英語教學 *English Teaching & Learning*
38.3 (2014 Special Issue): 71-101
DOI: 10.6330/ETL.2014.38.3.03

The Effectiveness of Listening Support on L2 Learners' Listening Comprehension Ability: A Meta-Analysis

Natsuko Shintani[a,*], Matthew P. Wallace[b]

Abstract

This article reports on a meta-analysis of studies that investigated the effectiveness of listening support on L2 listening comprehension ability. Altogether 18 studies were retrieved and coded for effect sizes from pre-to-post change. We first analyzed the effects of listening support by comparing listening practice with support to listening practice without support (i.e., listening exposure only). We then compared the effects of two main types of listening support: linguistic support (e.g., vocabulary teaching) and contextual support (e.g., pictorial information). The subsequent moderator analysis tested the impact of three mediator variables: the timing of listening support, the length of instruction, and the agent of input control (either teacher or students). The results showed that (1) listening support increased the effectiveness of listening practice; and (2) linguistic support showed a larger effect size than contextual support. The moderator analysis showed that there were no significant group differences in the timing of listening support or the length of listening practice. However, listening practice where students controlled the listening materials had a significantly larger effect size than teacher-controlled input practice. We discuss the theoretical and pedagogical significance of these findings and the role of small-scale meta-analyses.

Key Words: second language listening instruction, meta-analysis, second language listening support

[a] School of Curriculum and Pedagogy, Faculty of Education, University of Auckland. E-mail: n.shintani@auckland.ac.nz
[b] National Institute of Education, Nanyang Technological University. E-mail: mpw2121@gmail.com
[*] Corresponding Author

INTRODUCTION

Listening instruction has been defined as a "pedagogic plan that focuses on ... improving language learner comprehension of spoken language" (Rost, 2006, p. 47). Historically, listening instruction methodology has evolved along with language-teaching trends. Three types of listening instruction in the history of second language acquisition research have been identified by Vandergrift and Goh (2012): *Text-oriented*, *communication-oriented*, and *learner-oriented*. *Text-oriented* listening instruction gained popularity when Grammar Translation was the dominant language-teaching methodology. This approach used written texts to raise awareness of language features by emphasizing bottom-up processing skills (e.g., decoding words and sentences). Common activities of the *text-oriented* approach involved teachers reading aloud written texts as learners completed dictation exercises or cloze-type tasks. With the rise of Communicative Language Teaching (CLT), listening instruction began to involve the use of more authentic listening materials. In these activities, teachers typically activate background knowledge of the topic prior to listening by discussing the listening topic or pre-teaching vocabulary, then guiding learners through communicative tasks (e.g., information-gap activity) that require the use of information from the input to complete (Vandergrift & Goh, 2012).

More recently, based on research of successful language learners, a *learner-oriented* approach to instruction has emerged, which specifically focuses on teaching the skills and strategies employed by successful listeners. Research on listening strategies has shown that successful listeners use more varied listening strategies than unsuccessful learners (Goh, 2000, 2002; Rost & Ross, 1991; Vandergrift, 2003). In particular, they make greater use of metacognitive strategies (i.e., techniques that enhance learning such as planning and monitoring) (Goh, 2008; Vandgergrift, 2002; Vandergrift, Goh, Mareschal, & Tafaghodtari, 2006). Surprisingly, however, studies investigating the effect of strategy training on learners' L2 proficiency have shown relatively modest effects (Carrier, 2003; Graham & Macaro, 2008; Thompson & Rubin, 1996). Plonsky's (2011) meta-analysis of studies that investigated strategy instruction reported a much smaller effect (d = 0.06) when the focus was on training listening strategies than on other skills (e.g., speaking: d = 0.97, reading: d = 0.74, and writing: d = 0.42).

Another type of listening instruction which has been widely used in language pedagogy but remains relatively under-researched is listening practice. Listening practice in this study refers to any pedagogical

activity that engages learners in listening to L2 input in order to develop their listening comprehension ability. Listening practice activities consist of two principle elements: listening exposure and listening support. The former involves giving learners exposure to oral L2 input, which might be followed by comprehension tasks that require an understanding of the content of the listening texts. Listening support refers to the provision of some kind of information, either linguistic (e.g., vocabulary list) or contextual (e.g., pictures), about the listening materials to assist learners' comprehension.

Figure 1 summarises the terms reviewed above. Listening instruction is an umbrella term including both listening strategy training and listening practice. The current meta-analysis investigated the effects of (1) listening practice, (2) listening support, and (3) linguistic and contextual support on L2 comprehension skills.

Investigating listening practice is of considerable pedagogical importance as listening skills can be a challenge for foreign language learners to develop (e.g., the participants in our meta-analysis) (Farrell & Mallard, 2006). Goh (2000) reported multiple problems that foreign language learners face in listening, such as the inability to recognize word meaning, maintaining the meaning of the message in working memory, and making the connection between word-by-word meanings and the actual ideas the text is conveying. L2 listening, thus, requires the utilization of multiple skills, especially when the learners cannot control the speed of the listening text (Farrell & Mallard, 2006). The following provides an overview of the research that has investigated the effects of listening practice.

The majority of studies investigating listening practice have involved instruction consisting of both listening exposure and support (e.g., Jensen & Vinther, 2003; Thompson & Rubin, 1996). The type of support varies according to the type of information provided, such as key vocabulary (e.g., Chung, 2002; Jafari & Hashim, 2012; Pan, 2012), availability of a transcription of the text (e.g., Chang, 2010,

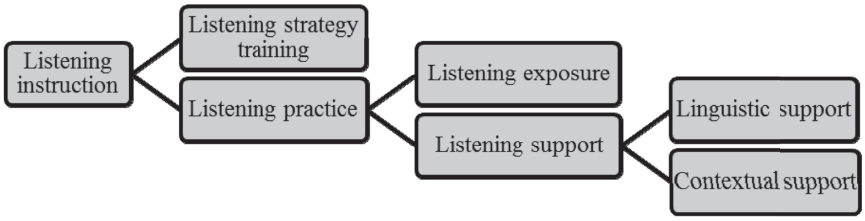

Figure 1
Types of Listening Instruction

2011; Chang & Millett, 2013), video captions (e.g., Hayati, 2010; Hsu, Hwang, Chang, & Chang, 2013), phonetic features (Brown & Hilferty, 1986; Matsuzawa, 2006; Underwood & Wallace, 2012), accompanying pictures (e.g., Gilakjani & Ahmadi, 2011; Othman & Vanathas, 2005; Teichert, 1996), and videos (e.g., Soureshjani & Etemadi, 2012; Verdugo & Belmonte, 2007). These studies have shown that providing support for listening practice has positive effects on the development of listening comprehension ability. On the other hand, only a few studies have examined the effects of a listening-only treatment—those which typically involve extensive listening (Chang, 2010, 2011; Chang & Millett, 2013). Other studies have used a listening-only condition for the control group (e.g., Ghorbani, 2011; Verdugo & Belmonte, 2007). Such studies generally show smaller effects of the listening-only conditions compared to the listening-plus-support condition. To confirm whether these findings were robust, the present meta-analysis compares the two conditions in the studies reviewed.

Linguistic Versus Contextual Support

As previously mentioned, listening support can be categorised as either linguistic support or background-contextual support. Linguistic support provides information via words or phonological features that appear in the listening texts. In contrast, contextual support addresses the content of the listening texts by providing visual and/or background information. Contextual information may involve a written summary of a story the learners will be asked to listen to. Though these summaries may sometimes include words or sentences that appear in the listening text, we have chosen to code them as "contextual" information, because their main purpose is to provide learners with content schema for the listening text rather than to pre-teach the linguistic forms that appear therein.

The studies investigated in the present meta-analysis provided linguistic support in three principal ways: (1) key vocabulary, (2) a written version of the whole text, and (3) phonological information. A common way to provide key vocabulary is to present the learner with a word list and corresponding definitions. Jafari and Hashim (2012) gave Iranian university EFL learners key words on cards before they listened to short (1 ~ 4 minutes) passages. The cards provided definitions in English, two sentences with each word in context, and translations in the learners' native language. Other studies included activities for teaching key words. For example, Pan (2012) gave Taiwanese university learners a list of key words from an upcoming text, followed by vocabulary-building activities where they had to repeatedly use the words in

context. Chung (2002) provided Taiwanese college EFL students a list of key words with a teacher-led explanation of their meanings in their native language, together with opportunities to say the words aloud before they watched a video. All these studies reported positive effects for pre-listening vocabulary teaching.

Support involving a written version of a text can take the form of a complete transcript (Chang, 2010, 2011; Chang & Millett, 2013) or video captions while listening (Hayati, 2010; Hsu et al., 2013; Soureshjani & Etemadi, 2012). Some studies reported the effects of reading-while-listening (RWL), where learners read graded readers while listening to audio recordings of the text (e.g., Chang, 2010, 2011; Chang & Millett, 2013). The findings of these studies generally indicate that RWL improved learners' comprehension skills. Other studies investigated the effects of video captions. Hayati (2010) provided video subtitles and allowed university-level learners to listen to the input at different speeds. Hsu et al. (2013) provided elementary school learners with captions on instructional videos. Soureshjani and Etemadi (2012) provided elementary learners with captions as they watched stories. All the above studies reported a positive effect of using captions in promoting learners' comprehension skills.

Phonological support often involves instruction on "reduced forms" of speech (i.e., phonetic forms that have been changed, minimized or removed from a text). The goal of this type of support is to familiarize learners with forms of input that they might otherwise have difficulty processing. This has been a popular research topic in university foreign language contexts, with studies indicating that helping learners identify reduced forms can improve comprehension skills, as demonstrated in dictation tests (Brown & Hilferty, 1986; Matsuzawa, 2006; Underwood & Wallace, 2012) and comprehension tests with (Dai & Liu, 2012) and without (Brown & Hilferty, 1986) dictation tasks. In addition to raising awareness of reduced forms, instruction aimed at helping learners recognize specific phonemes has also been found to improve comprehension (Ghorbani, 2011). Ghorbani (2011) reported that elementary-aged learners who received such instruction scored higher on a comprehension test after listening to a text than those who did not.

Contextual support, on the other hand, provides information involving visual (e.g., picture or videos) or written background information about the listening materials. Pictures are typically given before listening and are used to activate prior knowledge to prepare learners for listening activities. Studies have reported positive effects of providing visual support prior to listening, whether it be pictures (Gilakjani & Ahmadi, 2011; Othman & Vanathas, 2005) or videos

(Jensen & Vinther, 2003; Verdugo & Belmonte, 2007). Other studies have investigated the effects of watching videos while listening, and have reported positive effects in the case of children (Hsu et al., 2013; Soureshjani & Etemadi, 2012; Verdugo & Belmonte, 2012) and university learners (Hayati, 2010; Jensen & Vinther, 2003). A few studies have included activities to elicit background information about the topic of the listening text. Othman and Vanathas (2005) asked students to participate in semantic mapping activities prior to listening, and observed positive effects.

As can be seen, contextual support has, overall, been found to be effective in improving listening comprehension; however, some researchers have questioned the value of contextual support. For example, Rubin (1994) warned that if visuals provide too much contextual information, learners may not need to listen at all, as they will be able to derive meaning entirely via top-down processing. Thus, taking this into consideration, we investigated the relative effects of contextual and linguistic listening support in developing learners' listening comprehension.

Methodological Variations

The way in which listening support is implemented varies across studies. For example, the length of treatment can vary from a few hours (e.g., Matsuzawa, 2006) to a whole year (e.g., Chang, 2010). This understandably might influence the effects of listening instruction. In general it can be assumed that the longer the listening instruction, the more effective it will be in developing listening comprehension ability, since longer treatments give learners greater exposure to input and thus are more likely to facilitate L2 acquisition (Ellis, 2002). The current meta-analysis thus compared studies with longer length of instruction to studies with shorter instructional length.

Another variation is the point at which listening support is provided, e.g., before or while learners listen. Pre-listening support typically involves pre-teaching vocabulary (Chung, 2002; Jafari & Hashim, 2012; Pan, 2012) and/or raising learner awareness of language features in the upcoming text (Brown & Hilferty, 1986; Matsuzawa, 2006; Underwood & Wallace, 2012). Within-listening support draws learners' attention to the audio-visual connection of words in the listening text, for example, when they read a transcript while listening (e.g., Chang & Millett, 2013) or complete a dictation activity (e.g., Kuo, 2010). Some studies have provided both pre- and within-listening support, such as reduced-form instruction before listening followed by a dictation activity (Brown & Hilferty, 1986; Matsuzawa, 2006; Underwood & Wallace, 2012).

Vandergrift (2004) suggested that pre-listening support is beneficial because prior knowledge of the topic reduces cognitive load and frees up attention to focus on listening for details. Vandergrift (2003) argued that learners with lower proficiencies might have difficulty in performing dual or multiple tasks while listening, such as decoding the listening text while also monitoring their listening strategy use. This suggests that pre-listening support might be especially beneficial for low-proficiency learners. However, there is no single study that has compared the effects of providing support prior to and during listening. To examine this, we analysed the moderating effect of pre-listening and within-listening support on listening comprehension ability in studies that included either type of support.

Studies also differ in terms of who controls the input in the listening practice. In most studies investigating listening comprehension in the classroom, input is primarily controlled by the teacher, who decides when and how often the listening text is played, stopped, or repeated. However, there have been a few studies in which students were given the opportunity to control the input. Verdugo and Belmonte (2007) investigated the effects of allowing elementary school learners to progress at their own pace while listening to digital stories in class, stopping and restarting as they wished. Results indicated that giving learners control over the input led to improved listening comprehension. In the university context, Chang (2011) compared the effects of extensive listening (EL) with typical listening instruction. Each learner in the EL group was given audio recordings of graded readers to play by themselves, which allowed them to play, stop, and restart the audio at their discretion during class. Audio recordings in the typical-instruction group were controlled by the teacher. Results revealed higher gains on listening comprehension post-test scores for the EL group than the typical group. Though Chang attributed the comprehension gains to the opportunities for extended exposure to input in the reading-while-listening condition, the fact that learners controlled the input also may have contributed to their improved performance.

Other studies have given learners control of the input outside of class. In a study investigating the effects of pre-listening activities on listening comprehension in the German foreign language context, Teichert (1996) gave university students pictures prior to listening, and had them complete comprehension tasks based on common themes (e.g., friends and family) while listening to audio or watching videos. After these lessons, the audio and video materials from class were made available in a language laboratory so that learners could independently revisit and re-play them. Learners who were given access to these

materials outperformed those who were not. Though none of these studies explicitly investigated the effects of student-controlled input, they do suggest that giving learners control over the input contributed to students' improved comprehension.

Measuring listening comprehension also varies across studies. Most listening instruction studies employ some kind of comprehension test without providing any support for the input. Such tests typically require learners to demonstrate their comprehension skills using multiple-choice questions (e.g., Chang, 2011; Hayati, 2010) or short-answer responses (e.g., Othman & Vanathas, 2005). Some studies used dictation tests requiring learners to provide the entire listening text (e.g., Underwood & Wallace, 2012) or a part of the text using gap-fill activities (e.g., Dai & Liu, 2012). Some studies employed the listening section of an established standardised test such as The Test of English for International Communication (TOEIC®). We considered all the above-mentioned tests in our analysis, and used their effect sizes to synthesize the studies.

Small Scale Meta-Analysis

One challenge with using the meta-analytic approach in SLA research is the lack of replication studies (Li, Shintani, & Ellis, 2012). Topics that have attracted interest tend to have been investigated in relatively few empirical studies (as Jackson & Suethanapornkul, 2013; Lyster & Saito, 2010). Meta-analyses in SLA typically include a far smaller number of studies compared to meta-analyses in other domains such as medical science (Li, 2010). Li et al. (2012) reported a number of problems of SLA meta-analyses, including small sample sizes, sample bias, lack of statistical robustness, and the so-called "apples and oranges" and "garbage in garbage out" problems. However, as Oswald and McCloy (2003) suggest, meta-analysis on a small set of studies with a narrowly focused topic can still be beneficial, as such syntheses provide insight on specific research interests. Some examples of such meta-analyses in our field are Jackson and Suethanapornkul (2013), who investigated the influence of increasing task complexity on the quality of second language production; Keck, Iberri-Shea, Tracy-Ventura, and Wa-Mbaleka (2006), which investigated the effects of task-based instruction on L2 learning; and Lyster and Saito (2010), which investigated the effects of oral corrective feedback in the classroom context.

There are some methodological issues to consider when conducting small-scale meta-analyses. One is the choice of effect size aggregation—either a fixed-effect model or a random-effect model. A fixed-effect model assumes that the population mean effect size is the same in all included studies, and that the variation among effect sizes is only

attributable to sampling errors. A random-effect model allows the population effect to vary among studies, and effect size variation results from both within-study and between-study variability. While fixed-effect models have been the default meta-analyses of recent years researchers have started to adopt random-effect models because of the potential Type I error rates and misleading narrow confidence intervals associated with the former (Cooper & Hedges, 1994; Hunter & Schmidt, 2004). For these reasons, we suggest it is generally preferable to employ the random-effect model for meta-analysis in SLA research. A *Q* test, which indicates the statistical significance of the homogeneity hypothesis at a certain confidence level (Lipsey & Wilson, 2001), is a robust tool for determining whether there is a statistically significant difference between the effect sizes of different studies and for investigating the relative effects of moderating variables, because it takes variation within the sample size into consideration. It is also important to indicate how the studies included in a meta-analysis represent the wider population in order to inform any publication bias (i.e., the selectiveness of publications according to outcomes such as positive results or large sample size). One way to examine publication bias is the funnel plot, which visually displays the distribution of treatment effects estimated from individual studies against a measure of study size. This is the most common method used in SLA meta-analyses (Jackson & Suethanapornkul, 2013; Li, 2010; Miller & Pan, 2012; Norris & Ortega, 2000). Other means, such as the failsafe *N* or trim-and-fill analyses, have also been used in SLA meta-analyses (e.g., Li, 2010; Spada & Tomita, 2010). However, most small-scale meta-analyses have typically failed to report publication bias, although there are exceptions (Abraham, 2008; Jackson & Suethanapornkul, 2013; Miller & Pan, 2012; Yun, 2011), possibly due to the difficulty in interpreting the funnel plot with a small number of studies. The meta-analysis reported in this article thus employed the measures described above—namely the random effect model and *Q* tests—and addressed publication bias using a funnel plot.

Research Questions

The present meta-analysis investigated three research questions:

(1) Does listening practice have an effect on learners' listening comprehension ability?
(2) Do different types of listening practice have different effects on learners' listening comprehension ability?
(3) Which variables moderate the effects that listening instruction has on learners' listening comprehension ability?

METHOD

Meta-analyses aim for replicable and systematic research synthesis by making the selection criteria and the coding scheme transparent, and also by using statistical analyses (see Norris & Ortega, 2006). This section describes (1) the selection criteria, (2) the coding scheme, (3) the reliability of the coding scheme, and (4) the method of data analysis.

Selection of Studies

The present meta-analysis included studies published between 1980 and 2013. The year 1980 was chosen as the starting point because most earlier studies do not provide sufficient information for conducting a meta-analysis. The literature search involved identifying published experimental studies that investigated the effects of listening instruction on improving students' listening comprehension ability. The meta-analysis was carried out by inspecting electronic databases such as Google Scholar, the Education Resources Information Center (ERIC) and the Linguistic and Language Behaviour Abstracts (LLBA). The key words used in the database search were "second language listening instruction," "EFL and/or ESL listening intervention," and "reduced-forms instruction." Fugitive studies (e.g., unpublished doctoral dissertations and conference presentations) were not included due to the difficulty of retrieving the materials. A total of 73 published studies were initially retrieved. These studies were then further screened based on the following criteria:

(1) Only studies that included listening practice as defined above were included in the meta-analysis. Studies that did not report the type of listening support were excluded.

(2) Only studies that reported test scores measuring the effects of the treatment on listening comprehension were included. Listening comprehension includes both understanding and interpreting audio input. It is typically measured by asking the learners to listen to an audio-recorded passage and then demonstrating their understanding of the contents of the passage either by answering questions (e.g., short-answer or multiple-choice comprehension test) or by reproducing the passage (e.g., dictation test, oral elicited imitation test). Penno, Wilkinson, and Moore (2002) was not included because it only measured the learners' knowledge of vocabulary in the listening text (i.e., it did not provide a measure of their comprehension).

(3) When the same experiment was reported in more than one published article, it was only included once in the meta-analysis. In such cases, the study providing the fullest report of the experiment was chosen.

(4) Only studies that provided sufficient data (i.e., mean, SD and the sample size of each group) to enable calculation of effect sizes were included.

As a result, a total of 18 unique sample experiments in 18 published studies were ultimately included in the meta-analysis.

Coding

The 18 research experiments were coded for independent, dependent and moderating variables. Independent variables were the instructional treatments investigated in the experiments. In order to answer research question 2, each experimental group was first coded as either "–listening support" (learners only received listening exposure) or "+listening support" (learners received listening exposure with some kind of support). The studies classified as "+listening support" were further categorised according to the type of information given as support: "linguistic," "contextual," or "both." A study was categorized as "linguistic" when the treatment provided at least some part of the listening text or explanation/practice of the linguistic forms in the listening text either before or during the listening practice. This included pre-teaching vocabulary items, phonological segments, and transcripts/captions of the listening text. Other treatments that provided pictorial and/or video support about the topic of the texts, but did not pre-teach any of the linguistic forms in the listening text, were coded as "contextual." A treatment involving both linguistic and contextual support (e.g., captions and videos as in Hsu et al., 2013) was coded as "both."

Dependent variables were the learners' test scores reported in the studies. All except one study (Dai & Liu, 2012) administered only one post-test. Thus, we did not analyse the impact of the timing of the post-tests. All the analyses were conducted using only the first post-test scores. When a study provided combined mean scores of multiple listening comprehension tests, we used this score to calculate one effect size. None of the studies provided any kind of listening support during the tests.

Other variables which were coded include type of publication, instructional context, target language, participants' age, timing of listening support, the duration of instruction and the agent of input control. These factors were subjected to the moderator analysis, but we report only three of these factors that showed some differences in the sub-variables: (1) the timing of the listening support, (2) the duration of instruction, and (3) who controlled the input. Other variables were reported in Appendix.

Duration of instruction. According to this variable, the instruction was categorised as either "short" (0 to 999 minutes), "medium" (1,000 to 1,499 minutes) or "long" (lasting over 1,500 minutes) based on the total time of instruction reported in each study. The cut-off points were decided based on the distribution of the instruction time (see Figure 2).[1] Hsu et al. (2013) was excluded from this analysis because it failed to report the duration of the instruction.

Timing of listening support. Each experimental group was classified as belonging to one of four categories depending on the presence or absence of a pre-listening activity (PL) or a within-listening (WL) activity; namely: (1) +PL–WL, (2) –PL+WL, (3) +PL+WL, and (4) –PL–WL. As the purpose of this moderator analysis was to compare the impact of the timing of listening support, it only included categories (1), (2) and (3) for the Q_b (where b indicates "between") test.

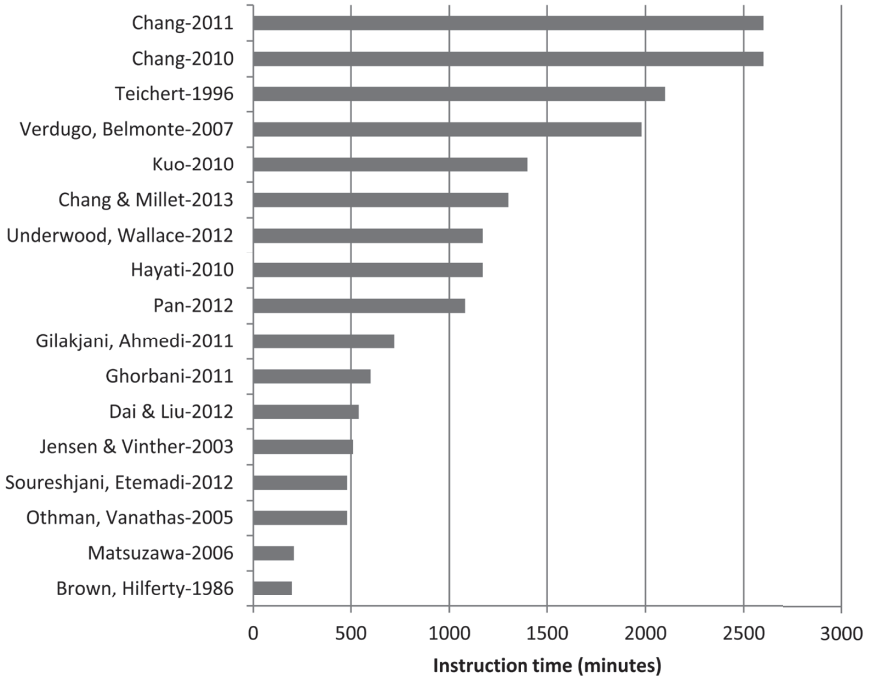

Figure 2
Instruction Time for the Studies

[1] An anonymous reviewer pointed out that the cut-off points are very different from those used in Norris and Ortega (2000). The reason for this is because, as Figure 2 shows, listening instruction studies tend to provide longer periods of time to assess the effectiveness of intervention than do studies on grammar instruction (e.g., in Norris & Ortega, 2000).

Control of the input.[2] This category is concerned with whether the teacher or students controlled the delivery of the listening text. Experiments were categorized as "teacher" if the input was controlled by the teacher (e.g., starting, stopping or replaying the audio input) and "student" if at any point during the treatment students were able to control the input.

Reliability of Coding

After one of the authors completed coding all the studies, a subset of 7 studies (over one-third of the total 18 studies) was coded by the other author. The overall agreement rate was 96%, and disparities were resolved through discussion. A final round of coding was then performed with special attention to the disputed coding.

Data Analysis

The meta-analysis followed a three-step procedure: calculation of effect sizes, effect size aggregation and moderator analysis (see Li et al., 2012 for detailed procedures involved in data analysis).

Calculation of effect sizes. In the current meta-analysis, the effect size for a treatment was calculated as the contrast between pre- and post-test scores (i.e., within-group effect size).[3] Although some meta-analyses aggregate both within- and between-group effect sizes (e.g., Norris & Ortega, 2000), we considered it best to examine only within-group effect sizes in the present study. We did not examine the experimental-control contrast because in a number of included studies (e.g., Soureshjani & Etemadi, 2012; Teichert, 1996; Verdugo & Belmonte, 2007), the control groups received "regular lessons" rather than the same listening practice as the experimental groups without support; thus, the difference between the experimental and the control group was not the provision of listening support. Excluding such studies, there were only three studies that reported between-group scores (i.e., an experimental-control group comparison), which we consider insufficient to justify meta-analysis.

[2] An anonymous reviewer pointed out that input control can be a type of listening support. However, we used this factor as a moderating variable because in the three preliminary studies, this factor was not used as a variable. It has not, to date, been considered as a kind of listening support, but we think it should be in the future.

[3] Some of the studies had different but similar tests for the pre- and post-tests, for example, the same kind of listening comprehension tests with different topics (in the case of Othman and Vanathas [2005] and Gilakjani and Ahmedi [2011]) and different versions of the TOEIC listening section, an established standardised test (in the case of Pan [2012]).

Cohen's *d* was used as the index of effect-size estimation.[4] Effect sizes were interpreted with reference to the 95% confidence interval and related *p* values. A 95% confidence interval suggests that if the same treatment is repeated many times, 95% of the time the effect would fall within that range. A narrow confidence interval indicates a robust finding. If a confidence interval crosses zero and the *p* value is above .05, the analysis fails to reject the null hypothesis (i.e., the effect of a treatment or intervention is significantly different from zero).

Effect size aggregation. Effect size aggregation generates an average effect size that indicates the overall effect of an instruction type or treatment. When a study had more than one effect size, the average of the multiple effect sizes was used for the aggregation. The random-effect model was chosen because of the diverse research designs of the studies included in the analysis.

Moderator analysis. The purpose of a moderator analysis was to identify variables that potentially mediate the effect of listening treatment, and that may account for within-subject variability among the effect sizes. *Q* tests were performed to determine whether a variable was a significant moderator (see Reinard, 2006 for the formula). A significant $Q_{b[\text{between}]}$ would mean that the moderator variable accounts for a significant portion of the variance of the effects of the treatment or intervention (see Li et al., 2012). All analyses were conducted using a professional meta-analysis software, the *Comprehensive Meta-Analysis* package (Borenstein, Hedges, Higgins, & Rothstein, 2006). Cohen's (1988) benchmarks were used to interpret the magnitude of effect size: .80 constitutes a large effect, .50 a medium effect, and .20 a small effect.

RESULTS

Descriptive Results

A total of 18 studies were published between 1980 and 2013. As shown in Figure 3, there has been a rapid growth in the number of studies that have investigated listening instruction in the past few decades.

[4] The equation used in effect size calculation was:

$$d = \frac{\text{Mean difference}}{\text{Pooled } SD}$$

where mean difference refers to the difference between the mean of pre-test and that of post-test. The pooled standard deviation (*SD*) was calculated based on the standard deviations of the pre-test and post-test means. The pooled standard deviation (S_p) is given as:

$$\text{Sp} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

where n_1 and n_2 are the sample sizes of the two tests (pre- and post-tests), and s_1 and s_2 are the standard deviations.

Altogether 1,582 participants were included in the 18 studies. All studies were conducted in a foreign language classroom setting. All except one study (Matsuzawa, 2006) were published as journal articles. Table 1 shows a summary of the methodological features of the studies.

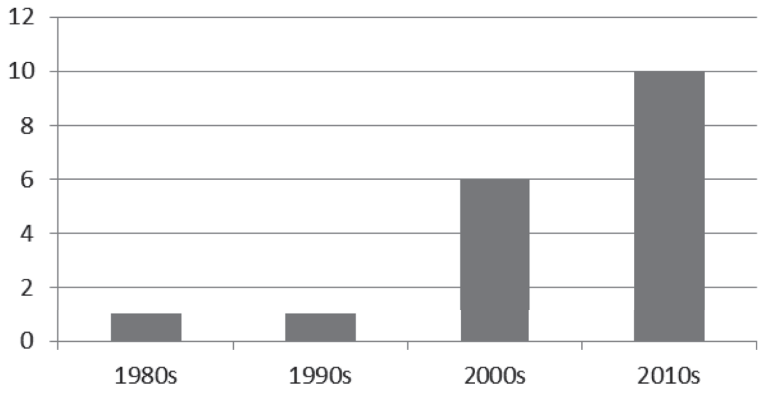

Figure 3
Number of Studies Investigating Listening Instruction

Table 1
Methodological Features of Included Studies

Aspects	Subcategories	*k*
Publication	Journal article	17
	Book chapter	1
Instructional context	Foreign language	18
	Second language	0
Research setting	Classroom	18
	Laboratory	0
Target language	English	16
	German	1
	Spanish	1
Instructional setting	University	11
	High-school	1
	Pre-school/Elementary	4
	Miscellaneous	2
Age	Adult (18+ y/o)	13
	Adolescent (13 ~ 14 y/o)	1
	Child (12- y/o)	4

Note. *k* = Number of unique sample studies.

Publication Bias

Figure 4 shows the funnel plot of effect against sample sizes. A funnel plot provides scatter plots of the treatment effects estimated from individual studies against a measure of study size (Sterne, Becker, & Egger, 2005). In the absence of publication bias, the effect sizes would be evenly distributed around the mean effect size, with small-sample studies at the bottom and large-sample studies at the apex. The asymmetry of the plot in Figure 4 is caused by the fewer effect sizes at the bottom left part of the funnel, indicating that there was a lack of studies with small sample sizes and small effect sizes, while on the other hand, studies with medium and large sample sizes were well represented in the data. We then conducted a trim-and-fill analysis, which creates imaginary data points to mirror the extreme values of the data set (see Duval & Tweedie, 2000 for the formula). An adjusted mean effect size was then calculated to show how the estimated population effect would change if the missing values were added (Duval & Tweedie, 2000). The analysis indicated that there were six effect sizes missing in the left part of the funnel and that adding the eight effect sizes would change the overall effect size from d = 1.269 (95% CI = .908, 1.622) to d = .611 (.223, .999). No missing studies were indicated by the analysis for the right side of the plot.

The effect sizes for individual experiments are reported in Appendix.

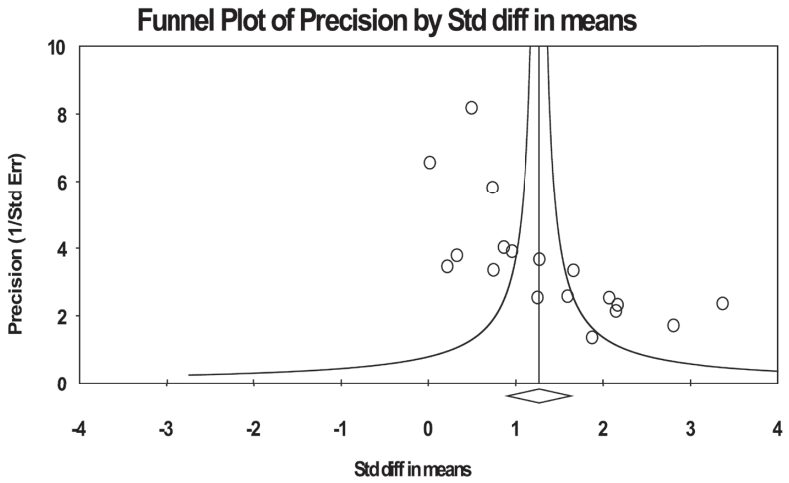

Figure 4
Funnel Plot of Precision by Effect Sizes

Note. X-axis: effect size index; Y-axis: the standard error.

Meta-Analytic Results

The analysis involved comparing the mean effect sizes of the pre-test and post-test obtained from groups receiving listening practice in each study. Table 2 shows the number of contrasts (*k*), and the mean effect size (together with the related *p* value, standard error, confidence interval). The mean effect size for the 18 experiments was quite large (*d* = 1.269), with a significant *p* value. However, as noted above, if the publication bias was resolved by including studies with a small sample size and small effect sizes, the adjusted effect size calculated by the trim-and-fill analysis would be only *d* = .611.

The independent variables were analysed by computing the mean effect size for each variable and performing *Q* tests to compare the contrasted variables (Table 3). The between-group *Q* (Q_b) indicates the significance of the effects of the treatment between the scores for two or more variables. The analyses first investigated the presence or absence of listening support (+ or – listening support). As Table 3 shows, 17 studies provided some kind of support prior to or while the oral input was given, while three studies involved groups that only listened to the text. The between-group *Q* showed that there was a significant group difference between the two conditions, indicating that listening support had a greater effect on listening comprehension ability than only providing listening practice.

Table 2
Overall Effects of Listening Practice

Variables	*k*	Mean *d*	*p*	SE	95% CI	
					Lower	Upper
Listening practice	18	1.269	.000	.184	.908	1.622

Note. *k* = Number of unique sample studies.

Table 3
Effects of the Provision of Intervention on Listening

Variables	*k*	Mean *d*	*p*	SE	95% CI		Group contrast	
					Lower	Upper	Q_b	*p*
+/–Listening support							**4.285**	**.038**
+Support	17	1.317	.000	.203	.919	1.715		
–Support	3	.820	.000	.128	.568	1.072		

Note. *k* = Number of unique sample studies.

Table 4 shows the effects of linguistic and contextual support. Eight studies provided some kind of linguistic support while six provided contextual support. Three studies provided both types of support. All three types of support showed a significant effect, which was greater when there was linguistic support. The Q_b value indicates that there were statistically significant differences. When only "linguistic" and "both" groups were compared, the Q_b value was non-significant (Q_b = .145, p = .703), indicating that linguistic support was more effective than contextual support, but that combining linguistic and contextual support did not confer a significant advantage over linguistic support by itself.

Moderating Analysis

The four moderating factors were analysed by computing the mean effect size for each variable and performing Q_b tests to compare the contrasted variables (Table 5).

Timing of listening support. Four studies provided both pre-listening (PL) and within-listening (WL) support (+PL+WL), four provided only pre-listening support (+PL–WL) and eight studies provided only within-listening support (–PL+WL). The Q_b value was marginal and was not statistically significant, indicating that the timing of listening support did not have any impact on the effect of the listening instruction.

Duration of treatment. Table 5 shows that all three lengths of treatment (i.e., Short, Medium, and Long) showed significant effects with large effect sizes. However, the Q_b test did not show a significant difference among them (Q_b = 1.940, p = .379), indicating that the length of treatment did not have any impact on the development of listening comprehension ability.

Input control. In most of the studies (n = 12), the teacher controlled the listening materials. However, studies where the learners were

Table 4
Effects of the Linguistic and Contextual Support

Variables	k	Mean d	p	SE	95% CI		Group contrast	
					Lower	Upper	Q_b	p
Linguistic/Contextual support							**11.259**	**.004**
Linguistic	8	1.890	.000	0.313	1.276	2.504		
Contextual	6	0.719	.001	0.211	.305	1.132		
Both	3	1.686	.000	0.436	.831	2.540		

Note. k = Number of unique sample studies.

Table 5
Moderator Analysis for the Effects of Listening Instruction

Variables	*k*	Mean *d*	*p*	SE	95% CI		Group contrast	
					Lower	Upper	Q_b	*p*
Timing of listening support							**.168**	**.919**
+PL+WL	4	1.600	.002	.510	.601	2.599		
+PL–WL	4	1.343	.016	.559	.247	2.438		
–PL+WL	8	1.376	.000	.280	.827	1.924		
Duration of treatment							**1.940**	**.379**
Short	8	1.049	.000	.262	.537	1.562		
Medium	5	1.647	.000	.383	.896	2.398		
Long	4	1.006	.009	.386	.250	1.762		
Input control							**9.396**	**.002**
Teacher	12	1.154	.000	.202	.759	1.549		
Student	4	2.105	.000	.236	1.643	2.568		

Note. *k* = Number of unique sample studies.

allowed to control the input showed a significantly greater effect size than studies where the teacher was in control (Q_b = 9.396, p = .002), indicating that listening practice was more effective when students controlled the listening materials.

DISCUSSION

Research question 1 asked whether listening instruction had an effect on L2 learners' listening comprehension ability. Results showed that the overall effect size for listening instruction was found to be large (d = 1.269). The adjusted value taking publication bias into account, however, showed only a medium effect (d = .611). This effect size nevertheless indicates that L2 listening instruction was in general effective in developing learners' listening comprehension ability.

The effect size of the current meta-analysis is in marked contrast to the effect of listening strategy instruction reported in Plonsky (2011), which was marginal (d = .06). One explanation for this contrast is that strategy training requires learners to utilise the strategies being trained in actual listening situations, where they usually do not have any control over the speed of the listening text. Strategy application involves a conscious action (Oxford, 1990) which might place a burden

on learners' working memory capacity. Listening practice in our meta-analysis generally included some kind of listening support (in 17 studies out of 18 that provided information on the meaning of the listening material; this might have contributed to the strong impact on learners' listening comprehension ability. Considering that all the studies in our meta-analysis were conducted in foreign language contexts, the learners investigated were likely to have had limited exposure to the L2 outside the classroom. For such learners, direct assistance on how to map linguistic form onto meaning in the listening texts might be more beneficial than strategy training. Also, the larger effect sizes might have been resulted from the use of pre-post effect sizes in this study, whereas only studies with between-subject designs were included in Plonsky (2011). This might be one reason for the lack of primary studies with small effect sizes in the current meta-analysis.

Research question 2 asked whether different types of instruction had different effects on learners' listening comprehension ability. To answer this question, we analysed two pairs of instructional variables: (1) listening practice with and without listening support, and (2) linguistic support versus contextual support.

The first contrast indicated that although both types of listening practice had large effect sizes (–listening support: d = .820; +listening support: d = 1.317), the +support condition was found to be more beneficial than exposure only. This finding indicates that listening support helps learners improve their listening skills presumably because it helps learners to overcome various problems that they face in listening, for example by reducing the burden on their working memory capacity (see Goh, 2000). Listening support may enable learners to process the input more easily and subsequently facilitate comprehension.

The second analysis—linguistic support versus contextual support—showed that linguistic support (either by itself or in combination with contextual support) had a larger effect size than contextual support by itself. This indicates that providing linguistic information about the text in the form of vocabulary teaching, grammar explanations, or a transcript of the text was more beneficial in developing listening comprehension ability than assistance in the form of pictures, videos, and background information. The key difference between these two types of support is that while linguistic support includes information about both form (e.g., how a word sounds) and meaning (what the word means), contextual information only provides information about the meaning of the text (e.g., visual images). In other words, contextual support may or may not lead to learners making a

successful connection between linguistic form and meaning. Another problem with contextual support is that, as Rubin (1994) pointed out, learners might be able to comprehend the message content of the listening text without having to engage in any bottom-up linguistic processing. In other words, contextual support may obviate the need for form-meaning connections while listening. This also explains why the effect of the combination of linguistic and contextual support (d = 1.686) was slightly smaller than that of linguistic support alone (d = 1.890). It has to be noted, however, that only three studies provided such a combination. Thus, further research is needed to clarify this point.

To answer research question 3, three moderating factors were examined: the timing of listening support, the duration of listening instruction, and who controlled the input. First, three types of timing of the listening support were analysed: (1) +pre-listening (PL) +within-listening (WL) instruction; (2) +PL–WL; and (3) –PL+WL. All three conditions showed significant and large effect sizes. The Q test did not indicate significant differences among the three conditions. Thus, the results did not support the general assumption in the literature that pre-listening support is more beneficial than within-listening support because it reduces demands on working memory while listening. Interestingly, the condition involving both PL and WL did not show a statistically significant advantage over the other conditions. However, again the number of studies for each category was small, and therefore it is premature to conclude that differences in the timing of support have no effect on listening comprehension.

The duration of treatment also did not show any significant impact on the development of listening comprehension ability. All three variables (short, medium and long) showed large effect sizes, and the Q value did not reach a level of significance. This contradicts the general assumption that longer practice leads to better acquisition. However, the limited sample size did not allow us to investigate the length of the practice provided in greater detail. The length categories we used were quite basic; there is again, then, a need for further research.

Who controlled the input, however, proved to be a significant moderating factor. The effectiveness of the treatment was greater when students had control. The most likely explanation for this result is the burden that listening places on cognitive processing. As many researchers have pointed out, one reason for the difficulties in L2 listening is that learners cannot control the speed of the input (Farrell & Mallard 2006; Goh, 2000). Not being able to control the input places a heavy burden on the learners' working memory capacity as they are faced with having to process both linguistically and for meaning

(Field, 2002; Lynch, 2002; Rost, 2002). Giving learners control over the input may not only help to reduce the processing load associated with listening, but also help them maintain their motivation by reducing anxiety. Such affective factors need to be investigated in the future.

It has to be noted, however, that in all of the above moderator analyses, each of the variables included different proportions of linguistic and contextual, as well as both types of support. For example, the eight "pre-listening" studies actually consisted of seven "linguistic" and only one "contextual" support, while the 12 "within-listening" treatments consisted of seven "linguistic," three "contextual" and two "both" support (see Appendix). Moderator analyses often involve confounding effects, and in many cases it is not possible to proceed with cross-variable analysis.

The descriptive results of this meta-analysis indicate that the selection of studies involved a potential publication bias. The funnel plot revealed that the meta-analysis lacked studies with small sample sizes and small effect sizes. Such studies typically do not feature in journal articles. Appendix shows that the methodology involved in the 18 studies varied considerably. This suggests that the results of the meta-analysis should be interpreted with caution.

CONCLUSION

The current article had two goals—to systematically synthesize studies that have investigated the effectiveness of listening practice on L2 learners' listening comprehension ability, and to demonstrate the value of small scale meta-analysis. Results of the meta-analysis indicated an overall positive effect for listening practice and providing listening support, particularly linguistic support. It was also found that giving learners the opportunity to control the listening materials led to a significant advantage in learning. However, the meta-analysis did not indicate any significant impact of either the timing of listening support or the length of listening practice on developing listening comprehension ability. We have argued that listening support might be more effective when it directly helps the learner to map linguistic form onto meaning (i.e., bottom-up processing). We also suggest that giving learners control over the listening materials is helpful because it aids in reducing their processing load and increases task motivation.

There are a number of pedagogical implications regarding how to conduct effective listening practice to be drawn from the meta-analysis. First, listening support, either before or during listening, should focus on assisting form-meaning mapping by the learners. This might be especially beneficial for learners who have limited opportunities for

exposure to the L2 outside of the classroom (i.e., FL learners). Second, teachers should consider giving learners control over the input materials, for example, by providing learners the opportunity to revisit the listening materials online either during or after class.

As discussed earlier, the methodological limitations of the meta-analysis include the small sample size and publication bias. Including unpublished studies might have increased the robustness of the meta-analysis; however, results of the present study demonstrate that small-scale meta-analyses can still be insightful. We believe that providing transparent information about the limitations of the analysis—such as publication bias—together with detailed information about the studies included in the analysis (see Appendix) will help readers to appropriately interpret the results. We reject the view that small-scale meta-analyses have no role to play in SLA. As Norris and Ortega (2006) pointed out, the limitations of such meta-analyses apply equally to traditional narrative reviews. Meta-analysts aim to provide transparent information regarding studies' limitations to enable readers to make their own judgments about how the results should be interpreted and the extent to which the results are generalizable.

Clearly, there is a need for further primary studies that investigate the effects of listening support. In particular, studies investigating contrasting types of listening instruction, such as strategy training versus listening support, linguistic support versus contextual support, and listening exposure versus exposure plus support, are needed. Additionally, studies investigating the influence of the timing of support and the length of practice are needed as well. Future studies should also examine the effect of giving students control over the listening materials, both on the immediate comprehension of a text and on long-term listening skill development. Such studies need to provide detailed information about the materials and procedures of testing instructional effects. In this way, it will be possible to identify the impact of the test type on the effectiveness of listening support.

ACKNOWLEDGMENT

The authors would like to thank the guest editors, Yo In'nami and Rie Koizumi, and the anonymous reviewers for their insightful comments.

REFERENCES

Abraham, L. B. (2008). Computer-mediated glosses in second

language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning*, *21*, 199-226.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2006). Comprehensive Meta-Analysis (Version 2.2.027) [Computer Software]. Englewood, NJ: Biostat.

Brown, J. D., & Hilferty, A. (1986). Listening for reduced forms. *TESOL Quarterly*, *20*, 759-763.

Carrier, K. A. (2003). Improving high school English language learners' second language listening through strategy instruction. *Bilingual Research Journal*, *27*, 383-408.

Chang, A. C.-S. (2010). Second-language listening anxiety before and after a 1-yr. intervention in extensive listening compared with standard foreign language instruction1. *Perceptual and Motor Skills*, *110*, 355-365.

Chang, A. C.-S. (2011). The effect of reading while listening to audiobooks: Listening fluency and vocabulary gain. *Asian Journal of English Language Teaching*, *21*, 43-64.

Chang, A. C.-S., & Millett, S. (2013). The effect of extensive listening on developing L2 listening fluency: Some hard evidence. *ELT Journal*, *68*, 31-40.

Chung, J.-M. (2002). The effects of using two advance organizers with video texts for the teaching of listening in English. *Foreign Language Annals*, *35*, 231-241.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). San Diego, CA: Academic Press.

Cooper, H., & Hedges, L. V. (1994). *Handbook of research synthesis*. New York: Sage.

Dai, C., & Liu, L. (2012). The effectiveness of explicit instruction of certain decoding skills in improving Chinese EFL listeners' general comprehension performance. *Chinese Journal of Applied Linguistics*, *2*, 243-255.

Duval, S., & Tweedie, R. (2000). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, *95*(449), 89-99.

Ellis, N. C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition*, *24*, 143-188.

Farrell, T. S., & Mallard, C. (2006). The use of reception strategies by learners of French as a foreign language. *The Modern Language Journal*, *90*, 338-352.

Field, J. (2002). The changing face of listening. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An*

anthology of current practice (pp. 242-247). Cambridge, UK: Cambridge University Press.

Ghorbani, M. R. (2011). The impact of phonetic instruction on Iranian students' listening ability enhancement. *Asian EFL Journal*, *52*(2), 24-35.

Gilakjani, A. P., & Ahmadi, S. M. (2011). The effect of text familiarity on Iranian EFL learners' listening comprehension. *Journal of Language Teaching and Research*, *2*, 783-789.

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

Goh, C. (2002). Exploring listening comprehension tactics and their interaction patterns. *System*, *30*, 185-206.

Goh, C. (2008). Metacognitive instruction for second language listening development: Theory, practice and research implications. *RELC Journal*, *39*, 188-213.

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

Hayati, A. (2010). The effect of speech rate on listening comprehension of EFL learners. *Creative Education*, *1*, 107-114.

Hsu, C.-K., Hwang, G.-J., Chang, Y.-T., & Chang, C.-K. (2013). Effects of video caption modes on English listening comprehension and vocabulary acquisition using handheld devices. *Education Technology and Society*, *16*, 403-414.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Newbury Park, CA: Sage.

Jackson, D. O., & Suethanapornkul, S. (2013). The cognition hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning*, *63*, 330-367.

Jafari, K., & Hashim, F. (2012). The effects of using advance organizers on improving EFL learners' listening comprehension: A mixed method study. *System*, *40*, 270-281.

Jensen, E. D., & Vinther, T. (2003). Exact repetition as input enhancement in second language acquisition. *Language Learning*, *53*, 373-428.

Keck, C. M., Iberri-Shea, G., Tracy-Ventura, N., & Wa-Mbaleka, S. (2006). Investigating the empirical link between task-based interaction and acquisition: A meta-analysis. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 91-131). Amsterdam, The Netherlands: John Benjamins.

Kuo, Y. (2010). Using partial dictation of an English teaching radio program to enhance EFL learners' listening comprehension. *Asian EFL Journal*, *47*, 4-29.

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, *60*, 309-365.

Li, S., Shintani, N., & Ellis, R. (2012). Doing meta-analysis in SLA: Practice, choice, and standards. *Contemporary Foreign Languages Studies*, *384*(12), 1-17.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Lynch, T. (2002). Listening: Questions of level. In R. B. Kaplan (Ed.), *The Oxford handbook of applied linguistics* (pp. 39-48). New York: Oxford Univeristy Press.

Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition*, *32*, 265-302.

Matsuzawa, T. (2006). Comprehension of English reduced forms by Japanese business people and the effectiveness of instruction. In J. D. Brown & K. Kondo-Brown (Eds.), *Perspectives on teaching connected speech to second language speakers* (pp. 59-66). Honolulu, HI: University of Hawaii, Foreign Language Resource Center.

Miller, P. C., & Pan, W. (2012). Recasts in the L2 classroom: A meta-analytic review. *International Journal of Educational Research*, *56*, 48-59.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, *50*, 417-528.

Norris, J. M., & Ortega, L. (2006). The value and practice synthesis for language learning. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 3-52). Amsterdam, The Netherlands: John Benjamins.

Oswald, F. L., & McCloy, R. A. (2003). Meta-analyses and the art of the average. In K. R. Murphy (Ed.), *Validity generalization: A critical review* (pp. 311-336). Mahwah, NJ: Erlbaum.

Othman, J., & Vanathas, C. (2005). Topic familiarity and its influence on listening comprehension. *The English Teacher*, *8*, 19-32.

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

Pan, Y.-C. (2012). Effects of multi-faceted lexical instruction on the TOEIC listening performance of Taiwanese EFL college students. *International Journal of English Linguistics*, *2*(6), 71-79.

Penno, J. F., Wilkinson, I. A. G., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology*, *94*, 23-33.

Plonsky, L. (2011). The effectiveness of second language strategy instruction: A meta-analysis. *Language Learning*, *61*, 993-1038.

Reinard, J. C. (2006). *Communication research statistics*. Thousand Oaks, CA: Sage.

Rost, M. (2002). *Teaching and researching listening*. London: Longman.

Rost, M. (2006). Areas of research that influence L2 listening instruction. In E. Usó-Juan & A. Martínez-Flor (Eds.), *Current trends in the development and teaching of the four language skills* (pp. 47-74). Berlin, Germany: Mouton de Gruyter.

Rost, M., & Ross, S. (1991). Learner use of strategies in interaction: Typology and teachability. *Language Learning*, *41*, 235-268.

Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal*, *78*, 199-221.

Soureshjani, K. H., & Etemadi, N. (2012). Listening comprehension success among EFL preschool children using internet-based materials. *e-BANGI: Jurnal Sains Sosialdan Kemanusiaan*, *7*(1), 243-251.

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, *60*, 263-308.

Sterne, J. A. C., Becker, B. J., & Egger, M. (2005). The funnel plot. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 75-98). Chichester, UK: Wiley.

Teichert, H. U. (1996). A comparative study using illustrations, brainstorming, and questions as advance organizers in intermediate college German conversation classes. *The Modern Language Journal*, *80*, 509-517.

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

Underwood, P., & Wallace, M. (2012). The effects of instruction in reduced forms on the performance of low-proficiency EFL university students. *Asian EFL Journal1*, *4*(4), 1-24.

Vandergrift, L. (2002). "It was nice to see that our predictions were right": Developing metacognition in L2 listening comprehension. *Canadian Modern Language Review/La revue canadienne des langues vivantes*, *58*, 555-575.

Vandergrift, L. (2003). From prediction through reflection: Guiding

students: Through the Process of L2 Listening. *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, *59*, 425-440.

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

Vandergrift, L., & Goh, C. (2012). *Teaching and learning second language listening: Metacognition in action*. New York: Routledge.

Vandergrift, L., Goh, C., Mareschal, C. J., & Tafaghodtari, M. H. (2006). The metacognitive awareness listening questionnaire: Development and validation. *Language learning*, *56*, 431-462.

Verdugo, D. R., & Belmonte, I. A. (2007). Using digital stories to improve listening comprehension with Spanish young learners of English. *Language Learning & Technology*, *11*(1), 87-101.

Yun, J. (2011). The effects of hypertext glosses on L2 vocabulary acquisition: A meta-analysis. *Computer Assisted Language Learning*, *24*, 39-58.

ABOUT THE AUTHORS

Natsuko Shintani is a Senior Lecturer at University of Auckland, New Zealand. Her research interests encompass roles of interaction in second language acquisition. She has worked on several meta-analysis studies including a published work in *Language Learning* and *Applied Linguistics*.

Matthew P. Wallace is a doctoral student at the National Institute of Education, Nanyang Technological University. His research interests include second-language listening and the role of working memory in second-language comprehension.

APPENDIX

Included Studies

Study	Year	N	Pre-listening	Within-listening	Linguistic/ Contextual	Treatment length	Input control	Context	Language	Setting	Age	Test timing	Publication	Effect size (*d*)
Brown & Hilferty	1986	32	+	+	L	200 Short	T	FL	English	University	Adult	I	J	1.261
			–	–	NA	NR -								0.782
Chang	2010	92	–	+	L	2,600 Long	S	FL	English	Miscellaneous	Adult	I	J	2.154
Chang	2011	19	–	+	L	2,600 Long	S	FL	English	High-school	Adoles	I	J	1.883
Chang & Millett	2013	113	–	+	L	1,300 Mid	T	FL	English	University	Adult	I	J	1.670
			–	–	NA	NR -								0.983
Dai & Liu	2012	117	–	–	NA	540 Short	T	FL	English	University	Adult	I & D	J	0.741
Ghorbani	2011	45	+	–	L	600 Short	T	FL	English	Elementary	Child	I	J	2.811
Gilakjani & Ahmedi	2011	60	+	–	L	720 Short	T	FL	English	University	Adult	I	J	0.024
Hayati	2010	62	–	+	B	1,170 Mid	T	FL	English	University	Adult	I	J	0.962
Hsu et al.	2013	81	–	+	B/C	NR -	S	FL	English	Elementary	Child	I	J	2.077
Jensen & Vinther	2003	65	–	+	C	510 Short	T	FL	Spanish	University	Adult	I	J	0.334
Kuo	2010	31	–	+	L	1,400 Mid	T	FL	English	University	Adult	I	J	0.869
Matsuzawa	2006	16	+	+	L	210 Short	T	FL	English	Miscellaneous	Adult	I	C	0.751
Othman & Vanathas	2005	34	+	–	L	480 Short	T	FL	English	University	Adult	I	J	1.281

Study	Year	N	Pre-listening	Within-listening	Linguistic/ Contextual	Treatment length		Input control	Context	Language	Setting	Age	Test timing	Publication	Effect size (*d*)
Pan	2012	47	+	–	L	1,080	Mid	T	FL	English	University	Adult	I	J	1.603
Soureshjani & Etemadi	2012	50	–	+	B	480	Short	T	FL	English	Pre-school	Child	I	J	2.173
Teichert	1996	50	+	+	C	2,100	Long	T (S later)	FL	German	University	Adult	I	J	0.224
Underwood & Wallace	2012	52	+	+	L	1,170	Mid	T	FL	English	University	Adult	I	J	3.378
Verdugo & Belmonte	2012	220	–	+	C	1,980	Long	T	FL	English	Elementary	Child	I	J	0.501

Note. N = Number of participants; Linguistic: L = linguistic, C = Contextual, B = both; Treatment length: Long = 1,500 min or more, Mid = 1,000 ~ 1,499 min, Short = 999 or less, NR = not reported; Input Control: T = Teacher, S = Student; Test timing: I = immediate, D = delayed; Publication: J = Journal article, C = book chapter.

聽力輔助對第二語言學習者聽力理解的效用：後設分析研究

摘要

本文旨在報導一個後設分析研究，其目的在探討聽力輔助對第二語言學習者聽力理解的效用。我們共取得了 18 篇相關研究，且以前後測聽力理解變化的效果量進行編碼。首先，我們藉由比較聽力輔助的有無，剖析聽力輔助的效果；接著，我們探究語言輔助（如，單字教學）及語境輔助（如，圖像訊息）兩種主要輔助方式的效果。後續的調節變項分析則檢定三個中介變項的影響力：聽力輔助使用時機、輔助教學時間的長短、聽力素材輸入主控者（教師或學生）。根據研究結果，我們得到以下結論：（一）聽力輔助能增進聽力練習的成效。（二）語言輔助的效果量大於語境輔助。而調節變項分析的結果顯示，聽力輔助使用時機與輔助教學時間，無顯著調節效果。然而，聽力素材輸入主控者之差異則具顯著調節效果：學生為聽力素材主控者的練習，其效果量顯著高於以教師為主控者的練習。文末，我們討論本研究發現的理論與教學意義以及小規模後設分析的角色。

關鍵詞：第二語言聽力教學 後設分析 第二語言聽力輔助