



**MALMÖ HÖGSKOLA**  
Faculty of Education and Society

Department of Skolutveckling och Ledarskap

Degree Project with Specialisation in  
English Didactics  
30 Credits - advanced studies

# Effects of a Metacognitive Approach to Teaching L2 Listening

Effekter av Metakognitiv Undervisning i L2 Hörförståelse

Tina Webb

Master of Science in Education, 120 Credits  
Opposition Seminar: 31 January 2017

Examiner: Björn Sundmark  
Supervisor: Anna Wärnsby

# Abstract

Metacognitive listening instruction is the method recommended to Swedish teachers by the Swedish National Board of Education (*Skolverket*) in a document authored by Lena Börjesson (2012) found in the commentary material to the steering documents. This method is based on a metacognitive pedagogical sequence of L2 listening instruction suggested by Vandergrift and Goh (2012). In this study, I test this method using action research. The participants of the study were first year upper secondary school students from a vocational program, the control group consisted of students from a preparatory program. In general, the treatment group exhibited low motivation to study, while the second group had higher motivation. Both groups attended an upper secondary school in the South of Sweden. During seven classes, the treatment group (n=16) received training in the method, and the control group (n=21) was given more traditional tests during six classes. In this study, I used the following methods to obtain my data: the PET listening test, the listening segment of the Swedish National Test of English and the Metacognitive Awareness Listening Questionnaire (MALQ).

The results demonstrated that both groups improved their results on the listening aptitude test significantly; however, the treatment group did not with a statistical significance improve more than the control group. Secondly, the students did not perceive that they were using more strategies after the explicit strategy training they had received; both groups reported to using strategies less, as the listening texts became increasingly difficult. Thirdly, the students from the two groups did not report perceiving any difference in learning how to listen, despite one of the groups receiving explicit instruction in listening strategies. Finally, the students both in the treatment group and in the control group have reported to increasing listening anxiety after the instructional period, but the levels of anxiety increased less in the treatment group.

The results of this study thus do not unequivocally suggest the effectiveness of the method for teaching listening recommended by *Skolverket*. In particular, it is questionable whether the method is at all suitable for students with low motivation as those who have participated in the study.

Keywords: Action research, Classroom research, Listening comprehension, Metacognitive awareness, Metacognition in listening, Teaching listening in L2





# Table of Contents

<b>1. Introduction</b>	<b>7</b>
1.1 Aim and research questions	9
<b>2. Background</b>	<b>11</b>
2.1 Understanding the Listening Process - A Model of Listening	11
2.1.1 Bottom-up and Top-down Processes	12
2.1.2 Automatic and Controlled Processing	13
2.1.3 Perception, Parsing and Utilization	14
2.2 Strategies for Listening Comprehension	15
2.2.1 Some Cognitive Competences Applied in the Listening Process	17
2.2.2 Metacognitive Strategies	18
2.2.3 Affective Factors in the L2 Listening Process	22
2.3 Metacognitive Instruction	23
2.4 Recent Studies on Metacognitive Strategy Instruction	25
<b>3. Method and Design</b>	<b>30</b>
3.1 Site and Participants	30
3.1.1 Homogeneity and Comparability of the Treatment- and Control groups	31
3.2 Why Action Research?	32
3.2.1 The Mixed Method Approach	33
3.3 The Quantitative Method	34
3.3.1 The Pre- and Post-Tests	34
3.3.2 MALQ – Metacognitive Awareness Listening Questionnaire	36
3.3.3 The Anonymous Questionnaire (3 Questions)	37
3.4 The Qualitative method	38
3.4.1 The Think-Aloud Protocols	38
3.4.2 The Field Notes	39
3.5 Classroom Treatment Design	39
3.5.1 Treatment Group Intervention	40
3.5.2 Control Group Intervention	41
3.6 Selection of Listening Material	43
3.7 Validity and Reliability	43
3.8 Ethical Considerations	44
<b>4. Results</b>	<b>46</b>
4.1 To What Extent Will Students’ Listening Skills be Affected by a Four Week Period of Metacognitive Listening Strategy Instruction?	
- Quantitative Data	46
4.1.1 Pre- and Post-Testing, Treatment- and Control groups	47
4.2 To What Extent Will Students’ Listening Skills be Affected by a Four Week Period of Metacognitive Listening Strategy Instruction?	
- Qualitative Data	48
4.2.1 Field notes, Treatment- and Control Groups	49
4.3 How Will Students’ Percieved Strategy Use be Affected by Explicit Metacognitive Listening Strategy Instruction?	50
4.3.1 Results of the MALQ	50

4.4 Students' Perception of Strategy Use: Qualitative Data – Treatment Group Only	52
4.5 Students' Perception of Learning	56
4.5.1 Students Perception of Learning: Quantitative Data – Treatment- and Control Group	57
4.6 How Will Students' Anxiety Towards Listening be Affected by Metacognitive Listening Strategy Instruction?	58
4.6.1 How did you feel about listening today? Quantitative Data, Treatment- and Control Group	59
4.6.2 Do you know the reason you feel this way about listening today? Qualitative Results, Treatment- and Control Group	60
4.7 Limitations of the Study	61
<b>5. Discussion and Future Research</b>	<b>63</b>
5.1 To What Extent Will Students' Listening Skills, Measured by a Standardized Listening Aptitude Test, be Affected by a Four Week Period of Metacognitive Listening Strategy Instruction?	63
5.2 How Will Students' Perceived Strategy Use be Affected by Explicit Metacognitive Listening Strategy Instruction?	67
5.3 How Will Students' Perception of Learning be affected by Metacognitive Listening Strategy Instruction?	70
5.4 How Will Students' Anxiety Towards Learning be Affected by Metacognitive Listening Strategy Instruction?	72
<b>6. Summary and Conclusions</b>	<b>75</b>
<b>References</b>	<b>77</b>
Appendix I: Test Descriptions	79
Appendix II: MALQ	80
Appendix III: Anonymous Questionnaire	81
Appendix IV: Results of the t-tests	82

# 1. Introduction

Traditionally the focus on studying reception has been on reading rather than on listening. Reading instruction is often based on teaching the students how to use strategies to sort out the increasingly complex input. The skills needed for reading and listening are, however, rather similar. The same development, from decoding to comprehension that occurs in reading, must also occur in L2 students when it comes to listening. However, the listening skills can be dissimilar in a first and second language; as word order, stress and intonation patterns as well as grammatical markers differ, the listening skills have to be re-learned to some extent to bridge the gaps between L1 and L2 listening. Moreover, the listening input is much faster and often less structured than written text. To understand aural input in a second language, L2 students need to have automated knowledge of the L2. Both teachers and students need to be made aware of the different conditions of these skills. In spite of the differences between listening- and reading input the strategy-based instruction that over the years has been proven effective in reading - may well be effective for teaching listening also.

Research on listening instruction and on metacognitive strategy training have become increasingly common over the last few decades. During the last 20 years, some research reviews (e.g. Flowerdew & Miller, 2005) have become the basis for new evidence-based ways of teaching listening (Goh 2008, Al-Alwan, Asassfeh & Al-Shboul 2013). Vandergrift and Cross (2015) claim that the three most studied paths of L2 listening instruction are; strategic instruction, metacognitive instruction and using multimedia applications with CALL (computer assisted language learning).

Cross (2015) suggests a difference between strategy instruction and metacognitive instruction. Both forms of instruction draw on the work of Flavell (1979). However, strategy instruction has the rather narrow focus to teach strategies be they cognitive, metacognitive or socioaffective (e.g. the teacher presents a strategy, models how to use it and provides guided activities for practising it). Metacognitive instruction, on the other hand, is process-based and has a wider purpose, “it targets the development of 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, p. 886).

Individual strategy instruction (e.g. listening for gist or applying schemata) has only been proven to be effective short term and may not lead to overall listening improvement at all (Field 2001).

Even though there are difficulties with teaching and learning listening, as mentioned above, and in spite of the short-term value of strategy instruction, this type of training may well be an effective way of teaching listening. Metacognitive listening instruction may increase listening proficiency in testing, and this type of instruction may also help learners decrease their anxiety towards listening and help them become more independent. Goh (2008) states that the studies she reviewed:

“indicate that metacognitive instruction in listening can be beneficial in at least three ways: (1) It improves affect in listening, helping learners to be more confident, more motivated and less anxious; (2) It has a positive effect on listening performance; (3) Weak listeners potentially benefit the greatest from it” (p. 196).

In the Swedish curriculum for English 5, the first of the upper secondary English courses, one of the knowledge requirements is that the students must be able to “choose and with some certainty use strategies to assimilate and evaluate the content of spoken [...] English (Skolverket, 2011b, p 4)<sup>1</sup>. By this phrasing the Swedish National Board of Education implies that explicit strategy training in the receptive skills is mandatory, since teachers cannot grade what they do not teach. The grading system dates from 2011, and in 2012 the Swedish National Board of Education published Börjesson’s document on their website, in order to clarify what strategies were aimed at in the steering documents. In this document, Börjesson suggests the Vandergrift seven-step model (Börjesson 2012, p 7-8). Börjesson claims that the method is based on recent research that indicate that teachers should focus on the process rather than on the product. Despite the obvious potential impact of this recommendation, Börjesson, however, does not mention any studies where the method has been tested in the Swedish school setting.

For that reason I decided to test whether the method of metacognitive strategy instruction is indeed useful in Swedish upper secondary English 5 classes by conducting action research in two classes, using one as a trial group and one as a control group. In order to test Goh’s statement that weaker listeners would benefit greatest from the method, I decided to use the class with the lowest English scores from grade 9 as the trial group. I designed two different pedagogical interventions, one testing the model of listening instruction described in section 2.3 and another, with an equal amount of listening, for the

---

<sup>1</sup> Skolverket offers an English translation of the curriculum, this translation of the knowledge requirements is taken from there.

control group. I taught both groups of students. The rationale for the study and its design are discussed in sections 2 and 3. The methods for collecting and analysing data are discussed in sections 3.3 to 3.5. The results are displayed in section 4, and discussed in relation to the results of other studies in section 5. The results are finally summarized in chapter 6.

### *1.1 Aim and Research Questions*

Firstly, this study aims to investigate whether the method created by Vandergrift *et al* and suggested by Börjesson (2012) can increase students' listening proficiency in a Swedish classroom. Further, the study aims to investigate whether explicit strategy training increases students' perceived use of some listening strategies and lends these students, as suggested by Goh (2008), a greater sense of agency in learning how to listen. The idea is to see that if the students think that they are learning a useful skill they will be more motivated to participate in this kind of instruction. Finally, this study aims to investigate whether this teaching method helps decrease the students' levels of anxiety as it is reasonable to assume that when they master the skill of listening, they will also be more confident using it. Thus, the research question is as follows:

- How and to what extent will students listening skills, measured by standardized listening aptitude tests, be affected by a four-week period of metacognitive listening strategy instruction?

To get a more holistic perspective on the results of the pre- and post-tests, and answer the *how*-part of the research question above, the aim is to follow up those results using the following sub questions:

- I. How will students' perceived strategy use be affected by explicit metacognitive listening strategy instruction?
- II. How will students' perception of learning be affected by metacognitive listening strategy instruction?
- III. How will students' anxiety towards listening be affected by metacognitive listening strategy instruction?

Metacognitive strategy instruction has been investigated previously but, in other settings than the Swedish upper secondary school. Swedish students of English 5 are to be graded on their use of strategies on the receptive skills. Teachers and students alike need to be aware of what the Swedish National Board of Education means by strategies. Since the metacognitive strategy instruction is suggested in the Swedish steering documents, its efficiency in a Swedish setting ought to be investigated.

## 2. Background

This section starts with an overview of research on the actual listening process in segment 2.1, to give an understanding of the demands as well as skills and subskills of listening. In chapter 2.2, I describe the different kinds of strategies available to the language learner, and discuss how the different kinds of strategies are intertwined. In section 2.3, I present the model of metacognitive instruction that forms the basis of this study. Finally, in 2.4, I summarize earlier, recent studies that have investigated the effectiveness of metacognitive strategy instruction.

Understanding the listening process, the L2 listening process in particular, was the key to create the appropriate material for the sequence of listening classes. The model of listening presented below was used as the common framework for me and the students, so we could discuss what strategies would be appropriate at particular stages of listening. The idea was that if the students could understand what they were doing and why they were doing it, their motivation to complete the tasks would increase.

### *2.1 Understanding the Listening Process - A Model of Listening*

Listening is made up from many processes that the listener must deal with - simultaneously to comprehend the aural input. Vandergrift and Goh (2012) describe the different kinds of processes as 1) bottom-up and top-down processing, 2) controlled and automatic processing, 3) perception, parsing, and utilization and 4) metacognition. These processes are based on the model of listening proposed by cognitive psychologist J. R. Anderson (1985). The figure below illustrates the interrelationships between the processes. These are the overarching definitions of processes, but they help create a visualization of the listening process that can be used as a platform for the discussion of listening with students.

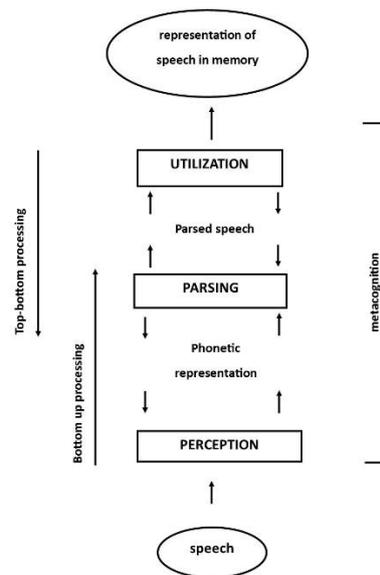


Figure 1: A visualization of the cognitive processes in L2 listening and how they interrelate (Vandergrift & Goh 2012, p. 18).

This figure illustrates how top-down and bottom-up processing start to work together from the time speech is perceived until the input has reached the utilization phase: when it is comprehensible enough for the listener to use properly. According to this model, other, more specific, phases also interrelate, and the input bounces back and forth between the different phases until the listener has made a mental representation of what has been heard: until the input has been understood by the listener. Metacognitive strategies can be applied on all phases of listening. The different phases of listening are described below.

### *2.1.1 Bottom-up and Top-down Processes*

Bottom-up and top-down processes are fundamental to understanding, and listeners generally use both types of process. The bottom-up process involves segmentation of the language that is heard. The stream of connected language is divided into smaller segments (words, phonemes or individual sounds) and suprasegmentals (intonation, stress, tone and rhythm). These segments and suprasegmentals are then built up to increasingly larger meaningful units: phrases, sentences and chunks of discourse (Vandergrift & Goh 2012, p.18).

The top-down process involves students' application of context to what they have heard. Listeners can apply different kinds of knowledge to their understanding a text. If they know

they are listening to an announcement at an airport, they can apply all their knowledge (prior knowledge, pragmatic knowledge or cultural knowledge about the target language) about airport announcements to help them make sense of what they hear. Listeners apply schemata that are stored in the long-term memory (Vandergrift & Goh 2012, p.18). The top-down and bottom-up strategies are overarching types of strategies. Genre-recognition and applying schemata are top-down strategies, while detail identification and discourse marker identification are bottom up strategies (Siegel 2015, p. 66). Neither top down nor bottom-up processing is adequate on its own to achieve comprehension. Which of the processes is used more will depend on the purpose of listening (Vandergrift & Goh 2012, p. 18-19).

Interestingly, Flowerdew & Miller (2005, p. 69-70) state that effective listeners use both bottom-up and top-down processes and ineffective listeners only bottom-up (based on a study from O'Malley and Chamot 1990, p.132). Field (2008), however, disagrees with this statement, referring to a 1998 study conducted by Tsui and Fullilove, who found that skilled listeners outperformed the less skilled listeners because the former correctly answered questions where they could *not* draw upon their world knowledge: the more skilled listeners could answer the questions because they had superior decoding skills (Field 2008, p. 131). Teaching students decoding skills thus seems important. Tsui and Fullilove also found that L2 learners more often used bottom-up processing in formal testing situations (Flowerdew & Miller 2005, p.70).

### *2.1.2 Automatic and Controlled Processing*

The processes with which spoken language is understood becomes automated with practise. The listening process in L1 listening is almost fully automated: listeners do not have to consciously work to interpret what they hear. L2 listeners, on the other hand, are not able to automatically process everything they hear (depending on language proficiency, of course). L2 listeners can then resort to controlled processing of some aspects of the input, such as content words. Controlled processing is a conscious effort to process parts of what they hear (Vandergrift & Goh 2012, p.19). Since controlled processing is not very effective, it takes too long and the listeners may not be able to keep up with the incoming input if they have to think about everything they hear (or mentally translate what they hear, mental translation being such a controlled processing strategy). When forced to use controlled processing, because the automated processing is insufficient, L2 listeners have to resort to other compensatory strategies, such as contextual factors (top-down processing (Vandergrift & Goh 2012, p. 19-20)). As learners becomes more and more familiar with the target language (vocabulary,

grammar, stress & intonation), more and more of the processing will become automated and happen unconsciously, more like it is in L1 listening.

### *2.1.3 Perception, Parsing and Utilization*

Anderson (1985) suggested a model of the listening process with three separate phases: perception, parsing and utilization. In the perception phase: listeners use bottom-up and top-down processes and decode the input; they attend to the speech: exclude other sounds, note similarities, pauses and stresses in the speech and they group the sounds they have heard into categories of identified language (Vandergrift and Goh 2012, p 21). In the parsing phase the listener fragments larger units of speech into words (or smaller units) in connected speech. When reading, one can easily separate one word from another as words are separated by spaces in written text, but when listening to connected speech such separation may be more difficult since “word segmentation skills are language specific and acquired early in life” (Vandergrift & Goh, p. 21) When parsing, listeners try to find potential candidates for words they know in their long-term memory, using cues such as word onset or knowledge of the topic or context. (Vandergrift & Goh 2012, p. 22). L2 listeners have to become efficient in how they identify content words of the input since working memory capacity hinders them from processing every word they hear. In the utilization phase, listeners hopefully have meaningful units of language (after having perceived and parsed what they have heard) and at this point they can go to the appropriate schemata, in the long-term memory, to interpret the intended or implied meanings of what they heard: listeners apply top-down processing to the input (Vandergrift & Goh 2012, p. 22). This process is automatic in fluent listeners; however, if inferences (the application of prior knowledge to monitor what they heard) do not work automatically, listeners will have to make a conscious effort to solve problems in comprehension (controlled processing).

Because the listening process is so rapid, listeners report that they have difficulties remembering what they have heard and creating a mental representation needed to help them store what they have heard in their long-term memory (Vandergrift & Goh, p. 44). This is important for teachers to know as some listening tests may test listeners’ memory rather than actual listening skills (Buck 2001). In the method that is under investigation, the discussion steps of the pedagogical sequence help the students create and develop such mental representations, and the students are given the opportunity to see how others create their mental representations, which in turn can be beneficial all students in the learning situation.

Flowerdew and Miller describe, based on O'Malley, Chamot & Kupper (1989), how effective and ineffective listeners differ in the perceptual process, parsing process and utilization phase. Effective listeners have the ability to keep attending to a text where ineffective listeners are unable to redirect their attention to a text when they have lost attention after being put off by the length of a text or by the number of unknown words in it. Listeners who are effective in parsing use intonation and pauses to deal with chunks of text, where ineffective listeners go word-by-word. In the utilization phase, effective listeners engage and use different kinds of knowledge (e.g. world-knowledge or self-knowledge), while less effective listeners are more passive and engage less in what they hear (2005, p. 70). According to Vandergrift (2003) O'Malley et al. (1987) studied the issue of strategy use in L2 listening by using think aloud methodology, to decipher skilled listening behaviour. O'Malley et al. came to the conclusion that skilled listeners use the task (e.g. the questions on a written listening comprehension exam) (strategy: planning) to establish the topic of the listening text and thus apply the proper prior knowledge (strategy: elaboration) and to predict possible content (strategy: inferencing). When students listen, they focus on important content (to free working memory capacity they do not focus on everything they hear) (strategy: selected attention). After listening, they keep using relevant prior knowledge (strategy: elaboration) to help with overall comprehension. If necessary, skilled listeners revise their predictions and keep working on comprehending (strategy: monitoring). Knowing the strategies of effective listeners may help teachers make students aware of what aspects of listening they need to work with specifically in order to become more efficient listeners.

This model of listening described above was presented to the students of the treatment group before the instruction started, and it was used as a platform for discussion throughout the instructional period.

## *2.2 Strategies for Listening Comprehension*

O'Malley, Chamot, Stewner-Manzanares, Russo & Kupper (1985) divided the learning strategies into three groups (based on the work in cognitive psychology by Palinscar & Brown from 1982); cognitive, metacognitive and socioaffective strategies.

<p><b>Cognitive strategies:</b> strategies that are used for processing information, solve problems through analysis and make the new knowledge part of the pre-existing body of knowledge. Cognitive strategies can be controlled or automatic.</p>	<p><b>Metacognitive strategies:</b> strategies we use to consciously plan, monitor and evaluate our learning. Metacognition means that the learner plays an active part in the learning.</p>	<p><b>Socioaffective strategies:</b> strategies used for cooperation and interaction. Learners ask for help and they listen to each other. The socio affective strategies are important to classroom atmosphere.</p>
<p><b>Inferencing:</b> (linguistic inferencing, voice inferencing, extra linguistic inferencing, between parts inferencing).  <b>Elaboration:</b> (personal elaboration, world elaboration, academic elaboration, questioning elaboration, creative elaboration)  <b>Imagery</b>  <b>Summarization</b>  <b>Translation</b>  <b>Transfer</b>  <b>Repetition</b></p>	<p><b>Planning for learning:</b> listeners assess their prior knowledge, understanding of the task and draw from internal and external resources to engage actively in the task  <b>Monitoring comprehension:</b> listeners evaluate the effectiveness of their performance as it takes place  <b>Self-evaluating:</b> listeners evaluate their performance after an activity has been completed  <b>Self-testing:</b> listeners test the effectiveness of their language use (or lack of language use).</p>	<p>Social-mediating activity  Transaction with others</p>

Figure 2: Strategies for listening comprehension. Adapted from: Al-Alwan, Asassfeh & Al-Shboul (2013), Buck (2001), Tornberg (2009), Vandergrift (2003).

This investigation focuses on the metacognitive strategies: the strategies that manage the cognitive strategies. However, the distinction between cognitive and metacognitive strategies is not always clear (Field 2008, p. 204). For example, if one uses a strategy in a controlled manner, it is metacognitive, but if the same strategy is automatically used, it is deemed to be cognitive.

In the pedagogical sequence of this study all three kinds of strategies are used by the students of the treatment group. They use cognitive strategies to make inferences about the circumstances of the texts before and during listening, they use metacognitive strategies when they monitor their understanding and since they have to mediate their understanding with their peers they a required to use socio-affective strategies also.

### 2.2.1 Some Cognitive Competences Applied in the Listening Process

L2 listeners must use an array of cognitive strategies to interpret aural input and become efficient listeners in the L2. Listeners apply different kinds of knowledge or different cognitive competences (Vandergrift & Goh 2012, p. 24; Buck 2001, p. 104; Flowerdew & Miller 2005, p. 30), such as vocabulary, stress (to identify important words), grammatical knowledge (e.g. to parse out tense) and pragmatic knowledge (e.g. to help the listener interpret the speaker's intent) (Vandergrift & Goh 2012, p. 24) and sociolinguistic knowledge (Buck 2001, p. 25). Sociolinguistic knowledge refers to the understanding of how certain utterances work in certain contexts and settings. Even kinetic knowledge can be added to cognitive strategies, according to Flowerdew & Miller (2005, p. 30, 45). Kinetic knowledge, the knowledge about how to interpret facial expressions or gestures, is more important to interactional listening as it is only applicable when the speaker is visible to the listener. Its importance to individual listening should not, however, be underestimated, due to the advances in technology. Students may, for example, be asked to watch video recordings of speakers to train this ability.

Language competence also includes discourse knowledge, that may help listeners to direct their attention to certain parts of the text when listening by using the appropriate rhetorical schemata. Discourse knowledge also helps listeners with certain signals of, for example, a beginning (*firstly*) or an opposing argument (*on the other hand*) etc. (Vandergrift & Goh 2012, p.25-6). Prior knowledge (e.g. world knowledge, discourse knowledge) also plays an important role in the utilization phase (top-down processing): students should receive contextual information when they are taught how to listen. When prior knowledge, and the appropriate schemata is applied, listeners can use information stored in their long-term memory to understand what they hear. Using more of the information stored in long-term memory frees up working memory resources allowing the L2 listener to process more information (Buck 2001, p.25).

All of these competences help listeners to apply the cognitive strategies in figure 1. Applying the proper schemata allows listeners elaborate on what they have heard or create mental images. Applying their different kinds of language knowledge can also help them, for example, transfer what works from their L1 listening --for example, raising intonation on questions-- or help them translate words. The list of cognitive competences involved in listening can be made much longer, but, this suffices to show how complex second language listening is, and how much an L2 learner must know to become an efficient listener in the

language he or she is learning. It also shows how complex teaching listening is as these cognitive competences are used interrelatedly and simultaneously.

2.2.2 Metacognitive Strategies

As discussed earlier, metacognitive strategies can be defined as the strategies that manage the cognitive strategies. In addition, as stated above, some metacognitive strategies, are cognitive strategies that are used unconsciously. In this section I describe the metacognitive strategies that are taught in the pedagogical sequence of listening instruction used in this study. Primarily, the focus in the instruction was on the five categories of strategies measured by the MALQ, planning and evaluation, problem solving, person knowledge, directed attention and mental translation. Because the strategies are so intertwined, more strategies were used to reach the object of each class.

Vandergrift and Goh (2012) suggest a metacognitive framework that serves two different functions in language learning, firstly *self-appraisal*- or *knowledge* about cognitive states and progresses and secondly *self-management* or *control* of cognition (p. 85). Self-appraisal is a learner’s reflections about his or her listening abilities and abilities to meet a cognitive goal. Self-management helps a learner to control the way he or she thinks.

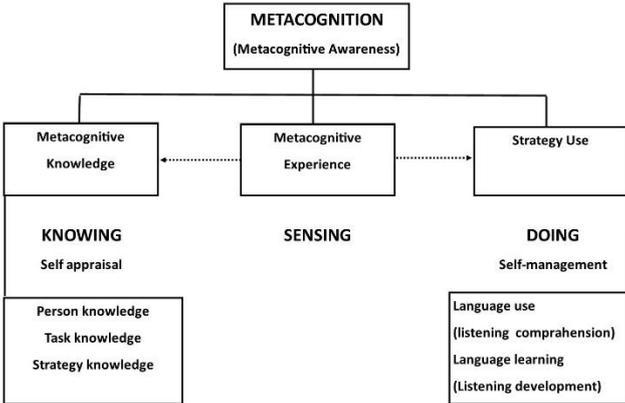


Figure 3: How metacognition works, from Vandergrift & Goh 2012, p. 85

These functions in turn rely on three components: metacognitive experiences (what happens in exposure to aural input -- fleeting or remaining -- these experiences shape the metacognitive knowledge), metacognitive knowledge (the remaining metacognitive experiences) and strategy use (the ability to use appropriate strategies to solve a task). There are three categories of metacognitive knowledge: person knowledge, task knowledge and

strategy knowledge (e.g. Cross 2015, p. 884, Farhadi, Zoghi & Talbei 2015, p. 141). Person knowledge is knowledge about how individual learners learn, and what factors affect the learning. It also includes the learner's beliefs about what leads to success or to failure. If a learner believes that listening is too hard, the learner may try to avoid situations where he or she is required to listen and show understanding, such as a testing situation. Task knowledge refers to the understanding of the purpose and demands of a task. It may also refer to the learners' knowledge of features of spoken language, such as discourse knowledge or grammatical knowledge. Lastly, strategy knowledge is the knowledge of which strategies can be employed to reach the goal of the task, and knowing about certain features of different text-types etc, so that learners can employ the appropriate strategies and knowledge (Vandergrift & Goh 2012, p. 86-87).

Person knowledge	Task knowledge	Strategy knowledge
Aims to develop better knowledge of self as an L2 listener.	Aims to understand the nature of L2 listening and the demands of learning how to listen.	Aims to understand the role of cognitive, metacognitive and socioaffective strategies.
Objectives: Examine personal beliefs about self-efficacy and self-concepts with regards to L2 listening. Identify listening problems, causes and possible solutions.	Objectives: Experience mental, affective, and social processes involved in listening. Differentiate between types of listening skills (e.g. listening for details, listening for global understanding, listening to infer information). Analyze factors that influence listening performance (e.g. speaker, text, interlocutor, strategy). Compare and evaluate ways to improve listening abilities outside of formal education. Examine phonological features of spoken texts that influence perceptual processing.	Objectives: Identify strategies that are appropriate for specific types of listening tasks and problems. Demonstrate use of strategies. Identify strategies that may NOT be appropriate for learning style or culture.

Figure 4: Aims and objectives of metacognitive instruction (Vandergrift & Goh 2012, p. 99).

This study focuses specifically on trying to promote the use of four of the five categories of strategies mentioned above, strategies measured in the MALQ, that is, planning & evaluation, problem solving, person knowledge and directed attention. Mental translation was set aside as it is considered an ineffective strategy. However, it was discussed with the students so that the students would be aware of doing it.

According to Vandergrift (2003) planning, in general terms, simply refers to figuring out what exactly needs to be done to understand or to complete a task. In the school setting, this means that the students become aware of the problem to be solved; in real life, the L2 listeners have to understand what information they need to understand and for what reason. There are four sub-divisions of planning:

PLANNING:

Advance organization	Clarifying the objectives of an anticipated listening task and/or proposing strategies for handling it.
Directed attention	Deciding (in advance) to attend <i>in general</i> to the listening task and to ignore irrelevant distractions; maintain attention while listening.
Selected attention	Deciding to attend to specific aspects of language input or situational details that assist understanding (and/or task completion).
Self-management	Understanding the conditions that help one successfully accomplish a listening task and arrange for the presence of those conditions.

Figure 5: Planning (Vandergrift 2003, originally from Vandergrift 1997).

Planning (see fig. 3) includes several of the phases of listening that were described above as there are links between directed attention and top-down processing, and selected attention and bottom-up processing. Moreover, planning includes more than one of the strategies of the MALQ (Directed attention, Planning and Evaluation).

In the planning phase of the pedagogical sequence, students are asked to look at the problem they are to solve (e.g. read the questions they are to answer), to see what they need to do to prepare for the task (advance organization), The students listen to the text three times, they are told to use the first time to attend to the text in general. The second time, they are to listen for details they have missed (directed attention). The students are asked to use selected attention during the first listen when they are asked who is speaking, and what role the speaker plays in the text (interviewer, respondent, audience member). Knowing who's speaking and why may help them solve the problems and to understand the text as a whole. Self-management can be suggested to the students by giving them ideas to improve their listening, these ideas can range from sitting next to someone who will not disturb them to studying convenient vocabulary prior to listening.

The second metacognitive strategy is monitoring, which refers to when a listener checks, verifies or corrects his or her comprehension during the course of a listening task. There are two types of monitoring:

**MONITORING:**

Comprehension monitoring	Checking, verifying or correcting one's understanding at the local level.
Double-check monitoring	Same as comprehension monitoring - but during the second listen

Figure 6: Monitoring (Vandergrift 2003, originally from Vandergrift 1997).

Monitoring is a strategy that is used unconsciously in L1 listening. This means that it is not easily transferrable to when students learn how to listen in an L2, therefore, it has to be taught. In the pedagogical sequence, monitoring is used in/after all three listens. In the first listen, students check their predictions; in the second listen students check if they have resolved the problems in comprehension they had after the first listen; during the third and final listen, when they have access to the written text, the students verify their solutions and details such as tense, word segmentation in connected speech etc. This was explained to the students in the treatment group in every listening class.

The third metacognitive strategy is evaluation. Evaluation occurs when listeners check the outcomes of their “listening comprehension against an internal measure of completeness and accuracy” (Vandergrift 2003, p. 494). Evaluation is closely related to monitoring, with the addition that the listeners have to check not only *if* they understood, but *how well* they understood. In a pedagogical sequence, students can be asked to see how many of the questions on a test were solved, but, as Vandergrift states above, it is also a way of thinking for listeners as the checking is internal. It requires a listener to understand how much or how little they not understand. This is an internal process that is difficult to evaluate on its own since listeners have to be aware of what strategies they used in order to evaluate them. In the MALQ evaluation is paired with planning.

The fourth and last metacognitive strategy discussed in this chapter, is problem identification, which refers to when a listener explicitly identifies “the central point needing resolution in a task or identifying an aspect of the task that hinders its successful completion” (Vandergrift 2003, p. 494), i.e, exactly what it is that is not, but needs to be, understood. When listeners have identified the problem, they have to figure out a way to solve it, for example, by directing their attention at the next listen, or asking for clarification.

### *2.2.3 Affective Factors in the L2 Listening Process*

L2 listening is also affected by affective factors such as anxiety, self-efficacy and motivation. In traditional teaching of L2 listening, the skill is more often tested than taught. Some of the students' anxiety comes from the fact that listening instruction often is covert testing (Buck 2001), and that listening receives the least attention in the classroom. For teachers, it may be unnerving that listening is the least understood of the four main language competences when it comes to assessing (Buck 2001). Good, authentic listening material that is on the right level of complexity is difficult to make, in the same way as tests that assess listening skills rather than, for example, retention capacity. The notions that L2 listening is perceived as one of the most difficult skills to learn as well as to teach (e.g. Vandergrift & Goh 2012) and that students traditionally sense that there is little they can do to learn how to listen on their own further add to this anxiety. Vandergrift and Goh (2012) refer to a study by Graham (2006) when they suggest that students' anxiety levels were high because they "attribute L2 listening success to factors outside their control" (p. 71). Giving students the means to control their learning of L2 listening should by the reasoning of metacognitive theory, decrease the students' level of anxiety about L2 listening. The pedagogical sequence created for this study was designed to give the students the sense of agency suggested by Goh (2008).

Self-efficacy refers to listeners' belief in their ability to deal properly with the listening situation (Vandergrift & Goh 2012, p. 71). Listeners with low self-efficacy may feel a lack of confidence in their abilities and may try to not participate in listening exercises to avoid revealing their weaknesses. They may also feel that they cannot improve their L2 listening abilities: as it is out of their control (Vandergrift & Goh 2012, p. 71). According to Vandergrift and Goh this has implications to teaching because teaching students to regulate their learning may improve students' self-efficacy and, in turn, increase their motivation.

Motivation is believed to be a factor in L2 listening success, even though empirical evidence may be weak when it comes to motivation in L2 listening specifically. Learners with low motivation "perhaps because of a lack of self-confidence and self-efficacy, demonstrate [...] a passive attitude towards L2 learning, and also report [...] using less effective listening strategies" (Vandergrift & Goh 2012, p. 72). Even if empirical evidence is scarce, there is a high probability that motivation affects the learning process, which is why a strategy that provides the students with a way to raise their motivation may be a good idea for the classroom. However, it is uncertain that a poorly motivated student would go through the trouble of rethinking the way that their personal learning process works, which in and of itself can be an arduous task.

### 2.3 Metacognitive Instruction

Now that the metacognitive strategies have been established, I will present the seven- step pedagogical sequence of listening instruction suggested by Vandergrift, and in turn by Börjesson (2012) who was published by the Swedish National Board of Education on their website.

Metacognitive instruction, refers to pedagogical procedures that enhance the awareness of metacognitive strategies in learners, while at the same time, learners are acquainted with ways to plan, monitor and evaluate their comprehension efforts and their listening development (Vandergrift & Goh 2012, p. 97, Goh 2008, p 192). Since the students are allowed to listen to the aural input several times, and they are instructed to listen in different ways at different stages in the process, they may become better at perceiving, parsing and utilizing what they hear, and they “strengthen their ability to engage in parallel processing (see Fig. 1), including both bottom-up and top-down processes” (Vandergrift & Goh 2012, p. 101).

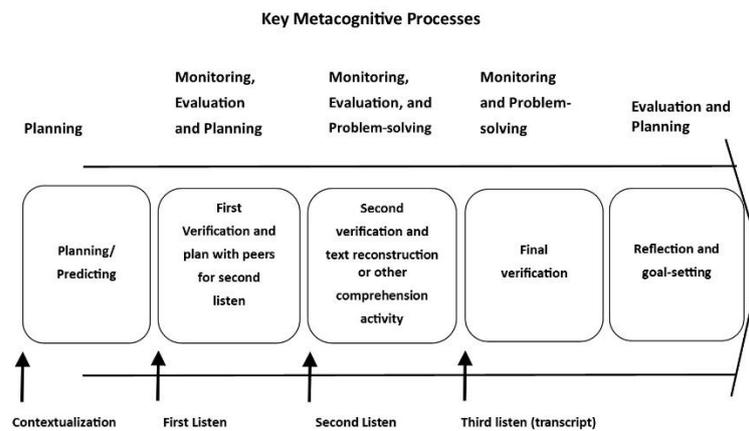


Figure 7: The seven-step model (Vandergrift & Goh 2012, p. 109).

In a metacognitive pedagogical sequence, as suggested by Vandergrift and Goh 2012 (p. 110), the underlying metacognitive processes - or strategies (planning, monitoring, evaluation and problem-solving) will be used in the following way:

## Stage of instruction

### Metacognitive strategies:

---

1 – Pre-listening (planning/predicting) Teacher informs of topic and text-type - Students predict possible words	1. Planning
2 - First listen (first verification stage) Learners verify initial guesses Write down additional info Learners compare notes with partner - Establishes what still needs resolution	2 Monitoring, evaluation  3 Monitoring, evaluation, planning
3 - Second listen (second verification stage) Learners verify points of disagreement Write down additional info Class discussion - all members contribute to the Reconstruction of the text.	4 Monitoring, evaluation, problem-solving  5 Monitoring, evaluation, problem-solving
4 - Third listen (final verification stage) Learners listen specifically for info they have not yet heard (listen may be accompanied by transcript)	6 Monitoring, problem-solving
5 - Reflection and goal setting Learners write goals for next listening activity Based on earlier discussions	7 Evaluation and planning

Figure 8: The pedagogical sequence (Vandergrift & Goh, 2012).

While carrying out these pedagogical sequences, students should not feel tested, these exercises should be understood as formative assessment. Learners do these activities to learn how to listen, and they should understand that purpose. For that reason, it is pertinent to create an environment where students feel safe, for example discussing their predictions with each other.

Field (2008) suggests that a good listening lesson today consists of four parts - a pre-listening phase where the context is established, and the teacher helps the students create motivation for the task. Finally, the teachers acquaint the students with critical words, at most four or five words per listening sequence. Establishing the context of an aural text is a way for teachers to mimic real life. Language is rarely heard in situations where the listeners have no idea of the context of the spoken output. Creating motivation for listening is important to the students, since the quality of the listening may be enhanced if the listener has a purpose for listening and the right mental set. There are thus several reasons for teachers not to teach all unknown words before listening, as this may impede them from listening purposefully and effectively applying listening strategies (Field 2008, p. 17). The second phase, according to Buck (2008) is the extensive listening phase, where students listen for the answers to general

questions on context and about the attitudes of speakers. Students learn to listen for gist. The third phase is the intensive listening where students learn to listen for details. Students get to listen for answers to pre-set questions or checking answers to questions (as in the third listen, suggested by Vandergrift and Goh 2012). The fourth and final phase is the post-listening phase, where the teacher can focus on functional language (ways of greeting, refusing or apologising), learners can also be asked to infer the meaning of words from the sentences where they appear, and, they could also get to see the transcript of the text.

Revisiting the four (Field 2008) or five (Vandergrift & Goh 2012) phases, it is also important to understand what knowledge source students need to draw from to understand what they hear. During the first phase, the pre-listening stage, students are invited to activate certain schemata: they are invited to use a top-down model of processing the input, where they draw on their prior knowledge on a subject. According to Flowerdew and Miller “The basic idea is that human knowledge is organized and stored in memory according to re-occurring events” (2005, p. 25). Morra de la Peña & Soler (2001) explain schema theory as “comprehending a text implies more than merely relying on one's knowledge of the linguistic system. It involves interaction between the reader's knowledge structures or schemata (knowledge of text content and form) and the text” (p. 219). In the Anderson (1985) model of listening, this activation of schemata plays an important part in the utilization phase of the listening process. Teachers, however, have to instruct students with care, so that they keep monitoring what they have understood, since their prior knowledge can lead them in the wrong direction.

#### *2.4 Recent Studies on Metacognitive Strategy Instruction*

I have taken part of nine other studies, as they have been presented in scientific articles: Al-Alwan, Assafeh & Al-Shboul (2013), Bozorgian (2012), Coscun (2015); Farhada, Zoghi & Talebi (2015), Graham & Macaro (2008), Khonmari & Ahmadi (2015), Movahed (2014) Vandergrift & Tafghodtari (2010), and Wang (2015). These studies have all in some way evaluated the efficiency of either process based metacognitive strategy training on listening proficiency, or the results of the MALQ in connection with listening proficiency. None of these studies are based on action research and none of them consider the teacher perspective. However, several of them state that teachers need to educate themselves within this field.

The most famous of the nine studies is Vandergrift & Tafaghodtari (2010). In their study the treatment group significantly outperformed the control group. The students who scored below the mean on the pre-test, outperformed their peers in the control group by

statistically significantly more, suggesting that less proficient students have more to gain from this method. The students, and their listening behaviour was followed up by randomly selecting six students and let them participate in stimulated recall sessions on their MALQ performances. In the Vandergrift & Tafaghodtari study the subjects were 106 university students (59 in the treatment group and 47 in the control group) of French as a second language. Those students studied their regular course during the 13 week semester (it is unclear how many lessons a week the students had), and did a listening sequence once a week, on a topic that related to what they were studying the rest of the time.

The argument that this type of instruction is better if it is integrated with the normal instruction is strengthened by the results of Coskun (2010) who did a study on beginner preparatory students at a university in Turkey. In this study, also, the treatment group statistically significantly outperformed the control group in the post test. Two beginner groups, one control group (n=20) and one experimental group (n=20) received five weeks of metacognitive training embedded into their texts books. The instruction was based on CALLA (Cognitive Academic Language Learning Approach) (from O'Malley and Chamot 1994), not Vandergrift's seven step model. Notably, in this case too, the students were university students.

Another of the studies that has gained some importance is the Graham & Macaro study from 2008. The results of this study too are that metacognitive listening strategy instruction makes the treatment group outperform the control group. Graham and Macaro did not only measure the impact of the metacognitive listening strategy on listening comprehension, but, also if the amount of scaffolding given to the students under the intervention would vary the outcome. They also measured the students level of self-efficacy. One thing that is significant with this study is that they measured the students' difficulties beforehand, and, they tailored the instruction towards the known needs of the students, rather than trying to show all strategies at once. This kind of tailoring is very likely to affect the results, which is why they cannot be a part of a generalized consensus of that this method is effective. Most teachers who are supposed to be the consumers of these findings, will not have the tools to properly do this on their own. The results are interesting no less because this is a longitudinal study. The intervention lasted from October to April and the results were followed up the following October – the results were that the high scaffolding group as well as the low scaffolding group significantly outperformed the control group at the last measure. At the second measure, the low scaffolding group even outperformed the high scaffolding group. This may suggest that the students of my study, who received high scaffolding, may

have done equally well with less, something that would could potentially be much more labour economic, in terms of how useful this method could be for practising teachers. This could also mean, that too much help, here in the form of scaffolding, counteracts making the students independent learners, and rather makes them relying too much on the help they know they will receive.

The subjects of the Graham & Macaro study were 68 lower intermediate learners of French as a foreign language in the UK. The students were 16-17 years of age and had studied French for 5 years. They had elected to continue studying French after finishing the GCSE. All of the instruction given to students within this study was given within the regular course. As in this case, it was presented as a theme that stood on its own, this too makes the results difficult to compare with the results of the present study. It seems evident that integrated instruction would be better for learning, but, as far as measuring the effectiveness of a method – it could be argued that more factors could affect the longer studies and in a way make them more unreliable.

The Bozorgian study from 2011 is another study that has received some attention, that is, it is mentioned in some other studies that are in this discussion. He measured the effect of four 70-minute long lessons where the metacognitive strategies advance organization, directed attention, selective attention and self-management were taught to 28 Iranian male (17-24 years old) high intermediate English students. The results of the study show that the students made progress, the less skilled listeners more so. There was no control group in this study. The Bozorgian study, and its results, are very similar to mine, but, where he comes to the conclusion that the method is effective, I cannot, since the control group of my study had similar results, indicating that just listening more and answering different kinds of questions would give the same results as listening with high scaffolding metacognitive strategy training.

Two other small scale studies that both state that this type of instruction may have a positive effect on listening performance are the Khonmari & Ahmadi study from 2015 and the Farhadi, Zoghi & Talbei study, also from 2015. Khonmari & Ahmadi (2015) investigated a group of 40 female Iranian intermediate students of English from a language institute. The 40 participants of the study were selected from 80, through the results on the PET, to ensure the homogeneity of the study-group. The study was conducted with a treatment group and a control group with 20 women each. The treatment took seven lessons to complete. The students were given a TOEFEL and then an MALQ before starting the intervention. The intervention was not based on Vandergrift's seven-step model, but on CALLA, which is very similar. The students were asked to do homework assignments as a part of the intervention.

The results were followed up by listening logs. Their results show that there was a statistically significant difference between the means of the pre-and post-tests of the treatment group, but, that there is no statistically significant difference between the pre- and post-tests in the control group. The control group treatment allowed the students to listen to the same texts, twice, but they had no formal discussions about strategy use. Regarding the MALQ the results of the study show that there was a slight increase in most of the strategies in the treatment group. The results of this study, that is very similar to mine in many ways, is the fact that the control group did not have the same success as the control group of my study did. Both of our studies are too small to draw any generalizable conclusions, but, it would be interesting to know what it was in the difference of the control group treatment that gave such different results. Was it the third listen? The written manuscript that gave ample opportunity to monitoring towards the end?

Almost simultaneously with Khonmari &Ahmadi, Farhadi, Zoghi & Talbei (2015) conducted a very similar study, with the hypothesis that “the instruction of metacognitive strategies does not have any impact on listening performance” (142). They selected 60 women, from the original 80, by the results of their PET scores, to ensure homogenous groups. The 60 were then randomly selected to a treatment group and a control group. The same CALLA method of instruction was used. The participants filled out a pre- and post-MALQ. The same type of t-tests on the mean results of the tests were used as in the previous study, and the results are similar. The use of Chamot and O’Malley’s strategy training “can play a significant role in in enhancing intermediate EFL learners’ listening comprehension (144). The results of both the Khonmari &Ahmadi and the Farhadi, Zoghi & Talbei studies from 2015 are based on all female subjects. Again, this raises the question of whether this kind of instruction is more suitable to some audiences than others. It cannot be disregarded that the results of homogenous groups of women who voluntarily study English, could differ from the results of a mixed gendered group of younger students who are not certainly voluntarily in school, and who do not study English from their own choice, just on account of the difference of the groups.

Al-Alwan, Asassfeh & Al-Shboul made a larger study, of 386 10<sup>th</sup>-grade EFL students from public schools in Amman. The participants were on average 16-years old. Their study explores the relationship between listening comprehension and metacognitive awareness, by use of a language comprehension test and the MALQ. The result shows that there is a statistically significant correlation between listening comprehension and an overall high score on the MALQ, in fact, they show that 56% of the total variance on listening comprehension

can be explained by the MALQ results, however considering the rather large standard deviations on the subscales, these results should be read carefully. These results are interesting in comparison to mine, as my results show that my subjects almost invariably scored higher on the final proficiency test, but, the mean scores of the MALQ sank, on all subscales except mental translation, where a decrease was desirable. This undoubtedly raises questions on implementation of the metacognitive strategy instruction and, how statistics are used in this kind of studies.

The Wang study (2016) investigates 100 (n=45 treatment group, n=55 control group) first year university students at a “key university in Mainland China” (81). The students had 9 years of previous English studies. The intervention lasted for 10 weeks, for one hour a week. The pedagogical cycle is based on Vandergrift’s model. The tests were followed up by reflective listening journals, and the entries were coded into three categories; person, task and strategy knowledge. The results of the proficiency study show that both the intervention group and the control group increased their scores, and, there was no statistically significant difference. Wang still claims that the significant gains on the post-test may indicate a positive impact of the pedagogical cycle since the listening journals show an increase in strategy use in the treatment group. There are many similarities with my results and those of Wang. My subjects too showed some improvement in expressing strategy use, however, more research is needed in the field of conscious and unconscious strategy use, and, since listening is such an internal affair, more research is needed into what strategies are actually used by students, and what the students can express with words. Letting them write dairies may just show how much or how little they enjoy writing.

Finally, Movahed (2014) did a study on the effect of metacognitive strategy instruction on listening performance, metacognitive awareness and listening anxiety. The participants were 55 (homogenized from an original 65) (n=30 experimental, n=25 beginner students of English at a University in Iran. The students received 8 sessions of strategy instruction based on Vandergrift’s seven-step model. The results show that the experimental group significantly outperformed the control group on the post TOEFL test, and that the anxiety levels of the students decreased significantly. The results of this study in comparison to mine are interesting because Movahed has investigated the students’ level of anxiety, and he has done so, during a pedagogical sequence that is similar in length to the one used in my study.

### 3. Method and Design

In this section, I describe the context of this investigation, in addition to the methods used for the collection and analysis of the data. I have designed two intervention sequences, one experimental, testing the seven-step model suggested by Vandergrift and, by the Swedish National Board of Education on their web-site, and one traditional sequence, where students listened to aural input and were asked to individually answer questions.

The students were administered both pre- and post-tests (section 3.4.1 below) and they were asked to fill out the MALQ (section 3.4.2 below) both prior to and after the intervention. After all the listening classes students of both groups were asked to fill out anonymous questionnaires about how they perceived their learning experience and their levels of anxiety (section 3.4.3 below). The students of the treatment group also filled in think-aloud protocols as a part of their instruction (section 3.5.1 below).

#### 3.1 Site and Participants

This investigation was conducted at a small upper secondary school in two groups of first year students of English 5. Since the study was conducted before the students' possibility to change schools freely had closed, there were some changes to the original groups.

The participants of the treatment group are described in the figure below:

Number of students that did pre- and post-tests and MALQ	Number of students who were absent during one of the tests	Number of students who started in the class after the initial test - but that participated in the lessons and answered the questionnaires
n=16	3	3

The average grade in English of the test group (from column 1) was: 12.8<sup>2</sup>

Figure 9: Participants of treatment group.

<sup>2</sup> The grades A generates 20 points, B 17,5 points, C 15 points, D 12,5 points and E 10 points. The grades were added and divided by the number of participants.

The participants of the control group are described in the figure below:

Number of students that did pre-and post-tests and MALQ	Number of students who were absent during one of the tests	Number of students who started in the class after the initial test - but that participated in the lessons and answered the questionnaires
n=21	5	3

The average grade in English of the control group (from column 1) was: 15.6

Figure 10: Participants of the control group.

The participants of the study then are for the quantitative study  $N=37$  ( $n=16 + n=21$ ) and for the qualitative study  $N=51$  ( $(n=16+3+3) + (n=21+5+3)$ ).

This study was conducted within the framework of the course English 5. The treatment-group studied English 5 as a part of a vocational program, the control group studied the course within a program that is preparatory for further studies. The course English 5 is mandatory for all upper secondary school students. The completion of the course is needed for the students to graduate. The knowledge requirements for the course within the different programs are identical, even if the course content may differ. The students were informed that they participated in a study, and that their test-results and questionnaires would be used for research. No students in either of the groups declined to participate.

### 3.1.1 Homogeneity and Comparability of the Treatment- and Control Groups

The participants of the study were selected through convenience sampling, I had to ensure the two groups were really comparable to each other. There was a difference in the previous English-grades of the students in the classes, as can be seen in the figure below, and the control group had achieved higher grades in English in grade 9. Thus, there was a concern that the groups would be heterogeneous and their results would be difficult to compare, without having to factor in their original differences.

Grade	Points per grade	Treatment group: n=16	Control group: n=21
		number of students who received grade	number of students who received grade
A	20	0	4
B	17,5	0	4
C	15	9	8
D	12,5	2	3
E	10	5	2
Mean:		13.1	15,2

Figure 10: Previous English grades of the students.

Since the students were not tested beforehand, a Levene's test<sup>3</sup> was executed on the results of the pre-tests of the groups. The test performed showed (see Figure 11 below) the null hypothesis could not be rejected, since the difference in variance was not significant. The p-value was calculated at  $0,926447 \geq \alpha=0.05$  so the variances could be assumed to be homogenous with 95% certainty. This homogeneity does not mean that the groups are equal in proficiency, but that their bell curves are similar, and they both have a normal distribution of the scores: so, the groups are similar in composition and thus comparable.

Single factor ANOVA

S				
<i>Groups</i>	<i>n</i>	<i>Sum</i>	<i>Mean</i>	<i>Variance</i>
Treatment group	16	82,25	5,140625	13,52891
Control group	21	110,2857	5,251701	12,5311

ANOVA	
<i>p-value</i>	
0,926447	

Figure 11 – Results of the Levene's test on the pre-tests of the treatment and control groups

### 3.2 Why Action Research?

This project was driven by a desire to explore the possibilities of a recommended teaching method. Moreover, one of the aims was to gain an authentic situational understanding of how the method may work in the specific setting of this study, but equally importantly, to test this theory of listening strategy instruction suggested by Börjesson on the Swedish National Board of Education's website, in the Swedish setting and potentially initiate a change in the way listening is taught in Swedish language classrooms.

Action research, according to Cohen, Manion and Morrison, is a "powerful tool for change and improvement at the local level" (2005, p. 226). The goal is not to create positivistic evidence that is applicable to all students and situations, but to give other practising teachers a chance to see what was done, what theories were behind the actions, and what the effects were for the teacher as well as for the students. The goal is that practising

---

<sup>3</sup> Levene's test for equality of variances tests the null hypothesis that the population variances are equal, that is, it two (or more) groups are similar in composition with regards to some specific variable, in this case listening comprehension. Since the t-tests I have used later in the study assume homoscedasticity this was done on the results of the pre-tests.

teachers can use the results of the study. The results then, go beyond the mere statistical evidence of the pre- and post-tests and the MALQs. Siegel (2015) suggests that action research carried out by a fellow teacher, rather than by a researcher, that may not be in tune with the present realities of teaching, may be easier to take to heart (p. 78).

There are two main reasons for conducting action research, 1) improving workplace practice and 2) advancing knowledge and theory (McNiff & Witehead 2005, p. 3). As was stated above, the results can be used by practising teachers, and as well, the results add to the existing body of data within this burgeoning field of study. The experimental study is similar to other studies (e.g. Coscun 2010 and Vandergrift & Tafaghodtari 2010) so that the results of the studies can be discussed, and the differing variables can be identified.

### *3.2.1 The Mixed Method Approach*

As quantitative (tests and questionnaires) as well as qualitative (students' think aloud protocols) data is explored in this study, a mixed method approach has been used (Efron & Ravid 2013, p. 47). The mixed method approach is philosophically appropriate in this case. Cross and Vandergrift (2015) observe that a mixed method approach that include "a qualitative, naturalistic and/or ethnographic stance may provide deeper understanding of how L2 listening is generally dealt with (if at all), and how L2 learners react to listening pedagogy" (Siegel 2015, p. 324). Since the research questions of this research paper range from measuring effect to seeing if the students think they are learning how to listen, it can be argued that the different methods can interact to provide comprehensive answers to the research question and sub questions:

- How and to what extent will students listening skills, measured by standardized listening aptitude tests, be affected by a four-week period of metacognitive listening instruction?
- I. How will students' perceived strategy use be affected after receiving explicit metacognitive listening strategy instruction?
  - II. How will students' perception of learning be affected by metacognitive strategy instruction?
  - III. How will students' anxiety towards listening be affected with metacognitive strategy instruction?

### *3.3 The Quantitative Method*

In this study three different quantitative tests were used, the pre-and post-tests (described below, there are links to the actual tests in the references), the MALQ (described below, can be found in full in appendix II) and the questionnaires given to the students.

#### *3.3.1 The Pre- and Post-Tests*

To measure the students' aptitude in listening, a parametric test was administered to the students before and after the pedagogical sequence had been performed. Ready-made commercial aptitude tests were used for the study, namely the Cambridge PET test (Preparatory English Test) and parts of the listening part of the Swedish National Test for English 5. There are several benefits with using commercial tests: they have been piloted, refined and standardized (Cohen, Manion & Morrison 2005, p. 319). Since the pre- and post-testing is an essential part of this experimental study, a test with high validity and reliability that was at the right level of difficulty was needed. The National test suits the level of the students, according to the Swedish National Board of Education, and the Cambridge PET test is rated to be at the CEFR (Common European Framework of Reference) B1 level (low intermediate).

At the beginning and at the end of the intervention, changes in listening achievement were measured by using versions 1 and 2 of the PET, Preliminary English Test, which is a part of the Cambridge test battery. The PET was selected as it meets the criteria stated above. Since the test is available in two versions, the idea is that the tests are both of similar if not identical difficulty, although no information on the exact internal consistency of the tests is posted on the Cambridge website. Finally, the PET was chosen since it is at a CEFR B1 level (low intermediate). According to the Swedish National Board of Education, the course English 5 is equivalent to level B.1.2 (the upper level of B1) (*Skolverket* 2011d). As this investigation took place only three weeks into a year-long course, the aim was to stay at the lower spectrum of the course to avoid creating more anxiety about listening and listening instruction.

To challenge the students who might already be strong listeners, one part of the listening section of a Swedish National Test was also added as it is a test just like the one the students meet at the end of the English 5 course and represents the higher end of the spectrum. The maximum score on this test was 16 (total maximum score 42). The listening texts of the national tests are longer in duration than the texts of the PET, and consist of multiple choice

questions as well as some short answer questions. Two similar tests from the Example tests that are published on the Swedish National Board of Education's website were chosen. The subtests from the National tests require students to listen to interviews, with realistic language. Students are to understand linguistic information that is unequivocally included in the text, as suggested by Buck (2001), although, in a few items the students needed to make inferences to answer the questions. The exact data regarding the tests' internal consistency is unknown, but as the Swedish National Board of Education uses these tests to evaluate students' aptitude and vouch for their validity and reliability on their website.

The pre- and post-tests were not similar at all to the texts that were part of the listening instruction. The reasoning for using a parametric test, (Cohen, Manion & Morrison 2005, p. 317) such as the national tests of Sweden, or the Cambridge exams, was that I was not aiming to measure the students' ability to recollect any information. The goal of the pre- and post-testing was rather to measure the students' overall listening ability, and their abilities to take on listening tasks about a variety of subjects.

Even though the Swedish system calls for criterion-based testing and assessing, a norm-referenced, rather than a criterion-referenced test was used. In essence, these tests were not given to the students for the purpose of grading them, but rather to measure their increase in aptitude, if any. The aim of the pre- and post-tests was also not to test what strategies students employ and how efficient each strategy was (cf. Cross 2015, on metacognitive instruction vs. strategy instruction), but to measure these possible changes in aptitude.

The tests from the national tests are published and accessible to anyone on the Swedish National Board of Education's website. For that reason, the risk is that one or some of the students might have heard the tests before. When asked, one student in the control-group said that he or she had done the test before; it was the pre-test. This student did, however, not outperform the others, and the results are a part of the n=21 group. None of the students of the treatment-group reported to having seen the test before.

The tests took approximately 65 minutes to complete. All texts were played to the students twice. Buck (2001) argues against this in as it does not resemble real-life listening; however, test users have come to see this as standard practice. Listening twice also allows the students to apply the strategies that learned during instruction.

The tests have all been scripted: in the first test, there was a certain lack of false starts, hesitations and phonological modification, such as assimilation, elision and intrusion (see Buck 2001, p. 33). The speakers in the dialogues, however, aim at authentic, but clear language. The speakers are animated, so different prosodic features can be picked up. The

speech rate, and the level of authenticity increases over the test and, from the PET to the National test. The PET is in RP (received pronunciation) and the national tests were in RP (pre-test) and GA (General American) (post-test).

Since some of the tasks require reasoning, it is questionable if they only measure listening comprehension, and not also cognitive abilities. However, the students, through this double focus are able to use the strategies they have been taught. In the tests, the students have to process language automatically and efficiently. The shorter tasks in the beginning allow for more controlled processing. For a detailed description of the tests see Appendix II.

The parts of the test taken from the national test were similar, but not identical in structure, and they did not have the same final score, they differed by one point. To be able to use these two tests to compare the students' results with a paired T-test, the scores had to be normalized. To solve that problem, I decided to exclude question 6 on the NT post-test, a short answer question where many of the students had misinterpreted the question. Three students from the treatment group and six students from the control group had one point subtracted from their overall test score on the post-test.

### *3.3.2 MALQ - Metacognitive Awareness Listening Questionnaire*

To see if the students' perceived use of strategies and processes related to L2 listening comprehension increased after the four-week instruction period, a questionnaire was used: the MALQ - Metacognitive Awareness Listening Questionnaire. The MALQ was constructed by Vandergrift, Goh, Mareschal and Tafaghodtari to find out whether metacognitive awareness has positive effects on language learners' listening development (Vandergrift et al, 2006, p. 432). The questionnaire and its development is discussed in detail in Vandergrift et al (2006) but for the purpose of this paper, it suffices to state that the questionnaire is reliable, has strong psychometric properties and, was specifically "designed for researchers and instructors to assess the extent to which language learners are aware of and can regulate the process of L2 listening comprehension" (Vandergrift et al. 2006, p. 432). In the present study, the questionnaire was used as a pre-and post-test to measure if the instruction had any effect on the students' perceived use of strategies.

The MALQ contains 21 statements where students can grade their answers on a 6 graded Likert scale from Strongly disagree (1) to Strongly agree (6) (see Appendix: II). In the MALQ, there are 6 items representing problem solving (5, 7, 9, 13, 17, 19), 5 items representing planning and evaluation (1, 10, 14, 20, 21), 3 items representing mental

translation (4, 11, 18), 3 items representing person-knowledge (3, 8, 15) and 4 items representing directed attention (2, 6, 12, 16). Items 3, 4, 8, 11, 16 and 18 have reverse coding, so that a high score is not always what is desirable. The items representing mental translation (all reversed) should be particularly noted here, as mental translation is not considered an effective skill, so high scores, or an increase of use of this specific strategy is not desirable. Students were given as much time as they needed to fill in the questionnaire.

The first time the students filled it in, the questions were read to them out loud, and after each question, an explanation was given if needed. The questionnaire was given to them in English, and since the student's proficiency in English was unknown to the researcher at the time, this had to be done to ensure that the students knew what they were answering. In the treatment group, the questionnaire was preceded by a discussion on how they listened - if they were aware of any strategies that they used. The second time the students filled in the questionnaire, it was just before the post test, that is, during the same class. The control group also filled in the test before doing the post-test.

All students wrote their names on the questionnaires, so that each student could see if there had been any change after the six listening lessons. I stressed that there were no right or wrong answers, so that the students would not feel like in a high stake listening situation.

The MALQ contains questions about five distinct factors related to L2 listening comprehension success, discussed above in section 2.3.2. In the analysis of the results, the questions about each of these factors are separated, and high or low scores in each factor suggests information about the students perceived abilities within this factor. A number of the questions are inverted, so that a high score is not always most desirable.

### *3.3.3 Anonymous Questionnaires (3 Questions)*

The third and fourth of the research questions were about the students experiences of the listening tasks, the students were asked if they thought they were learning how to listen, and, how they felt about listening. To answer those questions, I asked the participants of the study (both of the control group and the experimental group) to fill out an anonymous questionnaire after each session. The questionnaire was comprised of three questions (see Appendix III). The aim of this questionnaire was to collect data on whether there were any differences between the groups both with regards to the students' level of anxiety about listening, and whether the students perceived that they were learning how to listen through the method that was applied to them. The students of the experimental group were asked the questions twice,

as the questions were also in their named questionnaires to see if they would answer differently if they had to give their names.

McNiff and Whitehead (2005) state that one should be careful of using questionnaires as they “are notoriously hard to construct” (p. 65). This questionnaire has not been piloted or tested in any way before it was used in the study. However, since the aim was to know the students’ experience of the classes right then, I decided that the best way was to ask them what they thought. I wanted the sentiments of the whole group rather than those of just a few students, which is why I opted for the questionnaire rather than interviews. Even though the validity of the questionnaire could be questioned, and the sample is very small, it may be possible to see trends that may have had an impact on the other results. This is further discussed in section 5.

### *3.4 The Qualitative Method*

The qualitative study was comprised of the students’ think-aloud protocols from the lessons and the teacher’s field notes.

#### *3.4.1 The Think-Aloud Protocols*

The students in the treatment group were asked questions that they answered at the end of the work-sheet for the class. The students were asked 4 free-text answer questions;

- 1) How did you think during the first listen? Was your strategy successful?
- 2) How did you think when you listened to the text the second and third time? Did you listen differently? If so, how?
- 3) How did you feel about listening today - describe in your own words:
- 4) What strategies will you use next time? Use the strategies from the list, or, use your own words to describe what you will do.

I asked the first and second questions to see if the students could identify their strategy use. The lessons led the students through the aforementioned pre-listening activities, where they were told the subject matter of the text, and what text-type to expect- Then the class brainstormed together to gather all the information possible about both issues. Following the brainstorm, I asked the students to try to predict the text, vocabulary as well as content. The

second listen was a chance for them to monitor their predictions, the third time, the students could monitor their comprehension as they had a written transcript of the text at hand.

I asked the third question to see if there would be any difference in trends between the anonymous questionnaire and the named lesson material. The fourth and final question was designed to remind the students of the strategies that we had been working with, to offer another chance to look at the list of strategies (and the model of listening) they all had received (digitally, so they all had it accessible at all times).

These student protocols, as well as their written answers gave a fairly clear picture of students' general language proficiency, willingness to communicate in English and motivation to participate. They did not give so much information on students' strategy use, which, however, in some ways is corroborated by the results of the MALQ.

#### *3.4.2 Field Notes*

From the systematic field notes I gathered two types of data, descriptive and reflective. After each class I filled in a semi-structured observation form. The notes were short, as suggested by McNiff and Whitehead (2005, p. 65), and the reflections were about student behavior that stood out, or, the teacher's behavior and how it might have affected the students. These reflections serve as a compliment to the student questionnaires that were filled in during each of the classes, to capture the students as well as possible (Cohen, Manion & Morrison 2005, p. 306). Finally, no hypothesis was tested in this study, so the observations were rather to see if any connections could be made between the class and the material the students were subjected to (Cohen, Manion & Morrison 2005, p. 306).

### *3.5 Classroom Treatment Design*

During six of the nine (eight for the control group) classes of the four-and-a-half-week intervention, students listened to an authentic-like text, related to the topics in the *Core Content* in the curriculum. The text types differed from informative texts, lectures, and radio interviews. The texts were gradually both more difficult and longer. The first three texts were from the BBC, the speech was in RP and Australian English -- and the three latter texts were from the CBC, and in Canadian English.

### *3.5.1 Treatment Group Intervention*

In the experimental group, the procedure was basically the same during all six listening classes. The three (two for the control group) remaining classes of the intervention consisted of testing and filling out the MALQs and consent forms, as well as the introduction lesson, where the treatment group received information about the listening process and discussed strategies and their use as an incentive to take in the instruction. They were also informed that strategy use was one of the knowledge requirements for the course.

The material for the listening classes in the intervention group follows the 7-step pedagogical circle suggested by Börjesson (2012). Students received a hand-out, new material was given each time, so there would be no problems with students forgetting their things. At the start of each class the students were asked to remember the strategies we had talked about earlier (focusing on the strategies measured in the MALQ, and on top-down and bottom-up strategies). Students were also reminded of the Anderson model of listening (see Figure 1). Before the work with the hand-outs started, students were informed of the topic of the text, and the text-type and the fact that they would listen to the text three times.

1. Based on topic and text-type - students were the asked to brainstorm first on what (information, words etc.) could be expected from the topic, then from the text-type (speaker(s), order of information, linking words etc.). For two lessons, the entire class brainstormed as a group, and thinking behaviour was modelled to the class, as well as note taking (on the whiteboard). It turned out to be very difficult for the students to guess, so they were given much support. Then for two classes, the students were allowed to brainstorm in pairs, and finally, for two classes they were asked to try brainstorming individually.
2. After the predictions, we listened to the text for the first time. The students then marked the information that they had guessed correctly in their predictions, and, they were allowed to jot down any additional information that they understood during the first listen (in Swedish or English, whichever the students were comfortable with).
3. After the students had marked what they had guessed correctly, they were asked to compare notes with a partner. The students switched seats each time, so that they would not sit with the same partner every time, to reduce the risk that sitting with the same partner would change the outcome for any students: for example, if a less proficient student was always paired with a more proficient student – that student could feel insecure, and be reluctant to guess, and just wait for the more proficient

students to provide the answers. During the final two classes the students were asked to reflect on what they understood individually: they did *not* compare notes. The students could, during the first 4 classes, discuss what they had understood, points of confusion, disagreements and give each other logical explanations of why they had heard it right (the thinking behaviour was modelled by the teacher by way of thinking-out-loud during the first two classes). The pairs should also write down what parts of the text they needed to listen to more carefully during the second listen

4. After comparing notes, students were introduced to the assignment - with questions to answer. They could then monitor their comprehension and see where they had to direct their attention. Then they listened to the text a second time. Afterwards the students had some time to enter the new information on their answer sheets (everything was done by hand, with pen and paper, to avoid the disturbance of the computers in class), the class was asked to discuss the texts, during the first two classes, out loud in the group, then in pairs, and finally, by writing the answer to the question – “What was this text about?”. This was done primarily to get the students to confirm their comprehension and to share how they had managed to succeed in comprehending the texts (what strategies they had used).
5. A final, third listen allowed students to verify their perceptions, to listen for information they now knew was in the text, and to hear where they had possibly misconceived information. They were also asked to pay close attention to parsing and perception and, to read/listen for clues to why they might have missed any information.
6. The students finally filled in a personal reflection on the activity – “How did they feel about listening today? Why? Was this activity helping you learn how to listen in English? What strategies were used during the three listens?. Was the text too easy? Too difficult? Why? Lastly, the students were asked to fill in: “What strategies would they try to use the next time?”.

### *3.5.2 Control Group Intervention*

The control group did not have the initial lesson with information about how listening works, which meant that they had only eight sessions of listening exercises and testing. However, they filled out the MALQ before starting with the listening sessions. In the MALQ, the

listening strategies are apparent, but there was no class discussion on strategy use. The control group also listened to the same texts three times.

1. Students in the control-group were asked to fill in a traditional short answer test, as well as questions where they were asked to summarize information and show if they could make inferences with the motivation that practice would enhance their listening skills. The tests were made according to Buck's (2001) guidelines: the questions for listen one and two were short answer questions, multiple choice questions or gap fills, where the questions were phrased differently from what was said in the text, so that comprehension could be deduced (rather than the students just scanning and copying language even if they did not understand).
2. After the first listen, the students had approximately two minutes to give the answers to some control questions, and read the questions for the following listen.
3. We listened to the same text once more, and again, the students were given a few minutes to fill out the answers to the questions. During the two first classes, they were not allowed to go back and correct if they heard anything new: they were to focus only on the questions on the second paper. Again, they were also asked to read through the questions for the third listen. These questions were of a different nature, they required longer answers, the students were asked to summarize information or give opinions for example.
4. After the third listen, they were given as much time as they needed to write the answers to the questions and to fill out the anonymous questionnaire regarding how they felt about listening on this day.
5. When the test had been handed in, the class was given oral assignments that had to do with the topic that had been talked about in the listening text. For example, they were asked to do job interviews and act/speak appropriately for the situation; this, rather than listening strategies, became the focus of the classes. In that way, the control group also received instant feedback on how much of the texts they had understood. The control group was told they were in the experiment, but they were not made aware that they were a control group until afterwards. There were no discussions on strategies what during the sessions of the control group, cognitive, metacognitive or socio-affective.

The students were not given their test-results from the six listening classes during the testing period, unless they asked for them and, that was rare. They were only given the results of the pre- and post-tests, and, a number indicating how much they had improved.

### *3.6 Selection of Listening Material*

Kilic and Uckun (2012) conducted a study on how text-types affect learner anxiety in L2 listening. They investigated three text types -- informal dialogue, lecture and radio talk show - - and found that informal dialog created the least anxiety in students, while listening to lectures and radio talk shows created increasingly more anxiety. For that reason, the first part of the proficiency test, the PET was deemed appropriate as it consists of informal (but scripted) dialogue. The very first of the BBC texts was a lecture, with streaks of informal dialogue, and then the difficulty escalated with radio segments with increasing difficulty.

Kilic and Uckun (2012) claim that authentic texts tend to cause more anxiety in language learners, but authentic listening material should nevertheless be used in instruction to prepare students for real-life listening (see Buck 2001, Field 2008, Vandergrift & Goh 2012). This is the reasoning behind selecting the last, rather difficult radio interviews.

### *3.7 Validity and Reliability*

The study I have conducted is completely replicable in design. All the lesson material for both the control-group and the treatment group has been saved. The construct of the research is somewhat controlled, and replicable. Siegel (2015) suggests that the notions of validity and reliability were drawn from other research paradigms and that some flexibility should be given to action research (81). Instead of the standards of validity and reliability of the positivistic research paradigm he suggests factors such as dependability, credibility, confirmability and transferability to be more appropriate matters to be concerned about (82). Additionally, the real-life aspect of action research makes all variables difficult to control.

There are many types of validity and reliability, and it would be naive to think that all threats to validity and reliability could be erased (Cohen, Manion & Morrison 2005, p. 105). The validity and reliability of some of the data collection instruments for the quantitative study, such as the pre- and post-tests and the MALQ have been addressed above in sections 3.4.1 and 3.4.2. For the questionnaires, that have not been piloted, and whose answers have only been interpreted by myself, the validity is lower even if the questionnaires are completely replicable.

To be able to reach the most comprehensive answers to my research questions quantitative results of the study more than one way of collecting data to answer the research questions have been used. The data has also been collected from different perspectives (teacher/student, formal testing/informal filling out of questionnaires, questionnaires/ free text questions).

In both the quantitative and the qualitative investigations, the extraneous variables of the teaching situations have been considered. Extraneous variables are variables that are not within the control of the teacher/researcher: time of day, how many classes the students have had before the listening class, or disturbances outside the classroom. Those variables have been documented in the field notes for each class. With regards to the qualitative data, I aimed at what Efron and Ravid (2013) call disciplined subjectivity. Preconceived notions about the results of using the method were acknowledged, and biases were monitored through on-going self-reflection with respect to setting, researcher's actions in the classroom and participants in the interpretation of the data (Efron & Ravid, 2013, p. 71).

The field notes were semi-structured to ensure comparability. The validity of the students' listening diaries is low, however. The students had a difficult time expressing themselves. Even if they were allowed to switch to Swedish, many of them did not see the point of the questions, which was reflected in their answers.

### *3.8 Ethical Considerations*

Before commencing the data collection, I was granted permission to conduct this research by the headmaster of the school. The problem with using one of the groups as a control group was discussed. Using a control group which is denied a treatment that is hypothesized to lead to higher results was problematic. The control-group were given a more traditional extensive listening approach, and it was decided that they would receive metacognitive instruction at a later point, which was possible since the trials took place at the very beginning of a year-long course.

According to the Swedish Research Council (*Vetenskapsrådet*) there are four main principles to protect individuals who participate in studies: the information requirement, the consent requirement, the confidentiality requirement and the rules regarding the use of personal information collected through the investigation.

The participants were informed of the object of the study, that it was completely voluntary for them to participate, and that they were free to end their participation at any time during the study without repercussions. Before the investigation started the students were

given information about the study, about the aims of this study and about their roles in it. The students were not told about the control group, but that they would receive two different types of listening instruction, and that both groups' progress would be measured.

There was no additional work for the students. The students all had to do the work for the course, but they were still given the right to decline participation at any time during the study if they did not want the results of their tests, or their actions in the classroom on record. They were also informed that they would be given a copy of the paper to read, and discuss with me before its handed in for publication.

The participants had to give their consent to taking part in the study. In connection to the oral information given in class, I handed out information letter, which the students were to sign and hand back to me. Parents were also informed of the study. No student declined participation.

Confidentiality of information as well as the anonymity of the participants was maintained throughout the study and in the paper. The test results, MALQ results, my observation sheets and the students' lesson material were all kept in my computer or home office at all times. No names have been used in the study, even if the students had to name their tests and lesson materials so they could get feedback on their progress. Neither the students nor the school can be identified through the published results of the study.

No personal information that was given for the purposes of this study will be used for any other reason than the study itself. The results of the tests and the lesson material, however, will be used in assessment as they are a part of the students' course.

Thus, all of the principles outlined by Swedish Research Council (*Vetenskapsrådet*) have been considered throughout the study.

## 4. Results

The aim of this study was to find out if and how students responded to metacognitive listening instruction compared to a control group. To get a more holistic perspective on the results of the pre- and post-tests I also worked with sub questions on the student's perception of the method.

- How and to what extent will students listening skills, measured by standardized listening aptitude tests, be affected by a four-week period of metacognitive listening strategy instruction?
- I. How will students' perceived strategy use be affected by explicit metacognitive listening strategy instruction?
  - II. How will students' perception of learning be affected by metacognitive listening strategy instruction?
  - III. How will students' anxiety towards listening be affected by metacognitive listening strategy instruction?

The presentation of the answers to these questions is divided into subsections primarily focusing on the research question and then at the sub questions. The results are discussed in section 5.

### *4.1 To What Extent Will Students' Listening Skills be Affected by a Four Week Period of Metacognitive Listening Strategy Instruction?*

#### *- Quantitative Data*

After having ascertained that the results of the pre- and post-tests of the treatment- and control-group were comparable through the Levene's test<sup>4</sup> (see 3.2.1), I compared the data from both sets of tests. Both the treatment- and the control-group had increased significantly on the post-tests.

---

<sup>4</sup> Levene's test for equality of variances tests the null hypothesis that the population variances are equal, that is, it two (or more) groups are similar in composition with regards to some specific variable, in this case listening comprehension. Since the t-tests I have used later in the study assume homoscedasticity this was done on the results of the pre-tests.

#### 4.1.1 Pre- and Post-Testing, Treatment- and Control Groups

The results of the pre- and post-tests indicated that both groups had made statistically significant progress after the intervention period. The progress of the treatment group was not shown to be greater than that of the control-group, with statistical significance. The answer to the first research question, no, the treatment-group did not outperform the control-group after the intervention.

In figure 12 below, the boxes represent the second and third quartile of the results of the treatment-group. The line in the middle is the median score and the two lines at the ends represent the highest and lowest scores of the group. The highest scorer did not improve much. The fact that the top scorer almost had a perfect score did not leave much room for improvement. However, the lowest scorer improved significantly, as did the two middle quartiles. The figure below visualizes the results and the improvement in the treatment group. The mean score increased with  $32,63 - 26,63 = 6$  points.

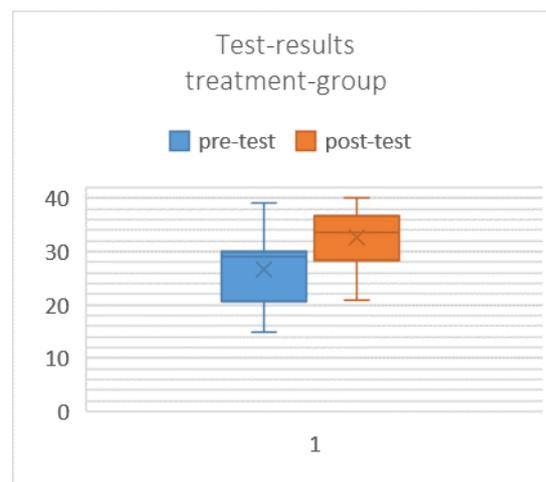


Figure 12: Box diagram of the results of the treatment group's pre- and post-apptitude tests.

To determine if the students' listening skills had improved at a statistically significant level, a paired two sample t-test was employed on the students' scores derived from the pre- and post-tests (see Appendix IV). The test showed that it could be concluded that there was an increase of the mean scores on the aptitude tests after the intervention. The standard deviation decreased from 6,46 to 5,19. The results of the control-group were very similar to those of the treatment-group. As can be seen in the box diagram below, this group too improved their results significantly. The results of post-test of the control-group were slightly higher than

those of the treatment-group. The mean of the treatment-group was 32,63 points compared to 34,76 points in the control-group.

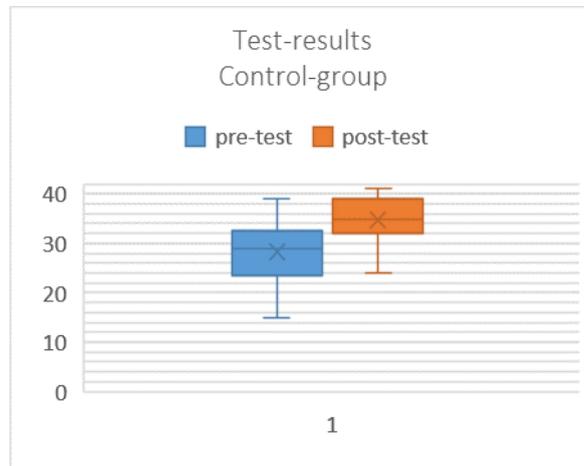


Figure 13: Box diagrams of the results of the control group's pre- and post-apitude tests.

The same procedure with the paired-two sample t-test was performed on the results of the control group. In this group of 21 students, the standard deviation had decreased even more than in the treatment group, from 6,44 to 4,72.

When the mean results of the test-results of the two groups are compared, the improvement of both groups is similar. The mean score of the control group increased by  $34,76 - 28,24 = 6,52$  points.

#### *4.2 To What Extent Will Students' Listening Skills be Affected by a Four Week Period of Metacognitive Listening Strategy Instruction?*

##### *- Qualitative Data*

The results of the field notes should be considered when interpreting the results above. To try to illustrate the difference between the classes and the way in which they have received the different types of instruction, some of the lessons will be briefly presented. Based on the qualitative data, the answer is that the students of the treatment group did indeed improve their results on the listening test with statistical significance, however, they did not improve more than the control group that was subjected to an equal amount of listening. In fact, the control group made slightly greater gains. I have used the field notes to try to discover why the results came out this way. In light of the field notes described below, the level of

motivation in the students and the way the groups were comprised may have affected the results. Additionally, if the students had had prior training in oral expression in general and if we had spent time, for example, playing games where they practised connecting different phenomena, the students of the treatment-group could possibly have made greater gains from the method. Their insecurity and inability to interact with each other was a hindrance to applying a teaching method that in many ways relies on interaction.

#### *4.2.1 Field Notes, Treatment- and Control Groups*

The first lesson of the treatment group was based on informing the students of the investigation, and presenting and discussing the metacognitive strategies and why they might be of help to them. The class was inspired; there were good discussions, where many students participated. Possibly, this lesson was affected by the fact that the principal was present.

During most of the lessons there were some absences, and, the mood of particular students affected the mood of the group in a negative way. During the first class when the first handouts were distributed some student behaviour stood out: one student became worried, and asked if this was going to be graded; another exclaimed that this could not be listening comprehension since the hand out did not look like a test. This illustrates how the students of the treatment group regarded listening comprehension instruction.

In the beginning, the students expressed that they thought the hand out was “messy”, even though it had numbered assignments that were carried out in order. This methodology was obviously new to them, and it took a few lessons for them to get used to the process. This affected the student’s opinion of the method negatively, and I had to reassure them that strategies would be good for them to know.

Another issue that became apparent, was that the students needed training in interaction, before they could fully make use of this method. They had never before been asked to guess about the content of texts they had not heard. They were so worried about being wrong that many were reluctant to participate in discussions at all. Even those who tried to participate, did not know how to speak about the subject. Almost all of the first lessons was spent by me modelling the desired behaviour.

Finally, when the students were asked why they felt like they felt about listening that day, 10 of the 19 students present answered that they were either tired or hungry, which must have, affected their ability to focus on the task. I asked the students at the beginning of each class, what strategies from last time they would try to use the present time. All classes also ended with them reflecting on what strategies they used were successful during the

present lesson, and that they could try to use them again the following lesson. However, a majority of the answers to that question was that they did not remember what strategies we talked about last time. This is indicative of the students' engagement in the instruction. This may also have affected the results.

The mood of the control-group was much more even, and there was not a small group of leading students who affected the mood of the group negatively as there was in the treatment group. However, particularly the weaker students started to complain that it was boring when the texts became more complex, and, most of the group claimed to be bored with listening comprehension tests during lesson 6. In the control group, there were no discussions about strategies, or anything else regarding the listening texts, however, there were never any questions about that.

The absence was similar in the two groups, on average two or three students per class. In both groups, there was one student who missed three of the six occasions. This may be indicative of avoiding-strategies. One student in the control group left school just before the pre-test, due to stress.

### *4.3 How Will Students' Perceived Strategy Use be Affected by Explicit Metacognitive Listening Strategy Instruction?*

The first of the sub questions was investigated through the pre- and post-metacognitive awareness listening questionnaires (MALQ). The results for both the treatment group and the control group indicated that the students perceived using all the strategies except mental translation with less frequency after the intervention. The items representing mental translation should be particularly noted here, as mental translation is not considered an effective skill; therefore, high scores, or an increase of use of this specific strategy is not a desirable result of any teaching intervention.

#### *4.3.1 Results of the MALQ*

In figure 14 below it can be seen that even though the treatment group received explicit strategy training their perceived use of strategies did not increase during the intervention period. However, the variances became smaller in all the strategies, suggesting that the students' answers, and thus likely, the students' way of thinking was homogenized.

### MALQ 1 Treatment group

strategy	Planning and evaluation	Problem solving	Person knowledge	Directed attention	Mental translation
number of items	4	6	3	5	3
mean	4,1	3,6	4,3	3,9	3,7

### MALQ 2 Treatment group

strategy	Planning and evaluation	Problem solving	Person knowledge	Directed attention	Mental translation
number of items	4	6	3	5	3
mean	3,5	3	14,2	3,6	3,8

Figure 14: The results of the pre- and post MALQ of the treatment group

The results of the control-groups questionnaires reveal the same results, that is a decrease in perceived strategy use in all the measured strategies, except for mental translation, the one strategy where an increase was not desirable. Since no explicit strategy training had occurred in this group, these results are not as surprising as in the treatment group. The fact that both groups perceived that they used less strategies could be because they felt more anxious and insecure as the texts became increasingly difficult.

### MALQ1 Control group

strategy	Planning and evaluation	Problem solving	Person knowledge	Directed attention	Mental translation
number of items	4	6	3	5	3
mean	4,0	4,2	3,9	3,6	3,3

### MALQ 2 Control group

strategy	Planning and evaluation	Problem solving	Person knowledge	Directed attention	Mental translation
number of items	4	6	3	5	3
mean	3,6	4,0	3,0	3,5	3,6

Figure 15: The results of the pre- and post MALQ of the control group

Just like in the case of the results from the pre- and post-tests, the standard deviations of the groups' results are so great that, I found no statistically significant difference, neither when it comes to the decreasing numbers or to the increasing numbers. The results to the second research question then is no, there is no statistically significant increase in the perceived use of metacognitive strategies in the treatment group compared to the control group after the 6-lesson intervention period.

#### *4.4 Students' Perception of Strategy Use: Qualitative Data – Treatment Group Only*

The MALQ measured that the students perceived using less of the effective strategies and more mental translation in both groups. For the treatment group, where I have collected additional qualitative data, the results could, however, be regarded as inconclusive. The students' think-aloud protocols could be showing that they use more metacognitive and cognitive strategies than they can express with words. In addition, in so far as they developed their skills during the four weeks of instruction, an increase in strategy use can be detected. However, more research with think-aloud protocols, including the control group, is needed to comprehensively show the students' perception of their strategy use.

The students in the treatment group were asked to fill out think-aloud protocols in connection to the tasks of each of the lessons in the instructional lesson series. At each point in the seven-step pedagogical cycle, I pointed out that they were using strategies and what strategies they were using. I sometimes used the vocabulary of the MALQ, but more often I used a more colloquial language, and sometimes even Swedish, to make sure every student was on board with the metacognitive instruction. At first, I asked the students to fill in the think-aloud protocols in English, however, after evaluating the students' protocols after the two first lessons of the first two lessons I decided to let them write in Swedish if they wanted to. This was an attempt to help students better reflect on what they were doing.

The two first questions of the think-aloud protocol were “How did you listen the first time? Was your strategy effective?” and “How did you listen the second and third time? Were there any differences from the first time?” It should be noted that at the third listen, they had access to a written manuscript. For the purpose of the analysis, the answers have been sorted as follows: the answers that do not specifically answer the questions have been omitted, and the answers that mention any kind of strategy have been divided into categories according to

strategy. Seven strategies had been presented to the students at the beginning of the pedagogical sequence, and several times during the classes that followed; advance organization (e.g. looking at the task), directed attention (listening for gist, top down), selected attention (focusing on details, elaboration), bottom up (focus on details they understand to build understanding), top-down (focus on other knowledge to guess at details), self-management (focus), monitoring (in first listen - how well the predictions fell out), evaluation (how much did they understand?) and problem identification (did they figure out what they did not understand, that they needed to understand?). Evaluation was a tag-on question to the question of how they listened during the first listen, so some students were counted twice, and many students were not counted at all as they did not report using any strategies at all.

	Planning/ Advanced organization	Directed attention	Selected attention	Self- managem ent	Monitoring	evaluation	Problem identification	I used no strategies
Lesson 1 19 respondents	0	1	3	1	0	4	0	3
Lesson 2 17 respondents	4	1	5	0	0	4	0	2
Lesson 3 22 respondents	3	4	0	5	0	8	0	1
Lesson 4 19 respondents	1	2	0	2	2	3	4	1
Lesson 5 18 respondents	2	4	2	2	4	1	2	0
Lesson 6 17 respondents	1	5	3	1	0	1	2	0

Figure 16: Students' responses to: How did you listen the first time? Was your strategy successful?

In these free texts answers, there is much uncertainty. The students simply did not know how to answer the questions. Many students gave answers that did not reply to the question, and they do not know how to express the way they think, neither in English nor in Swedish. The process of listening is still too covert or complex to them. These are some examples of answers to how the students listened the first time, to illustrate what I mean: "it worked and I were just listening to the text" "yes!", "Jag tycker att det var enklast första gången eftersom man läste det innan. det var svårare andra gången för man läste samtidigt. Strategin var lyckad." (I think it was easiest the first time because you had read it before. it was harder the

second time because you read it at the same time. The strategy was successful. My translation), “after my first listen I started to think more about it and it was a success”, “it was good it wasn’t too hard or too easy”.

To be able to illustrate the students’ strategy use in the matrices (figures 16 and 17) I had to code their answers. I have coded the answers alone, which can lead to a greater margin of error than if more than one person had interpreted the data.

If anything can be read from the numbers in figure 16, it is that the students were able to express some change in strategy use. There was a decrease of students who explicitly report using no strategies. Also, at the first text that was more difficult, students started identifying problems, such as rapid speech. During the last lessons, when the texts were both longer and had more rapid speech as well as on more complex subject matter, the students report to using more top-down strategies. They express trying to understand the context, so that they can fill in the details during the next listen: they monitor their comprehension. Therefore, even if the answers have been coded just once, many strategies still can be found: problem identification (text mass), problem solving (finding the right strategy), directed attention (focus on context), planning (planning to look for details in the next step) etc.

The difficulty for the students to express in words how they listen differently became apparent as they answered the second question too. Many students replied that they did not listen differently during the third listen, even if they had access to a written manuscript. “No I listened the same”, “yeah i thik it was easyer the second and thirt time (sic)”, “I were much more concentrated the first time than for example the third time”, “Nej där var ingen skillnad, jag tycker att det gick bra på alla och använde samma strategier” (No, there was no difference, I think all of them went well and used the same strategies. My translation). However, in the students’ answers after lesson six, some change in the way they express their listening behavior is apparent: “The second time you just fill, and you know pretty much. Now, the third time you understand more and it’s easier to listen becauser you see the text. (sic)”, “Det var en stor skillnad. Det gick snabbt 2:a gången men 3:dje gånger kändes det långsammare eftersom man läste/tittade på texten.” (It was a big difference. It went fast the 2<sup>nd</sup> time but the 3<sup>rd</sup> it felt slower becasue you read/looked at the text. My translation.)

I had given the students a direct instruction to practice the relationships between letters and sounds, and listen for cues when words, or sentences began and ended, as well as to for example, parse out tense when they listened the third time and had the text in front of them. Thus, the answers to this question were much more complex to code since most of the students’ express using more than one strategy. Sometimes the students motivated their

answers: “I paid more attention the second time because I had stuff I wanted to solve, the third time I became more unfocused on the text” (my translation) and sometimes they did not: “yes the third time I didn't listen as carefully”.

The matrix below (Figure 17) shows the frequency of the students’ expressions of using strategies when answering the second question: How did you think when you listened to the text the second and third time? Did you listen differently? If so, how? These results were also coded by me alone. Many of the answers expressed using more than one strategy, so these answers were coded twice. The answers and the coding are a matter of interpretation. In addition, this does not show what strategies the students actually used but rather what they were able to express. This leads to a decrease in validity, it is apparent that the students needed more training than they had received in expressing such thoughts. The students of the treatment group were coached during the third lesson, and they were shown their answers and given ideas on how to elaborate and express thoughts. The answers are slightly longer and more elaborate towards the end of the pedagogical cycle, perhaps due to the instruction, or perhaps because they became increasingly confident when they started to recognize the format. The answers in the matrix below should be read only for trends.

	Planning/ Advanced organization	Directed attention	Selected attention	Self- managem ent	monitorin g	evaluation	Problem identification	I used no strategies	I used the same strategies	I read the text
Lesson 1 19 respondents	0	0	8	1	4	2	1	1	3	3
Lesson 2 17 respondents	1	1	8	3	2	0	0	1	1	1
Lesson 3 22 respondents	0	0	8	1	1	2	1	0	4	5
Lesson 4 19 respondents	1	0	4	0	6	4	3	0	2	2
Lesson 5 18 respondents	0	1	7	2	3	2	2	0	2	6
Lesson 6 17 respondents	0	3	7	1	2	2	1	0	1	7

Figure 17: Students’ responses to: How did you think when you listened to the text the second and third time? Did you listen differently? If so, how?

During the second and third listens, it became apparent that the students had listened for gist the first time, and that they focused on the details they had not understood the first time during the second and third listens. This means that the students evaluated their performance during the first listen, even if they had not been able or willing to put that into words. They had also monitored their understanding (by acknowledging that they did not understand everything) and identified the problems (“stuff they wanted to solve”).

There is an increase in students who say they read towards the end of the listening instruction period; many students also express (not shown explicitly in the matrix) that they lose focus as it becomes boring to listen to the text. (“I were much more concentrated the first time than for example the last (sic)”). Some students also express that monitoring is not “an assignment to solve”, so they claim that they have “nothing to do” during the third listen, which makes it difficult to stay focused. This shows that the students are able to self-manage, evaluate their listening efforts and identify problems, however, this is still problematic since losing focus can possibly destroy their other efforts. This leads to the question of what strategies are most important to teach, and to learn.

Some students were reluctant to the instruction. They clearly expressed this at the beginning of the classes, as well as in the surveys after each class. One student reported, “I think that some of them [the strategies] werre good but i don't think that i will use them because it is to much for me to think about (sic)”. During all sessions, there was a feeling that the students were asked to do two separate tasks: learn how to listen, and solve listening comprehension questions. Some students kept answering that they used no strategies, or that they used the same strategies as they always did/like last time. During classes, they voiced their thoughts toward the instruction. However, based on the teacher’s field notes the open reluctance in the answers faded towards the end of the pedagogical cycle.

#### *4.5 Students’ Perception of Learning*

The second sub question “How will students’ perception of learning be affected by metacognitive strategy instruction?” was investigated by way of anonymous questionnaires. Even if explicitly I told the treatment group that I was testing a method designed to teach them how to listen, the students’ answers to the questionnaires, revealed that they did not perceive that they were learning how to listen to a greater extent in the treatment group than in the control group.

#### 4.5.1 Students' Perception of Learning: Quantitative Data, Treatment- and Control-Group

At the first point of measure, after the first lesson, there was a slight difference between the treatment group and the control group. As was stated earlier, the treatment-group were told that they would be taught how to listen, which may have affected their answers. However, in the control group, the most frequent answer was 6 after the first lesson, so they too perceived that they were learning, even though no strategy instruction or other instruction than the questions on the tests was given to them.

As can be seen in the graphs the groups' perception of learning follow each-other fairly closely, something that could be expected since both groups actually improved their scores on the post-proficiency-test at approximately the same rate.

There is a dip in the students' belief that they were learning after lesson 4, the first of the CBC texts that had a new type of questions and the texts themselves were longer and more complex. Even if the students of both groups to a very large extent could answer the control questions, they seemed to perceive that they were learning less when they were challenged with a more difficult text. It would seem that other factors may be at play than just how effective the students believe method is, for example affective factors.

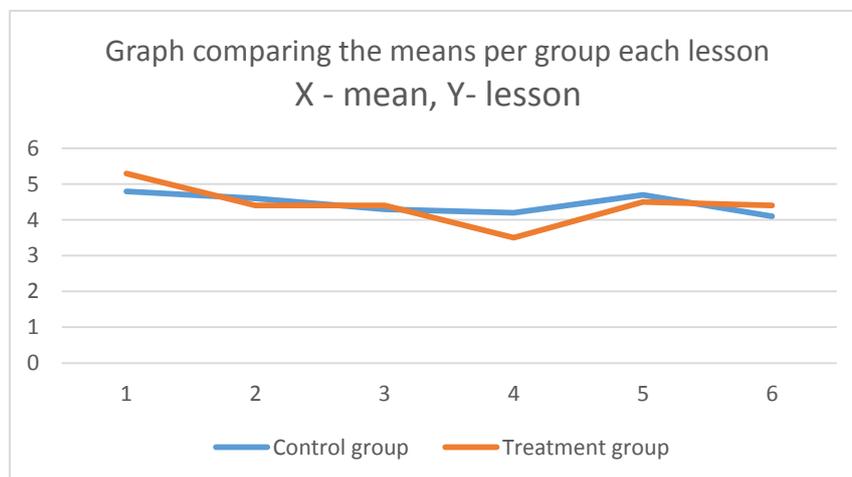


Figure 18: Treatment group- and control group students' answers to the question "In your opinion are you learning how to listen?" from lesson 1 to 6

The drop in the students' faith in their learning how to listen, in addition to their feelings of losing control over the situation when they didn't understand everything in the texts, may be indicative of them starting to become bored with doing the same practice during each class.

After lesson six, the scores of the treatment-group evened out even if text 6 was nearly nine minutes long and of a rather complex nature. None of the students in the treatment group gave the answer 1 or 2 on that questionnaire; 16 out of 20 students gave an answer that is higher than 3. Including the students' sense of boredom and loss of control, they vast majority still sensed that they were learning how to listen. The scores of the control group, decline a little after lesson 6. Of the 25 students who answered the questionnaire after lesson 6, 9 students gave an answer of 3 or lower. From this I conclude that more of the students in the control group seem dissatisfied with how much they were learning how to listen. Since the questionnaire lacks in validity, these results can be only used for comparing the two groups.

#### *4.6 How Will Students' Anxiety Towards Listening be Affected by Metacognitive Listening Strategy Instruction?*

The third sub question regarding the students' anxiety towards listening was also answered by way of anonymous questionnaire. Both groups reported to feeling more anxiety towards the end of the instructional period, however, the increase was considerably smaller in the treatment group. I interpret this data as the treatment group feeling less anxiety towards listening than the control group after the metacognitive listening strategy instruction.

I asked the students to answer two questions regarding their level of anxiety. The first question was; "How did you feel about listening today?" The question was answered on a 6 graded Likert scale (1: I was very anxious and knew it would go poorly - 6: I was not nervous at all; I feel confident in my listening abilities). The results of that question are presented in section 4.7.1 below. The second question was; "Do you know the reason you feel this way about listening today? (the text was difficult, I am tired, I feel very confident because it's easy)". The students were asked to answer in their own words. The students' replies to that question are presented in section 4.7.2 below.

##### *4.6.1 How did you feel about listening today? - Quantitative Data, Treatment- and Control Group*

The students of both groups were quite confident about listening to begin with. The mean score of the treatment group was 4.11 and the control group gave a mean score of 4.88. As the texts became more complex, the student's levels of confidence decreased, particularly in the control group, who received no metacognitive strategy training during their (lack of) listening instruction.

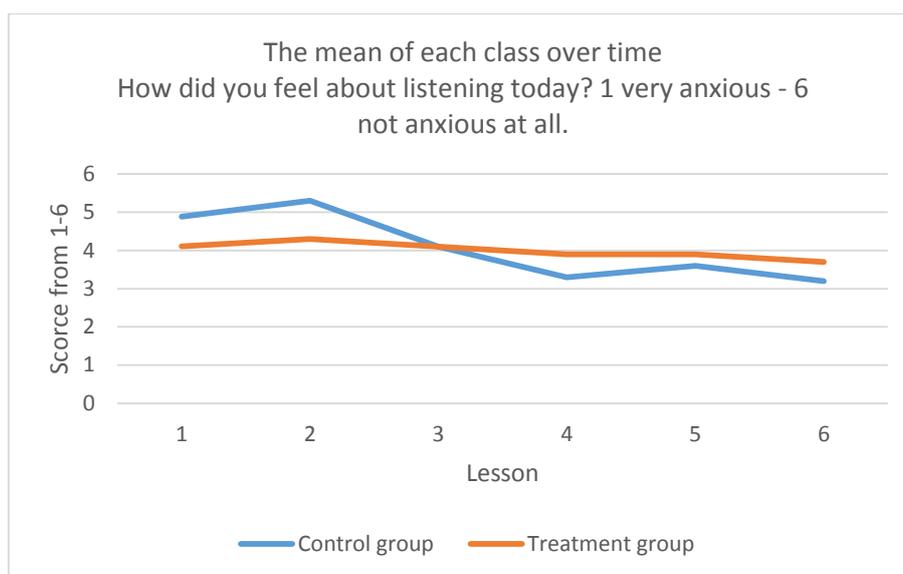


Figure 19 –Development of means in scores of anxiety-levels of the treatment- and control-groups per lesson.

The decrease in confidence is clearly larger in the control groups as can be seen in Figure 19. The students filled in the questionnaires as the last part of the listening lesson. For that reason, the level of difficulty of the text is likely a part of the student’ answers. Had I asked the students how they felt about listening today – before they listened to the lesson of the day, the answers might have been different. How students’ level of anxiety is affected by meeting texts that are too difficult, however, is not a part of this study. Another reason for the students’ of the treatment-group to feel less anxious is that, during lessons 3 and 4, we were still working together, and they did not perceive that they were being tested in the same way as the students in the control-group who were left to answer questions on their own. The students who were allowed to ask questions and monitor their answers felt less anxious – but, they did not perceive that they were learning more than the students of the control-group.

To be able to compare the mean scores of the treatment- and control- groups a few things must be taken into account. After the first lesson 2 students give the answers 1 and 2 on the questionnaire: they report that they feel high levels of anxiety. That number for the control-group is 0. After lesson 6, three students of the treatment group gave the answer 2, but no students gave the answer 1. In the control-group, two students gave the answer 1 after lesson 6, and four students gave the answer 2. At the higher end of the spectra, in the treatment group, after lesson 1, seven students gave the answer 5 and 2 the answer 6: they reported to high levels of confidence in their listening abilities. After lesson 6, these numbers were down to four students who gave the answer 5, 1 student answered 5,5 and 2 students

answered 6. In the control-group the decrease in students who gave the answers 5 or 6 was much bigger, the number of students who answered 5 dropped from eleven to four, and the number of students who gave the answer 6 dropped from eight to two. That is, compared to the treatment group, there was a significant increase in students who reported to high levels of anxiety and decrease in students who reported to high levels of confidence in the control-group.

After lesson 6, the mean score of the treatment group had increased by 0,41 points, to 3,7 and, the mean score of the control group had increased by 1,66 to 3,22, and the number of students who felt very anxious about listening increased more in the control group.

#### *4.6.2 Do you know the reason you feel this way about listening today? - Qualitative Results, Treatment- and Control Group*

The students of the treatment group reported feeling only slightly less confident after lesson 6 than after lesson 1, even though the 6<sup>th</sup> lesson was much more difficult. The drop in confidence was much larger in the control group. In addition, the students of the treatment group reported that their anxiety was due to factors outside of their control, such as being tired or hungry.

Listeners with low self-efficacy generally believe that it is outside their control to learn how to listen, and they will project failure on variables outside of their control (see 2.3.3). When asked the follow-up question “Do you know the reason you felt that way?” after lesson 1, 11 students (of 19) of the students of the treatment group answered that they were either very tired or hungry. This was after a lesson where they had had only one class prior to the English class, PE. After lesson 1, only one student answered that the text was difficult. Before lesson 1, when so many students reported to being tired or hungry, they had plenty of time to shower and eat the free lunch offered by the school, before the English class. It is difficult to see how the circumstances could have been better for learning. After lesson 6, however, the number of students who answered that they were either tired or hungry is down to 5, even though they had had two academic lessons before the English lesson that was just before their lunch break. The number of students who explicitly said that the text was difficult was 5.

In the control group the number of students who gave answers that their level of anxiety was due to them being tired or hungry was 4 (of 25) after the first lesson. They had had only had one class prior to the listening lesson (history). After lesson 6, when the text had become much more difficult, and longer, five students answered that they were tired, and one

student complained of hunger. The 6<sup>th</sup> lesson took place in the afternoon (lesson was 1 in the morning), the students had had two classes before the listening class, P.E. and mathematics and, they had a substantial lunch break. 13 students answered that the text was difficult after lesson 6, without giving an external cause.

The answers of the first questionnaire of the treatment-group were very short, on average 4,5 words. The answers of the questionnaires after lesson 6 were much longer: on average 11.5 words. The students seemed to become better at evaluating their experience. This could be because of the discussions that took place during the lessons, both regarding how the assignments went, but, also, due to the evaluating of understanding that took place in the opening segments of the listening classes. The students in the control group generally gave more elaborate answers; however, their answers progressively became shorter, after lesson 1 the answers were on average of 14, 9 words long and after lesson 6 the answers were on average of 11 words long.

#### *4.7 Limitations of the study*

The results of this action research are not meant to be generalized but, rather to be used by fellow teachers to evaluate whether Vandergrift's seven-step model for metacognitive instruction of listening can be of use to them in Swedish upper secondary schools.

Thus the first limitation that needs to be addressed is the small number of participants, this may have particularly affected the answers to the first two research questions. The abstraction rate due to absence during at least one of the tests was three students in each group. In the treatment group, two students were added to the group after the pre-test, the number of students added to the control-group after the pre-test was three. These students, even though they participated in the classes, could not be included in the results from the pre- and post-tests. Two students quit the study because a change in programs or schools. In particular, the number of students in the treatment group became so small, that the group do not correspond to a class of normal size in a Swedish upper secondary school. The abstraction rate also affects the relationship between the results of RQ1 and 2 and RQ 3 and 4, since it is not the results of the exact same groups that are being connected. Since the initial population in the study was so small, this may have had an effect on the results.

Secondly, results based on the answers to questionnaires can often be questioned. The MALQ has been tested and evaluated by professional researchers, which makes those results more generally legitimate than the questionnaire questions that I produced. However,

in the case of this action research, I thought it was of importance to test what the students' experience of the method was, to be able to discuss it in relation to my other results.

Thirdly, the fact that I only spent four weeks on the listening project may also have had an effect on the results. The students reported being bored with listening, especially in the treatment-group where we did nothing else for four weeks. If I had spent the entire semester conducting the study, and, made sure that the content of the listening texts was always in relation to something that we were currently working with, the students may have kept their interest up more, in addition, the long-term effects of the instruction, even if they were not measured in this study, may have been different.

I maintain, however, despite these limitations, the collected results may nevertheless be interesting to my peers who might also like to test the method recommended by the Swedish National Board of Education but which has never been researched in the Swedish upper secondary school setting of English 5 courses. After all, strategy use is of particular importance for English 5 since strategy use is to be graded in that course.

## ***5. Discussion and Future Research***

In this chapter I start by discussing the results of the research question. Then, in order to get a fuller understanding of the results I discuss the sub questions in the same order as they have been given throughout the paper. For each of the questions, there is a discussion of the results' implications for teaching, as well as a discussion of my results in comparison with the results of other similar studies. At the end of each section I suggest paths for future research on metacognitive instruction in listening in the setting of Swedish upper secondary school in light of the Swedish National Board of Education having suggested this method to be used in Swedish schools.

### *5.1 To What Extent Will Students' Listening Skills, Measured by a Standardized Listening Aptitude Test, be Affected by a Four Week Period of Metacognitive Listening Strategy Instruction?*

In my study, the treatment group did not outperform the control group. The standard deviation from pre- to post-test decreased more in the control-group, suggesting that a few students improved a lot, and a few considerably less. In the treatment group, the increase in scores was more equally divided. Pedagogical interventions based on Vandergrift's seven-step pedagogical sequence may, therefore, be more suitable, than the alternative of leaving the students to their own devices to solve listening problems as was done in the control-group. However, this difference, as indicated by the standard deviations of the results, is not statistically significant, which is remarkable. If this method is to be regarded effective, there needs to be a greater difference in listening performance between those students who have been subjected to it and those who have not.

Why then was the difference between how much the groups improved on the post-tests not more significant? The answer may partially depend upon how the groups were selected, by convenience sampling. The English 5 students were selected because English 5 is the only upper secondary English class where strategy use is to be graded and thus has to be taught. Also, English 5 is taught to all Swedish upper secondary school students, regardless of what program they study. In this case, the treatment group were a class studying a vocational program and the control group were studying a program that would prepare them for future academic studies: there may be a general difference in the groups' aptitude and motivation. It could be argued that this had an impact on the way the students of the treatment group

received the metacognitive instruction, and on the fact that many of the students of the control group did well on the listening tests even without particular instruction.

In addition, the results of my study go against most of the other similar international studies. I found 8 articles investigating the effectiveness of metacognitive strategy instruction in listening: Vandergrift and Tafaghodtari (2010), Coskun (2010), Graham and Macaro (2008), Bozorgian (2011), Khonmari and Ahmadi (2015), Farhadi, Zoghi and Talbei (2015), Movahed (2014). All these studies have reported their treatment groups to have outperformed their control groups. Only one study, Wang (2016), has not come to that conclusion.

The participants of the studies in Vandergrift and Tafaghodthari (2010), Coskun (2010), Bozorgian (2011), Movahed (2014) and Wang (2016) were all university students and the participants in Khonmari and Ahmadi study (2015) and the Farhadi, Zoghi and Talbei (2015) were from language institutes in Iran. All of these students, therefore, are older than my participants and can be assumed to study language out of their own volition. University students who voluntarily study language can be assumed to be more motivated, and less anxious towards learning language, than students for whom language study is obligatory. In Sweden, 98% of students with sufficient grades continue their upper secondary education after they finish the compulsory schooling (*Skolverket*, 2014). There really is no place in the current Swedish society for 16-year-olds without schooling. It can be argued that many of the students would not have continued studying if they had a viable alternative. Students who have elected to attend a vocational program may be more adverse towards school in general and, academic subjects such as English in particular. This may have been the case for some of the students of the treatment group of this study; therefore, comparing the results of this study, with the results of the above-mentioned studies that investigate voluntary and older students is difficult. However, voluntary participation and motivation seem to be of importance even with younger students. For example, Graham and Macaro (2008) test the method with 16-17-year-old students of French as a second language in Great Britain. This course was not mandatory, however, the researchers started with an initial sample of 151 students (120 girls and 31 boys), which in the results was decimated to 107 students. It can be assumed (although this is not explicitly stated in the article) that the students with the poorest results opted out of the course, which may have altered the researchers' results. This choice and, therefore, motivation may have caused the positive results of the studies above, while the lack of choice and, therefore, the lack of motivation may have caused the negative results in my study. Secondly, the time aspect also needs to be considered. I held nine listening classes over 5 weeks in my treatment group and eight listening classes during four weeks in my control

group (only one lesson differed: a class where the treatment group was introduced to the ideas behind metacognitive strategy instruction). The content of the listening material was selected with the course content in mind, but it was not related to anything else that we did. There was new, unknown content each lesson. However, there was a pedagogical purpose to the repetitive nature of the pedagogical sequence as the aim was to try to make the students change their automated listening behaviors. Interestingly, there were some complaints in the treatment group at lesson four when they were no longer allowed to discuss with a partner, but were instead asked to reflect on their own. The students expressed that not following the regular script was confusing, which may be indicative of learning of the new thinking patterns.

Vandergrift and Tafaghodtari (2010) conducted their investigation during the normal 13-week semester with the content that related to the topics of the units of study. The control group did the same amount of listening, but received no strategy instruction. Coskun (2010) used five weeks with material from a listening course book, while his control group did nothing. Graham and Macaro (2008) used three groups: a high scaffolding group, a low scaffolding group and a control group; all groups had the same amount of lesson time. Their study was conducted over the course of seven months. The Wang (2016) study lasted for only ten weeks, the students had one hour of listening instruction per week, while the students of the control group did not receive any instruction on metacognitive strategy use. Ten weeks is considerably more than the four weeks I spent on listening instruction with my students. The results of these studies cannot be directly compared to the results of mine as the conditions were different, not only because of the time spent on instruction but also because time had an effect on the listening material. Indeed, time is not a factor that can be disregarded when it comes to learning a language skill.

It would be interesting to find the ideal minimal length for listening instruction. Wang's results suggested that his treatment group did not outperform the control group in his shorter study, just as in my study. The other studies mentioned above, however, spent more time on the instruction and obtained positive difference between their treatment and control groups. The question of how much time needs to be spent with metacognitive strategy instruction before it has effects on language proficiency, or lasting effects on listening behavior needs to be further investigated. Interestingly, even if the participants of Wang's study were university students, the English course they were studying was compulsory. This lack of choice, rather than the time span, may have affected his results, similarly to my study.

In other studies, over short pedagogical interventions, the treatment groups became better at listening than the control groups. Movahed (2014), for example, used eight sessions for listening instruction, his control group received no listening instruction at all. Finally, Khonmari and Ahmadi (2015) used six lessons for their study while their control group received “conventional listening instruction” (p 23), Farhadi, Zoghi and Talbei (2015) used seven lessons for their study, and Bozorgian (2011) used only 4 lessons for his study. However, the results of these last studies cannot be directly compared with my results either: the participants were all single gendered and older than the students in my study. In addition, their participants studied language by choice. All in all, the presence or absence of choice and the resulting increase or decrease of motivation may be stronger predictors of the method’s efficiency than time.

Another factor that separates the results of some of the studies mentioned above from the result of the present study, is the selection of the participants. For example, in the Movahed (2014), Khonmari and Ahmadi (2015) and the Zoghi and Talbei (2015) studies, the participants were homogenized to fit the study. However, a method that is recommended for English 5, a compulsory course for all upper secondary school students of Sweden, needs to be proven to be effective in heterogenous groups too since upper secondary school groups comprised of students from various backgrounds are naturally heterogeneous. Contrary to the studies above the participants of my study were selected through a convenience sample, in part due to the fact that one of the aims was to see how useful this method was for a practicing teacher by conducting action research. The treatment group and the control group were not similar enough to really compare. Even though the Levene’s test for equality of variances performed on the means of the pre-test showed that the groups were homogenous, that test only shows that the bell curves of the groups are similar, making the groups comparable. It does not show that the groups are equal in proficiency, aptitude or motivation: in fact, based on the students’ previous grades of English, the groups were not completely homogenous. In a group where 38% received a B or higher, the students can be assumed to feel more confident and learn more regardless of the teacher intervention, which may not be the case in a group where none of the students reached B or higher as was the case in my treatment group.

Consequently, research suggesting that weaker students have more to gain from this method may well mean weaker students in a high performing, homogenous group. The composition of the groups may well be the reason why the results of my study vary from the results of the majority of the other studies investigating the effectiveness of metacognitive listening strategy instruction. Larger scale studies are, therefore, needed to see whether this

method should be used invariably also in heterogeneous setting such as English 5 courses in the Swedish upper secondary schools.

Graham and Macaro (2008) went even further than homogenizing the groups for the study: they pre-tested their subjects to see what strategies the students needed to work on and tailored the instruction towards the thus discovered needs of the students rather than trying to introduce all strategies at once as I did. This tailoring has very likely affected the results, which is why the results of Graham and Macaro's study should not be seen as strong evidence that the method *per se* is effective.

Consequently, more research is needed on the influence of aptitude, motivation, prior knowledge, etc. on the suitability of the method for listening instruction in the Swedish upper secondary schools. Questions should be raised and answered about how to use this method with students with low motivation, and how it can be used to increase students' motivation to learn. In addition, more research is needed to ascertain whether students' level of maturity affects their chances to benefit from this type of instruction. It would be interesting to see if this method indeed is more suitable for more cognitively mature students, particularly in light of previous claims that it is the less skilled listeners who have the most to gain.

The implications for using this method in teaching that can be derived from my study and from the previous research are that one ought to consider what group the method should be used in. How to work with self-knowledge, self-efficacy and motivation should also constitute part of the planning. How much time should be spent on using the method should also be considered: judging from the results of the above-mentioned studies, it seems wise to integrate the metacognitive work with listening into the normal course.

### *5.2 How Will Students' Perceived Strategy Use be Affected by Explicit Metacognitive Listening Strategy Instruction?*

The results of the pre- and post MALQs of both the treatment and the control group showed that the students perceived using strategies less after the interventions. Since I did not want the students of the control group to reflect on their strategy use, I could not ask them questions about their strategy use during the listening classes: the only data I have to support my claims for the control group are thus the results of the MALQs. On the other hand, the treatment group participated in guided reflection and reflected on their strategy use individually, in writing, during every lesson as part of the instruction. The results of their

think-aloud protocols, in addition, show that they report using more strategies after lesson 6 than after lesson 1. The MALQ results of the two groups were very similar, even though their treatment during the intervention was not. My conclusion is that another factor, the one the groups had in common, played a more prominent role for their perceived strategy use: the level of difficulty of the text. When the texts became more difficult, the students may have resorted more to bottom-up strategies, such as mental translation (the one strategy where their perceived use increased). The fact that many students reported that the speech was too rapid after lesson 6 supports this idea since mental translation is time-consuming and could have made the students miss out on what was said.

Looking only at the quantitative results, it would seem that the straight answer to the second research question is no. Both groups reported to declining perceived strategy use. However, looking at the answers of the questions that were given to the students of the treatment group after each listening session, I find some evidence of a change in either behavior or awareness of behavior. The evidence could also point, however, towards the students becoming better at expressing themselves with practice and teacher modelling. The students' significant improvement in the post-test scores may still suggest that their behavior changed.

During the pedagogical intervention the students were introduced to an array of strategies. Before starting each class the students were asked to recall what strategies were worked with on the previous occasion; they also had access to an online version of the list of strategies at all times. Yet, the students very frequently reported that they had forgotten about the strategies when they answered the first question of the lesson plans (What do you think about the strategies we talked about last time?) It is likely instead that they did not want to invest energy into recalling or writing the strategies down, as they saw no immediate gain from the activity. It could also mean that they did not possess the vocabulary to express it. As mentioned in section 4.4, the students became better at expressing themselves over time, which made strategy use more visible in the later stages of the intervention.

Similarly, Wang (2016) investigated students' strategy use through what he called student listening diaries (written by his students at the end of the listening sessions as a part of the procedure). Wang did not look for specific strategy use, but students' awareness of person knowledge, task knowledge and strategy knowledge.

Interestingly, he claims that it took his students between 10 and 20 minutes to write their listening diary entries and that the entries were on average 300 words. The students of the treatment group of my study spent considerably less time writing their think-aloud

protocols (I did not time them), even though they all received as much time as they wanted: the average number of words ranged between 3 and 14. This indicates how interested my students were of reflecting on their behaviour. The students in my treatment group became better at expressing their way of thinking with practise, however. Even if it was not measured in this investigation, my perception is that most of the students did not become better at orally expressing their way of thinking, while the students with low self-efficacy chose to not participate at all, missing out on an integral part of the seven-step model. Although the subjects of my study may have showed some improvement in expressing strategy use, more research is needed in the field of conscious and unconscious strategy use. Furthermore, since listening is such an internal affair, more research is needed also into what strategies are actually used by students, as opposed to what they can express with words.

The discrepancy between active and passive knowledge of language is not only seen in the way that the students express themselves. Many students expressed that they liked reading the text during the third listen and that they used the reading to check their understanding. To some extent this must mean that they did have the language skills to comprehend the language, but taking in the information aurally was a hindrance to them. The students expressed in their think-aloud protocols, that reading was easier, but no one explained why. One student, however, mentioned that it became more difficult to listen the third time, because the students were asked to read. This particular student was dyslexic. It would be interesting to research how to help students with dyslexia monitor their understanding effectively. It should be noted, however, that not all the students in the class who were dyslexic have mentioned difficulties to read and listen at the same time. It should also be noted that the method originally does not address language impediments, and no adaptation of this method, to, for example, dyslexic students is described in the previous studies.

The MALQ has been used in some of the studies as an awareness-raising tool. It was not used to raise awareness during my study but was administered twice: five weeks apart. Perhaps, the results would have been different had the students worked continuously with the material. Al-Alwan, Asassfeh and Al-Shboul (2013), for example, claim that there is a connection between students' awareness of strategy use and listening comprehension. They conducted their study on a convenience sample of 386 16-year-old native speakers of Arabic. Their study, however, did not entail any instruction: the students were just given a MALQ and a proficiency test, and the researchers used the results of the MALQ to see in what way the

result for each of the categories of strategies that were measured related to the scores of the proficiency test.

The implications for teaching that can be extracted from these results are that although the students of my study reported to be using strategies less, it is not necessarily the fact that they actually used strategies less. Working with strategies during the courses and giving the students tools may very well help students become independent learners of listening. Also, the MALQ can be used to help select what strategies to work with, so that the students do not have to work with all strategies at once. However, the MALQ measures the students' perceived strategy use, not their actual strategy use. Teachers should, therefore, spend lesson time on helping students become aware of their actual listening behaviors.

### *5.3 How Will Students' Perception of Learning be Affected by Metacognitive Listening Strategy Instruction?*

I found no other studies that investigated the students' perception of this method *per se*. The studies that used student listening diaries may have collected data on it, but those studies aimed primarily at figuring out either what categories of strategies that the students were using or what individual strategies they used, not whether the students perceived that the method worked for them, or what individual strategies were considered particularly effective. Therefore, this question is not discussed in those studies. My rationale behind posing this research question was that I hypothesized that the students would be more likely to learn how to use this method in learning how to listen if they thought of it as effective.

My results are interesting for a few different reasons. The students in the treatment group were told, during all of the lessons, that they were being taught how to learn to listen. In contrast, the control group were only administered listening comprehension tests. The latter group knew that they were a part of the study, and they had been told at the beginning of the study that the aural texts and comprehension tests would help them learn how to listen in the same way extensive listening would. Even if the groups in that respect were treated very differently, they still started out at approximately the same level, as measured by the questionnaires. Considering that this study is based on only one question, answered on a Likert scale, with no chance for the students to elaborate it is difficult to say what their answers are based on. It seems, however, that the teacher telling them the method was effective had no effect on the students' perceptions of it. On the other hand, this may well

have been due to the students' lack of motivation to learn the method. Again, more research is needed on how low motivation affects the effectiveness of this method. Also, more research is needed on how students' perception of learning is affected by motivation or lack thereof. The results of my study may have been different had I used the more motivated group as the treatment group.

Secondly, both the treatment group and the control group ended up claiming that they were learning how to listen to a lesser extent towards the end of the instructional period. Interestingly, the curves (see Figure 18) of the two groups, receiving very different instruction are almost identical. This suggests that there may be other factors at play, such as text difficulty. This claim is further supported by the fact that it was text 4 that had the biggest increase in difficulty, and the texts did not follow the exact same format (that is where I switched from the BBC texts to the CBC texts), resulted in the lowest scores from both groups. After text 5, both groups reported to having learned more. What is noteworthy here is that the treatment group ends up with a mean score on that question that indicates that they perceived that they were learning slightly more (but not at a statistically significant rate) than the control group, even though so many of the students were negative to the instruction. This could indicate that the students of the treatment group felt more self-reliant, and sure of their capabilities, because they had learned to unconsciously work with these strategies. More research is needed on the link between students' perception of learning and their learning success.

It could be considered logical that the groups have similar curves in Figure 18, since the groups' increase on the scores of the proficiency tests were very similar too. The results may indicate that the students did not respond well to being challenged with too difficult tasks. This may have implications for teaching, if the goal is to teach the students to use a method, teachers should not make the challenges too great too fast, then the students may lose faith in the method. This has implications for research into this method too. More comprehensive can be achieved results if all the texts are at the same general level throughout the study.

#### *5.4 How Will Students' Anxiety Towards Listening be Affected by Metacognitive Listening Strategy Instruction?*

When looking at the number of students who feel confident about their listening abilities in the treatment- and control group of my study, their prior grades in English and their GPA need to be taken into account. Likely, the students who have received a higher grade, also feel less anxious about listening, regardless of their actual skills. The investigation into the students' levels of anxiety during the intervention clearly indicated that the students of the treatment group gained in confidence compared to the control group: they felt less anxious when the texts became more difficult. This is surely to be considered a gain, even if the students in both groups report feeling more anxious after lesson 6 than after lesson 1.

Similarly to my study, Movahed (2014) investigated the effect of metacognitive strategy instruction on listening performance, metacognitive awareness and listening anxiety. His participants were 55 beginner students of English at a University in Iran. The students received 8 sessions of strategy instruction based on Vandergrift's seven-step model. His results show that the experimental group significantly outperformed the control group on the post TOEFL test, and that the anxiety levels of the students decreased significantly. Movahed has investigated the students' level of anxiety, and he has done so during a pedagogical sequence similar to the one used in my study. Therefore, the fact that in my study, the levels of anxiety have increased in both groups, contrary to Movahed's claims is interesting. Two issues need to be considered when comparing the results of our studies: the participants and the instruments of measurement. Firstly, as in section 5.2, the participants of Movahed's study were university students studying language voluntarily, and they were older and possibly more cognitively mature. All his participants were also male. Secondly, the participants of Movahed's study were beginner level students. It can be assumed that beginner level students, where the majority can be assumed to have Farsi as their L1 (a language that has considerably less in common with English than Swedish – the L1 of the majority of my treatment group), may experience higher levels of anxiety than the Swedish lower intermediate learners of English who have studied English for about seven years, and who are exposed to aural representations of the English language on a daily basis, such as the students of my treatment group. If a group starts out with very high levels of anxiety, and are exposed to easy texts (to fit a beginner level course), a decrease in anxiety may be easier to achieve than it was for my treatment group, a group with somewhat heterogeneous levels of anxiety to begin with. In addition, it should be noted that Movahed and I used different methods for measuring the

students' levels of anxiety. He used FLLAS (Foreign Language Listening Anxiety Scale), and I used a self-made questionnaire, where the students had the opportunity to elaborate on their answers in an open question. The FLLAS is a 33-item questionnaire, where students rate their answers on a five-graded Likert scale (Movahed 2014, p. 95). The different instruments may well have affected our results.

Further research on combinations of the different factors in the Swedish upper secondary school setting – choice, motivation, aptitude, prior knowledge etc. should be conducted before the method can be considered effective and, possibly before the method should be recommended by *Skolverket*.

The results of Movahed's study and the way I interpreted mine, after having taken into consideration the increase of difficulty of the texts that the students were subjected to can be interpreted as corroborating. This may indicate that the method helps decrease students' levels of anxiety. But what about using this method produces that effect? In my study, the students in the treatment group were allowed to work together, and test their ideas before having to reveal to me what they thought. If they had no ideas, they were allowed to listen to other students' ideas before answering. This may, of course, have led to the students of the treatment group feeling less anxious than the control group, where all the students were working individually in a test situation. In my study, then, I cannot clearly see that it is the method *per se* that decreases anxiety, and not the fact that one group writes a test, and the other group does not: more research is needed into how metacognitive listening strategy instruction helps decrease students' anxiety towards listening.

Finally, in my study, much time was spent on task knowledge and strategy knowledge during the lessons with the treatment group. There was a particular focus in the beginning of the classes on the strategy knowledge and knowledge about how listening works. During the listening part of the lessons, the focus was on task knowledge. The class discussions became general, and we were constantly trying to find "the best way" of solving the problems; however, we spent very little time figuring out the personal differences between the students, and their possible different needs. The students were left to do that on their own, and many students just gave up. In the treatment group, they were asked if they gained anything from discussing monitoring with a peer. It was clear from the answers and from looking at the students interact, that the weaker students just opted to not speak at all, and thus they did not take part in all parts of the instruction. For this method to work, it is important that the students feel comfortable with each other, which was not the case in my treatment group. The class discussions on how to solve problems may have led to the students'

increased self-efficacy with regards to listening but not in their oral communication. The implications for teaching that can be drawn from this study are that teachers should not focus on strategy knowledge and task knowledge alone; self-knowledge is an equally important part of the learning process and of this pedagogical sequence. Also, to make this metacognitive instruction as effective as possible, spending lesson time where students learn how to interact with each other in English will likely positively affect the outcome of the use of this type of instruction.

## ***6. Summary and Conclusions***

The primary result of this study was that I found no conclusive evidence that Vandergrift's pedagogical sequence of metacognitive listening strategy instruction was more effective than traditional listening instruction. The treatment group and the control group both improved their results, the difference in improvement was not statistically significant.

I found no evidence that the explicit strategy instruction increased the conscious use of strategies in the treatment group compared to the control group. In fact, both groups showed decreasing levels of awareness of strategy use as measured by the MALQ. This may indicate that there is a connection between perceived strategy use and the level of difficulty of the tasks given to the students. I also found no connection between increased perceived listening strategy use and listening proficiency.

I found no evidence that the students who received explicit strategy training perceived that they were learning how to listen to a greater extent than the control group, which was receiving traditional listening instruction.

On the other hand, I have found some evidence suggesting a link between metacognitive instruction and self-efficacy. The students of the treatment group felt less anxious about listening after lesson 6, than the control group. The control group started out with higher self-efficacy, feeling less anxious about listening than the treatment group, but after the six-lesson instructional period, this was no longer the case.

The results of this small-scale study cannot be generalized to a wider population, but, it raises questions as to what kind of students, at what level of proficiency and at what level of motivation this method is most useful for. This study focused on task knowledge and strategy knowledge, but, perhaps, self-knowledge regarding motivation to learn and of problem solving needs to be stressed more, than what was done in this study.

This method of teaching listening is recommended by Swedish National Board of Education. However, my study points out that this recommendation has not been sufficiently supported by previous research. Further, what research exists has not been conducted in the setting of the Swedish upper secondary school. The method is, therefore, not necessarily unquestionably useful for all teachers and students who teach or take English 5. Factors such as choice, motivation and time have been repeatedly implicated in this and other studies as being of importance to the success of the method, and more research needs to be conducted in the Swedish setting. However, even if I have yet to find empirical evidence for long-term positive effects of metacognitive strategy instruction, I would still recommend it. Working

with the students in a way that makes the teacher and the students aware of how their minds work provides valuable insights for how to tailor individual instruction and guide the students to lifelong learning.

# References

- Al-Alwan, Ahmed., Assafeh, Sahail. & Al-Sbboul, Yousef. (2013). EFL learners' comprehension and awareness of metacognitive strategies: How are they related?. *International Education Studies*, 6(9), 31-39. doi:10.5539/ies.v6n9p31
- Bozorgian, Hossein. (2012). Metacognitive instruction does improve listening comprehension. *International Scholarly Research Network ISRN Education*, vol.2012, 1-6. doi: 10.5402/2012/734085
- Buck, Gary. (2001). *Assessing Listening*. Cambridge University Press. Cambridge.
- Börjesson, Lena. (2012). *Om strategier i engelska och moderna språk*. 2016-05-11, from: [http://www.skolverket.se/om-skolverket/publikationer/visa-enskild-publikation?\\_xurl=http%3A%2F%2Fwww5.skolverket.se%2Fwtpub%2Fws%2Fskolbok%2Fwpubext%2Ftrycksak%2FBlob%2Fpdf3120.pdf%3Fk%3D3120](http://www.skolverket.se/om-skolverket/publikationer/visa-enskild-publikation?_xurl=http%3A%2F%2Fwww5.skolverket.se%2Fwtpub%2Fws%2Fskolbok%2Fwpubext%2Ftrycksak%2FBlob%2Fpdf3120.pdf%3Fk%3D3120)
- Cohen, Louis, Manion, Lawrence & Morrison, Keith. (2005). *Research Methods in Education*. 5th Edition. Taylor & Francis e-Library, London.
- Coskun, Abdullah. (2010). The effect of metacognitive strategy training on the listening performance of beginner students. *Novitas-ROYAL*, 4(1), 33-50.
- Cross, Jeremy. (2015). Metacognition in L2 listening: Clarifying instructional theory and practice. *TESOL Quarterly*, 49(4), 883-892. doi: 10.1002/tesq.250
- Efron, Sara Efrat., Ravid, Ruth. (2013). *Action Research in Education - A Practical Guide*. The Guildford Press, New York.
- Farhada, Sima., Zoghi, Masoud. & Talebi, Zahra. (2015). Intermediate EFL learners' listening performance viewed in the light of metacognitive strategy instruction. *Modern Journal of Language Teaching*, 5(3), 140-146.
- Field, John. (2008). *Listening in the Language Classroom*. Cambridge University Press, Cambridge.
- Flavell, John H. (1979). Metacognition and cognitive monitoring a new area of cognitive developmental inquiry. *American Psychologist*, 34(10) 906-911.
- Flowerdew, John & Miller, Lindsay. (2005). *Second Language Listening - Theory and Practice*. Cambridge University Press, New York.
- Goh, Christine. (2008). Metacognitive instruction for second language listening development: Theory, practice and research implications. *RELC Journal*, 39(2) 188-213. doi: 10.1177/0033688208092184
- Graham, Suzanne., Macaro, Ernesto. (2008). Strategy instruction in listening for lower-intermediate learners of French. *Language Learning*. 58(4), 747-783. ISSN: 0023-8333
- Khonmari, Fatme., Ahmadi, Faezeh. (2015). The effect of metacognitive strategy training and raising EFL learners' metacognitive knowledge on listening performance. *Indonesian Journal of Applied Linguistics*, 5(1) 19-28. doi:10.17509/ijal.v5i1.827.g575
- Kilic, Mehmet., Uckun, Berrin. (2012). Listening text type as a variable affecting listening comprehension anxiety. *English Language Teaching*. 6(2). Doi: 10.5539/elt.v6n2p55
- McNiff, Jean., Whitehead, Jack. (2005). *Action Research for Teachers – A practical Guide*. David Fulton Publishers Ltd., London.
- Morra de la Peña, Ana María & Soler, Lidia Rosa. (2001) Cognitive strategies for academic reading and listening in EFL, *Journal of College Reading and Learning*, 31:2, 217-232, Doi: 10.1080/10790195.2001.10850117
- O'Malley, Michael J., Chamot, Anna Uhl, Stewner-Manzanares, Gloria, Russo, Rocco. P, Kupper, Lisa. (1985). Learning strategy applications with students of English as a second language. *TESOL Quarterly*, 19(3), 557-584.

- PET - Listening test (parts 1-4) 2016-07-10 from:  
[http://www.examenglish.com/PET/PET\\_listening.html](http://www.examenglish.com/PET/PET_listening.html)
- Siegel, Joseph. (2015). Readers Respond: Research involving L2 listening instruction and instructors. *ELT Journal*, 69(3), 323-326. Doi: io.tog3/elt/ccvoi4
- Siegel, Joseph. (2015b). *Exploring listening Strategy Instruction Through Action Research*. Palgrave Macmillan, Hampshire.
- Skolverket. (2011a). *English*. 2016-05-11 from:  
[http://www.skolverket.se/polopoly\\_fs/1.174542!/English%20120912.pdf](http://www.skolverket.se/polopoly_fs/1.174542!/English%20120912.pdf)
- Skolverket. (2011b). *Alla kommentarer*. 2016-05-11 from: <file:///home/chronos/u-e68eae8eb483e95af1cd765c7475b81f86b47200/Downloads/Alla%20kommentarer.pdf>
- Skolverket. (2014). *Nästan Alla Grundskoleelever Fortsätter till Gymnasiet*. 2017-01-01 from: <http://www.skolverket.se/statistik-och-utvardering/nyhetsarkiv/nyheter-2014/nastan-alla-grundskoleelever-fortsatter-till-gymnasieskolan-1.223182>
- Skolverket (UÅ) Nationella prov – exempeluppgifter. 2016-05-11 from:  
[http://nafs.gu.se/prov\\_engelska/exempel\\_provuppgifter/engelska\\_5\\_exempeluppg](http://nafs.gu.se/prov_engelska/exempel_provuppgifter/engelska_5_exempeluppg)  
 Göteborgs Universitet.
- Tornberg, Ulrika. (2009) *Språkdidaktik*. 4th edition. Gleerups, Malmö.
- Vandergrift, L. (2003). *Orchestrating strategy use: Toward a model of the skilled second language listener*. *Language Learning*. 53:3. p.463-496.
- Vandergrift, Laurence., Baker, Susan (2015). Learner variables in second language Listening Comprehension: An Exploratory Path Analysis. *Language Learning* 65(2), 390–416.
- Vandergrift, Laurence., Cross, Jeremy. (2015). Replication research in L2 listening comprehension: A conceptual replication of Graham & Macaro (2008) and an approximate replication of Vandergrift & Tafaghodtari (2010) and Brett (1997). *Language Teaching*, 1-10. doi:10.1017/ S026144481500004X
- Vandergrift, Laurence & Goh, Christine. (2012). *Teaching and learning second language listening - Metacognition in action*. Taylor & Francis, New York.
- Vandergrift, Laurence; Goh, Christine; Mareschal, Catherine & Tafaghodtari, Marzieh. (2006). The Metacognitive Awareness Listening Questionnaire: Development and validation. *Language Learning* 56:3, 431-462.
- Vandergrift, Laurence. & Tafaghodtari, Marzieh. (2010) Teaching L2 learners how to listen does make a difference: An empirical study. *Language Learning* 60(02) 470-497. doi: 10.1111/j.1467-9922.2009.00559.x
- Vetenskapsrådet (20XX). *Forskningsetiska principer inom humanistisk-samhällsvetenskaplig forskning*. Vetenskapsrådet
- Wang, Weili. (2015), Learning to listen: The impact of a metacognitive approach to listening instruction. *The Asia-Pacific Education Researcher*, 25(1), 79-88. doi: 10.007/s40299-015-0235-4

## Appendix I: Test descriptions

### Test description PET (pre- and post-test - are identical in structure) (Cambridge)

part	Task type	No	Task description
1	Comprehension questions in short texts - choose the most appropriate image	7	Short monologues or dialogues, only between 10-30 seconds long. Answer by choosing one of three pictures.
2	Listen for information in a dialogue	6	Listen to a conversation or interview - find information and answer six three-option multiple choice questions (test 1: 4.34 min, test 2: 1.33 min)
3	Note taking, blank filling	6	Fill in gaps - information missing on a form. (test 1: 1.52 min, test 2: 2.29 min)
4	Comprehension questions (and inference) on a dialogue	6	Right or wrong questions on a longer text (test 1: 2.05 min, test 2: 1.11 min)

Figure XX: test description of the PET

### Test description National Test (The Swedish National Board of Education)

MC - Multiple Choice, SA - Short Answer (2xSA - 2 answers required - 1 point for each answer)

Pre-test: The River police (19 min) Listen twice - interview - divided in three parts. Accents: RP (both instruction and interview)

Above: question - below - type of answer required

1a	1b	2	3	4	5a	5b	6	7	8	9	10a	10b	11
2xSA	MC	2xSA	SA	MC	MC	SA	MC	2xSA	SA	MC	SA	MC	SA

Max 17p

Post -test Hurricane Hunters (20 min) Listen twice - interview - divided in three parts. Accents: RP (instructions) AmE (interview)

Above: question - below - type of answer required

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
SA	MC	MC	SA	2x SA	SA	MC	SA	MC	SA	SA	MC	SA	SA	SA	MC	SA

Max 18p (Question 6 was taken out in order for the tests to have the same score - the students were given the test WITH questions six, the score was taken out after the fact.)

Figure XXX: Test description of the listening tests from the National Test of English for English 5.

## Appendix II: MALQ

### Metacognitive Awareness Listening Questionnaire (MALQ)

The statements below describe some strategies for listening comprehension and how you feel about listening in the language you are learning. Do you agree with them? This is **not a test**, so there are **no “right” or “wrong” answers**. By responding to these statements, you can help yourself and your teacher understand your progress in learning to listen. Please indicate your opinion after each statement. Circle the number which best shows your level of agreement with the statement. For example:

	Strongly disagree	Disagree	Slightly disagree	Partly agree	Agree	Strongly agree
I like learning another language	1	2	3	4	5	6

1. Before I start to listen, I have a plan in my head for how I am going to listen.	1	2	3	4	5	6
2. I focus harder on the text when I have trouble understanding.	1	2	3	4	5	6
3. I find that listening is more difficult than reading, speaking, or writing English.	1	2	3	4	5	6
4. I translate in my head as I listen.	1	2	3	4	5	6
5. I use the words I understand to guess the meaning of words I don't understand.	1	2	3	4	5	6
6. When my mind wanders, I recover my concentration right away.	1	2	3	4	5	6
7. As I listen, I compare what I understand with what I know about the topic.	1	2	3	4	5	6
8. I feel that listening comprehension in English is a challenge for me.	1	2	3	4	5	6
9. I use my experience and knowledge to help me understand.	1	2	3	4	5	6
10. Before listening, I think of similar texts that I may have listened to.	1	2	3	4	5	6
11. I translate key words as I listen.	1	2	3	4	5	6
12. I try to get back on track when I lose concentration.	1	2	3	4	5	6
13. As I listen, I quickly adjust my interpretation if I realize that it is not correct.	1	2	3	4	5	6
14. After listening, I think back to how I listened, and about what I might do differently next time.	1	2	3	4	5	6
15. I don't feel nervous when I listen to English.	1	2	3	4	5	6
16. When I have difficulty understanding what I hear, I give up and stop listening.	1	2	3	4	5	6
17. I use the general idea of the text to help me guess the meaning of the words that I don't understand.	1	2	3	4	5	6
18. I translate word by word, as I listen.	1	2	3	4	5	6
19. When I guess the meaning of a word, I think back to everything else that I have heard, to see if my guess makes sense.	1	2	3	4	5	6
20. As I listen, I periodically ask myself if I am satisfied with my level of comprehension.	1	2	3	4	5	6
21. I have a goal in mind as I listen.	1	2	3	4	5	6

## **Appendix III: Anonymous Questionnaire**

**1) In your opinion, did the class today help you to learn how to listen?**

(1, no, I don't understand what this is for - 6 - yes, this is helping me becoming a better listener)

1 2 3 4 5 6

**2) How did you feel about listening today?** (1: I was very anxious and knew it would go poorly - 6 I was not nervous at all; I feel confident in my listening abilities.)

1 2 3 4 5 6

**3) Do you know the reason you feel this way about listening today?** (the text was difficult, I am tired, I feel very confident because it's easy) (Free text answer.)

## Appendix IV: Results of the t-tests

### The results of the t-tests

Paired two-sample t-test; treatment group

	<i>pre-test</i>	<i>post-test</i>
Mean	26,63	32,63
Variance	41,72	26,92
Observations	16	16
P(T<=t) one-tailed	0,000013	
P(T<=t) two-tailed	0,000025	
Standard deviation	6,46	5,19

Paired two-sample t-test;  
Control group

	<i>pre-test</i>	<i>post-test</i>
Mean	28,24	34,76
Variance	41,49	22,29
Observations	21	21
P(T<=t) one-tailed	0,000001	
P(T<=t) two-tailed	0,000002	
Standard deviation	6,44	4,72