Michael Yeldham*

Developing a framework for investigating L2 listeners' longitudinal development

Abstract: In research conducted by this author examining the development of adult second language listeners, one key aspect of that study was to evaluate the learners' growth in their listening strategy use. A search of the literature, however, found no reliable conception of listener growth to guide the evaluation. Consequently, this study sought to develop such a framework. To accomplish this, cross-sectional verbal report studies comparing the strategy use of proficient and less proficient listener groups were examined, and the strategies used frequently by the proficient listeners, but infrequently by the less-proficient listeners, served to indicate developmental trends. A second component of the study was then to order these strategies into the stages of Anderson's (1985, *Cognitive psychology and its implications*, 2nd edn. New York: Freeman, 2010, *Cognitive psychology and its implications*, 7th edn. New York: Freeman) three-stage comprehension model, those of perception, parsing and utilization. This second part of the study was conducted to address apparent flaws in past research categorizing listening strategies based on this model.

Keywords: second language listening, listening strategies, listener development

1 Introduction

In longitudinal research conducted by this author examining adult second language (L2) listeners in a strategies instruction course (Yeldham and Gruba 2016), a crucial aspect of the study was to evaluate how each of the learners progressed in the course. A metacognition framework (Flavell 1976; Vandergrift and Goh 2012) was used to guide the research, which encompassed the person-, task-, and listening strategy-related factors considered important to listener growth. While the nature of learner growth in the person-related and task-related aspects of the framework seemed relatively straightforward to determine, it seemed less clear in terms of learners'

*Corresponding author: Michael Yeldham, Centre for Applied English Studies, The University of Hong Kong, Hong Kong, China, E-mail: myeldham@hku.hk

use of listening strategies. Consequently, this study aims to provide a clearer picture of how learners tend to develop in their use of listening strategies.

Such a conception of adolescent/adult strategy development is limited largely because there have been no suitable longitudinal studies to provide such insight. The longitudinal studies which have been conducted have either (1) examined elementary school learners (Peters 1999)[1]; (2) elicited strategies used while listeners completed multiple choice comprehension questions (Graham et al. 2008, 2011) (while the task used in my research and most other listening studies is less specific, that of essentially just listening to a text and trying to understand its overall meaning); or (3) they have examined learner progress in a listening strategies course (Chen 2009; Mareschal 2007), thus not providing a 'general' view of strategy development unaligned to a particular instructional approach. It is therefore problematic as to what constitutes development. For example, is it mainly an increase (or a decrease) in certain strategies? And if so, which ones? Or is it largely how appropriately or how effectively certain strategies are used (e. g., Macaro 2006)?

This paper, therefore, addresses this uncertainty by analyzing the findings from a series of cross-sectional studies that compared the strategies used by more-proficient and less-proficient listeners, an approach which can indicate how listeners tend to develop. Note that Berne (2004) and Macaro et al. (2007) have conducted similar analyses of cross-sectional studies. These produced useful insights, although the analyses were somewhat limited as they only constituted a part of the researchers' more expansive overviews of listening research. In addition, in each analysis, the research techniques varied across the studies examined. In Berne's analysis, for example, these techniques included verbal reports, interviews, focus groups, observations of the two-way listening task of learners conversing with native speakers, and strategy checklist questionnaires (which tend to guide and limit listeners' choices in reported strategy use). On one hand, this diversity could be seen in a positive light, with data from the different sources complementing each other. More likely, though, as Berne (2004) concedes, such variation reduced comparability between the studies, and therefore the reliability of the findings. In fact, Macaro et al. (2007) responded to this limitation (and others such as variation in learner age, as adults and elementary school learners were both included in

---

1 Peters' (1999) study is not considered relevant because such young learners may not have acquired adult-like higher level cognitive abilities (Piaget 1957) and their (pre-puberty) language processes may differ from those of adults (DeKeyser 2000).

their analysis) by only listing a small number of strategies, ones they felt clearly distinguished listeners at the different levels. These were that proficient listeners used more metacognitive strategies (especially comprehension monitoring) – which the researchers concluded also served to orchestrate their cognitive strategies more effectively – processed language in chunks rather than word by word, and used less translation.

Given these limitations of the previous analyses, reliability in this present analysis is strengthened by limiting it to studies where data was gathered primarily through verbal reports (Bacon 1992; Goh 1998; Graham 1997; Mareschal 2007[2]; Murphy 1985; O'Malley et al. 1989; Vandergrift 1997, 1998a, 1998b, 2003).[3] Verbal reports are chosen because this author and others (e. g., Chamot 2005; Graham et al. 2011) consider the technique provides the best insights into listeners' strategy use – also note that most cross-sectional listening studies have used the technique, so the research is still accessing a large pool of research. The research also aims to provide a more reliable and expansive depiction of listener growth by examining the various studies in more detail and more systematically than the earlier two analyses.

As an overview of this article, first, relevant theories, processes and research are introduced, and the verbal report technique is also discussed. The participants, and the data collection and analysis procedures used in each of the studies examined is then outlined in detail. This is followed by an analysis and summary of the main findings from these studies, including the resulting listener framework. There is then a discussion of issues emerging from the research, including some recommendations for redefining some of the strategies commonly used to depict L2 listening.

2 Relevant listening theories, processes and background research

In effective listening, top-down and bottom-up processes operate automatically, and in harmony. Bottom-up processing provides data for the listener to build meaning, and also aids in constraining possible listener interpretations of the

---

2 Mareschal's (2007) study was conducted longitudinally, but the initial cross-sectional differences in verbal report results between the learners at different proficiency levels are included here.
3 There was no mention in any of the studies that the listeners had received any particular form of listening skills training, such as strategies instruction.

utterance (Wu 1998). Top-down processing facilitates interpretation by promoting an active construction of meaning (Richards 2003) through contextualizing, guiding and enriching the linguistic input (Field 2004; Wu 1998). By contrast, because L2 listening is characterized by partial understanding of a speaker's message (Rost 2002), L2 listeners commonly require strategies to help them understand an utterance. These strategies are deliberate, conscious procedures that compensate for actual or anticipated breakdowns in comprehension (Afflerbach et al. 2008; Field 2008).

The most common framework categorizes the various strategies into cognitive and metacognitive strategies.[4] This framework, introduced to L2 learning by O'Malley and Chamot (O'Malley et al. 1989; O'Malley and Chamot 1990), is used to guide the current analysis. Cognitive strategies include bottom-up processes, such as segmenting speech and utilizing prominent lexical and discourse cues from the text; they also include top-down strategies, such as inferring meaning and predicting what a speaker will say next. Metacognitive strategies coordinate the various strategies and help to manage one's performance before (planning), during (monitoring) and after (evaluating) listening (Goh 2005; Lynch 2006; O'Malley et al. 1989; Vandergrift 2003).[5]

Researchers (e. g., Cross 2010; Goh 2000; Mareschal 2007) are also increasingly examining listeners from the wider perspective of learner metacognition. A key component of metacognition highlights learners' person-, task- and strategy-related knowledge and the application of this knowledge (Vandergrift and Goh 2012; Vandergrift and Tafaghodtari 2010). Person-related knowledge involves learners' awareness of the affective and cognitive factors and abilities that they possess and which are required to facilitate their listening and development. Task-related knowledge includes awareness of the nature, purpose and demands of listening tasks, and of what is required to succeed in listening.

These aspects of metacognition delineate many of the key areas involved in listener development. As their listening improves, the way learners develop in some of these areas, particularly many person- and task-related areas, is either

---

4 Along with socio-affective strategies to manage one's emotions and interactions with others (these are not relevant to the current analysis).

5 Note that another emerging framework, by Field (2008), categorizing strategies as 1) those reactive to listener difficulties, and 2) those used by listeners to minimize anticipated problems, was not considered as an organizing structure for the current analysis chiefly because no studies of L2 listeners have yet used the framework.

fairly obvious and/or has been documented in past research. For example, some person-related improvements are that anxiety lessens (Bacon 1992), and self-efficacy (confidence in being able to perform a given task) and motivation both increase (Chen 2007; Graham and Macaro 2008). Task-related improvements include a heightened need to concentrate when listening (Chen 2007) and also to keep listening in the face of comprehension difficulties (Bacon 1992). Learners have also been found to do more extracurricular listening (Mareschal 2007), and to increasingly challenge their listening with more demanding texts (Thompson and Rubin 1996).

What is lacking in the literature at present, though, is a clear depiction of growth in the strategies aspect of metacognition, involving learners' knowledge of available strategies and when and how to use them. As mentioned earlier, such insight is needed to indicate in longitudinal studies whether a listener's use of such strategies is progressing, and if it is, to what extent and in what areas. This is the main purpose of the current analysis of past studies.

A second purpose of the research is then to categorize the strategies from the framework developed here in terms of Anderson's (1985, 2010) three-stage comprehension model, a well-known conception of how listeners construct meaning. This model involves three rapid-fire, overlapping and recursive processes known as perception, parsing and utilization. Note that these stages describe different phases in the comprehension process, rather than implying that comprehension occurs in a bottom-up manner. The first stage comprises the perceptual processes involved in identifying the sound signals in the utterance as words or as meaningful chunks of language. During the parsing stage, the words are combined and transformed into propositions for the listener to hold a meaning-based representation in WM as more of the text is processed. In the utilization stage, listeners integrate this propositional representation with information from long-term memory to interpret the text (Anderson 1985, 2010; Vandergrift and Goh 2012). Note that from perception through to utilization, top-down information takes on an increasing role. However, parallel bottom-up and top-down processing can take place at each of these phases, and the phases themselves occur simultaneously in fluent speech (Vandergrift and Goh 2012).

Three studies have previously categorized the strategies listeners use at these three phases. O'Malley et al. (1989) compared the strategies used by effective and less-effective listeners, Goh (2000) compared the difficulties encountered in the stages by listeners of different ability levels, and Bacon (1992) listed both listeners' reported strategies and their difficulties at each stage.

Of most relevance to the current research are the results by O'Malley et al. (1989), comparing the strategies used by effective and less-effective listeners at each stage. These strategies comprised a better ability by the effective listeners to direct attention to the task (at the perceptual stage), to chunk larger units of meaning, and show a greater tendency to infer unknown words (parsing stage), and to use elaborations more effectively to infer meaning and assist information recall as the elaborations strengthened schema in memory. Buck (1990), however, has criticized some of these categorizations. One criticism concerns O'Malley et al.'s equating of listener attention with perception. Buck points out that the two are separate processes, and that attention occurs first followed by perception, when sensory information is stored in echoic memory. Buck also argues that inferring the meaning of unknown words is misplaced in the parsing phase. Buck contends that such inferences are better positioned in the utilization phase because they need to be activated by previous knowledge – that is, after the listener has interpreted the text.

Because of such criticism, and also a concern (which is explained shortly) with how O'Malley et al. conducted their research, this current research will attempt to provide a more viable taxonomy of listener development through these three processing stages.

3 Verbal reports to examine listeners

Because the framework is developed from the results of verbal report studies, it is important to outline this technique and some of the issues associated with it. With this technique, the text is commonly stopped at various junctures for learners to provide immediately retrospective accounts of the mental processes they used while they were listening. The segments of text between these stoppages are usually about two sentences long. This is long enough to prevent simple verbatim recall of the text through echoic memory, which would yield little insight into strategic processes, but not too long for the listener to forget the strategies they used in their attempt to construct meaning from the text.

Other introspective methods, such as interviews, questionnaires, learner diaries and stimulated recalls, are usually further removed, temporally, from the listening event than verbal reports. Therefore, many researchers (Chamot 2005; Cohen 1998; Graham et al. 2011; Wu 1998) consider verbal reports to provide the most direct access to listeners' mental processes, resulting in more accurate insights into listener processes. However, there are criticisms

of the technique. These include that (1) the intermittent stopping of a text for individuals to report their thoughts, (2) the common use of prompts to elicit data, and (3) the actual act of learners verbalizing their thoughts, may alter the nature of the task and also lead to individuals forgetting, distorting or underreporting their thoughts (e. g., Seliger 1983). Nevertheless, an exhaustive examination of the technique led Ericsson and Simon (1993) to conclude that verbalizing one's thoughts – either concurrently or immediately retrospectively, where the task being reported on retrospectively is of relatively short duration – does not place extra demands on working memory. Consequently, Ericsson and Simon (1993) contend that verbal reporting does not alter the way the task is undertaken or the mental processes involved, except perhaps to slow these processes down. Bowles (2010) reached similar conclusions after an overview of verbal reports used in L2 research. However, Ericsson and Simon (1993) do not deny the likelihood of individuals underreporting their thoughts; they do claim, though, that when conducted with care, and particularly when individuals are not asked to explain or justify *why* they used certain processes, verbal reports can elicit "at least a subset of the thoughts heeded while completing a task" (p. xxxv). Language researchers (e. g., Cohen 1998) say this subset of thoughts comprises mostly the conscious, problem-solving processes involved.

In terms of data produced by verbal reports, Buck (1994) explains that they can provide insights into "some of the cognitive [i. e., mental] processes important in listening comprehension, and also of the interaction between these" (p. 163). These processes include an array of bottom-up and top-down strategies, the ongoing development and monitoring of the listener's mental model of a text, and insights into the listener's attention to the task (Yeldham and Chen 2016; Buck 1991, 1994). However, because of its 'online' nature, the technique does not readily elicit more 'offline' strategies involved in planning and evaluating one's comprehension (Chamot 2004).

4 Analysis: Developing the framework

The studies included in the analysis are summarised in Table 1. The general findings, from a combination of quantitative and qualitative analyses of the data (discussed shortly), are shown in the two right-hand columns of the table. Note that in two of the studies, while verbal reports provided most of the data, this was supplemented by other techniques. In Goh (1998) this technique was learner diaries. And in Graham (1997) data for the quantitative analysis was gathered

Table 1: Cross-sectional verbal report studies: more-proficient v less-proficient listeners.

Study	Learners	Less-proficient listeners: main findings	More-proficient listeners: main findings
Murphy (1985)	12 American university ESL learners	Greater reliance on either b-u[a] or t-d strategies (n = 6)[b]	More balanced and flexible use of t-d and b-u strategies (n = 6)
O'Malley et al. (1989)	11 American high school ESL learners	Greater use of b-u strategies (n = 3)	Greater use of t-d strategies. More balanced use of t-d and b-u strategies (n = 8)
Bacon (1992)	12 American university learners of Spanish	Greater reliance on either b-u or t-d strategies (n = 6)	More balanced and flexible use of t-d and b-u strategies (n = 6)
Graham (1997)	24 English high school (A-level) learners of French and German	Greater use of t-d strategies (n = 11)	More balanced combination of b-u and t-d strategies (attributed to better metacognitive control by Graham 2003) (n = 13)
Vandergrift (1997)	21 Canadian high school learners of French	Greater reliance on either t-d or b-u strategies (n = 14)	Greater use of metacognitive strategies (n = 7)
Goh (1998)	16 Chinese university-aged ESL learners in Singapore	Use of t-d and b-u strategies (n = 8)	Use of top-down and b-u strategies, but wider range of t-d strategies. Greater use of metacognitive strategies (n = 8)
Vandergrift (2003)	36 Canadian high school learners of French	Greater reliance on b-u strategies (n = 18)	Greater use of metacognitive strategies (n = 18)
Mareschal (2007)	8 Canadian adult learners of French	Greater use of b-u strategies (n = 5)	Greater metacognitive control of strategies (n = 3)

Note: [a]b-u = bottom-up; t-d = top-down.
[b]n = Number of participants.

through verbal report, but for the qualitative analysis the verbal reports were supplemented with retrospective interviews and learner diaries. Another point to note was that the studies varied in their terms used to classify the two listener groups, using for example, successful and less-successful (Bacon 1992), high-ability and low-ability (Goh 1998), and more-proficient and less-proficient listeners (Murphy 1985; Vandergrift 1997). For convenience, the terms *more-proficient* (sometimes the shorthand, *proficient*) and *less-proficient* are used for all the studies. Note that proficiency in this paper relates to the learners' listening proficiency rather than their general proficiency or their level of linguistic

knowledge,[6] as the emphasis is on the interaction between listening proficiency and strategy use.[7]

All of the studies organized their strategies into O'Malley et al.'s (1989) cognitive and metacognitive framework, except for Murphy (1985) whose research pre-dated O'Malley et al. (1989) – although it was relatively easy to identify which of Murphy's strategies fitted into these categories. In addition, many of the individual strategies in the taxonomies remained true to those from O'Malley et al. (1989), especially in the Vandergrift (1997, 2003) studies. However, there was a degree of variation, as some researchers added new strategies found in their data, such as 'contextualization', 'real-time assessment of input' (Goh 1998) and 'wild guess' (Graham 1997), and Bacon's (1992) strategies sometimes provided more information than other taxonomies about how strategies were enacted, as in 'listen for structure' and 'piece things together from the details'.[8]

Appendix 1 shows the listening proficiency levels of the two learner groups in each study, along with the basis for division into the groups. Here, the studies used varying criteria to assess their learners' proficiency, from scores on standardized tests (Goh 1998; Murphy 1985) to teacher judgements of listener ability (O'Malley et al. 1989). In addition, there was some variability in the listening proficiency levels of the two groups across the studies. However, in virtually all studies it appears (it was sometimes difficult to pin down the groups' exact levels) that the more-proficient group was slightly higher than, or in the upper reaches of, intermediate level, with the less-proficient group in the lower reaches

---

6 There were two exceptions here. One was Mareschal's (2007) division of the two groups based on general proficiency tests (see Appendix 1). However, there was such a large difference between the two groups on these tests that it seems clear their listening proficiency would have also differed. The second exception was Vandergrift's (1997) distinction between more-proficient and less-proficient learners based on the ACTFL Oral interview. This general proficiency distinction was used in place of Vandergrift's (1997) other distinction between 'successful and less-successful listeners', because he determined this latter distinction more by the listeners' strategy use than their listening proficiency. Note, though, there was a large amount of overlap between the two sets of classifications, with 6 of the 7 more proficient learners (ACTFL intermediate level) also classed as successful listeners, and 10 of the 14 less-proficient learners (ACTFL novice level) labeled unsuccessful listeners.

7 Macaro et al. (2007) took note of learners' linguistic proficiency in their research, where possible, to examine the relationship between linguistic knowledge, strategy use and listening success (e. g., whether effective strategy use was conditional on a high level of linguistic knowledge, or whether listeners could use strategies to overcome low linguistic levels). This issue was of minor concern in the present research (note also that only one study, by O'Malley et al. 1989, included details on linguistic proficiency, and only a small number of studies included information on learners' general proficiency).

8 Goh (2002) would call such detailed descriptions *tactics* used to implement strategies.

of intermediate level. The main exception was Vandergrift (2003), where both groups were sub-intermediate level listeners.

Appendices 2 and 3 also outline, respectively, the verbal report procedures used in the various studies, and measures applied to enhance the reliability of data coding. In relation to the latter (Appendix 3), most of the studies used inter-rater reliability procedures, although some of these studies did not report these figures. In addition, researchers regularly provided the definitions that guided their coding of the strategies, adding a further degree of confidence that the various strategies were coded consistently within and across the studies.

There was also a degree of consistency between how the verbal reports were carried out. Most researchers had their respondents report their thoughts after every few sentences of the passage. The main exception was Bacon (1992), who played the whole text, and had the listeners signal when they used a strategy and report these strategies later when the text had finished. Bacon reasoned that the listeners should have been able to remember many of their strategies because the texts were relatively short. Most researchers also allowed their respondents to report their thoughts in their native language, and in most studies respondents either received training or undertook a pre-sessional task to practice verbalizing their thoughts before their report began. However, the studies differed in whether prompts were used to elicit data, and also in the number of texts that were used, with some using one text and others using more (see Appendix 2).

To compensate, somewhat, for the different aspects of variation between the studies, I triangulated my analysis by giving prominence in the final framework to trends in strategy use that were identified in two or more of the studies. As shown in Table 1, the main developmental trends from these various studies were that less-proficient listeners relied more on either bottom-up strategies or on top-down ones than proficient listeners, who exhibited a more balanced use of these top-down and bottom-up strategies (Bacon 1992; Graham 1997; Murphy 1985; O'Malley et al. 1989). Proficient listeners also used more metacognitive strategies than less-proficient listeners – with this greater metacognitive control also seen as a reason why these proficient listeners were able to orchestrate their top-down and bottom-up strategies in more balanced fashion (Goh 1998; Graham 1997; Vandergrift 1997, 2003).

The body of research also provided a conception of how listeners develop in their use of specific strategies. This was indicated in the various studies by differences between the proficient and less-proficient groups from quantitative and qualitative analyses of the data. The quantitative analyses were in the form of strategy counts. Some would question this use of strategy counts, especially as it is often argued that verbal reports only capture a subset of the listener's

strategies (Ericsson Simon, 1993). However, these tallies were viewed as "suggestive of trends and patterns in strategy use rather than definitive" (Vandergrift 1997: 396): so, where possible, only clear intergroup differences were included here (e. g., where one group out-tallied the other by a substantial margin, or where the researcher concluded a strategy was used far more by one group than the other). The qualitative analyses, on the other hand, enabled insight into such aspects as the flexibility, effectiveness and appropriateness with which the strategies were used (Graham 1997; Vandergrift 1997), along with additional rough overviews of which strategies were used most often. It seemed important, then, to include both these forms of complementary data: the impressionistic, but insightful, views from the qualitative analysis, and the more objective strategy counts, which may have been incomplete, but still provided a point of comparison between the two listener groups.

5 Insights from quantitative findings

Table 2 shows the quantitative differences from the verbal report studies for top-down strategies, with accompanying quantitative figures (when they were provided by the researcher). As shown in Table 2, Vandergrift (2003) found that proficient listeners used the strategy of questioning elaboration more frequently than less-proficient listeners. Questioning elaboration involves using "a

Table 2: Top-down strategy differences from the quantitative data.

Proficient listeners (compared with less-proficient listeners)	Description of strategy
1. More questioning elaboration* (O'Malley et al. 1989; Vandergrift 2003)	Use "a combination of questions and world knowledge to brainstorm logical possibilities" (Vandergrift 2003: 495).
2. More prediction* (Goh 1998; Murphy 1985)	Predict content and details of a text (Goh 1998).
3. More inferencing (Murphy 1985)	Use information, mainly from within the text, to speculate meaning (Murphy 1985)
4. More personal elaboration (Murphy 1985)	Connect what is heard with what is already known at a personal level (called 'Personalizing') (Murphy 1985)
5. More contextualisation (Goh 1998)	"Relate new information to a wider context or situation" … to provide an acceptable interpretation of it (Goh 1998: 135).

Note: *Found in two or more of the studies.

combination of questions and world knowledge to brainstorm logical possibilities" (Vandergrift 2003: 495). Vandergrift (2003) adds that this strategy is "also described by O'Malley et al. (1989) as effective self-questioning" (p. 487), which O'Malley et al. found was used more frequently by their proficient listeners, as well. Proficient listeners were additionally found to make more predictions (Goh 1998; Murphy 1985) and personal elaborations (Murphy 1985), and to use more contextualisation (Goh 1998).

Vandergrift (2003) quantified the strategies as a percentage of the total number of strategies used by each proficiency group, with questioning elaboration constituting 13.13 % of the strategies used by the more-proficient group and 5.91 % by the less-proficient group. O'Malley et al. (1989) did not provide frequency figures for this same strategy. However, Murphy (1985) furnished strategy counts for his learners: these counts, with those for the more-proficient listeners given first, were 80 to 46 for prediction, 389 to 259 for inferencing, and 188 to 46 for personal elaboration. Goh's (1998) criteria for counting a strategy as a majority usage was when a strategy (and the tactics used to enact it) tallied higher for a given proficiency group than for the other proficiency group and when it was also used by five or more of the eight listeners from the group with the higher tally.

Note that findings by O'Malley et al. (1989) showing that proficient listeners used more inferencing and elaborating than less-proficient listeners were not added to Table 2. These omissions were a consequence of O'Malley et al.'s criteria for allocating their participants into the two proficiency groups. While four of these five criteria related to listening proficiency, the fifth was the "ability and willingness to guess at the meaning of unfamiliar words and phrases" (p. 425). If this fifth criterion had been emphasized, it would have brought circularity to the result that the more-proficient group used more inferences and elaboration, strategies which both involve guessing (Macaro 2003). As mentioned above, self-questioning *was* included from O'Malley et al.'s results even though Vandergrift sees it as a form of elaboration. It is included because it involves asking oneself questions in considering possible interpretations of the text, which is different from guessing information.

Table 3 shows the quantitative differences from the verbal report studies for bottom-up strategies. The main trend indicates that proficient listeners employ less translation (Vandergrift 1997, 2003), transfer, repetition (Vandergrift 1997) and fixation (Goh 2000; Graham 1997; Murphy 1985) than less-proficient listeners. Vandergrift (1998a) has called the first three of these (and the same would apply to the fourth, fixation) "inefficient surface processing strategies" that "squander precious time and attentional resources" while not building

Table 3: Bottom-up strategy differences from the quantitative data.

Proficient listeners (compared with less-proficient listeners)	Description of strategy
1. Less translation* (Vandergrift 1997, 2003)	Render "ideas from one language to another in a relatively verbatim manner" (Vandergrift 1997: 394).
2. Less transfer (Vandergrift 1997)	Use L1 knowledge (e. g., cognates) to assist comprehension (Vandergrift 1997).
3. Less repetition (Vandergrift 1997)	Repeat a word or phrase during a listening task (Vandergrift 1997).
4. Less fixation* (Graham 1997; Murphy 1985); also found by Goh (2000)[a]	Focus on a small part of a text to understand it (Goh 1998), generally to the detriment of comprehension (Goh 2000).
5. Use more intonation cues and discourse markers cues (Graham 1997)	Derive meaning from tone of voice; utilize discourse markers used by speaker for cohesion or emphasis (Graham 1997)

Note: *Found in two or more of the studies.
[a]Goh (2000) was a re-analysis of Goh (1998), comparing listening *problems* of the proficient and less-proficient groups.

significant meaning (p. 391). In reference specifically to translation, Vandergrift (2003: 486) claims the listener's "interaction with the text remains superficial, because translation, which involves only surface mapping between languages, generally fails to activate conceptual processes."

Figures for bottom-up strategy use from Vandergrift's studies (shown as a percentage of the total number of strategies used by each proficiency group), with those for the more-proficient group shown first, were: for translation, 4.73 % to 13.09 % (Vandergrift 1997) and 1.85 % to 5.21 % (Vandergrift 2003); for transfer, 2.90 % to 11.38 % (Vandergrift 1997); and for repetition, 5.03 % to 8.67 % (Vandergrift 1997). Figures illustrating lesser use of fixation by proficient listeners were not provided by Murphy (1985), or by Goh (2000) (which was a re-analysis of Goh 1998, comparing listening *problems* of the two proficiency groups). Graham (1997), however, based strategy differences on a 30 % usage differential between the two proficiency groups. Based on this criterion, Graham concluded that proficient listeners used less fixation and also utilized more intonation and discourse marker cues than less-proficient listeners.

The quantitative differences from the verbal report studies for metacognitive strategies are shown in Table 4, with the more-proficient groups showing greater use of comprehension monitoring (O'Malley et al. 1989; figures not provided; Vandergrift 1997, 11.41 % to 4.53 %, 2003, 4.84 % to 2.16 %), problem identification

Table 4: Metacognitive strategy differences from the quantitative data.

Proficient listeners (compared with less-proficient listeners)	Description of strategy
1. More comprehension monitoring* (O'Malley et al. 1989; Vandergrift 1997, 2003)	Check, verify or correct one's comprehension during a listening task (Vandergrift 2003).
2. More problem identification (Vandergrift 1997)	Explicitly identify "the central point needing resolution in a task" or "an aspect of the task that hinders its successful completion" (Vandergrift 1997: 393).
3. More comprehension evaluation (Goh 1998)	Determine "the accuracy and completeness" of one's comprehension after the listening task (Goh 1998: 137–8).
4. More real-time assessment of input (Goh 1998)	Make "on-the-spot decisions about the value of different parts of the input" for achieving comprehension goals (Goh 1998: 137).
5. More selective attention (Graham 1997)	Focus on "particular aspects of a task or the language involved in it" (Graham 1997: 179)

Note: *Found in two or more of the studies.

(Vandergrift 1997, 4.79 % to 2.30 %), comprehension evaluation and real-time assessment of input (Goh 1998) and selective attention (Graham 1997).

6 Insights from qualitative findings

Insights into listener development were also provided from the qualitative analyses from the various studies. Findings from Vandergrift (1998a, 1998b) were added at this stage, which supplied extended qualitative analyses of various listeners from Vandergrift (1997).

An observation by Graham (1997) and Vandergrift (1998a) was that because proficient listeners understood more of the linguistic content, they had less need than lower-level listeners to resort to the compensatory use of top-down knowledge. This was especially in regard to inferencing – with such a strategy often virtually thrust on less-proficient listeners in the absence of alternatives. There was little evidence from the studies reviewed here, though, that proficient listeners used fewer inferences than less-proficient listeners – and Murphy (1985) found they used more. This suggests that when these proficient listeners inferred meaning it was commonly based on more reliable information, and was thus more educated. In addition, some researchers (e. g., Mareschal 2007;

Vandergrift 1998a) found that proficient listeners applied their inferences more effectively. One take on this is that it resulted from these individuals' superior decoding abilities freeing their attention to better orchestrate their inferences. Another possibility stems from the fact that many of the studies used multiple texts in their verbal reports, with the difficulty of these texts tailored to the listeners' levels. Presumably, here, the more difficult texts for the proficient listeners made decoding a challenge for them as well, with better strategic abilities the decisive factor between the two listener groups' use of inferencing. Regardless of the reason, in one of these studies using multiple texts, Mareschal (2007) found that the proficient listeners applied their inferences more judiciously and that the inferences were subject to more "rigorous or systematic verifications" (p. 176). As a result of such factors, both Graham (1997) and Mareschal (2007) found that proficient listeners used inferencing more successfully, being correct with their inferences more often than the less-proficient listeners. This result is shown in Table 5.

Table 5: Top-down strategy differences from the qualitative data.

Proficient listeners (compared with less-proficient listeners)	Description of strategy/tendency
1. Inferences more accurate* (Graham 1997; Mareschal 2007; Vandergrift 1998a)	Inferences made more correctly
2. Inferences based on more reliable information* (Graham 1997; Vandergrift 1998a)	Inferences are more educated

Note: *Found in two or more of the studies.

Among the qualitative findings for bottom-up strategies, Graham (1997) found that effective listeners tended to focus more on content words, such as nouns and verbs, which provide the main semantic cues from the text. This compared with the tendency by less-effective listeners to rely on any words they could understand, a strategy which Graham says can lead to the use of words irrelevant to the central message of the text, thus hindering the listener from grasping its meaning. Graham also noticed that effective listeners recognise words from the text more rapidly than less-effective listeners. These findings by Graham concur with those by Vandergrift (1998a), who stressed that successful listening involves a "focus on semantic cues that can be encoded in memory quickly and efficiently" (p. 391). The differences in bottom-up strategies between more-proficient and less-proficient listeners are shown in Table 6.

Table 7 summarises the qualitative insights into the use of metacognitive strategies. In particular, Vandergrift (1998a, 1998b, 2003) highlights how

Table 6: Bottom-up strategy differences from the qualitative data.

Proficient listeners (compared with less-proficient listeners)	Description of strategy/tendency
1. Segment words better and faster* (Graham 1997; Vandergrift 1998a)	Better and faster ability to identify words in the stream of speech
2. Focus more on semantic cues* (Graham 1997; Vandergrift 1998a)	e. g., Focus on words such as nouns and verbs, which provide its main semantic content, rather than simply on any words understood (Graham 1997).
3. Chunk larger stretches of language* (O'Malley et al. 1989; Vandergrift 1998a, 1998b)	Process words as larger chunks of meaning (Vandergrift 1998a).
4. Piece words together more* (Bacon 1992; Graham 1997)	Rely more on piecing words together than on using isolated words (Bacon 1992)

Note: *Found in two or more of the studies.

Table 7: Metacognitive strategy differences from the qualitative data.

Proficient listeners (compared with less-proficient listeners)	Description of strategy/tendency
1. Form accurate mental model earlier in text* (Vandergrift 1998a, 1998b, 2003)	Form effective mental representation of the text
2. More effective comprehension monitoring* (Bacon 1992; Mareschal 2007)	Check, verify or revise one's interpretations more effectively
3. Greater ability to direct attention* (Bacon 1992; O'Malley et al. 1989).	Avoid distraction from the task by extraneous influences
4. Attend more to rhetorical organisation of the text (Murphy 1985)	Discern main ideas and details, and focus more on organisational structure; less-proficient listeners focus more on resolving difficulties at word level (Murphy 1985).

Note: *Found in two or more of the studies.

proficient listeners establish an accurate mental model early in the text which provides them with a reliable framework for interpreting upcoming content and also for suppressing irrelevant information. Vandergrift (1998a) also outlines how, conversely, a poorly-developed mental model, associated with less-effective listening, often leads to the learner becoming overwhelmed by the volume of information in a text.

A further general observation on strategy use, by Graham (1997), was that more-proficient listeners combined their strategies more effectively than less-proficient

listeners. In addition, Murphy (1985) found that proficient listeners used strategies more flexibly than less-proficient listeners. Both of these results fit in with findings mentioned earlier that proficient listeners deploy their strategies better metacognitively (Goh 1998; Vandergrift 1997, 2003). These results also support the contention by other L2 listening scholars that the flexibility and effectiveness of the strategies used discriminates between efficient and inefficient listening, more so than the number of strategies used (Graham et al. 2008; Macaro 2006; Macaro et al. 2007). Nevertheless, both Bacon (1992) and Murphy (1985) concluded that the proficient listeners from their research not only used a wider variety of strategies, but also used more strategies. As Murphy (1985) explains, "In other words, the more proficient listeners did more while interacting with the selections presented to them" (p. 37). Goh (1998) and Vandergrift (1997) also found their more-proficient listeners to use more strategies.

7 The resulting framework

In summarizing the resulting framework, its first aspect, the general trends in strategy development with increasing listener proficiency, is shown in Table 8. These findings were all identified in two or more of the studies.

Table 8: Part one of the framework: General trends in strategy development.

- use top-down and bottom-up strategies in more balanced fashion
- use more metacognitive strategies
- combine strategies better, use them more flexibly and deploy them more appropriately to the listening situation
- use more strategies[a] and a wider range of them.

Note: [a]This result seems open to question, as discussed later in this paper.

The second aspect of the framework involving findings specific to the development of individual strategies is shown in Table 9. The results in the left-hand column are given prominence because they are the 'primary' strategy use trends that were observed in at least two studies. Those in the right column were observed in one study only, but are included here for reference.

    Note that it is debatable whether some of the processes in Table 9 actually are strategies. For example, from the primary strategy column, segmenting words, chunking language, and forming an early mental model could simply be considered skills that all listeners use – 'skills' being processes that characterize fluent

Table 9: Part two of the framework: Development of individual strategies.

Primary trends in strategy use (identified in two or more studies)	Secondary trends in strategy use(identified in one study only; shown here for reference)
Bottom-up	
– Segment words (from the speech signal) better and faster	– Use more intonation cues and discourse markers
– Focus more on semantic cues, such as nouns and verbs	– Use less repetition and transfer
– Chunk larger stretches of language	
– Piece words together more	
– Use less translation and fixation	
Top-down	
– Use more prediction	– Use more contextualization
– Use more questioning elaboration	– Use more inferences
– Inferences are more accurate	– Use more personal elaborations
– Inferences are based on more reliable information	
Metacognitive	
– Form accurate mental model earlier in text	– Identify problems more often
– Use more comprehension monitoring	– Use more selective attention
– Comprehension monitoring is more effective	– Use more comprehension evaluation
– Direct attention more effectively	– Use more real-time assessment of input
	– Attend more to rhetorical organisation of the text

comprehension (Afflerbach et al. 2008). On the other hand, they *would* be considered strategies if listeners used them to pre-empt probable comprehension problems: for example, if listeners purposely chunked larger units of meaning, or if they deliberately focused more on the gist than local aspects of the text to develop a mental model earlier than normal (to interpret further information more effectively). And segmenting words would be strategic when listeners purposely used stress patterns from the L2 to identify, for example, where words began (in English) or ended (in French) (see Graham and Macaro 2008).

The second component of the study was to organize the strategies into the three processing stages of perception, parsing and utilization (Anderson 1985, 2010). These results are shown in Table 10, which only includes the primary strategy use trends from Table 9. There are some debatable choices here. One is including directed attention at the perception phase. It needed to be placed into one of the three stages and is put here because it seems best suited to this lowest (or earliest) of the stages; note that Bacon (1992), Goh (2000) and O'Malley et al.

Table 10: The strategies reorganized into Anderson's three processing stages.

Perception stage
– Direct attention more effectively
– Segment words (from the speech signal) better and faster
– Focus more on semantic cues, such as nouns and verbs (i. e., stressed content words)
Parsing stage
– Piece words together more
– Chunk larger stretches of language
– Use less translation and fixation
Utilization stage
– Inferences are used more accurately
– Inferences are based on more reliable information
– Use more prediction
– Use more questioning elaboration
– Form accurate mental model earlier in text
– Use more comprehension monitoring
– Comprehension monitoring is more effective

(1989) also placed it at this perception stage. However, one could also take Buck's (1990) view that attention is separate from perception, occurring before it, meaning the strategy would not easily fit into any of the stages.

Placing the strategy of chunking language at the parsing stage was also slightly contentious. It is placed here mainly because O'Malley et al. (1989) and Vandergrift (1998a, 1998b) (the studies which found that proficient listeners used larger chunks than less-proficient listeners) emphasized that the proficient listeners used these larger chunks mainly in forming better meaning representations of the text. Note that the strategy, though, could also have slotted into the perception stage. This is because, like two of the strategies placed at the perception stage (segmenting words and identifying stressed semantic cues), chunking also involves using perceptual cues to identify the units.

The positioning of 'less fixation' at the parsing stage is also arguable. One perspective here is to categorize it in the perception stage, based on the idea that it represents overcoming an attention problem as listeners more readily divert their attention from a fixated word back to the text (Goh 2000). A different perspective is that because many of the proficient listeners possess various decoding advantages over their less-proficient peers (e. g., better abilities to segment words, identify stressed semantic cues, and piece words together; Graham 1997) the proficient listeners probably use fixation less than lower-level listeners because they have less need to. That is, their superior decoding abilities simply lead to better construction of meaning at the parsing stage (with

less straying of attention from the task). Given Buck's (1990) reservation over equating attention with perception, I sided with this second view.

Another debatable choice was to include prediction at the utilization stage. Some might argue it is better placed at a lower stage, firstly, because listeners often anticipate words that come next in an utterance (parsing stage) (Anderson 2010), and secondly, because listeners can predict even before an utterance begins. However, the strategy is placed at the higher, utilization stage because the strategic use of predicting key words and information first requires activation of the listener's prior knowledge (even contextual information, such as from advance organizers in classroom listening activities, would stimulate a tentative hypothesis of the text from which to deploy such predictions).[9]

Table 10 indicates the following notion of how the use of strategies at the three comprehension stages tends to develop. At the *perception stage*, listeners are better able to direct their attention to the task. They also segment words from the speech signal better and faster (using various phonological and phonotactic cues), and focus more on semantic cues, especially stressed content words such as nouns and verbs. In the *parsing stage*, the listeners tend to piece such content words together, and also chunk larger stretches of language (phrases, clauses and sentences), in building their meaning representations of the textual information. In addition, they use less translation, and fixate less frequently on local textual aspects to understand them. Then in the *utilization stage*, inferences are used more accurately, partly because they are often based on more reliable information. Also, listeners predict upcoming content more often, a function of their more active approach to comprehension. Listeners also use questions and world knowledge to brainstorm logical possibilities for their interpretations (questioning elaboration). In addition, they form an accurate mental model of the text earlier in the passage, and maintain it better by the use of more frequent and effective comprehension monitoring.

8 Discussion

A major reservation about the results, one relating to the general trends in listener development (Table 8), is the idea that an increase in the total *number* of strategies listeners use represents progress. Earlier, it was mentioned how this reservation

---

9 The perception and parsing stages would probably include top-down processing through automatic mechanisms such as spreading activation, and the use of context or prior knowledge to assist or modify listeners' decoding of the words (for example, see Field (2008: 133), on how top-down information may assist in decoding the word 'vegetables').

was also expressed by other researchers, who suggest that how appropriately the strategies are used better indicates listener success (e. g., Graham et al. 2008). This latter sentiment also underscores many of the findings in the framework developed here in listeners' use of individual strategies, shown in Tables 9 and 10 (proficient listeners' inferences are more accurate, their monitoring is more effective, and so forth). However, a number of the studies reviewed in this analysis *did* find that more-proficient listeners used more strategies (e. g., Murphy 1985). From this perspective, it is understandable that greater strategy use may indicate more active processing by the listener. However, a better approach is to distinguish between strategies that assist comprehension (chunking information, piecing content words together, monitoring, and so forth), and those that commonly hinder comprehension (such as translation, fixation, and trying to understand every word). Then an increase in the former group of strategies rather than the latter ones would better signal listener progress – as would a decrease in the latter type of strategies that hinder understanding. (A further take on this is that ultimately as learners reach quite advanced levels they would be expected to need *fewer* strategies in most listening situations, as they could come to rely more on automatic processes and less on strategic ones.)

After examining the strategies taxonomies from the various studies, there are also some recommendations for research, mainly with the aim of improving how some strategies are defined. One area, in particular, concerns the definitions of inferencing and elaborating. Commonly, the two strategies are distinguished based on the sources of knowledge the listener uses. This sees inferencing as using information from *within* the text, and elaborating as using information from *outside* it (e. g., Goh 1998; O'Malley et al. 1989; Vandergrift 2003; Vandergrift and Goh 2012). It is recommended here that a better approach is to redefine these two strategies according to their purpose. A common purpose for both is guessing, which is simply working out meaning stated by the speaker that the listener does not understand (Graham 1997; Lynch 2006). Once the strategy is defined as guessing, the researcher can then choose to provide more information about the guess: what sources of knowledge were primarily used to make it; whether the strategy was used to guess a word or guess a stretch of information; whether the guess was based on reliable information (an educated guess) or less reliable information (a wild guess), and so on.

Other purposes for inferences and elaborations can then also be distinguished from guessing, providing better insights into the listener's behaviours. As used by cognitive psychologists, *inferencing* commonly involves working out meaning not directly stated by the speaker (Kintsch 1998; Rost 2002). Characteristically, such inferences are regular listening processes rather than strategies, and they include deducing a speaker's intentions (Field 2008) and

making bridging inferences, which provide links between ideas to maintain coherence (Anderson 2010). For example, the listener would have to make bridging inferences to connect the following two statements: "Tom walked into the kitchen. The refrigerator was empty." In addition, cognitive psychologists point to another common form of elaboration, known as elaborative inference, which adds extra information to enrich an interpretation. Anderson (2010) points out that such elaborations are not essential to understanding the text. However, they may sometimes serve the strategic purpose of strengthening schema in memory, thus assisting the listener's subsequent recall of information (O'Malley et al. 1989); note that O'Malley et al. (1989) add that the success of this recall strategy may be conditional on how strongly the schema is activated by the listener. The main point here, though, is to recommend that researchers differentiate processes associated with inferencing and elaborating by purpose, rather than simply by knowledge source, in order to provide a clearer picture of a listener's behaviors.[10]

Another recommendation is for researchers to report their participants' use of some strategies in more detail, particularly that of comprehension monitoring. While examining the various studies during the analysis, it was often unclear whether the researchers defined monitoring as checking interpretation at both the discourse and local levels or simply at the discourse level. There is a difference between the two, which needs to be acknowledged in the research. The strategy used at the discourse level, which involves checking one's ongoing mental model, is usually more crucial to comprehension than checking whether a word or minor detail has been correctly understood. Another problem observed through the studies was that, sometimes because the definition used for monitoring included overseeing one's *comprehension* or *performance* (e. g., Vandergift 1997, 2003), it was unclear whether this involved only monitoring one's understanding of the text, or whether it also included monitoring one's strategy use.

There was also little clarification through the studies of what form the monitoring took, with most studies just tallying the total number of times monitoring occurred. However, O'Malley and Chamot (1990: 137) and Vandergrift (1997, 2003) define the strategy as "checking, verifying, or correcting one's understanding". Each of these different functions indicates a great deal about how the strategy is used by the listener, and it is suggested that these functions be

---

10 Questioning elaboration, outlined earlier, is already seen as serving a distinct strategic purpose, unlike 'regular' elaborations (personal, world, academic, etc.) which simply relate to the knowledge source used. Vandergrift (2003) describes questioning elaboration as "more metacognitive [than cognitive] in nature" (p. 495), adding that the strategy "demonstrates a flexibility essential to success in comprehension" (p.488).

acknowledged in research results. Additionally, Goh (2002) and Graham et al. (2010) have pointed out the need to acknowledge whether comprehension has actually taken place – whether the monitoring was successful or not.[11]

While making these recommendations, I also acknowledge it is often more convenient to include a minimal number of strategy-types in one's research in order to make analysing and reporting the research less complicated. However, illuminating the various functions of an 'overarching' strategy in one's results can present a more complete picture of listeners' behaviours to the reader. This argument is also used by Goh (2002) for breaking strategies down into the tactics used to operationalize them. In fact, Goh's use of tactics would be a useful solution for clarifying another fuzzily-defined strategy through the various studies, that of selective attention. Goh's (2002) tactics for enacting this strategy include, listen to words in groups, listen for gist, and listen for familiar content words (p. 193). Field's (2008) strategy taxonomy also usefully distinguishes between different forms of selective attention.

This analysis of such a number of studies was not without its limitations, with the main ones discussed earlier. Measures to address some of these limitations were also outlined, as were aspects of the study that enhanced its reliability. In addition, results from two further studies add reliability to the findings. The first was the initial cross-sectional aspect of a longitudinal study by Chen (2009) who collected data through learner diaries (had the analysis been expanded beyond verbal report studies, Chen's was the only other study considered suitable for inclusion.[12]) Chen found that less-proficient listeners have a more bottom-up approach to listening than proficient listeners, who use a better balance of top-down and bottom-up strategies. Chen also found that proficient listeners use more metacognitive strategies, use more directed attention, less fixation, and attempt less to understand each word. These results from Chen's study corresponded with the findings here, lending weight to them. The framework here also subsumed all of the strategies Macaro et al. (2007) found that clearly distinguished the two proficiency groups (listed in the introduction to this paper). This correspondence between their analysis and mine was not overly

---

[11] In my research, where I profiled individual listeners (Yeldham 2009; Yeldham and Gruba 2016), I indicated when they did not adequately monitor their mental model. This was to identify a discourse-level dysfunction known as perseverative text processing, where the listener perseveres with an incorrect mental model despite encountering subsequent evidence that it is faulty.

[12] Had the analysis included other techniques, then other cross-sectional studies that used close-ended questionnaire (e. g., Vogely 1995) and structured interview questions (Chien and Wei 1998) to collect their data would not have been considered because these procedures tend to guide and limit listeners' choices in reported strategy use.

surprising, given that we examined many of the same studies. Nevertheless, there were also a number of studies examined distinct to each analysis, and the fact that all the strategies in Macaro et al.'s results were included here adds further credence to the findings.

9 Conclusion

In conclusion it must be stressed that the framework developed here is not advised as a guide for listening instruction. Instructors are not advised simply to teach their less-proficient listeners the strategy use patterns found here to be used by proficient listeners. As scholars have pointed out (e. g., Graham and Macaro 2008), because such findings are derived from cross-sectional studies, they do not imply causality. That is, the strategies used by more-proficient listeners could be strategies these listeners use *when* they become more advanced-level learners, rather being a cause of their advanced ability.

Nor is it suggested that all listeners will develop in the manner indicated by the framework. Individual listeners will have different developmental paths. The framework is foremost to signpost researcher evaluations of listeners' progress in longitudinal studies. Note also that in using the framework for this purpose, researchers must leave open the possibility of refining the framework to make it more applicable to the data from their own participants (Miles and Huberman 1994).

References

Afflerbach, Peter P., David Pearson & Scott G. Paris. 2008. Clarifying differences between reading skills and reading strategies. *The Reading Teacher* 61(5). 364–373.

Anderson, John R. 1985. *Cognitive psychology and its implications*, 2nd edn. New York: Freeman.

Anderson, John R. 2010. *Cognitive psychology and its implications*, 7th edn. New York: Freeman.

Bacon, Susan M. 1992. Phases of listening to authentic input in Spanish: A descriptive study. *Foreign Language Annals* 25(4). 317–333.

Berne, Jane E. 2004. Listening comprehension strategies: A review of the literature. *Foreign Language Annals* 37(4). 521–531.

Bowles, Melissa A. 2010. *The think-aloud controversy in language acquisition research*. New York: Routledge.

Buck, Gary. 1990. *The testing of second language listening comprehension*. Lancaster: University of Lancaster dissertation.

Buck, Gary. 1991. The testing of listening comprehension: An introspective study. *Language Testing* 8(1). 67–91.

Buck, Gary. 1994. The appropriacy of psychometric measurement models for testing second language listening comprehension. *Language Testing* 11(2). 145–170.

Chamot, Anna U. 2004. Issues in language learning strategy research and teaching. *Electronic Journal of Foreign Language Teaching* 1(1). 14–26.

Chamot, Anna U. 2005. Language learning strategy instruction: Current issues and research. *Annual Review of Applied Linguistics* 25. 112–130.

Chen, Ai-hua. 2009. Listening strategy instruction: Exploring Taiwanese college students' strategy development. *Asian EFL Journal* 11(2). 54–85.

Chen, Yiching. 2007. Learning to learn: The impact of strategy training. *ELT Journal* 61(1). 20–29.

Chien, Ching-ning & Li Wei. 1998. The strategy use in listening comprehension for EFL learners in Taiwan. *RELC Journal* 29. 66–91.

Cohen, Andrew D. 1998. *Strategies in learning and using a second language*. London: Longman.

Cross, Jeremy. 2010. Raising L2 listeners' metacognitive awareness: A sociocultural theory perspective. *Language Awareness* 19(4). 281–297.

DeKeyser, Robert M. 2000. The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22(4). 493–553.

Ericsson, K. Anders & Herbert A. Simon. 1993. *Protocol analysis: Verbal reports as data*, Revised edn. Cambridge, MA: MIT Press.

Field, John. 2004. An insight into listeners' problems: Too much bottom-up or too much top-down? *System* 32. 363–377.

Field, John. 2008. *Listening in the language classroom*. Cambridge: Cambridge University Press.

Flavell, John H. 1976. Metacognitive aspects of problem solving. In Lauren B. Resnick (ed.), *The nature of intelligence*, 231–235. Hillsdale, NJ: Erlbaum.

Goh, Christine. 1998. How ESL learners with different listening abilities use comprehension strategies and tactics. *Language Teaching Research* 2(2). 124–147.

Goh, Christine. 2000. A cognitive perspective on language learners' listening comprehension problems. *System* 28(1). 55–75.

Goh, Christine. 2002. Exploring listening comprehension tactics and their interaction patterns. *System* 30(2). 185–206.

Goh, Christine. 2005. Second Language Listening Expertise. In K. Johnson (ed.), *Expertise in second language learning and teaching*, 64–84. Basingstoke, UK: Palgrave Macmillan.

Graham, Suzanne. 1997. *Effective language learning: Positive strategies for advanced level language learning*. Clevedon Avon, England: Multilingual Matters Ltd.

Graham, Suzanne. 2003. Learner strategies and advanced level listening comprehension. *Language Learning Journal* 28(1). 64–69.

Graham, Suzanne & Ernesto Macaro. 2008. Strategy instruction in listening for lower-intermediate learners of French. *Language Learning* 58(4). 747–783.

Graham, Suzanne, Denise Santos & Robert Vanderplank. 2008. Listening comprehension and strategy use: A longitudinal exploration. *System* 36(1). 52–68.

Graham, Suzanne, Denise Santos & Robert Vanderplank. 2010. Strategy clusters and sources of knowledge in French L2 listening comprehension. *International Journal of Innovation in Language Learning and Teaching* 4(1). 1–20.

Graham, Suzanne, Denise Santos & Robert Vanderplank. 2011. Exploring the relationship between listening development and strategy use. *Language Teaching Research* 15(4). 435–456.

Kintsch, Walter. 1998. *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.

Lynch, Tony. 2006. Academic listening: Marrying top and bottom. In Esther Usó-Juan & Alice Martínez-Flor (eds.), *Current trends in the development and teaching of the four language skills*, 91–110. Berlin: M. de Gruyter.

Macaro, Ernesto. 2003. *Teaching and learning a second language: A review of recent research*. London, NY: Continuum.

Macaro, Ernesto. 2006. Strategies for language learning and for language use: Revising the theoretical framework. *The Modern Language Journal* 90(3). 320–337.

Macaro, Ernesto, Suzanne Graham & Robert Vanderplank. 2007. A review of listening strategies: Focus on sources of knowledge and on success. In Andrew D. Cohen & Ernesto Macaro (eds.), *Language learner strategies: 30 years of research and practice*, 165–185. Oxford: Oxford University Press.

Mareschal, Catherine. 2007. *Student perceptions of a self-regulatory approach to second language listening comprehension development*. Ottawa: University of Ottawa dissertation.

Miles, Matthew B. & A. Michael Huberman. 1994. *Qualitative data analysis: An expanded sourcebook*, 2nd edn. Thousand Oaks: Sage Publications.

Murphy, John M. 1985. *An investigation into the listening strategies of ESL college students*. ERIC Document Reproduction Service (No. ED 278 275).

O'Malley, J. Michael & Anna U. Chamot. 1990. *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.

O'Malley, J. Michael, Anna U. Chamot & Lisa Küpper. 1989. Listening comprehension strategies in second language listening. *Applied Linguistics* 10(4). 418–437.

Peters, Martine. 1999. *The listening comprehension strategies of students in a French immersion program*. Ottawa: University of Ottawa dissertation.

Piaget, Jean. 1957. *Construction of reality in the child*. London: Routledge & Kegan Paul.

Richards, Jack C. 2003. Current trends in teaching listening and speaking. *The Language Teacher* Online. 27(7). http://jalt-publications.org/old_tlt/articles/2003/07/richards (accessed 14 January 2008).

Rost, Michael. 2002. *Teaching and researching listening*. Harlow: Longman.

Seliger, Herbert W. 1983. The language learner as linguist: Of metaphors and realities. *Applied Linguistics* 4(3). 179–191.

Thompson, Irene & Joan Rubin. 1996. Can strategy instruction improve listening comprehension? *Foreign Language Annals* 29(3). 331–342.

Vandergrift, Larry. 1997. The strategies of second language (French) listeners. *Foreign Language Annals* 30(3). 387–409.

Vandergrift, Larry. 1998a. Successful and less successful listeners in French: What are the strategy differences? *The French Review* 71(3). 370–395.

Vandergrift, Larry. 1998b. Constructing meaning in L2 listening: Evidence from protocols. In Sharon Lapkin (ed.), *French second language education in Canada: Empirical studies*, 89–119. Toronto: University of Toronto Press.

Vandergrift, Larry. 2003. Orchestrating strategy use: Toward a model of the skilled second language listener. *Language Learning* 53(3). 463–496.

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

Vandergrift, Larry & Christine Goh. 2012. *Teaching and learning second language listening*. New York: Routledge.

Vogely, Anita. 1995. Performance of three authentic listening comprehension tasks. *The Modern Language Journal* 79(1). 41–56.

Wu, Yi'an. 1998. What do tests of listening comprehension test? – A retrospection study of EFL test-takers performing a multiple-choice task. *Language Testing* 15(1). 21–44.

Yeldham, Michael. 2009. *Approaches to second language listening instruction: Investigating the 'top-down/bottom-up debate'*. Melbourne: University of Melbourne dissertation.

Yeldham, Michael & Rainbow T.-H. Chen. 2016. Investigating mediation styles of second language listener verbal reports. *Applied Linguistics Review* 7(2). 203–233.

Yeldham, Michael & Paul Gruba. 2016. The development of individual learners in an L2 listening strategies course. *Language Teaching Research* 20(1). 9–34.

Appendix 1. Proficiency levels of learner groups

Study	Basis for proficiency division	Less-proficient listeners	More-proficient listeners
Murphy (1985)	Michigan Test of Aural Comprehension; Listening rating scale; Reading test	Less-proficient intermediate level listeners	More-proficient intermediate level listeners
O'Malley et al. (1989)	Teacher judgement based on various criteria, including classroom listening ability	"Ineffective listeners" at intermediate level	"Effective listeners" at intermediate level
Bacon (1992)	Listening recall protocol scores	Approx. lower-intermediate	Approx. intermediate
Graham (1997)	Teacher judgement; Researcher judgement from performance in verbal report	Five years or more of learning the language (approx. intermediate)	Five years or more of learning the language (approx. higher intermediate)
Goh (1998)	SLEP listening test	Mean score of 49/75 (approx. lower-intermediate to intermediate)[a]	Mean score of 65/75 (approx. higher-intermediate)
Vandergrift (1997, 1998a, 1998b)	ACTFL Oral Proficiency Interview	ACTFL Novice levels 1,2,3 (Beginner to lower-intermediate)	ACTFL Intermediate levels 1,2,3

(continued)

(continued)

Study	Basis for proficiency division	Less-proficient listeners	More-proficient listeners
Vandergrift (2003)	Listening test (with its reliability established for the study)	Less-proficient listeners (based on listening test) from ACTFL Novice-level (Listeners approx. High-beginner)	More-proficient listeners (based on listening test) from ACTFL Novice-level (Listeners approx. Lower-intermediate)
Mareschal (2007)	Oral interview; Language aptitude test	"Beginner- intermediate" (from lowest 15 % of learners at school)	"Intermediate- advanced" (from top 5 % of learners at school)

Note: [a]SLEP proficiency conversions here are based on comparisons with scores from other standardized tests, and also on the author's experience using the test in research (Yeldham 2009), equating scores and associated proficiency levels to those of his participants who took the SLEP.

Appendix 2. Verbal report process characteristics of the studies

Study	Pre-session training	Number and nature of texts	Segmentation of text(s)	Report in L1 or L2?	Mediation method used
Murphy (1985)	Training sessions (2) (listening think-aloud tasks)	Six texts: Academic lecture format	Segmented by listener	L2	Respondents guided by pre-sessional instructions sheet.
O'Malley et al. (1989)	Training session (listening think-aloud tasks)	Large number of texts, of varying types	Researcher segmented, every two or three sentences	L1 or L2	Respondents asked how they made sense of the text; what was unclear; and what images, if any, occurred to them.
Bacon (1992)	Pre-session warm-up (listening think-aloud task)	One text used from choice of two	Whole text played[a]	L1	Open and close-ended questions asking respondents how they had tried to understand.
Graham (1997)	Not specified	One text each for French and German learners	Researcher segmented, every two or three sentences	L1	Respondents guided by pre-sessional instructions sheet

(continued)

(*continued*)

Study	Pre-session training	Number and nature of texts	Segmentation of text(s)	Report in L1 or L2?	Mediation method used
Goh (1998)	Pre-session warm-up (listening think-aloud task)	One text used from choice of two	Short segments read by the researcher	L1 and L2	Respondents guided by pre-sessional instructions to verbalise how they had tried to understand
Vandergrift (1997, 1998a, 1998b); Vandergrift (2003)	Training session (various think-aloud tasks)	Texts of varying types, tailored to listeners' levels	Researcher segmented, every two or three sentences	L1	Use of prompts such as: What are you thinking? How do you know that? What else are you thinking?
Mareschal (2007)	None	Texts of varying types, tailored to listeners' levels	Researcher segmented, every one or two sentences	L1	Use of prompts such as: "What are you thinking now?", "How did you come to this understanding?"

Note: [a]In Bacon (1992), the text was played without breaks. Listeners raised a finger to indicate when they were using strategies, explaining these to the researcher after the text was completed. The researcher says this method was used as a compromise between inserting unnatural breaks in the text and the more natural process of listening to the whole text.

Appendix 3. Reliability of verbal report strategy coding

Study	Inter-rater reliability index	Intra-rater reliability index
Murphy (1985)	0.81	0.80
O'Malley et al. (1989)	0.96 (and 'uncertainties', constituting 7 % of dataset, also coded by both raters in tandem)	Not provided
Bacon (1992)	0.88	0.93
Graham (1997)	Not provided	Not provided
Goh (1998)	Yes, but figure not provided	0.88
Vandergrift (1997; 1998a, 1998b)	0.78	Not provided
Vandergrift (2003)	Yes, but figure not provided	Not provided
Mareschal (2007)	Not provided	Not provided