness Raising	3.36	.924	11
	Conventional Approach	3.38	1.302	8
	Total	3.44	.982	32
Test 7	Metacognitive Pedagogical Sequence	3.46	1.198	13
	Awareness Raising	3.36	.674	11
	Conventional Approach	3.25	1.488	8
	Total	3.37	1.100	32
Test 8	Metacognitive Pedagogical Sequence	3.31	.751	13
	Awareness Raising	3.27	.647	11
	Conventional Approach	3.63	.916	8
	Total	3.38	.751	32

The mean scores of these tests do not show substantial differences between the groups. For example, the mean scores for Test 1 are virtually the same for the three groups (i.e., Metacognitive Pedagogical Sequence = 4.23; Awareness-Raising = 4.36; Conventional

Approach = 4.25). In Test 2, the mean score difference between the highest score (Metacognitive Pedagogical Sequence = 4.69) and the lowest score (Awareness-Raising = 3.82) was 0.87. That is, in test two, participants in the Metacognitive Pedagogical Sequence group scored 17.4% higher than participants in the Awareness-Raising group. In Test 3, the mean score difference between the highest score (Awareness-Raising = 3.09) and the lowest score (Metacognitive Pedagogical Sequence = 2.38) was 0.71 indicating that the Awareness-Raising group scored 14.2% higher than the Metacognitive Pedagogical Sequence group. The mean score differences between the highest and lowest scores in Test 4, 5, 6, 7, and 8 (i.e., 0.27, 0.52, 0.18, 0.21, and 0.36, respectively) were even smaller (i.e., 5.4%, 10.4%, 3.6%, 4.2%, 7.2%, respectively).

The scores of the in-class listening tests show how similarly or how differently the groups performed at any single test. While these scores also show how the groups performed across time (i.e., from Test 1 to Test 8), it is difficult to tease this apart because the level of difficulty of the audios and the in-class listening tests varied across the intervention. In this sense, although the mean scores of the groups decreased from Test 1 to Test 8 as can be observed in Figure 7 below, this cannot be interpreted as a negative effect because the level of difficulty of the audios increased as they progressed from Test 1 to Test 8. That is, the audio text for intervention session one (Test 1) was less difficult than the audio text for intervention session eight (Test 8). Thus, what can be examined from these scores is how the groups performed at any single in-class listening test.

Figure 7: Mean Scores Comparison of the In-class Listening Tests



The mean scores of the three groups are nearly the same across most of the tests (i.e., Test 1, Test 4, Test 6, Test 7, and Test 8). As for Test 2, Test 3, and Test 5, the mean scores of the groups were more spread out from each other. However, this difference is not very large. That is, in Test 2, the Metacognitive Pedagogical Sequence group scored 17.4% higher than the Awareness-Raising group; in Test 3, the Awareness-Raising group scored 14.2% higher than the

Metacognitive Pedagogical Sequence group; and in Test 5, the Metacognitive Pedagogical Sequence group scored 10.4% higher than the Conventional Approach group.

An important outcome that the scores of the in-class listening tests portrayed refers to how the groups reacted to the intervention during the study. That is, the Metacognitive Pedagogical Sequence group scored highest in four of the in-class listening tests (i.e., Tests 2, Test 5, Test 6, and Test 7), the Awareness-Raising group scored highest in three of these tests (i.e., Tests 1, Test 3, and Test 4), and the Conventional Approach group scored highest in one of these tests (i.e., Test 8). These results are consistent with the previous analysis (repeated measures ANOVA) where the Metacognitive Pedagogical Sequence group and the Awareness-Raising group significantly improved from pre-test to post-test. That is, the scores of the in-class listening tests also indicate that participants in the Metacognitive Pedagogical Sequence group and the Awareness-Raising group did better than participants in the Conventional Approach group.

Also, it is interesting to observe that Test 1 (Figure 7 above) ranked the groups in a similar way as the pre-test did (Figure 5 above). That is, in both the pre-test and the first in-class listening test, the Metacognitive Pedagogical Sequence group scored last, the Conventional Approach group scored second highest, and the Awareness-Raising group scored the highest. This consistency in the ranking of the groups strengthens the confidence that the groups started at different levels of L2 listening. In the absence of random assignment, this consistency provides support to the previous analysis and strengthens the validity of the results.

4.2 Delayed Post-Test

This second part of the results section concerns research question two (RQ-2). This question refers to whether or not there are long-term effects associated with the different instructional approaches. To answer this question, a repeated measures ANOVA was performed with a new dataset that included the scores of the delayed post-test. This analysis is the most appropriate since it permits to observe changes over time and across groups (Gamst, Meyers & Guarino, 2008). The independent variables were *Time* with three levels (i.e., Time 1 at Pre-Test; Time 2 at Post-Test; and, Time 3 at Delayed Post-Test) and *Teaching Approach* with two levels (i.e., Metacognitive Pedagogical Sequence and Conventional Approach). The Awareness-Raising group was not included in the analysis of this new dataset because only two participants from this treatment group were able to take the delayed post-test (see section 3.2).

The dependent variable was listening performance and was measured with two versions of a standardized listening test (i.e., DELE practice test). Test Version A was used to measure listening ability at pre-test and post-test, and Test Version B was used to measure listening ability at delayed post-test. In this regard, although the two test versions had similar characteristics and were meant for the same level and listening proficiency, the two versions employed to measure listening ability differed in terms of level of difficulty. This is further elaborated below in section 4.2.1.

In the previous analysis (section 4.1), it was found that the groups started at different levels of listening ability. Thus, in this second analysis, the first thing to consider is whether or not the groups should be statistically equated before running the analysis. As mentioned earlier, statistical control is often a limitation in quasi-experimental designs. In the previous analysis,

statistical control (ANCOVA) was used under specific considerations. Statistical controls should be used carefully and "as the last resort" in quasi-experimental designs (Shadish, Cook & Campbell, 2002: 503). In this second part of the results, controlling for initial differences at pre-test is not necessary. The main concern in this second part of the results section is whether or not there are long-term effects (RQ-2). Thus, tracking and assessing changes through Time will suffice to answer research question 2 (RQ-2). Instead, what becomes more relevant is whether the test versions are equivalent or not since Test Version A was employed at pre-test and post-test, and Test Version B was employed at post-test. In this sense, this section is concerned with the results of the delayed post-test analysis and the results of the parallel test version analysis.

Therefore, to perform this second repeated measures ANOVA with the new dataset that included the scores of the delayed post-tests, assumptions associated with the analysis were evaluated first. The alpha level for the statistical analysis was set at 0.05 (i.e., $\alpha = .05$). A preliminary examination of the data showed one outlier. This outlier was removed from the data set and, hence, the number of cases in the analysis was reduced to 14—eight in the Metacognitive Pedagogical Sequence group and six in the Conventional Approach group (see section 3.2). The assumption of normality of the sampling distribution was confirmed by the non-significant *p-values* in the Shapiro-Wilk test (0.942, 0.626, and 0.369), as can be observed in Table 9 below.

Table 9: Tests of Normality for the Repeated Measures ANOVA: Analysis of Delayed post-tests.

	Shapiro-Wilk		
	Statistic	df	Sig.
Time 1 (Pre-Test)	.976	14	.942
Time 2 (Post-Test)	.954	14	.626
Time 3 (Delayed Post-Test)	.936	14	.369

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

The assumption of homogeneity of variance was also confirmed for all three levels of the within-subjects variable. That is, the non-significant *p-values* of Levene's test indicated that the assumption of homogeneity of variance was met (Pre-Test: $F(1, 12) = .302$, $p > .05$; Post-Test: $F(1, 12) = .036$, $p > .05$; and Delayed Post-Test: $F(1, 12) = 4.115$, $p > .05$). Also, the non-significant value of Box's M test ($p > .05$) indicated that the equality of covariance matrices assumption was not violated. As for the assumption of Sphericity, Mauchly's test was insignificant ($p > .05$) indicating that this assumption was not violated.

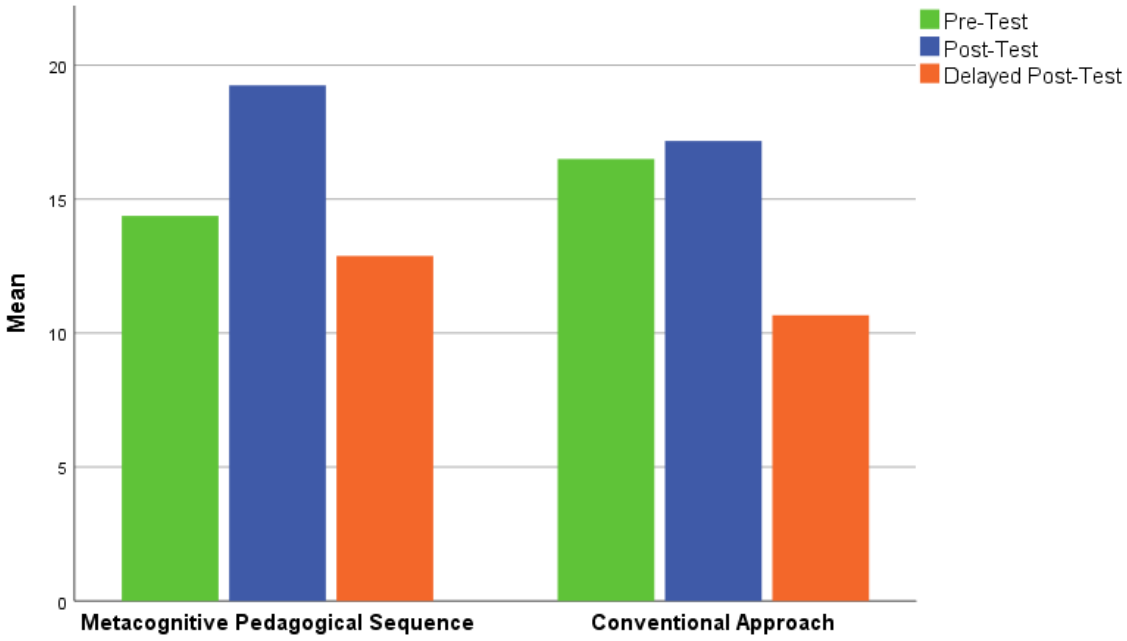
After evaluating these assumptions, the analysis was performed. Table 10 below presents the means, standard deviations and number of participants of the two groups at three different times (Time 1: pre-test; Time 2: post-test; and, Time 3: delayed post-test). As observed in this table, from Time 1 to Time 2, the Metacognitive Pedagogical Sequence group exhibited a larger gain (19.25 – 14.38 = 4.87) than the Conventional Approach group (17.17 – 16.50 = 0.67). On the other hand, from Time 2 to Time 3, a large decrease occurred in both the Metacognitive Pedagogical Sequence group (12.88 – 19.25 = – 6.37) and the Conventional Approach group (10.67 – 17.17 = – 6.50).

Table 10: Descriptive Statistics for the Repeated Measures ANOVA: Analysis of Delayed post-tests.

	Teaching Approach	Mean	Std. Deviation	N
Time 1 (Pre-Test)	Metacognitive Pedagogical Sequence	14.38	2.504	8
	Conventional Approach	16.50	3.017	6
	Total	15.29	2.840	14
Time 2 (Post-Test)	Metacognitive Pedagogical Sequence	19.25	3.536	8
	Conventional Approach	17.17	3.312	6
	Total	18.36	3.478	14
Time 3 (Delayed Post-Test)	Metacognitive Pedagogical Sequence	12.88	4.257	8
	Conventional Approach	10.67	2.503	6
	Total	11.93	3.668	14

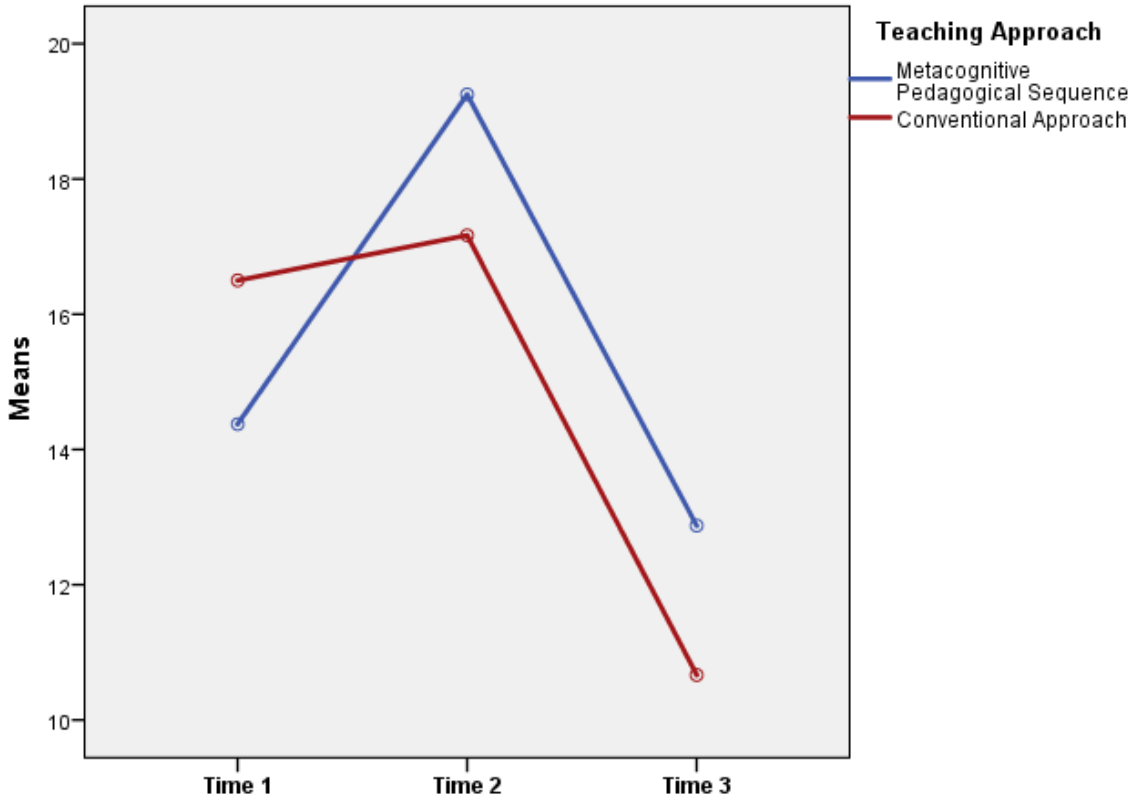
Interestingly, if the mean scores between Time 1 and Time 3 are compared, this decrease is not as large for the Metacognitive Pedagogical Sequence group (14.38 – 12.88 = 1.5) as it is for the Conventional Approach group (16.50 –10.67 = 5.83). This can also be observed in Figure 8 below. These changes in the mean scores, however, need to be tested for statistical significance.

Figure 8: Comparison of the mean scores for the Pre-/Post-/Delayed Post-Test.



The analysis yielded significant main effects for *Time* ($F(2, 24) = 36.717$, $p < .05$), but insignificant main effects for Teaching Approach ($F(1, 12) = 0.217$, $p > .05$). It also yielded significant effects for the interaction between Time and Teaching Approach ($F(2, 24) = 5.356$, $p < .05$). The effect size for the interaction was large with a Partial Eta Squared of .309. Since the interaction between the two independent variables was significant with a large effect size explaining 30.9 % of the variance and associated error, the significant main effect for Time was not further examined. Instead, the interaction effect was examined in more detail. Therefore, to further examine the effects of this interaction, Bonferroni multiple comparison tests were performed and the plot below was generated to describe how the interaction between the two independent variables operates.

Figure 9: L2 Listening performance as a function of Time and Teaching Approach across Pre-test, Post-test and Delayed Post-test.



As can be observed in Figure 9, the Metacognitive Pedagogical Sequence group improved much more than the Conventional Approach group from Time 1 to Time 2 (i.e., pre-test to post-test). This improvement was statistically significant for the Metacognitive Pedagogical Sequence group (Mean difference: 4.875, $p < .05$), but not for the Conventional Approach group (Mean difference: 0.667, $p > .05$). However, from Time 2 to Time 3 (i.e., post-test to delayed post-test), both groups plummeted considerably. This decrease was significant for both the Metacognitive Pedagogical Sequence group (Mean difference: -6.375, $p < .05$) and the

Conventional Approach group (Mean difference: -6.500, $p < .05$). Additionally, when comparing the mean scores from Time 1 to Time 3 (i.e., pre-test and delayed post-test), only the Conventional Approach group showed a statistically significant decrease (Mean difference: -5.833, $p < .05$). The Metacognitive Pedagogical Sequence group showed a decrease, but this decrease was not statistically significant (Mean difference: -1.500, $p > .05$). This is further elaborated in the section below.

4.2.1 Parallel Test Forms

Considering that the decrease experienced from Time 2 to Time 3 is practically the same for both groups (– 6.37; – 6.5), and that L2 listening performance was measured with Test Version A at Time 2 and with Test Version B at Time 3, it might be the case that Test Version B had a greater level of difficulty. Therefore, to further examine if the two test versions had the same level of difficulty, the tests were administered to a different group of participants (see section 3.5.3). Then, a correlation (i.e., reliability coefficient) was computed to determine if the two versions were parallel (Thorndike & Thorndike-Christ, 2010). In this sense, two parallel versions of a test should yield very similar scores for the same individuals. In other words, the scores of two parallel versions of a test should be highly correlated (i.e., $r > .70$).

The correlation between the scores of Test Version A and Test Version B was very low ($r = .143$). This indicated that the test versions were not parallel. Table 11 below presents the means and standard deviations of the two test versions that were administered to the same group of participants.

Table 11: Descriptive Statistics: Comparison of Test Version A and Test Version B.

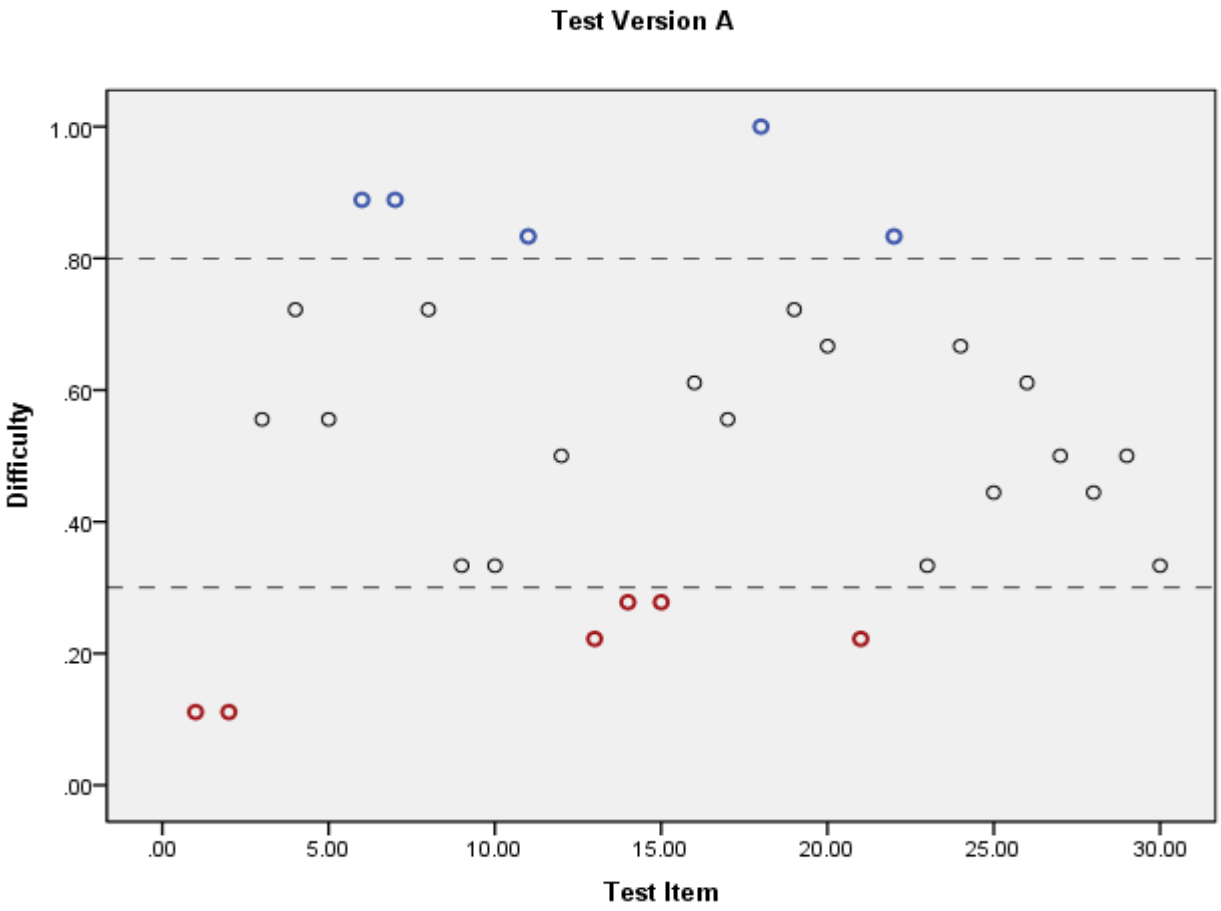
	Mean	Std. Deviation	N
Test Version A	16.3077	3.42502	13
Test Version B	9.9231	3.12147	13

As can be seen in Table 11, the mean score for Test Version A is higher than the mean score for Test Version B. The mean difference generated by the same group of participants in these two test versions is large (i.e., 16.3077 - 9.9231 = 6.4246). Interestingly, this mean difference (6.42) is very similar to the decrease that the groups experienced from Time 2 to Time 3 (i.e., – 6.37; – 6.50). This decrease can be observed in Figure 9 above.

To further examine the test versions, an item difficulty analysis was conducted. This analysis was based on "the proportion of examinees who answered the item correctly" (Thorndike & Thorndike-Christ, 2010: 306). It consisted of computing the number of examinees who answered the item correctly divided by the total number of examinees. Thus, an item that is difficult will be characterized by its low proportion (i.e., 0.3 or lower), and an item that is easy will be characterized by its high proportion (i.e., 0.8 or higher).

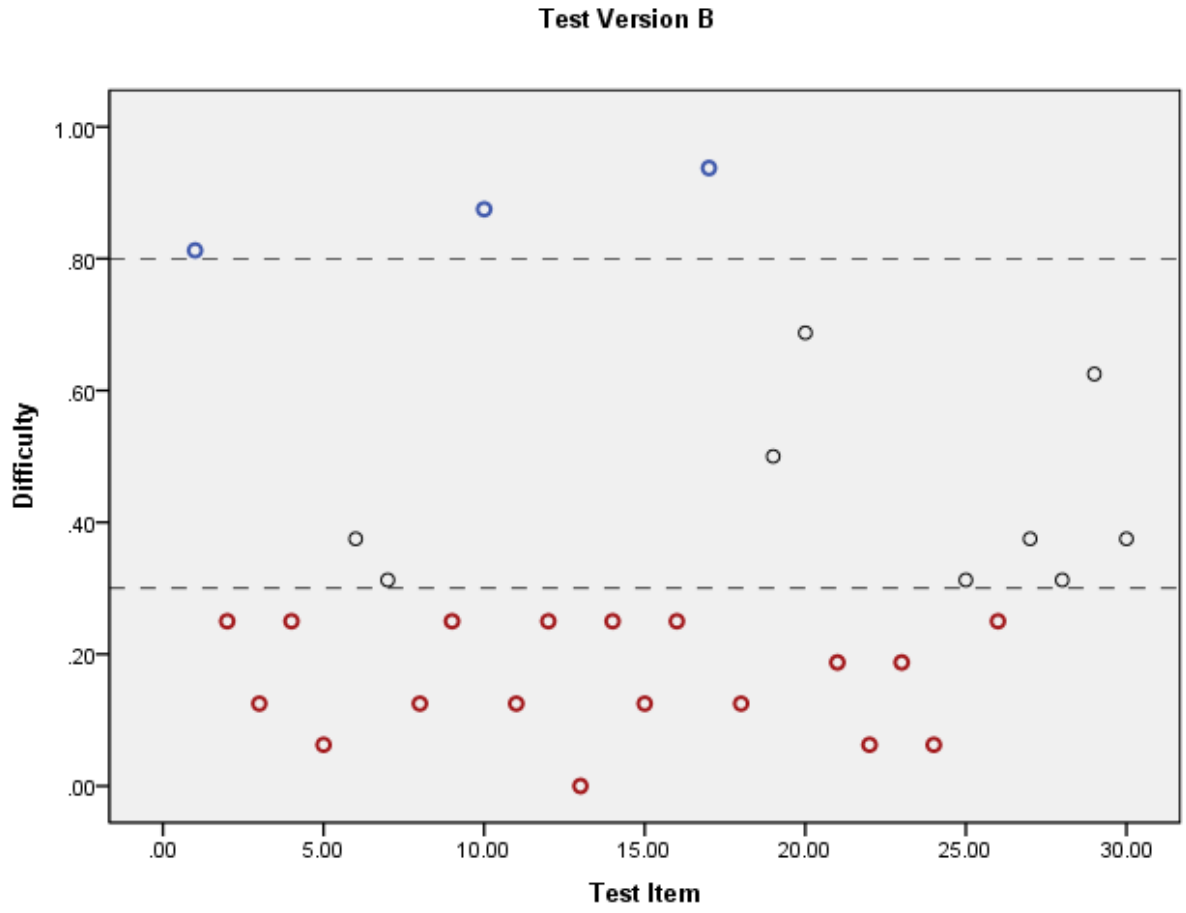
This analysis indicated that the test versions differed in level of difficulty. That is, out of the 30 items, Test Version A had six items characterized as difficult and Test Version B had 18 items characterized as difficult. This can be observed in Figures 10 and 11 below,

Figure 10: Item Difficulty Analysis: Test Version A.



In the lower section of Figure 10, it can be observed that six items were identified as difficult, and in the upper section of the same figure, one can observe that five items were identified as easy. In the middle section, 19 items fall in the category of suitable test items. The mean score generated by the 18 participants on this test version was 15.77.

Figure 11: Item Difficulty Analysis: Test Version B.



In Figure 11, most of the items fall in the lower section of the graph. There are 18 test items that fall in the category of difficult items. In the upper section of the graph, there are three items characterized as easy. And there are nine items that were identified as suitable. The mean score generated by the 16 participants on Test Version B was 9.44. Therefore, the low correlation and mean differences between the two versions of the test discussed above, as well as the item

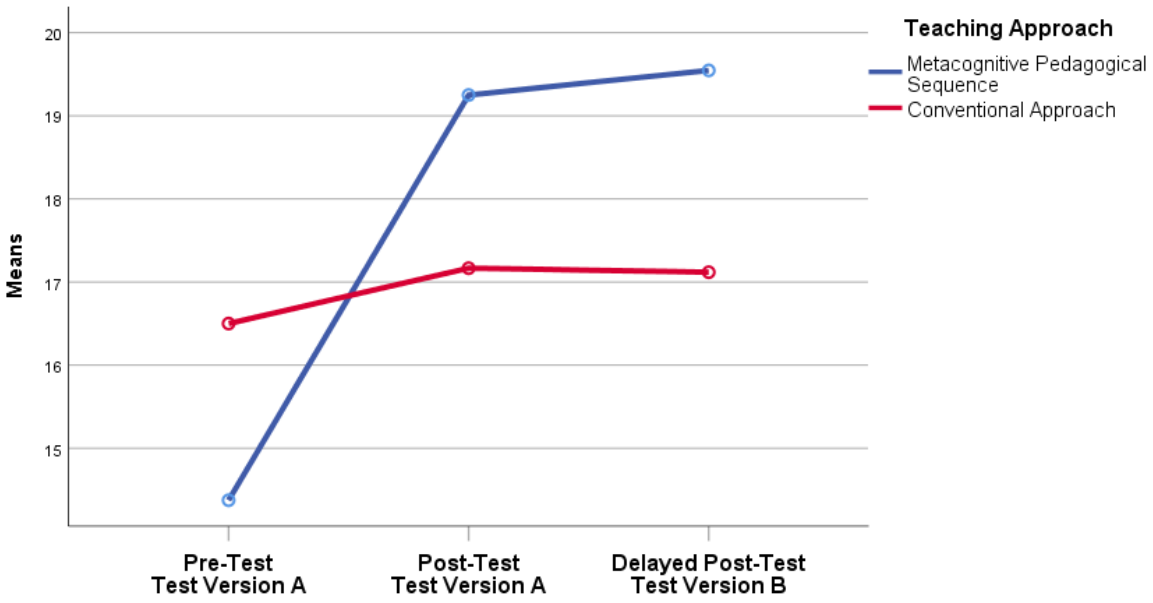
difficulty analysis shown here, indicated that Test Version B had a greater level of difficulty than Test Version A.

4.2.2 Delayed Post-Test: Test Version Difficulty taken into account

The results of the analysis that included the scores of delayed post-tests (section 4.2 above) indicated that the Conventional Approach group and the Metacognitive Pedagogical Sequence group decreased significantly from post-test to delayed post-test (Time 2 to Time 3). However, these results did not take into account that the delayed post-test was a different test version which had a greater level of difficulty as indicated in the item difficulty analysis (see section 4.2.1). This reasonably explained the significant decrease that both groups experienced from post-test (Test Version A) to delayed post-test (Test Version B). Interestingly, as mentioned earlier, this decrease was parallel in the two groups as can be observed in Figure 9 above.

Even more interesting, the analysis of the test versions indicated that Test Version A (15.77) differed from Test Version B (9.44) in a very similar magnitude (6.33) as the decrease experienced by the two groups (6.37 and 6.50). Therefore, if Test Version B is adjusted in the level of difficulty to match the same level of difficulty as Test Version A, the groups no longer exhibit a decrease from post-test (Time 2) to delayed post-test (Time 3). In fact, after adjusting for test difficulty, the groups exhibit two fairly horizontal lines from post-test to delayed post-test as can be observed in Figure 12 below.

Figure 12: Delayed post-test scores after adjusting for test difficulty.



To adjust the scores of Test Version B, linear equating was employed. This procedure allows to equate the scores of two different test forms. Linear equating is calculated with the mean scores (M) and standard deviations (SD) of both test versions. The formal definition of linear equating is "if X represents a score on the new form and Y represents a score on the reference form, then X and Y are equivalent in a group of test takers" (Livingston, 2014: 14); that is, Y – M(y)/SD(y) = X—M(x)/SD(x). Thus, the mean scores and standard deviations of the 13 participants (see Table 11 above) who took both tests were employed to adjust the scores of the delayed post-test (Test Version B). These scores were 19.55 for the Metacognitive Pedagogical Sequence group and 17.12 for the Conventional Approach group.

Thus, it appears that the gains that the groups experienced during the intervention remained after five months of the intervention. On the one hand, although the Conventional Approach group improved from pre-test to post-test, this improvement was not significant (Mean difference = 0.667, $p > .05$). On the other hand, the Metacognitive Pedagogical Sequence group exhibited a significant improvement from pre-test to post-test (Mean difference = 4.875, $p < .05$) which, as observed in Figure 12 above, seems to have remained. This answers research question 2 (RQ-2) regarding whether or not there are long-term effects. Thus, it seems that the effects of the approach implemented in the Metacognitive Pedagogical Sequence group remained as measured five months later.

4.3 Participants' perceptions regarding the L2 Listening activities

This last part of the results concerns research question three (RQ-3). This question refers to the perception of the participants regarding the L2 listening activities implemented in the different groups. These results compile data from the focus group interviews and the questionnaires (see section 3.4.3). These data were transcribed and segmented into categories based on how frequent the comments were. That is, the data were analysed by reading the participants' comments, grouping similar comments according to emerging themes, segmenting the comments based on how frequent they were within each group, labeling these segments, and collapsing overlapping labels. This process identified three categories: 1) positive aspects of the listening activities, 2) challenging aspects in the listening activities, and 3) suggestions for improvement.

4.3.1 Positive aspects of the listening activities

4.3.1.1 Metacognitive Pedagogical Sequence group

The approach implemented in the Metacognitive Pedagogical sequence group was positively received by the participants. Comments such as [1] "The listening activities were good, I learned how to listen more carefully." indicated that the approach had a positive impact on the development of L2 listening skills. This was echoed by comments such as [2] "It was good. I noticed my Spanish listening improved significantly. At first I would become exhausted from the strain quite quickly and later my endurance improved."

These participants reported that providing the topic of the audio text at the beginning of the listening activity as well as devoting some time for group discussions with the aim of sharing information and reconstructing the audio text were helpful. This can be observed in comments such as: [3] "I like that you can build up by thinking about the topic and hearing the audio once while jotting down notes. Talking to a partner and comparing notes helped, too."

The most salient aspect in this treatment group is the way participants learned to approach L2 listening. It seems that these participants learned to approach L2 listening more strategically as observed in comments such as the one below,

[4] Now, kind of like, I write down key words that I hear, that I recognize**,** and then I kind of like, slowly kind of communicate that to each thing and I try to see a pattern. I think I write a lot like I try to catch certain words and phrases. And then, I think, last time I even tried like to just listen once and then get down the words and then kind of put

them together.

Also, these participants reported that the L2 listening activities were helpful and contributed to the development of their listening skills in terms of their focused attention and anxiety. This was observed in comments such as [5] “It helped because it was definitely more practice than usual which increases our attention and skills in listening.”; also, [6] “It definitely helped with the midterm. I was a lot more used to listening activities and was much less nervous for that component of the midterm.” These participants commented that the approach boosted their confidence,

[7] Yeah, I'm not sure why I could understand things better the second time around. Maybe it's the daily Spanish, maybe it's the fact that I was already exposed to that same listening. I'm not sure, but seeing that improvement in itself, like just realizing that I understood more words, that made me like happier. It gave me more confidence. So, sometimes when you, when you realize there's been a little bit of an improvement like that helps so much, it makes you so happy.

4.3.1.2 Awareness-Raising group

Participants in the Awareness-Raising group also reported that the L2 listening activities contributed positively to the development of their L2 listening ability. They commented that the activities increased their awareness of L2 listening and helped them reflect critically on how to approach L2 listening. This can be observed in the comment below,

[8] I think it's interesting to be prompted by questions that allow to think how to conduct listening prior to doing them. It increases awareness of what you're doing and allowed me to think critically about how I was going about the activities. I think it helped me become more focused during these activities.

The listening activities implemented in this group seemed to have influenced participants to use contextual information to build a better understanding of the listening. For example, one participant said, [9] "Before, I used to get caught up when I lost track of what they're talking about, so like, after the activities I think, then I changed my approach. Now, I kind of look at the context instead." This can also be observed in the comment below,

[10] Before, I used to look for words I knew, but now like, I pay more attention to the context. And then, once you have the context, you can go to what you already know, and then you can, like, deduce the best in a better way you feel.

Thus, it seems that these participants, similar to participants in the Metacognitive Pedagogical Sequence group, learned to approach L2 listening more strategically. In fact, these participants commented that the discussion questions (i.e., guiding questions about L2 listening from the Metacognitive Awareness Listening Questionnaire—MALQ) as well as the small group discussions were helpful because they allowed them to plan and get ready for the listening activity. This can be observed in comments such as the ones below,

[11] I really liked the teaching style. I enjoyed how there was a different point to talk about each time we had a listening activity and also that we were given a chance to talk with our classmates about the questions. I'm pretty sure I gained some ideas from them.

[12] I find that the discussion questions before listening are effective because they force you to plan ahead (plan what kind of strategies you're going to use while listening). I also find that listening once before the [comprehension] questions are handed out is effective.

Moreover, the instructional approach employed in this treatment group seemed to have helped participants become more confident with L2 listening: [13] "I now feel more confident listening since I am used to it. I find it easier to find what I was looking for. I felt good about the listening component in the midterm and I think I did well." These participants commented that the listening activities helped them reduce their anxiety and better manage the listening component in the course midterm exam: [14] "Yes, I think it helped me. Firstly, I wasn't as nervous for it [midterm] as I usually am because we have been practicing. Secondly, I felt more organized and was able to plan the way I was going to listen and read the [comprehension] questions before." This can also be observed in comments such as the one below,

[15] I believe the listening practice helped with the listening on the midterm. I felt calm going into the listening section and was able to remain focused, as I had been practicing in class. I also did not get frustrated when I did not understand particular

words. I have been practicing using context to help with comprehension, and I believe I was able to achieve this on the exam.

4.3.1.3 Conventional Approach group

Participants in the Conventional Approach group also perceived the listening activities as positive and beneficial. They appreciated the listening activities implemented in their class: [16] "I appreciated the listening activities, they felt fun and engaging. I felt like I accomplished something, even if that wasn't the case." These participants reported that the listening activities provided an opportunity to practice their listening skills as can be observed in comments such as the one below,

[17] I think that the constant practice and exposure to the listening activities has helped me improve my listening comprehension more than my ability to speak. I find that I am becoming more fluent in listening than in speaking.

Participants in this treatment group also mentioned that presenting the topic of the audio text at the beginning of the activity helped in building a better understanding of the L2 listening context. Particularly, they commented that introducing and explaining vocabulary before listening to the audio was helpful: [18] "I liked how terminology was explained prior to the listening activities. I also liked how the themes were discussed to provide a better understanding of the context." In fact, it seems that the feature of vocabulary was the most salient aspect of this approach since these participants consistently reported this aspect.

Another aspect that was positively viewed was the time participants were given to read and get familiar with the comprehension questions in the in-class listening tests. They mentioned that these questions provided guidance on what to listen for in the audio input: [19] “The questions were useful cause they guided me to focus my attention.”; also, [20] “It was well done. I liked that vocab was taught before the listening. There was also time to read over the [comprehension] questions and reflect which was nice.”

Although these participants did not report explicitly to feel more or less anxious and/or confident, they reported that the listening activities provided the necessary practice to feel more comfortable with the listening section in the course midterm exam: [21] “By listening to recordings and practicing these skills, I feel I have emerged from the midterm exam with better listening than I had at the beginning.” This can also be observed in the following comment,

[22] I think the listening practices helped me. This provided me with plenty of experience with listening and determining the answer based on what I heard. I was able to become more comfortable with the speaker's speed which helped me prepare for the [course midterm] exam.

4.3.2 Challenging aspects in the listening activities

While participants in the three groups commented positively regarding the listening activities implemented in the different groups, they expressed some discomfort associated with the audio material. Participants found the audio material challenging due to the different accents and speed of how people talked. Interestingly, participants in the three groups also acknowledged

that it was useful to be exposed to such challenging audio material. For example, participants in the Awareness-Raising group commented that the audios very much reflect how people normally talk: [23] “It would be nice if they weren't so fast, but I understand that is how the language is often spoken.”, and participants in the Metacognitive Pedagogical Sequence group commented:

[24] It would be better to not have that many different accents in the listening samples. However, it is good that it is so close to the actual Spanish speaking environment and can prevent the shock if we go there.

Likewise, while participants in the Conventional Approach group were not very pleased with how fast people in the audios talked, they recognized that it was useful to be exposed to these audios. This can be observed in comments such as the one below,

[25] Well, I think it was useful. To the extent that, like it gives you an idea of the speed to which your comprehension is going to have to operate. Ah, I'm sure that someone who doesn't speak English would also feel that we speak English very quickly. It was good to know the tempo, like how quickly native speakers are going to be speaking, sort of oh okay, I really got a lot to learn now.

Also, two participants in the Metacognitive Pedagogical Sequence group commented that reflecting about the way they were listening interfered with their attention. It may appear that the reflection component of this approach made participants overthink about the way they were

listening and diverted their attention from the act of listening. This can also be observed in the comment below,

[26] I really liked having that discussion. I think that helped me fill in a lot of the blanks cause I didn't catch certain words that somebody else would, so I liked that. I think as overthinking it though, because sometimes there's a lot of questions about like how are you listening and things like that, so I was starting to think more about how I listen than what I was listening to. I think so, yeah, I think that it made me overthink, but I really love top parts [predicting and verifying information, reconstructing the audio text in small group discussions, monitoring comprehension], like I love having a group discussion. I think also that it made me use my Spanish more, like after hearing Spanish that I want to communicate, yeah, but I didn't really like, yeah just too much thinking about how I was listening.

4.3.3 Suggestions for improvement

In terms of improving the listening activities, participants in the three groups commented that they could benefit even more if the audio transcript is incorporated in the listening activities. For example, participants in the Metacognitive Pedagogical Sequence group commented that the written transcription could help them match sound with meaning:

[27] So, ah, for me, personally, it really helps seeing the text. Sometimes you hear a word and you just you think about like oh I wonder how it is spelled out, or these kinds of

things. So, maybe like listening to it for the first time, just to kind of get that listening part in, and then the second time around or third time around having access to the written, and then you are like oh yeah like I didn't understand the accent or like the way they said it, but if I look down on the word I can now associate the sound with meaning.

Participants in the Conventional Approach group also commented that the inclusion of the audio transcription can be useful: [28] “A transcript of the listening text would have helped better cause I sometimes could not tease the words apart.” and [29] “Reading the transcript would be helpful in understanding the fast bits too.” Also, participants in the Awareness-Raising group commented in this regard:

[30] It might be interesting to sometimes give students a written version of the audio after the exercise, so that they can understand everything they missed and improved more without making the activity easier. That's because the one problem I found was I never understood my mistakes as I didn't learn the right answer if I missed it twice.

Additionally, participants in the Awareness-Raising group felt that they could benefit more if visual input is incorporated into the listening activities: [31] “I would maybe try to show videos with the listening. This way it is possible to get more information about the context which would lead to a better understanding.” Participants in the Conventional Approach also commented that visual input can be useful:

[32] I really like watching Spanish videos with Spanish subtitles, but I don't know if it helps or not, it seems to when the talking is very fast. Sometimes the choices of the answers are similar or tricky, it takes me a long time to read and then focus on the audio, or at the same time, and I miss information in the audio while reading the questions and answers.

To summarize, the comments revealed that participants from the three groups perceived the activities as mainly positive and beneficial. Providing the topic of the listening text as well as giving participants some time to get familiar with questions of the in-class listening tests were considered useful since this provided important contextual information to build a better understanding of the audio text. Also, participants in the three groups commented that incorporating the audio transcription would be beneficial since this could help in matching sound (i.e., audio input/spoken language) with meaning (i.e., audio transcription/written language). Additionally, although participants in the three groups experienced some discomfort associated with the different accents and fast talking of the people in the audios, they recognized that normal speech is fast.

The comments also revealed that the most salient learning outcome for the participants in the Conventional Approach group was vocabulary as these participants consistently reported this particular feature. Participants in the Metacognitive Pedagogical Sequence group and the Awareness-Raising group displayed a more systematic behaviour when approaching the task of L2 listening as observed in their comments. It may seem that these participants learned to approach L2 listening more strategically. Participants in these two groups also believed that the

listening activities helped them manage their anxiety better when doing L2 listening in the classroom.

4.4 Summary and Conclusion

This chapter presented the results of this study in relation to the research questions, which were concerned, on the one hand, with whether or not there were significant changes in L2 listening performance after implementing the three treatments, and, on the other hand, with the participants' perceptions regarding the listening activities in the treatments. In addition, the chapter also discussed the appropriateness of the statistical procedures employed to analyze the data, as well as the comparability of the two test versions used to measure listening performance. Finally, the chapter also examined the participants' comments in relation to the positive and challenging aspects of the listening activities, as well as suggestions for improvement of those activities. The next chapter discusses these results with respect to previous research and explores potential pedagogical implications stemming from this study. The next chapter also addresses the limitations of the present study as well as the directions for future studies.

5 Discussion

This study investigated the effects of three instructional approaches to L2 listening; that is, a metacognitive pedagogical cycle, an approach that raised learners' awareness of the factors associated with successful L2 listening, and a comprehension approach. The overarching question that guided the study was concerned with whether or not there are significant changes in L2 listening performance among three groups of participants who were exposed to these instructional approaches. In this sense, the research questions that guided this study were concerned with whether or not the groups experienced a significant change from pre-test to post-test (RQ-1a), whether or not the groups differed significantly at post-test (RQ-1b), whether or not there were long-term effects (RQ-2), and the participants' opinions about the L2 listening activities implemented in the different groups (RQ-3). This chapter is organized in the following manner: 1) Discussion of the findings, which is organized in relation to the research questions mentioned above, 2) Pedagogical implications, and 3) Research implications, where the limitations of this study and future research possibilities are discussed.

5.1 Discussion of the Findings

5.1.1 Research Question-1a: Are there significant changes in L2 listening performance within each group from pre-test to post-test?

The results discussed in the previous chapter showed significant improvement in L2 listening from pre-test to post-test for the Metacognitive Pedagogical Sequence group and the

Awareness-Raising group. The Conventional Approach group also exhibited improvement; however, this improvement was not statistically significant. The effect size for this factor (i.e., the within-subjects factor: Pre-test and Post-test) was large indicating that 50.1% of the variance and associated error in listening performance can be explained by the factor Time.

The significant improvement that the Metacognitive Pedagogical Sequence group ($F(1, 28) = 27.442$, $p < .05$, $r = .70$) and the Awareness-Raising group ($F(1, 28) = 15.666$, $p < .05$, $r = .59$) experienced from pre-test to post-test provides further evidence of the importance of metacognition in learning. Metacognition is essential in learning because it allows individuals to control their thinking and learning strategies (Thomas, 2003). Learners with good metacognitive control can better manage their cognitive skills in a way that performance directly improves (Schraw, 1998), as it was the case of the learners exposed to metacognitive instruction in this study (i.e., the Metacognitive Pedagogical Sequence group and the Awareness-Raising group). Metacognitive knowledge and the ability to apply this knowledge are key characteristics of successful learners: "learners with a high degree of metacognitive knowledge and the facility to apply that knowledge are better at processing and storing new information, finding the best ways to practice, and reinforcing what they have learned" (Vandergrift & Tafaghodtari, 2010: 473).

The significant improvement of the Metacognitive Pedagogical Sequence group and the Awareness-Raising group also provides evidence of the importance and usefulness of metacognitive instruction in L2 listening, and contributes to the body of research that has found positive results regarding this type of instruction (e.g., Cross, 2011; Vandergrift & Tafaghodtari, 2010). Metacognitive instruction enables learners to take control of their learning as they learn to tune up their strategies and approach L2 listening more systematically. It raises awareness of the factors that affect L2 listening and positively "influences the manner in which learners approach

the task of listening and learning to listen" (Goh, 2008: 196). In this study, the learners in these two groups seem to have developed a sense of how to address L2 listening more strategically by developing a repertoire of strategies congruent with their personalities and the type of task.

These results provide support to the notion that listening practice should focus on the process, and not just on the outcome of the listening activities (Vandergrift & Goh, 2009). In fact, L2 listening instruction has been found to be more beneficial to learners when it is carried out as a learning process (e.g., Graham & Macaro, 2008; Mareschal, 2007; Thompson & Rubin, 1996; Vandergrift, 2004). In this study, listening instruction was carried out systematically and applied regularly in the language classroom as learners in the Metacognitive Pedagogical Sequence group and the Awareness-Raising group engaged in a learning process that allowed them to interact with peers, construct knowledge and reflect on their performance. The learners in these two groups were guided through a learning process which gave them opportunities to synchronize their strategies, try out new strategies, and experience the demands and nature of L2 listening. In this regard, Goh (2008) noted that learners who develop an awareness of the demands and nature of L2 listening are in a better position to plan, monitor and evaluate how to approach L2 listening.

The significant improvement of the Awareness-Raising group suggests that merely raising learners' awareness of factors such as problem solving, planning and evaluation, mental translation, person knowledge, and direct attention (MALQ factors) impacts L2 listening positively and substantially. By introducing these factors in the form of question prompts and having learners discuss these L2 listening aspects as well as by having learners reflect on their L2 listening experience, the learners in this treatment group improved their L2 listening skills significantly. It seems that raising the awareness of these factors as well as providing

opportunities for reflection enabled the L2 learners in this study to self-regulate and deploy the strategies that were more appropriate to their personalities and to the type of task (e.g., Vandergrift, et al., 2006).

Although there is not a current study that has attempted to systematically raise metacognitive awareness in this particular way, it seems that L2 listening instruction that combines process-based discussions and reflections brings potential benefits to learners when developing L2 listening skills (Goh & Taib, 2006). These results indicate that instructional practices that combine instruction, learners' reflection, and group activities in which learners discuss aspects associated with metacognition, such as the MALQ factors, impact learning in a positive and significant way (Schraw, 1998). Learner reflection in particular seems to be an important element in developing L2 listening skills. Reflecting about the L2 listening experience seems to have helped learners think about the way they were approaching the task of L2 listening and encouraged them to devise ways to improve their listening skills. In this sense, these results also provide support to the usefulness and effectiveness of encouraging and providing learners with opportunities to reflect about their listening experience (e.g., Goh, 2008; Graham & Macaro, 2008; Kohler, 2002).

The Conventional Approach group, conversely, did not improve as much as the other two groups. This treatment group improved from a mean score of 15.88 to a mean score of 16.38. However, this improvement was not statistically significant ($F(1, 28) = 0.298$, $p > .05$) and its effect size ($r = .10$) only accounted for 1% of the variance. Interestingly, the scores of the in-class listening tests also depicted a similar situation. That is, in the statistical descriptive analysis of the in-class listening tests, the Conventional Approach group scored the highest in only one of these tests. The other two groups (Metacognitive Pedagogical Sequence & Awareness-Raising)

scored the highest in four and three of these tests respectively. In other words, the improvement of the Conventional Approach group was not as good as the improvement of the other two groups as indicated by the in-class listening tests and by the non-significant results of the pre/post-test analysis.

The most salient learning outcome in the Conventional Approach group appears to be associated with vocabulary, as it was consistently reported by the learners in this group. However, it is worth noting that this outcome is only based on the participants' perceptions, as vocabulary learning was not measured in the study. Therefore, it is not possible to know whether these participants actually learned vocabulary. In any case, in the Conventional Approach group, unlike in the other two groups, vocabulary was introduced as a pre-listening activity. Then, participants were asked some comprehension questions intended to build a better understanding of the audio text. Thus, a possible explanation why this group might not have experienced a larger improvement in L2 listening may be that participants in this treatment condition were more concerned about the vocabulary than on building an understanding of the listening text. It may seem that these learners were more focused on decoding and parsing the speech stream than on applying prior knowledge (Vandergrift & Goh, 2009). In fact, the treatment condition in the Conventional Approach group seems to have recreated the situation described in Graham (2017): "Teachers in England also seem to direct learners' attention more to individual items of vocabulary than to broader themes that might rise in a listening passage, possibly because they see it as a surer way of helping learners obtain 'correct answers' to accompanying comprehension questions" (p. 114). Thus, it may seem that explaining vocabulary—or explaining it in this particular teacher-centered manner—did not help these learners focus on building an understanding of the listening text.

L2 Listening instruction in the Conventional Approach group was delivered in a more teacher-centered manner. That is, at the beginning of the lesson, the researcher drew the students' attention to the board, provided the topic of the audio text, and introduced some vocabulary from the audio text. Next, the researcher played the audio and subsequently asked some comprehension questions. Finally, the researcher handed out an in-class listening test, and the audio input was played again. Instruction in this group was delivered with little or no participation from the students. Learners just sat and listened to the researcher explain some vocabulary and sometimes interacted with the researcher to answer comprehension questions or to clarify meanings. This is certainly a teacher-centered approach in which learners take a passive role in the class activity and are not encouraged to interact with their peers or to make choices or to assess and reflect on their learning (Benson, 2012). In contrast, the other two approaches were more learner-centered and led to better results, thereby highlighting the need to focus instruction on the listening process and not just the product. By guiding and helping the learners in the two metacognitive instruction groups explore aspects of their personalities, strategies, the demands and nature of L2 listening, and the benefits of collaborative learning, these learners seemed to have developed a sense of how to become better L2 listeners as can be observed in the significant L2 listening improvement.

5.1.2 Research Question 1b: Are there significant differences in L2 listening performance between the groups at post-test?

To observe whether or not the groups differed significantly at post-test, initial differences between the groups were statistically controlled. The results of this analysis were significant

($F(2, 27) = 3.94$, $p < .05$) with a large effect size which indicated that 23 % of the variance and associated error in listening performance can be explained by the instructional approach. Further analyses indicated that only the Metacognitive Pedagogical Sequence group differed significantly from the Conventional Approach group (Mean difference = 3.269). That is, participants in the Metacognitive Pedagogical Sequence group significantly outperformed participants in the Conventional Approach group. Importantly, the effect size for this comparison ($r = 0.46$) indicated that 21% of the variance in L2 listening performance can be explained by the approach implemented in the Metacognitive Pedagogical Sequence group. These results are very similar to those reported in Vandergrift and Tafaghodtari's (2010) study in which participants in the experimental group significantly outperformed participants in the control group after controlling for initial differences.

These results provide empirical support to the benefits and effectiveness of the metacognitive pedagogical cycle proposed by Vandergrift (2004). Although a shorter version of this metacognitive cycle was implemented in this study, the fact that the Metacognitive Pedagogical Sequence group significantly outperformed the Conventional Approach group suggests that the metacognitive cycle has more benefits to offer learners. In this regard, Vandergrift and Tafaghodtari (2010) argued that the effectiveness of this approach might be associated with "the implicit knowledge about L2 listening acquired by learners through task performance" (p. 490). Thus, it seems that the metacognitive cycle is indeed a potentially useful approach when developing L2 listening skills in the classroom, especially when it is applied systematically and regularly.

While the Awareness-Raising group was not found to be significantly different from the Conventional Approach group, its improvement in relation to the Conventional Approach group

was substantial (Mean difference = 2.638). In fact, the effect size for this comparison (i.e., r = 0.39) indicated that the improvement of the Awareness-Raising group was relevant and considerable. In other words, although the Awareness-Raising group did not reach statistical significance, the magnitude of its improvement in relation to the Conventional Approach group is of practical significance. In fact, the effect size indicates that 15 % of the variance in listening performance can be accounted for by the instructional approach implemented in the Awareness-Raising group.

The Awareness-Raising group and the Metacognitive Pedagogical Sequence group were not found to be significantly different from each other. In fact, their mean scores were very similar (18.63 and 19.26 respectively). While both of these instructional approaches constitute metacognitive instruction in L2 listening, these approaches were not delivered in the same way. That is, learners in the Metacognitive Pedagogical Sequence group were guided through a more implicit learning process in which learners engaged in prediction, verification, and reconstruction of the audio text (Vandergrift & Tafaghodtari, 2010), whereas learners in the Awareness-Raising group more explicitly engaged in discussing factors that affect L2 listening (MALQ factors). Thus, the fact that the Metacognitive Pedagogical Sequence group exhibited a slightly higher mean score in listening performance than the Awareness-Raising group might be an indication that implicit metacognitive instruction that focuses on practicing strategies is somewhat more effective than explicit metacognitive instruction that focuses on talking about strategies. In any case, it is worth emphasizing the fact that no previous study has compared an implicit metacognitive approach with a more explicit metacognitive approach. This is hardly surprising since, as noted in chapter 2 of this dissertation, there is a general lack of intervention studies that compare instructional approaches in L2 listening.

Despite the slight non-significant difference between the Metacognitive Pedagogical Sequence group and the Awareness-Raising group, what is important to emphasize is the fact that both metacognitive instruction treatments led to a significantly better performance in L2 listening. Thus, the results of this study indicate that the metacognitive cycle proposed by Vandergrift (2004) indeed enhances L2 listening in a way that performance improves significantly. These results also show that raising awareness of the factors that affect L2 listening by promoting guided discussions and reflections impacts L2 listening substantially as indicated by the effect size for the comparison of the Awareness-Raising group and the Conventional Approach group (i.e., $r = 0.39$). While the effect size for the comparison of the Metacognitive Pedagogical Sequence group and the Conventional Approach group was larger (i.e., $r = 0.46$), both of these effects fall in the category of medium to large effect sizes (i.e., 0.3 to 0.5) (Field, 2009: 57). This suggests that both instructional approaches (Metacognitive Pedagogical Sequence and Awareness-Raising) offer substantial and better benefits to learners. These results respond to previous concerns regarding the lack of intervention studies that compare instructional approaches in order to determine which approach offers better benefits to learners (Berne, 2004; Macaro et al., 2007).

5.1.3 Research Question 2: Are there long-term effects associated with the different types of listening instruction?

As for whether or not changes in listening performance remained after the intervention, the results in the previous chapter indicated that the significant improvement that the Metacognitive Pedagogical Sequence group experienced from pre-test to post-test remained as

measured five months after the intervention. Although these results need to be interpreted cautiously due to participant attrition (see section 3.2) and the adjustment in the level of difficulty of the two test versions employed at post-test and delayed post-test (see section 4.2), it seems that the metacognitive cycle employed in this study produced long-term effects.

These results provide preliminary evidence of long-term effects attributed to metacognitive instruction in L2 listening. In this regard, in their intervention study about L2 listening strategy instruction, Graham and Macaro (2008) found evidence of long-term effects that they attributed to factors such as clustering of cognitive and metacognitive strategies and learner engagement in the strategy instruction process. They argued that strategic behaviour can only be measured within a model "of what a strategy is and does, how it operates in a cluster of strategies in relation to a task, and how it operates within an individual learner over a period of time" (Graham & Macaro, 2008: 774).

Metacognitive instruction encompasses these pedagogical principles. Metacognitive instruction in L2 listening is not concerned with teaching one or two strategies in isolation, but with providing opportunities in which learners develop person knowledge, task knowledge, strategy knowledge and strategy use. The instructional approaches (Metacognitive Pedagogical Sequence group & Awareness-Raising group) implemented in the present study are quite congruent with approaches that emphasize learner-centeredness and learner autonomy. These approaches encourage and empower learners to self-regulate, reflect, and ultimately take control of their learning (Wenden, 1998). Thus, while the positive results regarding long-term effects in the present study provide support to a type of listening instruction that encompasses such pedagogical principles, it also provides preliminary evidence of long-term effects attributed to the metacognitive cycle proposed by Vandergrift (2004).

5.1.4 Research Question 3: What are the learners' perceptions about the L2 listening activities implemented in the three different groups?

Regarding participants' opinions about the L2 listening activities implemented in this study, participants' comments in the three groups showed that the listening activities were perceived as useful, beneficial, and productive. Regardless of the instructional approach, participants in this study really appreciated that once a week, class time was uniquely devoted to L2 listening. While this may show how L2 listening is taken for granted in the language classroom (White, 2006), it is fascinating what eight 25-minute listening sessions can do.

Providing the topic of the listening text was positively viewed by the participants in the three groups. This is an important aspect in L2 listening because listeners do not exactly know what the audio input is going to be about, and the topic of the listening text contains contextual information that can help in making inferences to interpret the message better (Basnakova et al., 2013). In fact, L2 listeners rely on topic knowledge more than L1 listeners. When L2 listeners have access to the topic of the audio text, differences in working memory consumption between native and experienced non-native listeners are not significant (Tyler, 2001). Thus, providing the topic of the listening text is an important aspect in L2 listening instruction and learners benefit from it. As discussed in chapter 4, learners in this study reported that providing the topic of the listening at the beginning of the activity was useful because it helped them visualize the context of the listening.

Also, participants in the three treatment groups reported that the audio material was challenging. Learners reported experiencing some discomfort associated with the accents and

speed of the people talking in the audios. They felt that people in the audios talked too fast and the different accents in the audio texts made it difficult to understand. In this regard, despite the significant improvement in listening performance of the Metacognitive Pedagogical Sequence group and the Awareness-Raising group, the participants in these two group still perceived the audio material as challenging. This lends support to the argument that listening instruction needs to be carried out in the language classroom more often.

Observations in previous studies also indicate that the audio material is associated with distress and anxiety (e.g., Graham, 2006; Hasan, 2000). In fact, when the linguistic input is perceived as too fast and incomprehensible, listeners can experience high levels of anxiety (Vogely, 1999) which may in turn cause learners to avoid L2 listening situations. The learners in this study, however, acknowledged that native speakers have different accents and speak fast and, hence, that the audios very much resembled natural speech. That is, while feelings of distress and discomfort were associated with the speed and accents in the audios, learners acknowledged that the audios replicated authentic language.

Additionally, participants in the three groups believed that the written transcription of the audio text could potentially enhance their listening skills. They felt that the audio transcription could have helped them match sound with meaning in a better way. The incorporation of the written transcription in the listening activities can increase comprehension by helping learners tease words apart from the stream of speech. Recognizing words and identifying word boundaries is quite challenging for L2 listeners (Goh 2000), especially when the audio material is challenging as it was reported to be in this study. Nevertheless, while the audio transcription can certainly help in teasing words apart, it can also make the listening activity easier and perhaps not very useful as L2 listeners might engage in reading comprehension more than in

listening comprehension. It can also make the listening activity less authentic as an audio transcription may not be available when engaging in L2 listening outside the classroom. In fact, Vandergrift (2004) argued that the use of captions and written support in listening activities might encourage “word-by-word decoding rather than foster the development of compensatory strategies that can help students cope with the demands of real-time listening” (p. 10). Thus, the audio transcription needs to be carefully considered since it might transform the listening activity into a reading activity or perhaps into a translation exercise.

More specifically, in relation to the different instructional approaches, participants in the Conventional Approach group seemed to have mostly benefited from learning vocabulary and from exposure to the audio input. In this regard, Field (2012) noted that conventional practices in L2 listening bring benefits such as exposure to different voices and oral interactions as well as practice in different types of L2 listening. Interestingly, despite the positive comments from participants in this group, their L2 listening performance did not improve. This may be an indication that simply providing practice and exposure to oral input is not enough to improve one's ability in L2 listening. As Graham (2006) notes, "practice in itself does not address the issue that learners need to feel a sense of control over their listening, that improvement is possible" (p. 178).

Participants in the Metacognitive Pedagogical Sequence group and the Awareness-Raising group seemed to have developed a sense of approaching the task of L2 listening more strategically. This is consistent with what Schraw and Dennison (1994) indicated regarding metacognitively aware learners: “metacognitive awareness allows individuals to plan, sequence, and monitor their learning in a way that directly improves performance” (p. 460). Also, participants in the Metacognitive Pedagogical Sequence group and the Awareness-Raising group

more explicitly reported that the L2 listening sessions helped them become more confident, experience less anxiety, and better manage L2 listening in testing situations. L2 listening causes a lot of anxiety and frustration to L2 listeners (Graham, 2006) since it is often associated with evaluation in the language classroom (Mendelsohn, 1994). Furthermore, although L2 learners allocate a lot of effort to the task of L2 listening, they end up believing that L2 listening is impossible because they do not know how to address it more tangibly. In this study, participants in these two groups felt that the listening activities actually helped them increase their confidence and reduce their anxiety.

5.2 Pedagogical Implications

The most salient pedagogical implication that stems from the results of this study is that L2 listening instruction is more effective when it is student-centered and carried out as a learning process. In this study, L2 listening instruction in the Metacognitive Pedagogical Sequence group and the Awareness-Raising group was delivered in a student-centered manner and carried out systematically as a process-based approach. This learning process provided learners in these two groups with opportunities to practice L2 listening in the classroom. This listening practice offered the necessary scaffolding opportunities which enabled these learners to approach L2 listening more strategically.

As discussed above, the learning process through which the learners in the Metacognitive Pedagogical Sequence group were guided helped them improve their L2 listening skills significantly and ultimately outperformed participants in the Conventional Approach group. Participants in the Metacognitive Pedagogical Sequence group were systematically guided

through a set of stages that required them to predict information on the basis of the topic (before the activity), confirm and/or disconfirm these predictions on the basis of the audio input (during the first listening), share information and reconstruct the audio input in the small group discussions (after the first listening), assess their understanding of the listening text (during the second listening), and reflect about their listening experience (at the end of the activity). Guiding learners through these stages is feasible. In fact, having learners predict, confirm or disconfirm information, reconstruct the audio text, assess their understanding, and reflect about their listening experience is attainable in the classroom. Learners might need more guidance at the beginning, but it can become common practice as this cycle is applied regularly.

Similarly, participants in the Awareness-Raising group significantly improved from pre-test to post-test and performed substantially higher than the Conventional Approach group. Thus, having participants discuss the factors associated with successful L2 listening (MALQ factors) prior to the listening activity and having learners reflect about their listening experience at the end of the listening activity showed to be effective for the participants in the Awareness-Raising group. This, in particular, appears to be even more feasible and attainable in the language classroom. In fact, guided discussions and reflections can be easily incorporated in traditional listening activities (Goh & Taib, 2006), as it was the case of the activities implemented in this group. For example, the prompts in the metacognitive awareness listening questionnaire (MALQ) can be easily transformed into questions to promote small group discussions before the listening activity. Then, learners can do the regular listening activities, and at the end of these activities, learners can write a reflection about their listening experience in a listening journal. As Graham (2017) noted, "Giving learners opportunities to reflect on, and talk about how they listen and how they arrive at understanding might seem to be a relatively uncomplicated thing to do in

the classroom" (p. 111). This also seems to be a very practical way to encourage learner participation in the listening activities, especially in contexts where language teaching exhibits a strong teacher-centered tendency as well as in situations in which teachers do listening activities in the classroom but rarely teach L2 listening (Graham, 2017).

In this regard, although L2 listening instruction has highlighted the importance of pre-listening activities, contextualization of the listening, activation of schemata, and follow-up activities, conventional practices in listening instruction exhibit a strong teacher-centered tendency. In fact, Field (2012) noted that conventional practices in L2 listening need to move away from their current teacher-centered format by encouraging learners to be more actively involved and by encouraging teachers to be less interventionist. In this sense, while these conventional practices are a very important development in L2 listening, learners might not benefit as much as instructors do. In other words, while these conventional practices provide guidance to instructors on how to develop and structure their listening lessons (e.g., pre-listening activities, intensive listening, extensive listening, follow up activities), they do not necessarily provide learners with guidance on how to address L2 listening more strategically and systematically.

L2 listening instruction is more beneficial and productive when it is tailored around learners. A student-centered approach to L2 listening is definitely more beneficial than an approach in which learners just listen and answer comprehension questions. Engaging learners in an active process in which they interact with their peers, make choices, reflect, and assess their learning will certainly bring more benefits to learners than a teacher-centered approach. In this study, participants in the Metacognitive Pedagogical Sequence group and the Awareness-Raising group were guided through a process in which they interacted with each other, solved problems,

found out which strategies worked best and tried out new strategies, experienced the demands and challenges of L2 listening, and reflected about their listening experience. In this sense, L2 listening instruction brings learners more benefits when it is student-centered and when the development of listening skills is viewed as a learning process. L2 listening instruction also needs to be carried out regularly and systematically in the language classroom so that learners can develop a procedure to address L2 listening more strategically.

Importantly, developing listening skills only takes 25 minutes per class. In this study, eight L2 listening lessons were implemented. These listening lessons were conducted once a week and lasted for approximately 25 minutes each. The cost-benefit ratio of teaching learners how to listen in the second language certainly justifies its applicability in the language classroom. In fact, 25 minutes of class time devoted to developing listening skills translated into better L2 listening performance and more confident and motivated L2 listeners. Additionally, teaching learners how to listen in the second language has the potential of going beyond the classroom as listeners learn how to handle L2 listening more strategically, which in turn "boosts their confidence when they come to real-life listening encounters" (Field, 2012: 211).

Participants in the Conventional Approach group, unlike participants in the other two groups, were explained vocabulary as a pre-listening activity. While this study is not suggesting that explaining vocabulary is not beneficial, introducing vocabulary at early stages of the listening activity might cause learners to focus their attention on such vocabulary and not necessarily on building an understanding of the audio text. Vocabulary is a very important aspect in L2 listening (Rost, 2005); however, if the goal is to develop L2 listening skills and ultimately help learners address L2 listening more strategically, it is important to consider that learners might center their attention around the vocabulary introduced as a pre-listening activity and

might focus more on finding this vocabulary in the audio text than on trying to build an understanding of the audio input (Graham, 2017). Thus, it may be more appropriate to introduce vocabulary at later stages in the listening activity, once learners have had a chance to apply their strategies. In fact, it might make more sense to introduce vocabulary with the audio transcription at the end of the listening activity.

Regarding the audio transcription, participants in the three groups believed that having the transcription available can be helpful. The audio transcription can indeed help learners in building a better understanding of the audio text, especially in situations where the audio material is challenging as L2 listeners might find it difficult to tease words apart from the stream of speech. Nevertheless, it is also important to consider that the audio transcription can transform the listening activity into a reading activity or a translation exercise as learners might engage primarily in reading the audio transcription rather than paying attention to the audio material. Therefore, the incorporation of the audio transcription in the listening lesson should be considered carefully.

In this regard, it might be more appropriate to provide the audio transcription after the reflection stage in the Metacognitive Pedagogical sequence group and the Awareness-Raising group (i.e., as a follow-up activity) since learners would have already had a chance to practice their strategies. Likewise, at this time, it may also be appropriate to address vocabulary. In this particular way, learners would have a chance to develop their listening skills even more because they would be approaching L2 listening from a top-down perspective by applying metacognitive knowledge as well as from a bottom-up perspective when the audio transcription is provided and vocabulary is explained. Thus, adding a new phase at the end of the listening activities in which learners are given the audio transcription and the audio input is played one more time can

potentially enhance the approaches implemented in the Metacognitive Pedagogical Sequence group and Awareness-Raising group. Also, explaining vocabulary might be more appropriate at this time, turning it into a post-listening activity, rather than a pre-listening activity, when developing listening skills.

Another aspect that arose as an important pedagogical implication in this study is the provision of the topic of the audio text. The topic of the audio text contains contextual information that can guide learners in interpreting the message. It also helps listeners use their attentional resources more efficiently (Tayler, 2001). Providing details associated with the topic of the audio text at the beginning of the listening activity is indeed an important aspect in developing listening skills in the classroom because it helps learners establish connections with their background knowledge.

Finally, L2 listening instruction needs to incorporate authentic audio material. In this study, participants in the three group experienced some discomfort associated with the speed and accents of the audio material. However, at the end of the learning experience, they acknowledged that the audio material was helpful because it portrayed the reality of how people normally talk. In this sense, while exposure to challenging audio material can cause discomfort and distress to learners, exposing L2 listeners to authentic audio material is important and useful when the objective is to develop listening skills. This can make learners—as it was the case of the Metacognitive Pedagogical Sequence group and the Awareness-Raising group—become aware of the challenges and demands of L2 listening and help them devise ways to address L2 listening more strategically, which in turn reduces anxiety.

5.3 Research Implications

5.3.1 Limitations

One of the limitations in this intervention study is the lack of random assignment. The groups were already formed and it was not possible to assign participants to the treatment conditions randomly. In this sense, this intervention study is subject to biases inherent to quasi-experimental designs. These biases are mostly associated with selection (Christensen et al., 2011). Thus, although each group was randomly assigned to the treatment condition and the pre-test scores as well as the in-class listening tests were monitored and examined carefully, selection biases such as selection-maturation and selection-regression could have interacted differentially in the groups. For instance, it might be that participants in one or two of the treatment groups got tired and/or bored more easily than participants in the other groups. It is also possible that participants in one or two of the treatment groups were extremely different in terms of personal characteristics such as age and/or aptitude from those in the other groups.

Another limitation in this study has to do with the unequal number of participants in the groups. In this study, the main analysis was conducted with 32 participants. This sample size seems to be an adequate number for the factorial repeated measures ANOVA performed in this study. G*Power calculations yielded a sample size of 33 participants with an actual power of 75% (Faul et al., 2007). Nevertheless, the unequal number of participants in each group is problematic, especially because the Conventional Approach group is the one with the fewest participants (i.e., eight participants). Also, participant attrition at delayed post-test led to the

elimination of one treatment group (Awareness-Raising). Thus, long-term effects were assessed with two groups only (Metacognitive Pedagogical Sequence and Conventional Approach).

Vocabulary knowledge plays an important role in L2 listening (Mecartty, 2000; Staehr, 2009). In this study, vocabulary knowledge was not measured. This poses a limitation to the present study, especially because Wang and Treffers-Daller (2017) found that vocabulary knowledge is a more important predictor of listening comprehension than metacognitive awareness. Thus, although this study was carefully conducted and listening ability was statistically controlled, vocabulary knowledge might have affected the groups differentially. That is, it may well be that one or two of the groups had a broader knowledge of vocabulary than the others.

Also, although eight listening sessions in this study were enough to observe significant changes in L2 listening, more sessions might have yielded even more substantial changes in L2 listening. This may be so particularly for the Awareness-Raising group, whose improvement in relation to the Conventional Approach group was substantial with a medium effect size (i.e., r = .39), but did not score significantly higher in relation to the Conventional Approach group. Participants in Vandergrift and Tafaghodtari's (2010) study reported that they started to become bored when they reached thirteen sessions. Thus, an appropriate number of listening sessions would have fallen between eight and thirteen. That is, ten or eleven listening sessions might have been a more appropriate number of interventions.

Another limitation in this study has to do with the teacher variable. Although the interventions in the three groups were conducted by the same researcher, the three groups were taught by three different teachers. An ideal situation would have shown the interventions carried out by the same researcher and the three groups being taught by the same teacher. While this

ideal situation might be unlikely to occur, it is important to highlight that the teacher variable plays an important role in language teaching (Cross & Vandergrift, 2015).

5.3.2 Future Research

This study has identified some venues for future research. These new directions emerged from the findings in this intervention study and might contribute to a better understanding of how instruction in L2 listening can be more effective in developing L2 listening skills.

In this study, the prompts from the Metacognitive Awareness Listening Questionnaire (MALQ) were used as discussion questions in the Awareness-Raising group to systematically raise learners' awareness of factors such as problem solving, planning and evaluation, mental translation, person knowledge, and direct attention. Future studies need to look at different ways of incorporating these factors into listening activities that foster the development of L2 listening skills. Perhaps, these factors can be incorporated into a checklist that learners can use as a way of raising their awareness of the factors that affect L2 listening. In turn, incorporating such a checklist in the listening activities can help learners monitor their use of strategies. Thus, it might be interesting to observe how a checklist of the MALQ factors contributes to the development of L2 listening. A study that compares the use of guided discussions of the MALQ factors to the use of a checklist of those MALQ factors would be worth conducting since language instructors might find using a checklist more appealing and practical.

Future studies also need to investigate how the audio transcription contributes to the development of L2 listening skills. In this study, participants' comments from the three groups suggested that the incorporation of the audio transcription may be beneficial. As mentioned

earlier, written support in the listening activities might not encourage the development of compensatory strategies necessary to cope with real-time listening (Vandergrift, 2004). In fact, the audio transcript might transform the listening activity into a reading activity. However, incorporating the audio transcription at the end of the listening activity may contribute positively to the development of listening skills since learners would have already had a chance to practice their top-down strategies. Thus, an intervention study looking at how learners perform when given the audio transcription will certainly provide insights into whether or not the audio transcription contributes positively to the development of L2 listening skills.

Also, metacognitive instruction in L2 listening in this study was implemented using an implicit metacognitive cycle (Vandergrift, 2004) (i.e., Metacognitive Pedagogical Sequence group) and a more explicit approach that aimed at raising metacognitive awareness (i.e., Awareness-Raising group). Although no substantial and/or significant differences were found between these two groups, future studies need to investigate the interaction between implicit metacognitive instruction and explicit metacognitive instruction in L2 listening. Vandergrift and Tafaghodtari (2010) associated the effectiveness of the metacognitive cycle with its implicit nature. However, in the present study, both groups (Metacognitive Pedagogical Sequence & Awareness-Raising) performed similarly. Thus, it may well be that the effectiveness of the metacognitive cycle is not a consequence of its implicit nature. Therefore, a potentially interesting study could examine the differential effects of implicit and explicit metacognitive instruction, as well as a treatment that combines both approaches. Such a study might help to understand specific ways in which metacognitive instruction contributes to the development of L2 listening (Goh, 2008).

Participants in the Conventional Approach group consistently reported to have benefited from the vocabulary introduced at the beginning of the listening activity. These participants felt that vocabulary was an important learning outcome of the listening activities. However, these participants did not show improvement in L2 listening. It may seem that the feature of vocabulary in this treatment group might have diverted these learners' attention from building an understanding of the listening text. Therefore, because vocabulary knowledge plays an important role in L2 listening (Rost, 2005), future studies need to look at the interaction of explaining vocabulary as a pre-listening activity and as a post-listening activity. This is particularly important because it may occur that explaining vocabulary functions better as a post-listening activity. Thus, it might be worth conducting an intervention study that compares the metacognitive pedagogical cycle (Vandergrift, 2004) with a treatment that combines some features of the comprehension approach such as vocabulary. For instance, a study that compares some or all of the following treatments: a metacognitive pedagogical cycle, a metacognitive pedagogical cycle in which vocabulary is explained as a pre-listening activity, a metacognitive pedagogical cycle in which vocabulary is explained as a post-listening activity, a comprehension approach in which vocabulary is explained as a pre-listening activity, and a comprehension approach in which vocabulary is explained as a post-listening activity.

Also, it might be worth conducting an intervention study in which a comprehension approach is implemented in a more student-centered manner as Field (2012) noted that L2 listening needs to move away from its current teacher-centered format. Thus, it might be interesting to conduct a study that compares a student-centered version of the comprehension approach with a traditional teacher-centered comprehension approach. For instance, in the student-centered version, students could explain vocabulary and formulate general questions

about the audio text in small groups. This might help learners take a more active role in the listening activity.

Finally, in this study, low-intermediate learners of Spanish were able to improve their L2 listening significantly after receiving metacognitive instruction. Future studies need to look at how learners at different levels of language proficiency react to metacognitive instruction in L2 listening. That is, a study comparing L2 listening instructional approaches with beginner learners and/or advance learners would certainly provide more insights on how to teach L2 listening. Also, future studies may need to look at how metacognitive instruction in L2 listening prepares language learners for real-life listening. Although Field (2012) argued that "metacognitive training has the potential benefits beyond the classroom. Enabling learners to handle listening exercises strategically boosts their confidence when they come to real-life listening encounters" (p. 211), future studies need to look at how this type of listening instruction impacts L2 listening outside of the classroom.

5.4 Summary and Conclusion

This chapter has discussed the results of this intervention study in relation to the research questions that it sought to address, as well as in relation to previous studies about metacognitive instruction in L2 listening. In addition, the chapter also discussed pedagogical implications, limitations, and future research. The findings strongly support an approach to L2 listening instruction that is student-centered and focuses on the process of L2 listening. The findings in this study indicate that L2 listening instruction is more beneficial when learners are given opportunities to explore aspects of their personalities, strategies, the demands and nature of L2

listening, and the benefits of collaborative learning. The findings also indicate that guided discussions and reflections are an effective and practical way to teach L2 listening and help learners develop their listening skills. All in all, despite the limitations discussed above, the results of this study constitute a meaningful contribution to the field of L2 listening instruction.

Regarding future research, the chapter identified some possibilities for future studies such as comparing L2 listening instructional approaches that include the audio transcription in the listening activities, comparing L2 listening instructional approaches that combine guided discussions and checklists of the MALQ factors, and comparing L2 listening instructional approaches that incorporate vocabulary as a pre-listening and post-listening activity.

The next and final chapter of this thesis provides a summary of this intervention study and of its results. The chapter finishes with some concluding remarks that highlight the study's implication for research and pedagogy.

6 Conclusion

6.1 Summary of the Study

This intervention study was conducted to assess the effects of three instructional approaches in L2 listening; that is, a metacognitive pedagogical cycle, an awareness-raising approach, and a conventional approach. The metacognitive cycle was implemented in the Metacognitive Pedagogical Sequence group and was a shorter version of the metacognitive cycle proposed by Vandergrift (2004). This metacognitive cycle guides learners through stages such a prediction, verification, problem-solving, and reflection. Learners exposed to the Awareness-Raising approach engaged in discussing strategies associated with factors such as problem solving, planning and evaluation, mental translation, person knowledge, and direct attention (i.e., MALQ factors). Learners in these two treatment groups were asked to write a reflection about their listening experience at the end of each listening session. These two approaches aimed at developing metacognitive knowledge and regulatory skills; however, the metacognitive pedagogical cycle is an implicit instructional approach to L2 listening (Vandergrift & Tafaghodtari, 201210) whereas the Awareness-Raising approach was conducted in a more explicit manner. The approach implemented in the Conventional Approach group was a shorter version of what Field (2012) called the comprehension approach. In this group, vocabulary was introduced and explained as a pre-listening activity and comprehension questions were asked during the activity to build a better understanding of the audio text.

This study responds to concerns associated with the lack of intervention studies in L2 listening instruction, especially studies that compare instructional approaches in L2 listening

(Berne, 2004; Macaro, et al., 2007). To address these concerns, the three instructional approaches described above were implemented in three intact classes of intermediate Spanish. Thus, this intervention study constitutes a quasi-experimental design in which random assignment of participants to the different treatment conditions was not possible, but each instructional approach (i.e., each treatment condition) was assigned randomly to each of the intact classes. During the study, eight listening lessons designed specifically for each approach and that lasted for approximately 25 minutes were carried out in each treatment condition.

The research questions that guided this study were concerned with whether or not there are significant differences in L2 listening ability that could be attributed to the different instructional approaches. That is, whether or not there are significant differences within each group from pre-test to post-test (RQ-1a), whether or not there are significant differences between the groups at post-test (RQ-1b), and whether or not there are long-term effects (RQ-2). In addition, the study also examined the participants' opinions about the listening activities implemented in their classes (RQ-3). Thus, in order to answer research questions 1 and 2, the L2 listening ability of the learners was measured before the treatment conditions were applied at pre-test and after the last intervention session at post-test. L2 listening ability was also measured five months after the last intervention session at delayed post-test. Furthermore, to address research question 3, the participants' perceptions about the instructional treatments were collected from the three groups on week seven during the intervention and on week 10 after the intervention. Also, participants' perceptions were collected during focus group interviews after the post-test. To analyze these data, two factorial repeated measures ANOVA's and one analysis of covariance (ANCOVA) were conducted. Participants' opinions about the listening activities were coded, labeled and analysed.

6.2 Summary of the Results

Participants in the Metacognitive Pedagogical Sequence group and the Awareness-Raising group exhibited a significant improvement in listening performance from pre-test to post-test. Also, the Metacognitive Pedagogical Sequence group outperformed the Conventional Approach group at post-test after statistically controlling for initial differences in listening ability. These results are attributed to the clustering of cognitive and metacognitive strategies as well as to the student-centered nature of these two instructional approaches to L2 listening (e.g., Graham & Macaro, 2008; Vandergrift & Tafaghodtari, 2010).

Also, evidence of long-term effects was found. The significant improvement that the Metacognitive Pedagogical Sequence group experienced during the intervention seemed to have remained five months after the last intervention session. Although assessing long-term effects in this study was affected by participant attrition, these results are consistent with the results in Graham and Macaro's (2008) study regarding L2 listening strategy instruction. These results provide preliminary evidence of long-term effects that can be attributed to the metacognitive pedagogical cycle (Vandergrift, 2004) implemented in this study.

Participants in the Conventional Approach group also showed improvement; however, this improvement was not significant. While the listening activities implemented in this study were positively perceived by the participants in the three groups, participants felt they benefited in different ways. For instance, participants' comments in Conventional Approach group were mostly centered around vocabulary. It seems that vocabulary was the most salient learning outcome in this group since participants consistently reported to have benefited from this feature.

Participants' comments from the Metacognitive Pedagogical Sequence group and the Awareness-Raising group exhibited a more strategic behaviour. It may seem that these participants developed a sense of approaching L2 listening more strategically. Additionally, participants in the Metacognitive Pedagogical Sequence group and the Awareness-Raising group explicitly reported to feel less anxious, more confident, and more organized when doing the listening section in the course midterm exam. Participants in the Conventional Approach group did not report to have experienced more or less anxiety, confidence and/or motivation. These participants mostly reported to have benefited from the vocabulary introduced at the beginning of the activity as well as from the exposure to the audio material.

6.3 Concluding Remarks

This study makes significant contributions to the field of second language teaching and learning. It shows that developing metacognitive knowledge is important in language learning, particularly in teaching learners how to listen in the second language. The significant results in this study underscore the importance and usefulness of developing person knowledge, task knowledge, strategy knowledge and strategy use in teaching listening skills. This study provides further empirical support to the benefits of focusing on the process of L2 listening, and not just on the outcome (e.g., Graham & Macaro, 2008; Vandergrift & Tafaghodtari, 2010). Teaching learners how to listen in the second language is attainable when instruction is conceived as a learning process in which learners are encouraged to interact with their peers, make choices, reflect, and assess their learning. In this sense, L2 listening instruction is more effective and useful to learners when the learning process is centered on the learners.

This study also provides empirical support to the effectiveness and usefulness of combining process-based discussions and reflections when developing L2 listening skills (Goh & Taib, 2006). Providing learners with opportunities to discuss aspects such as problem solving, planning and evaluation, mental translation, person knowledge, and direct attention (MALQ factors) as well as providing opportunities for learner reflection impacted L2 listening positively and significantly in this study. Moreover, this seems to be a very practical way to help learners improve their listening skills in the classroom, particularly because discussing the MALQ factors and reflecting about the L2 listening experience can be easily incorporated in traditional listening activities (Graham, 2017).

Providing the topic of the listening text at the beginning of the activity was beneficial to the learners in this study. This helped them visualize the context of the listening text. Also, while exposure to authentic audio material can cause discomfort and distress associated with the different accents and speed, as it was the case in this study, exposing L2 listeners to authentic audio material is important and useful when teaching how to listen. This can make learners become aware of the challenges and demands of L2 listening. Another important aspect in this study is the fact that introducing vocabulary as a pre-listening activity might cause learners to focus their attention on individual words and not on building an understanding of the audio text. In this regard, Graham (2017) noted that "Teachers in England also seem to direct learners' attention more to individual items of vocabulary than to broader themes that might rise in a listening passage, possibly because they see it as a surer way of helping learners obtain 'correct answers' to accompanying comprehension questions" (p. 114). Thus, while providing the topic of the listening text and exposing learners to authentic audio material appear to be important aspects in teaching listening skills, introducing vocabulary should be carefully considered when

teaching learners how to listen since this might divert learners' attention from building an understanding to the audio text.

Finally, while exposure to oral input and listening practice is important, exposure and practice is not enough to help learners develop their listening skills. In this study, the learners that experienced significant improvement in listening performance were also those who reported to feel more confident and more motivated. This indicates that listening instruction is indeed beneficial to second language learners. However, in order for instruction in L2 listening to be beneficial to learners, it needs to be framed within a learning process in which learners are encouraged to discuss aspects associated with L2 listening and reflect on their listening development. Also, instruction in L2 listening needs to foster the development of person knowledge, task knowledge, strategy knowledge and strategy use. Metacognitive control over strategies is important to be able to address L2 listening more strategically. Moreover, listening instruction in L2 listening needs to consider cognitive and metacognitive strategies together. Teaching individual and isolated strategies does not appear to impact L2 listening as substantially as teaching clusters of cognitive and metacognitive strategies (Graham & Macaro, 2008).

While this study provides support to the short and long-term benefits of the metacognitive cycle proposed by Vandergrift (2004), it also indicates that L2 listening instruction is more effective when it is carried out systematically in the use of cognitive and metacognitive strategies (Thompson & Rubin, 1996). In this study, eight listening lessons were implemented regularly once a week for approximately 25 minutes each. This indicates that teaching learners how to listen does not take much time. It only needs to be systematic, process-oriented, and student-centered.

To conclude, the study reported in this thesis has demonstrated the importance of metacognition in the development of L2 listening skills. In this regard, it has highlighted the need for a systematic approach to L2 listening instruction that focuses on the listening process rather than on its outcomes, and that ultimately enables learners to address L2 listening more strategically. Although the findings of this thesis need to be considered within the context of this research, including its limitations, they make a meaningful contribution to the body of research regarding L2 listening instruction. At the same time, these findings underscore the need for future research that explores the benefits of both explicit and implicit instructional approaches to L2 listening, and that ultimately might provide valuable insights for the day-to-day practices in second language classrooms.

References

Anderson, J. R. (1995). *Cognitive Psychology and its Implications* (4th ed.). New York: Freeman.

Anderson, N. (2005). Learning Strategies. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 757-771). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Aron, A., Aron, E., and Coups, E. (2005*). Statistics for the Behavioral and Social Sciences: a brief course*, (3th ed.). Upper Saddle River, N.J.: Pearson/Prentice Hall.

Arnold, J. (2000). Seeing through listening comprehension exam anxiety. *TESOL Quarterly, 34*, 777-786.

Asher, J. (1977). *Learning Another Language Through Actions: the complete teacher's guidebook.* Los Gatos, CA: Sky Oaks Productions.

Barsties, B. & De Bolt, M. (2015). Assessment of Voice Quality: Current State-of-the-art. *Auris Nasus Larynx*, *42*, 183-188.

Basnakova, J., Weber, K., Petersson, K., Van Berkum, J. & Hagoort, P. (2013). Beyond the language given: The neural correlates of inferring speaker meaning. *Cerebral Cortex*, *24*, 2572-2578.

Benson, Ph. (2012). Learner-Centered Teaching. In Anne Burns & Jack C. Richards (Eds.), *The Cambridge Guide to Pedagogy and Practice in Second Language Teaching* (pp. 30-37). Cambridge: Cambridge University Press.

Berne, J. E. (2004). Listening Comprehension Strategies: A Review of the Literature. *Foreign Language Annals*, *37*(4), 521-533.

Best, C. (1995). A Direct Realist View of Cross-Language Speech Perception. In W. Strange

(Ed.), *Speech Perception and linguistic experience: Theoretical and methodological issues.* Timonium, MD: York Press.

Best, J. W. & Kahn, J. V. (2006). *Research in Education* (3rd ed.). Boston, M.A.: Pearson Education.

Borkowski, J. G. (1996). Metacognition: Theory or Chapter Heading? *Learning and Individual Differences, 8*, 391-402.

Broersma, M. & Cutler, A. (2008). Phantom word activation in L2. *System, 36*, 22-34.

Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L.F. & Palmer, A. (1996). *Language testing in practice: designing and developing useful language tests*. Oxford: Oxford University Press.

Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*, 1-47.

Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards & R. W. Schmidt (Eds.). *Language and communication* (pp. 2-27). London: Longman.

Celce-Murcia, M., Dornyei, Z. and Thurrell, S. (1995). A pedagogical framework for communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics, 6* (2), 5-35.

Christensen, L., Johnson, R. B., & Turner, L. A. (2011). *Research Methods, Design, and Analysis* (11th ed.). Boston, M.A.: Pearson Education.

Cross, J. (2011). Metacognitive Instruction for Helping Less-Skilled Listeners. *ELT Journal*, *65*(4), 408-416.

Cross, J. (2015). Metacognition in L2 Listening: Clarifying Instructional Theory and Practice.

TESOL Quarterly, 49(4), 883-892.

Cross, J., & Vandergrift, L. (2015). Guidelines for Designing and Conducting L2 Listening Studies. *ETL Journal, 69*(1), 86-89.

Curran, C. (1972). *Counseling-Learning: a whole-person model for education.* New York: Grune and Stratton.

Cutler, A. (2001). Listening to a Second Language through the Ears of a First. *Interpreting*, *5*, 1 - 23.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. Doughty, & M. Long (Eds.), *Handbook of second language acquisition* (pp. 589–630). Oxford, UK: Blackwell.

Fahim, M., & Fakhri, A. (2014). Maximizing learners' metacognitive awareness in listening through metacognitive instruction: An empirical study. *International Journal of Research Studies in Education, 3*(3), 79–91.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Field, A. (2009). *Discovering Statistics Using SPSS* (3rd ed.). London: SAGE.

Field, J. (1998). Skills and strategies: Towards a new methodology for listening. *ELT Journal, 52*(2), 110-118.

Field, J. (2008). *Listening in the Language Classroom.* Cambridge: Cambridge University Press.

Field, J. (2012). Listening Instruction. In Anne Burns & Jack C. Richards (Eds.), *The Cambridge Guide to Pedagogy and Practice in Second Language Teaching* (pp. 207-217). Cambridge: Cambridge University Press.

Flavell, J. (1976). Metacognition aspects of problem solving. In L. B. Resnick (Ed.), *The Nature of Intelligence* (pp. 231-235). Hillsdale, NJ: Erlbum.

Feyton, C. (1991). The power of listening ability: an overlooked dimension in language acquisition. *The Modern Language Journal*, *75*, 173-180.

Gamst, G., Meyers, L. S., & Guarino, A. J. (2008). *Analysis of Variance Designs: A Conceptual and Computational Approach with SPSS and SAS*. NY: Cambridge University Press.

Gattegno, C. (1972). *Teaching Foreign Languages in School: The silent Way* (2nd ed.). New York: Educational Solutions.

Goh, C. (1997). Metacognitive Awareness and Second Language Listeners. *ELT Journal*, *51*, 361-369.

Goh, C. (2000). A cognitive perspective on language learners' listening comprehension problems. *System, 28*, 55-75.

Goh, C., & Taib, Y. (2006). Metacognitive instruction in listening for young learners. *ELT Journal, 60*, 222–232.

Goh, C. (2008). Metacognitive Instruction for Second Language Listening Development: Theory, Practice and Research Implications. RELC Journal: *A Journal of Language Teaching and Research, 39*(2), 188-213.

Goh, C., & Hu, G. (2014). Exploring the relationship between metacognitive awareness and listening performance with questionnaire data. *Language Awareness, 23*(3), 255-274.

Graham, S. (2006). Listening Comprehension: The learner's perspective. *System, 34*, 165-182.

Graham, S. & Macaro, E. (2008). Strategy Instruction in Listening for Lower-Intermediate Learners of French. *Language Learning*, *58*(4), 747-783.

Graham, S., Santos, D. & Francis-Brophy, E. (2014). Teacher beliefs about listening in a foreign

language. *Teaching and Teacher Education, 40*, 44-60.

Graham, S. (2017). Research into Practice: Listening Strategies in an Instructed Classroom Setting. *Language Teaching, 50*(1), 107-119.

Griffiths, C. (2013). *The strategy factor in successful language learning.* Bristol, UK: Multilingual Matters.

Griffiths, C., & Oxford, R. (2014). The twenty-first century landscape of language learning strategies: introduction to this special issue. *System*, *43*, 1-10.

Guellai, B., Langus, A., & Nespor, M. (2014). Prosody in the hands of the speaker. *Frontiers in Psychology*, 5: 700.

Hasan, A. (2000). Learners' Perceptions of Listening Comprehension Problems. *Language, Culture & Curriculum, 13*(2), 137-153.

Hymes, D. H. (1972). On Communicative Competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected Readings* (pp. 269-293). Great Britain: Penguin Education.

Keppel, G., & Wickens T. (2004). *Design and Analysis: A research's handbook* (4th ed.). New Jersey: Pearson Education.

Kohler, B. (2002). *The Effects of Metacognitive Language Learning Strategy on Lower-Achieving Second Language Learners.* (Doctoral Dissertation). Available from ProQuest Dissertation and Theses database. (UMI number: 3053132)

Kojic-Sabo, I., & Lightbown, M. P. (1999). Students' approaches to Vocabulary Learning and their Relationship to Success. *Modern Language Journal*, *83*, 176-192.

Krashen, S., & Terrell, T. (1983). *The natural approach: language acquisition in the classroom.* Oxford: Pergamon.

Kuhl, P. (2000). A New View of Language Acquisition. *Proceedings from the National Academy*

of Science of the United States, *97*, 11850-11857.

Lamboy, E. M. (2008). Me cuesta entender el español de mis amigos caribeños. ¿Por qué será? In Jennifer D. Ewald & Anne Edstrom (Eds.), *El español a través de la lingüística: Preguntas y respuestas* (pp. 153-165). Somerville, MA: Cascadilla Press.

Littlewood, W. (1981). *Communicative Language Teaching*. Cambridge: Cambridge University Press.

Livingston, S. A. (2014). Equating Test Scores without IRT. (2nd ed.). Princeton, NJ: Educational Testing Services.

Macaro, E., Graham, S., & Vanderplank, R. (2007). A review of listening strategies: focus on sources of knowledge and on success. In A. D. Cohen & E. Macaro (Eds.), *Language learner strategies: thirty years of research and practice* (pp. 165-185). New York: Oxford University Press.

Macaro, E., Vanderplank, R., & Graham, S. (2005). *A Systematic Review of the Role of Prior Knowledge in Unintentional Listening Comprehension*. London: EPPI-Center, Social Science Research Unit, Institute of Education, University of London.

Mareschal, C. (2007). *Student Perceptions of a Self-Regulatory Approach to Second Language Listening Comprehension development* (Doctoral Dissertation). Available from ProQuest Dissertation and Theses database. University of Ottawa.

Martin-Peris, E. & Sans-Baulenas, N. (2014). Gente Hoy 2. Barcelona: Difucion.

McGowan, K. (2015). Social Expectation Improves Speech Perception in Noise. *Language and Speech*, *58*(4), 502-521.

Mecartty, F. (2000). Lexical and grammatical knowledge in reading and listening comprehension by foreign language learners of Spanish. *Applied Language Learning*, *11*, 323-348.

Mendelsohn, D. (1994). *Learning to listening*. San Diego, CA: Dominie Press.

Mills, N., Pajares, C., & Herron, C. (2006). A re-evaluation of the role of anxiety: Self-efficacy, anxiety and their relationship to reading and listening proficiency. *Foreign Language Annals, 39*, 276-295.

Naiman, N., Frohlich, M., Stern, H., & Todesco, A. (1978). *The good language learner.* Toronto: Ontario Institute for Studies in Education.

Nunan, D. (2002). Listening in Language Learning. In J. Richards & W. Renandya (Eds.), *Methodology in Language Teaching: An anthology of current practice* (pp. 238-241). Cambridge: Cambridge University Press.

O'Malley, J., Chamot, A., Stewner-Manzanares, G., Kupper, L., & Russo, R. (1985). Learning Strategies used by beginning and intermediate ESL students. *Language Learning, 35*(1), 21-46.

O'Malley, J., Chamot, A., & Küpper, L. (1989). Listening Comprehension Strategies in Second Language Acquisition (English). *Applied Linguistics, 10*(4), 418-437.

O'Malley, J., & Chamot, A. (1990). *Learning Strategies in Second Language Acquisition*. Cambridge: Cambridge University Press.

Oxford, R. (1990). *Language learning strategies: what every teacher should know*. New York: Newbury House Publisher.

Oxford, R., Rubin, J., Chamot, A., Schramm, K., Lavine, R., Gunning, P., and Nel, C. (2014). The learning strategy prism: Perspectives of learning strategy experts. *System, 43*, 30-49.

Pintrich, P. R. (2000). The Role of Goal Orientation in Self-Regulated Learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451-502). San Diego, CA: Academic Press.

Plonsky, L. (2011). The Effectiveness of Second Language Strategy Instruction: A Meta-analysis. *Language Learning*, *61*(4), 993-1038.

Prator, C., & Celce-Murcia, M. (1979). An outline of language teaching approaches. In M. Celce-Murcia & L. McIntosh (Eds.), *Teaching English as a second or foreign language* (pp. 3-5). New York: Newbury House.

Rahimirad, M. (2014). The impact of metacognitive strategy instruction on the listening performance of university students. *Procedia: Social and Behavioural Science*, *98*, 1485-1491.

Reichardt, C. (1979). The Statistical Analysis of Data from Non-equivalent Group Designs. In Thomas D. Cook & Donald T. Campbell, *Quasi-Experimentation: design & analysis issues for field settings* (pp. 147-205). Chicago: Rand Mcnally College Publishing Company.

Richards, J. C., & Rodgers, T. S. (2001). *Approaches and Methods in Language Teaching* (2nd ed.). New York: Cambridge University Press.

Rees-Miller, J. (1993). A Critical Appraisal of Learner Training: Theoretical bases and teaching implications. *TESOL Quarterly, 27*(4), 679-687

Rost, M. (2005). L2 Listening. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 503-527). Mahwah, NJ: ErlBaum.

Rost, M. (2016). *Teaching and Researching Listening* (3rd ed.). New York: Routledge.

Rubin, J. (1975). What the "Good Language Learner" can teach us. *TESOL Quarterly, 9*(1), 41 - 51.

Rubin, J. (1987). Learner Strategies: theoretical assumptions, research history and typology. In

A. Wenden & J. Rubin (Eds.), *learner strategies in language learning* (pp. 15-30). NJ: Prentice Hall International.

Rubin, J. (1990). Improving foreign language listening comprehension. In J. E. Alatis (Ed.), *Georgetown University Round Table on Languages and Linguistics* (pp. 209-316). Washington, DC: Georgetown University press.

Schraw, G., & Dennison, R. (1994). Assessing Metacognitive Awareness. *Contemporary Educational Psychology*, *19*, 460-475.

Schraw, G. (1998). Promoting General Metacognitive Awareness. *Instructional Science, 26*(1-2), 113-125.

Seigel, J. (2014). Exploring L2 listening instruction: examinations of practice. *ELT Journal, 68*(1), 22-30.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont, CA: Wadsworth Cengage Learning.

Spada, N. (2007). Communicative Language Teaching: Current Status and Future Prospects. In J. Cummins & C. Davison (Eds.), *International Handbook of English Language Teaching* (pp. 271-288). New York: Springer.

Staehr, L. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, *31*, 577-607.

Stern, H. (1975). What can we learn form the good language learner? *Canadian Modern Language Review, 31,* 304-318.

Stern, H. (1992). *Issues and options in language teaching*, (P. Allen & B. Harley, Eds.). Oxford: Oxford University Press.

Tabachnick, B., & Fidell, L. (2013). *Using Multivariate Statistics* (6th ed.). Boston: Pearson

Education.

Thomas, G. (2003). Conceptualisation, Development and Validation of an instrument for investigation the metacognitive orientation of science classroom learning environments: the metacognitive orientation learning environment scale-science (MOLES-S). *Learning Environment Research, 6,* 171-197.

Thomas, G., & Mee, D. K. (2005). Changing the learning environment to enhance students' metacognition in Hong Kong Primary School Classrooms. *Learning Environments Research, 8,* 221-243.

Thompson, I., & Rubin, J. (1996). Can strategy instruction improve listening comprehension? *Foreign Language Annals, 29*(3), 331-342.

Thorndike, R., & Thorndike-Christ, T. (2010*). Measurement and Evaluation in Psychology and Education* (8th ed.). Boston, MA: Pearson Education.

Tyler, M. (2001). Resource Consumption as a Function of Topic Knowledge in Nonnative and Native Comprehension. *Language Learning, 51*(2), 257-280.

Vandergrift, L. (1997). The Cinderella of Communication Strategies: Reception Strategies in Interactive Listening. *The Modern Language Journal, 81*(4). 494-505.

Vandergrift, L. (2003). Orchestrating strategy use: Towards a model of the skilled L2 listener. *Language Learning, 53*, 461–494.

Vandergrift, L. (2004). Listening to learn or learning to listen? *Annual Review of Applied Linguistics, 24,* 3-25.

Vandergrift, L., Goh, C., Mareschal, C., & Tafaghodtari, M. (2006). The Metacognitive Awareness Listening Questionnaire: Development and Validation. *Language Learning*, *56*(3), 431-462.

Vandergrift, L., & Goh, C. (2009). Teaching and testing listening comprehension. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching* (pp. 395-411). Malden, MA: Wiley-Blackwell.

Vandergrift, L., & Tafaghodtari, M. (2010). Teaching L2 learners how to listen does make a difference: An empirical study. *Language Learning*, *60*(2), 470-497.

Vandergrift, L. (2011). Second Language Listening: Presage, Process, Product, and Pedagogy. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning*, *2*(1), (pp. 455-471). Florence, US: Routledge.

Vandergrift, L., & Goh, C. (2012). *Teaching and Learning Second Language Listening: Metacognition in Action*. NY: Routledge.

Vandergrift, L., & Baker, S. (2015). Learner Variables in Second Language Listening Comprehension: An Exploratory Path Analysis. *Language Learning*, *65*(2), 390-416.

Vogely, A. (1999). Addressing listening comprehension anxiety. In D. J. Yong (Ed.), *Affect in Foreign Language and second language learning: a practical guide to creating a low-anxiety classroom atmosphere* (pp. 106-123). New York: McGraw-Hill.

Wang, Y. & Treffers-Daller, J. (2017). Explaining Listening Comprehension among L2 Learners of English: the contribution of general language proficiency, vocabulary knowledge and metacognitive awareness. *System, 65*, 139-150.

Wang, W. (2016). Learning to Listen: The Impact of a Metacognitive Approach to Listening Instruction. *Asia-Pacific Education Researcher (Springer Science & Business Media B.V.)*, *25*(1), 79-88.

Wenden, A. (1987). Metacognition: An Expanded View of the Cognitive Abilities of L2 Learners. *Language Learning*, *37*(4), 573-594.

Wenden, A. (1998). Metacognitive Knowledge and Language Learning. *Applied Linguistics, 19*(4)*,* 515-537.

White, G. (2006). Teaching listening: Time for a change in methodology. In E. Uso-Juan & A. Martinez-Flor (Eds.), *Current trends in the development and teaching of the four language skills* (pp. 111–135). Berlin: Mouton de Gruyter.

Wilson VanVoorhis, C. and Morgan, B. (2007). Understanding Power and Rules of Thumb for Determining Sample Sizes. *Tutorials in Quantitative Methods for Psychology, 3*(2), 43-50.

Zentella, A. (1997). *Growing up Bilingual: Puerto Rican Children in New York*. Malden, MA: Blackwell Publishers.

Appendix A

Checklist: Metacognitive Pedagogical Sequence group

Date_____ Started at_____ Ended at_____

1. Distribute the Listening handout.
2. Provide the topic of the listening text.
3. Ask to fill in the 'prediction' column in the Listening handout.
4. Play the audio for the first time. Ask to fill in the 'listening' column in the Listening handout.
5. Ask students to pair up to discuss issues associated with the audio input, and ask to fill in the 'peer-discussion' column in the Listening handout.
6. Distribute the in-class listening tests.
7. Play the audio for the second time.
8. Ask students to fill in the 'reflection' section in the Listening handout.
9. Collect handouts and in-class listening tests.

Comments

_____

_____

_____

_____

Checklist: Awareness-Raising group

Date_____ Started at_____ Ended at_____

1. Introduce the metacognitive factor and distribute the handout with the discussion questions.
2. Ask students to discuss the questions in the handout in small group.
3. Provide the topic of the listening text.
4. Play the audio for the first time.
5. Distribute the in-class listening tests.
6. Play the audio for the second time.
7. Ask students to fill in the ‘reflection’ section in the handout.
8. Collect handouts and in-class listening tests.

Comments

_____

_____

_____

_____

_____

Checklist: Conventional Approach group

Date_____ Started at_____ Ended at_____

1. Provide the topic of the listening text.
2. Explain vocabulary.
3. Play the audio for the first time.
4. Ask general question about the listening text.
5. Distribute the in-class listening tests.
6. Play the audio for the second time.
7. Collect in-class listening tests.

Comments

_____

_____

_____

_____

_____

Appendix B

Listening Worksheet: Metacognitive Pedagogical Sequence group

Nombre_____

Prediction	Listening	Peer-discussion

Reflection:

Will you do anything differently the next time you listen to a text? If so, please explain what will you do? If there is nothing you will do differently, please describe what you do when listening.

_____

_____

_____

_____

Appendix C

In-Class Listening Tests

Nombre_____ **# 1**

Situación: Luis y Silvia han quedado para conocerse. Después de la cita, Luis le cuenta a una amiga como es Silvia.

- ✓ **Escucha la conversación y selecciona la respuesta correcta.**

1. ¿Qué le encanta a Silvia?

   a) Los motores
   b) La naturaleza
   c) La ciudad

2. ¿Qué no le gusta a Silvia?

   a) Los animales
   b) El ruido
   c) Las bicis

3. ¿Qué le apasiona a Silvia?

   a) El baile
   b) El baile y la música
   c) La música

4. ¿Qué instrumento toca Silvia?

   a) El violín
   b) El piano
   c) No toca ningún instrumento

5. ¿Con qué frecuencia va Silvia a conciertos de música clásica?

   a) Nunca
   b) Dos veces al mes.
   c) Seis veces al mes.

Nombre_____ **# 2**

Situación: Escucha lo que haría Carla en las siguientes situaciones imaginarias / hipotéticas.

- ✓ **Escucha la conversación y selecciona la respuesta correcta.**

1. En la situación imaginaria, el trabajo es…

    a) Bien pagado en otro país.
    b) Mal pagado en otro país.
    c) Bien pagado en el mismo país.

2. En la situación imaginaria, ¿En dónde encontraría Carla el teléfono móvil?

    a) En un ascensor.
    b) En un bar.
    c) En el ascensor de un bar.

3. ¿Aceptaría Carla un papel en una película?

    a) Sí, porque le pagarían muy bien.
    b) No, porque le daría vergüenza.
    c) No, porque sería en otro país.

4. En la situación imaginaria, si Carla se encuentra con su actor favorito…

    a) Carla le pediría un autógrafo.
    b) Carla le pediría una foto.
    c) Carla le pediría un autógrafo y una foto.

5. Si un desconocido le escribe a Carla…

    a) Carla le pediría una foto.
    b) Carla le preguntaría quien es.
    c) Carla no le escribiría.

Nombre_____ **# 3**

Situación: 3 personas describen lo que hacen los domingos.

- ✓ **Escucha lo que dicen y selecciona la respuesta correcta.**

Persona 1

1. Se levanta…

   a) Muy temprano y sale a pasear.
   b) Muy tarde y desayuna tranquilamente.
   c) Muy temprano y desayuna tranquilamente.

2. Le gusta…

   a) Charlar con los chicos.
   b) Comer a la 6 de la tarde.
   c) Desayunar con su chico.

Persona 2

3. ¿Qué hace los domingos?

   a) Lava las ventanas para que entre el sol y caliente su sofá.
   b) Limpia su sofá antes de leer algún libro pendiente.
   c) Compra el periódico muy temprano y sale a pasear.

4. ¿Qué es más probable que NO le guste hacer?

   a) Pintar y hacer cosas manuales.
   b) Comprar y leer el periódico en la tarde.
   c) Salir a pasear y hacer algo pendiente.

Persona 3

5. ¿Qué le gusta hacer los domingos?

   a) Desayunar tranquilamente y tomar el sol.
   b) Ir al cine por la tarde y comer.
   c) Levantarse tranquilamente y salir a tomar algo en el cine.

Nombre_____ **# 4**

Situación: Unos amigos están conversando acerca de que podrían comer.

- ✓ **Escucha lo que dicen y selecciona la respuesta correcta.**

1) ¿Qué quiere decir “Comer de pie”?

   a) Comer mientras paseas.
   b) Comer rápido.
   c) Comer parados.

2) ¿Qué hay en la zona del mercado?

   a) Muchos bares.
   b) Restaurantes Gallegos.
   c) Embutidos y queso Gallego.

3) ¿Qué restaurante está cerca de la catedral?

   a) Azador Gozabar
   b) Tarala
   c) Abastos 2.0

4) ¿Qué venden en Azador Gozabar?

   a) Carne Gallega
   b) Tapas degustación
   c) Comida Gallega

5) ¿Qué es común en el mercado de Abastos?

   a) Te preparan un chuletón de ternera
   b) Te preparan lo que compras
   c) Te preparan un pan con embutidos y queso

Nombre_____ **# 5**

Situación: Una reportera de noticias está dando detalles de un acto de corrupción.

- ✓ **Escucha lo que dice y selecciona la respuesta correcta.**

1) ¿Dónde encontró la policía la maleta misteriosa?

   a) En el Parque del Retiro.
   b) En un almacén de la ciudad.
   c) En la casa del diputado.

2) ¿A quién encontraron atado a un árbol?

   a) Al diputado.
   b) Al mafioso.
   c) Al mago.

3) ¿Qué se ha encontrado en el ordenador?

   a) Correos electrónicos.
   b) Fotografías de Eduardo Barril.
   c) Dinero electrónico.

4) ¿Qué pensaba llevarse Max Abra?

   a) A la señora Barril.
   b) Dinero.
   c) Billetes de avión.

5) ¿Por qué cree la policía que Max Abra y la señora Barril planeaban fugarse juntos?

   a) Por las fotografías de Max Abra y la señora Barril.
   b) Por las declaraciones de Eduardo Barril.
   c) Por los billetes de avión.

Nombre_____ **# 6**

Situación: Una persona está recordando los detalles de un acto de corrupción.

- ✓ **Escucha lo que dice y selecciona la respuesta correcta.**

1) ¿Cómo llegó al barrio de Salamanca?

   a) En taxi.
   b) En tren.
   c) En bus.

2) ¿Qué le entregó el diputado?

   a) Una maleta llena de dinero.
   b) Ropa de verano.
   c) Un sobre con dinero.

3) ¿A qué hora se reunió con Eduardo?

   a) Un poco después de las nueve.
   b) A las diez y cuarto.
   c) A las dos de la tarde.

4) ¿Dónde recogió la segunda parte del dinero?

   a) En la peluquería.
   b) En el restaurante.
   c) En el aeropuerto.

5) ¿Cuánto de dinero tenía en total?

   a) Medio millón de euros.
   b) Un millo de euros.
   c) Dos millones de euros.

Nombre_____ **# 7**

Situación: Una persona está hablando de trastornos alimenticios y la obsesión por los músculos.

- ✓ **Escucha lo que dice y selecciona la respuesta correcta.**

1) ¿Cómo se llama la obsesión por la buena alimentación?

   a) Dieting.
   b) Ortorexia.
   c) Anorexia.

2) ¿Qué es la vigorexia?

   a) Obsesión por el ejercicio físico con dietas desequilibradas.
   b) Obsesión por comer comida libre de toxinas.
   c) Obsesión por hacer dieta principalmente cuando el verano está próximo.

3) ¿Cómo sabes si tienes un problema de estos?

   a) Porque tu dieta o ejercicio físico afectan a tus amigos y familia.
   b) Porque tu autoestima es baja.
   c) Porque tu dieta o ejercicio físico te causan problemas.

4) ¿Qué es recomendable hacer?

   a) Cambiar la dieta y hacer mucho ejercicio.
   b) Buscar la opinión de tus amigos y familia.
   c) Dejar de comer lo que te gusta y hacer una dieta estricta.

5) ¿Por qué aparecen problemas como la vigorexia y el dieting?

   a) Por la obsesión de tener un cuerpo perfecto.
   b) Por desórdenes psicológicos.
   c) Por influencia de tus amigos y familiares.

Nombre_____ **# 8**

Situación: Una persona está hablando de adicciones tecnológicas.

- ✓ **Escucha lo que dice y selecciona la respuesta correcta.**

1) ¿Qué forma parte de la vida diaria de la mayoría de las personas?

   a) El celular.
   b) Internet y Facebook.
   c) La tecnología.

2) De acuerdo a la encuesta realizada, ¿Cuántas personas tenían ansiedad o enfado?

   a) Veinticuatro por ciento.
   b) Un mil.
   c) Dos Mil.

3) ¿Qué se ha observado en los jóvenes adictos al internet?

   a) Problemas que podrían afectar sus emociones, la toma de decisiones, y el autocontrol.
   b) Un desarrollo en sus habilidades para videojuegos e internet.
   c) Modificaciones cerebrales que los convierten en ciberadictos.

4) ¿Cómo saber si tienes un problema de adición de estos?

   a) Si sientes que estas gastando mucho en videojuegos.
   b) Si sientes que eres el mejor en tu juego preferido.
   c) Si sientes que prefieres a tu ordenador más que a tus amigos y familia.

5) ¿Cuáles son las recomendaciones?

   a) Aprender jardinería y manualidades.
   b) Hacer actividad física.
   c) Acostarse temprano, no más tarde de las ocho de la noche.

Appendix D

Handout: Discussion Questions for the Awareness-Raising group

Nombre_____ **# 1**

I. Discussion questions:

✓ When listening to a Spanish text/audio,

1. Do you focus on the listening text especially when it is hard to understand?
2. How fast do you recover your concentration when you get distracted?
3. What do you do to get back on track when you lose concentration?
4. How persistent are you when you find difficulties understanding the listening text?

II. Notes:

III. Reflection:

Will you do anything differently the next time you listen to a text? If so, please explain what will you do? If there is nothing you will do differently, please describe what you do when listening.

_____

_____

_____

_____

Nombre_____ **# 2**

I. Discussion questions:

✓ When listening to a Spanish text/audio,

1) Do you translate in your head as you listen?
2) Do you translate key words as you listen?
3) Do you translate word by word, as you listen?

II. Notes:

III. Reflection:

Will you do anything differently the next time you listen to a text? If so, please explain what will you do? If there is nothing you will do differently, please describe what you do when listening.

_____

_____

_____

_____

Nombre_____ **# 3**

I. Discussion questions:

✓ When listening to a Spanish text/audio,

1) Do you have a plan in your head for how you are going to listen?
2) Before listening, do you think of similar texts that you may have listened to?
3) Do you periodically ask yourself if you are satisfied with your level of comprehension?
4) Do you have a goal in mind as you listen?
5) Do you think back to how you listened and about what you might do differently next time?

II. Notes:

III. Reflection:

Will you do anything differently the next time you listen to a text? If so, please explain what will you do? If there is nothing you will do differently, please describe what you do when listening.

_____

_____

_____

Nombre_____ **# 4**

I. Discussion questions:

✓ When listening to a Spanish text/audio,

1) Do you think that listening is more difficult than reading, speaking, or writing?
2) Do you feel that listening comprehension is a challenge?
3) Do you feel nervous when doing listening comprehension?

II. Notes

III. Reflection:

Will you do anything differently the next time you listen to a text? If so, please explain what will you do? If there is nothing you will do differently, please describe what you do when listening.

_____

_____

_____

_____

Nombre_____ **# 5**

I. Discussion questions:

✓ When listening to a Spanish text/audio,

1) Do you use the words you understand to guess the meaning of the words you do not understand?
2) As you listen, do you compare what you understand with what you know about the topic?
3) Do you use your experience and knowledge to help you understand?

II. Notes

III. Reflection:

Will you do anything differently the next time you listen to a text? If so, please explain what will you do? If there is nothing you will do differently, please describe what you do when listening.

_____

_____

_____

_____

Nombre_____ **# 6**

I. Discussion questions:

✓ When listening to a Spanish text/audio,

1) As you listen, do you adjust your interpretation if you realize that it is not correct?
2) Do you use the general idea of the text to help you guess the meaning of the words that you do not understand?
3) When you guess the meaning of a word, do you think back to everything else that you have heard, to see if your guess makes sense?

II. Notes

III. Reflection:

Will you do anything differently the next time you listen to a text? If so, please explain what will you do? If there is nothing you will do differently, please describe what you do when listening.

_____

_____

_____

_____

Nombre_____ # 7

I. Discussion questions:

✓ When listening to a Spanish text/audio,

1) Which of the aspects below do you think are the MOST important when doing listening comprehension? Why?

1. Attention
2. Mental translation
3. Planning and evaluation
4. Person knowledge
5. Problem solving

II. Notes

III. Reflection:

Will you do anything differently the next time you listen to a text? If so, please explain what will you do? If there is nothing you will do differently, please describe what you do when listening.

_____

_____

_____

_____

Nombre_____ **# 8**

I. Discussion questions:

✓ When listening to a Spanish text/audio,

1) Which of the aspects below do you think are the LEAST important when doing listening comprehension? Why?

    1) Attention
    2) Mental translation
    3) Planning and evaluation
    4) Person knowledge
    5) Problem solving

II. Notes

III. Reflection:

Will you do anything differently the next time you listen to a text? If so, please explain what will you do? If there is nothing you will do differently, please describe what you do when listening.

_____

_____

_____

_____

Appendix E

Questionnaire 1

- ✓ **Please, do not write down your name.**

1) Thinking about the listening section in the midterm exam, do you think the listening practice you've been receiving helped you? Please, briefly explain how it's helped you or why it has not helped you.

Appendix F

Questionnaire 2

- ✓ **Please, do not write down your name.**

1) What do you think of the teaching method used during the listening activities?

_____

_____

_____

_____

_____

_____

2) If you could modify or change the listening activities you did in class, how and what would you modify/change to make them more useful and/or more effective?

_____

_____

_____

_____

_____

_____

Appendix G

Exposure to Spanish

Nombre_____

1) Is English your native language? Yes No

2) If not, what is your native language? _____

3) How much exposure to Spanish, **outside of the class time**, do you usually have? Please, write the number of hours per week.

   3.1 Outside of my class time, I am exposed to Spanish _____ hours per week.

   3.2 Please, briefly explain how you are exposed to Spanish:

_____

_____

_____

	Strongly disagree	Disagree	Slightly disagree	Partly agree	Agree	Strongly agree
4) I like learning Spanish	**1**	**2**	**3**	**4**	**5**	**6**