CHAPTER SEVENTEEN

Researching Listening

Larry Vandergrift

It is probably self-evident that, of the four foundational language skills, listening is the least understood and certainly the most difficult to investigate. The covert nature of the process and ephemeral nature of the input make the perceptual and interpretation processes involved in listening comprehension difficult to access. Research on listening in applied linguistics remains limited; however, recent studies have led to some new insights into the underlying cognitive processes, the teaching and the assessment of listening (see Vandergrift & Goh 2012). That being said, many questions still remain and, compared to the other skills, few theoretical models have emerged.

This chapter begins with a very brief synopsis of current research on listening in applied linguistics, followed by a critical overview of the range of methodologies that can be used to investigate the listening skill. It concludes with a sample study that illustrates the use of some of these research methodologies.

Brief overview of listening

Cognitive processes

A conceptual understanding of listening must acknowledge the interplay of physiological and cognitive processes at different levels, as well as the influence of contextual factors. Recent reviews of research on listening (e.g. Field 2008; Lynch 2009; Rost 2011; Vandergrift 2007; Vandergrift & Goh 2012) highlight two fundamental cognitive processes required for comprehension: bottom-up and top-down processes. Listeners use bottom-

up processes when they construct meaning from the incoming sound stream by gradually combining increasingly larger units of meaning from the phoneme-level up to the discourse-level to build comprehension of an utterance or a text. On the other hand, listeners use top-down processes when they use context, prior knowledge (topic, genre, culture and other schema knowledge in long-term memory) and listener expectations to build a conceptual framework around the specific listening task. Individual units of meaning retained from bottom-up processing are slotted into the framework to eventually arrive at a reasonable interpretation of the message. Although these processes occur simultaneously and in parallel fashion, the degree to which listeners may use the one process more than the other will depend on the task or the purpose for listening.

Research on these cognitive processes suggests that listeners need to learn how to use both processes to their advantage, depending on the purpose for listening. Although the literature on listening instruction over the past years has tended to focus on the development of top-down processes, there is currently a renewed interest in the contribution of perceptual processing (bottom-up processes) to successful listening comprehension (Field 2008).

The speed and effectiveness of these cognitive processes depend on the degree to which listeners can efficiently process what is heard. Native language listeners do this automatically, with little conscious attention to individual words. Listeners learning a new language, on the other hand, have limited linguistic knowledge; therefore, less of what they hear can be automatically processed. Depending on their level of proficiency or the difficulty of the text or task, these listeners may need to consciously process some of the input and, given the limitations of working memory, comprehension either turns out to be incomplete or breaks down. To compensate for gaps in comprehension, skilled listeners can use their metacognitive knowledge about listening to orchestrate appropriate metacognitive and cognitive strategies, contextual cues and other relevant information available to them to inference what was not understood. Listeners also use metacognitive knowledge for successful comprehension when they (1) analyse task requirements; (2) activate appropriate listening processes for the task; (3) make appropriate predictions; (4) monitor their comprehension; (5) problem solve to figure out the meaning of what they do not understand; and (6) evaluate the success of their approach (Vandergrift 1999).

Variables that may influence comprehension

Listening involves more than cognitive processing, however. It can be constrained by affective factors such as motivation, self-efficacy and anxiety, which further limit how much information short-term memory can process at a given moment. Other learner variables that may impact comprehension include background knowledge of the topic of the text, proficiency level in

the target language, age, metacognitive knowledge about listening, strategy use, native language listening ability, working memory capacity and auditory discrimination ability (see Vandergrift & Baker 2015). The listening task itself, that is, the type of information and level of detail required and what the listener is expected to do with that information, may also influence the outcome of listening (see Brunfaut & Revesz 2013). Text characteristics such as speed of delivery, frequency of pausing and hesitations, accent, and amount of informal language in the aural stream are further mitigating factors (see Bloomfield et al. 2011). In the case of texts that carry visual support, such as video, the speaker's actions and reactions as well as the degree of congruency between the spoken and the visual may also influence the level of comprehension (see Cross 2011). In the case of interaction with an interlocutor, the listener's sensitivity to paralinguistic cues (body language and non-verbal voice cues) as well as his/her power relationship to the interlocutor play a significant role in comprehension. Moreover, a correct literal understanding of a message does not necessarily ensure accurate comprehension. Successful comprehension also requires listeners to apply pragmatic knowledge (see Roever this volume) to interpret the speaker's implied meaning, which may go beyond the literal meaning of the utterance. An awareness of the processes and variables related to listening success, and their interaction, is fundamental to an understanding of the listening construct. Since space limitations preclude a more detailed discussion of the complexity of the listening construct and related variables, see Field (2008), Lynch (2009), Rost (2011) and Vandergrift and Goh (2012) for more detailed discussions.

Research methodologies for investigating listening

Rost (2011) suggests that two overlapping processes – learning to listen in a new language and listening to learn this new language – are involved in listening development. The distinction between product (listening to learn) and process (learning to listen) is a useful heuristic for the following discussion of the research methodologies for investigating listening. However, before exploring and discussing the different methodologies for investigating the product and the process of listening, I will briefly discuss the methods for measuring listening ability.

Measuring listening ability

Proficiency tests

Essential to product-oriented research, listening test scores provide baseline data from which to measure growth in listening ability over time and/or consequent to a pedagogical intervention. Test scores can also be used

to assign a level of listening proficiency to participants in studies where proficiency in the target language is a variable under investigation. Test scores may provide a more objective criterion for assigning proficiency level than a grade, course level or teacher assessment. Their generalizability, however, is limited because they are either based on an in-house placement test instead of a more broad-based, standardized objective measure (Berne 2004) or they may be assessing only particular aspects of listening ability (Buck 2001). Listening research can be strengthened and more broadly generalized through the use of valid and reliable tests referenced to internationally understood benchmarks such as the ACTFL Proficiency Guidelines (ACTFL 1999) or the Common European Framework of Reference for Languages (Council of Europe 2001), for example.

The advantage of listening tests as a comprehension measure for research purposes is that they are not labour-intensive and can be easily administered to large groups. The major challenge for listening test development and listening research, however, is the measurement of pure listening comprehension ability – that is, assessing comprehension without introducing reading or writing as confounding variables. Pure listening comprehension is most appropriately measured through aural prompts and verification options that are limited to non-verbal evaluation techniques such as selecting among a choice of pictures or objects, sequencing pictures or other graphics, drawing a picture, tracing on a map or performing a physical response. On the other hand, when using aural prompts to assess listening comprehension, memory becomes a confounding variable.

Free written recall protocols

Free recall protocols (or comprehension restitution tasks, see Prince 2012) represent an alternative but more labour-intensive method for measuring aural comprehension of text content. Participants listen to a text and immediately afterwards write, in their native language or L2, as much information as possible about what they understood. For example, Guichon and McLornan (2008) used recall protocols to measure differences in comprehension when an aural text was presented in four different conditions (audio alone, video with audio, video with audio and target language subtitles and video with audio and native language subtitles) to different groups. After the first presentation of the text, students in each group wrote notes in the language of their choice, and after the second presentation, they prepared a detailed summary (recall protocols). The protocols were then analysed for the number of correct idea units, resulting in a score for each student and each group, to determine the level of listening success for each condition. Calculating a comprehension score is more labour-intensive than administering a listening test, especially if a second rater independently verifies a certain percentage of the protocols to ensure reliability of the scores. However, recall protocols have the capacity to assess comprehension without using

question prompts. Memory will be a confounding variable here as well, but perhaps less so if note taking and approximate spellings are allowed. An additional advantage to recall protocols as a measure of comprehension is their capacity for providing insights into the comprehension process. In fact, Guichon and McLornan (2008) and Prince (2012) view free recall protocols (or comprehension restitution tasks) as tools to measure comprehension and to gain greater insight into comprehension processes.

Although listening test scores and recall protocols provide an objective measure for determining comprehension gains for research purposes, the reliability of the scores will be limited by the reliability of the test or, in the case of the written recall protocols, inter-rater reliability. In order for any measure of comprehension to be reliable for research purposes, scores may need to be controlled to account for any initial significant differences between groups, such as differences in listening ability or background knowledge of the text topic. Finally, although listening test scores may provide a product to identify a level of performance, they tell us nothing about the process, that is, how listeners arrive at the right answer or why comprehension breaks down.

Investigating the product of listening

Experimental research

Research on listening has traditionally focused on the outcome, or the product of listening. In this line of research, student performance on a listening test is used to measure the success of a particular intervention. Two groups complete a pre-test; the experimental group experiences a different condition (such as exposure to an alternative pre-listening activity) from the control group. After the intervention, which may be long or short term, both groups complete an immediate post-test to determine the effects of the experiment: did the intervention result in a statistically significant (see Phakiti this volume) higher score on the listening post-test by the experimental group over the control group? Sometimes, a delayed post-test is used to assess the long-term effects of the intervention, such as the retention of certain vocabulary items three months after the intervention. Generally, there is no interest in determining how the listeners used the experimental intervention to improve their comprehension.

Correlational research

Another line of product-oriented listening research involves determining the degree of relationship between different variables hypothesized to be related to listening success. In this case, listening test scores are correlated with the scores of an instrument that measures an individual variable in

order to determine if there is a relationship between the two variables and the strength of that relationship. For example, student scores on a listening anxiety questionnaire or a sound discrimination test are correlated with listening achievement scores to determine any potentially significant relationship. It is important to note, however, that although positive or negative correlations may point to interesting relationships between a given variable and listening success, it is not possible to claim causality between that variable and listening success. Uncovering the nature of the relationship between the variables requires careful interpretation and may be elucidated by other quantitative methodologies (see below) such as regression analysis, serial equation modelling or path analysis or more qualitative research methodologies such as interviews or stimulated recalls (of questionnaire data) that explore the listening process.

If the sample size is large enough, one can also simultaneously examine the relationship of a number of learner variables with listening success through multiple regression analysis (see Phakiti this volume for further discussion of this statistical test) and determine the effect of these variables on listening success (Vandergrift 2006). By regressing the scores of a number of independent variables (e.g. sound discrimination ability, first language listening ability, metacognitive awareness of listening) on the dependent variable (listening test scores), it is possible to determine the approximate percentage of variance that each independent variable might contribute to listening success (the dependent variable). A path analysis goes beyond a simple correlational design to explore causality and observe how the variables might interact in leading to successful L2 listening comprehension (Vandergrift & Baker 2015). This line of research, although interesting for what it might reveal about the predictive validity of hypothesized variables on listening ability, must be based on scores from reliable instruments and statistically significant relationships that are carefully interpreted.

Investigating the process of listening

Although a product-oriented approach to investigating listening can yield useful information about level of listening proficiency or potentially interesting relationships between listening ability and other learner variables, this line of research is not interested in exploring the problems listeners may experience in comprehension or the reasons motivating a student response on questionnaires. A process-oriented approach, on the other hand, can provide potentially useful insights into the cognitive processes underlying listening comprehension. Process-oriented methodologies, such as questionnaires, interviews, stimulated recalls, think-alouds, observation, diaries, tracking software and aural perception processing responses provide opportunities for participants to reveal, or researchers to uncover, listener decision-making processes during comprehension.

Questionnaires

Questionnaires (open- or closed-response) administered immediately after a listening activity can provide insights into student awareness of the process of listening and, when used as a repeated measure, track any changes in awareness of the listening process or listening attitudes over time (e.g. Vandergrift et al. 2006). Closed-response questionnaires, or surveys, are particularly useful for collecting large amounts of data that can be compiled and analysed quickly. The reliability of this data, however, is dependent upon the reliability and validity of the instrument. Greater insights into the reason for student response to questionnaire items or to changes in item responses over time can be achieved with a stimulated recall, where the listener is asked to reflect on any changes in response patterns on the same questionnaire (see the sample study below).

Interviews

Interviews, a more personable oral version of the open-ended response questionnaire, represent a more flexible but time-consuming method for eliciting data about listening. Individual or group interviews afford an opportunity to gain greater insights from a representative group of participants on important themes emerging from a large-scale survey, such as the difficulties students experience with listening in learning another language (e.g. Graham 2006). In particular, a semi-structured interview allows the researcher to diverge from the interview protocol to explore listener responses in greater detail. Interview sessions must be recorded, transcribed, coded and analysed. The reliability and validity of interview responses are enhanced when the participant is given the opportunity to read the transcripts to validate or correct them and, in the case of coding, when inter-rater reliability is at an acceptable level (0.90 or higher).

Stimulated recall protocols

Stimulated recall is a version of the individual interview in which the researcher and the participant focus on another data set involving the participant, such as a video recording or questionnaire responses. In order to gain greater insight into listener behaviour or response, the researcher asks the participant to comment on a particular event in the video or an interesting pattern in questionnaire response. The reliability of the data, recorded and transcribed, will be directly proportional to the proximity of the event under discussion and the interview. Although time consuming, stimulated recalls can also provide insights into listener decision-making processes, such as navigating through support options while engaged in multimedia listening (Grgurović & Hegelheimer 2007).

Diaries

Diaries are, first of all, tools for language learners to reflect on their language learning; however, they are also a useful research tool for gaining insights into learner awareness of listening processes, strategy development, listener anxiety, listening goals and actions taken to improve listening performance. Diaries can be entirely open-ended or guided by the researcher through given prompts, and they can be particularly fruitful when used over time. For example, over a ten-week period, Goh (1997) was able to document the development of metacognitive knowledge about L2 listening. Classroom discussions based on reflection probes (e.g. Goh & Taib 2006) can also serve as the stimulus for subsequent diary writing. Although data gleaned from diaries, and any other kind of listening notes (e.g. Mareschal 2007), are constrained by ethical considerations (student permission) and lack of generalizability, they do provide helpful insights into metacognitive awareness about listening, strategy use and overall listening success.

Observation notes

Since listening is a covert process, observation is generally of limited value for investigating listening processes in uni-directional listening. However, observation of interviews, or other interactive situations, can provide some insights into listener behaviour in bi-directional listening. When recorded, these interviews can be reviewed by the researcher for evidence of the variable under investigation (e.g. number and type of clarification strategies used). Review of the video recording with the research participant, using stimulated recall immediately after the interview, provides an opportunity to discover how the participant either clarified meaning or helped the interlocutor to advance the conversation (Farrell & Mallard 2006).

Think-aloud protocols

The research methodologies discussed so far have been retrospective in nature, that is, listeners comment and reflect on listening events that have taken place in the recent or distant past. Introspective methodologies, on the other hand, attempt to tap the thought processes of listeners while they are actually engaged in the listening event. After some initial training in the process, participants 'think aloud' at predetermined intervals while listening to a recorded aural text. The resulting data, the listener's voiced thoughts on what he/she is doing to comprehend, are recorded simultaneously and later transcribed. Introspection is likely the closest researchers can come to tapping thought processes while information is still available to the listener in short-term memory. Think-aloud protocols can be useful for shedding light on where and how listeners experience difficulties as they are listening (Goh 2002), the development of strategy use over time (e.g.

Graham et al. 2008) or the differential use of visual and aural information to understand online academic lectures (e.g. Smidt & Hegelheimer 2004). The validity and reliability of think-aloud protocols depend heavily on how the listener is prompted by the researcher.

Tracking software

New technologies can provide further insights into the process of listening by tracking and analysing listener use of help functions while listening to a recorded academic lecture (e.g. Grgurović & Hegelheimer 2007) or listener use of pauses, rewinds and fast forwards while working on an MP3 player (e.g. Roussel 2011). Even greater insights into the motivation for listener behaviour can be gained, however, by complementing these methodologies with a stimulated recall as soon as possible after the listening event.

Auditory perception processing

Aural perceptual processing investigates phenomena associated with the bottom-up dimension of listening, such as word segmentation skills. In this line of research, listeners hear a stimulus, a word or a short sentence (presented with no contextual cues), and are then asked to choose from among several options, the word or phrase they thought they heard or write down what they thought they heard (recall protocol). The choice options are often built on potential error patterns (see Field 2004). Of interest are the errors made and what they might reveal about the listener's word segmentation strategies (e.g. influence of a native language segmentation strategy) and/or listener expectations. This line of research tends to be very micro in nature, excluding the macro-context of a text or communication that listeners generally use in real-life listening to support their interpretation of words in context.

Triangulation of data

Although all of the research methodologies outlined above provide greater insights into the process of listening, their generalizability is often limited. In order to overcome this limitation, researchers investigating a construct as covert as listening should try to use multi-method assessment to collect convergent data. Reliability and validity can be enhanced when data from more than one source are triangulated to provide a more complete picture of the listening construct; for example, videotaped data with a stimulated recall on the video recording can be complemented by a questionnaire (e.g. Cutrone 2005). In her doctoral thesis, Mareschal (2007) collected and analysed data from a listening questionnaire, stimulated recalls on the questionnaire responses, listening notebooks, think-alouds, observation and a summative, open-response questionnaire to document how a self-

regulatory approach to listening instruction influenced the listeners' self-regulatory ability, strategy use, metacognitive knowledge and listening success. Although these methodologies may provide greater insight into the process of listening, they are more labour-intensive with regard to collection and analysis of the data and less conducive to parsimonious reporting in order to demonstrate the trustworthiness and credibility of the research.

A sample study

Background

The study I have chosen to illustrate some of the methodological considerations discussed in this chapter is a recent long-term, mixed methods study (Vandergrift & Tafaghodtari 2010) on metacognitive instruction in listening, using a pedagogical sequence grounded in metacognitive theory. My interest in conducting this study emerged from a number of observations in the research literature. First, the results of long-term listening strategy instruction appear inconclusive and instruction in individual strategies does not appear to lead to overall listening improvement (Field 2001). Second, a number of recent studies suggest that skilled listeners appear to orchestrate cognitive and metacognitive strategies in an interconnected fashion (Goh 2002; Graham & Macaro 2008; Vandergrift 2003). Third, the reading research literature is in general agreement that instruction in a repertoire of strategies is more effective than individual strategy instruction for teaching comprehension skills (Grabe 2009). Fourth, recent studies on listener metacognitive awareness suggest that systematically guiding language learners through the process of listening as part of regular listening practice (metacognitive instruction) can lead learners to practise the metacognitive processes involved in listening (see Vandergrift 2004). Finally, although the short-term effects of this process approach on student motivation and the development of metacognitive knowledge about listening have been demonstrated, the enduring effects of this approach on listening achievement still needed to be empirically demonstrated (Berne 2004; Vandergrift 2004).

Methodological considerations

The nature of this investigation led me, first of all, to design a carefully controlled quasi-experimental study to test the following hypotheses: (1) the experimental group receiving metacognitive instruction will outperform the control group; and (2) the less-skilled listeners in the experimental group would make greater gains in listening achievement than their more-skilled counterparts. This constituted the product focus of the study. Secondly, in

order to track the hypothesized growth in metacognitive knowledge about listening throughout the study, and to gain greater insight into participant awareness of the process of listening, both groups completed a valid and reliable listening questionnaire (see Vandergrift et al. 2006, for information on the development and validation of this instrument). Finally, in an effort to gain an even deeper understanding of the development of listening processes, six participants from the experimental group were randomly selected to participate in a stimulated recall to explore any differences in their questionnaire responses at the mid- and end-points of the study. These last two methodologies constituted the process focus of the study.

Participants were university students from six intact French as a second language classes, randomly assigned by class to either a control or experimental group. Participants were identified as less- or more-skilled listeners on the basis of their score on a listening test (Cronbach's alpha = 0.90) also used as a post-test to measure growth in listening ability. Two instructors participated – one taught four different classes of low-intermediate learners and the other taught two different classes of high-beginners. To control for potentially confounding teaching variables, the same teacher taught both the control and the experimental group, and both groups listened to the same texts. All teaching sessions were observed by a research assistant to verify that the instructor respected the designated listening pedagogy for that group.

All participants completed the listening questionnaire at the beginning, mid- and end-points of the study, immediately after a listening activity. We hypothesized that the less-skilled listeners in the experimental group would demonstrate greater growth in the five factors related to metacognitive knowledge (i.e. Problem-Solving, Planning and Evaluation, Directed Attention, Mental Translation and Person Knowledge) than their more-skilled counterparts, and both the more- and less-skilled participants in the control group. Participants selected for the stimulated recall sessions met with a research assistant twice: after the mid-point and at the end of the study. At the first session, the research assistant presented the participant with his/her beginning and mid-point questionnaire responses and then discussed major discrepancies (two point differences) in responses with the participant. During the second session, the participant was asked to discuss possible reasons for further discrepancies in responses based on the final completed questionnaire. All stimulated recall sessions were audio-recorded, transcribed verbatim, coded and analysed for emerging themes using the QSR-N*Vivo7.[1]

The study took place over a 13-week term. Each week, the classes listened to a different authentic-type text related to the topic of the teaching unit. The pedagogical sequence experienced by the experimental group and the metacognitive processes underlying each step in the sequence are described in detail in Vandergrift and Tafaghodtari (2010). Participants in the control group listened to the same texts the same number of times: however, they

did not engage in prediction; they were not given an opportunity to discuss, predict or monitor their comprehension with a classmate; and, they did not engage in any whole-class reflection on strategy use.

Results

In order to confirm the first two hypotheses, a 2-factor ANCOVA (analysis of co-variance) was performed. Independent variables consisted of group (treatment and control) and the level of listening ability (less- and more-skilled) factorially combined. To control for any initial differences in listening ability, pre-listening test scores were used as covariate in the analysis.

Results revealed that, overall, the experimental (metacognitive instruction) group outperformed the control group. Furthermore, the less-skilled listeners who received metacognitive instruction outperformed the less-skilled listeners in the control group. In addition, the less-skilled listeners in the experimental group showed greater improvement than their more-skilled counterparts (although the effect size was small). Evidence for the third hypothesis was not as clear in terms of demonstrated growth in metacognitive knowledge about listening. Although listeners in the experimental group evidenced greater growth in all areas of metacognitive knowledge about listening, only growth in Problem-Solving and Mental Translation was statistically significant. The surprising finding about translation was later elucidated through the stimulated recall where it became evident that students were misinterpreting the concept of translation. The recall protocols also provided additional evidence for growth in the five areas of metacognitive knowledge. This study is a good example of research that is both product- and process-oriented, using different research methodologies to explore the development of metacognitive knowledge for listening.

Conclusion

This chapter has presented an overview of methodologies for investigating listening in applied linguistics. The field is ripe for research. Emerging technologies and increased access to multi-modal online resources through high-speed internet and self-access centres, in particular, open up rich possibilities for teaching and researching listening (see Godwin-Jones 2012; Robin 2007). When researchers follow reliable guidelines for designing and conducting careful research, then small and large-scale studies, using various methodologies, can contribute to a growing body of evidence-based knowledge, developed by an engaged community of listening researchers and research-practitioners (see Cross & Vandergrift 2014). For more detailed information on areas for future research, see Rost (2006) and Vandergrift (2007).

Acknowledgements

The author wishes to thank the editors and Jeremy Cross for their comments and suggestions for revision.

Note

1 N*Vivo is a qualitative research tool for classifying, sorting and arranging data in order to analyse them for patterns and to identify themes.

Resources for further reading

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

This book is strong on bottom-up approaches to listening and provides plenty of ideas for the classroom. The research reported in this book generally uses psycholinguistic laboratory methods which can be helpful for the insights they offer into listener decoding processes; however, they lack ecological validity in that laboratory methods often rob listeners of the contextual supports that usually go along with real-life listening.

Lynch, T 2009, *Teaching Second Language Listening*, Oxford University Press, Oxford.

Lynch presents a good, readable overview of issues in second language listening. The book covers many important topics and is grounded in much of the author's own listening experiences.

Rost, M 2011, *Teaching and Researching Listening*, 2nd edn, Longman, London.

This comprehensive volume provides an overview of teaching and researching listening. The third section introduces a number of research methodologies and frameworks, as well as concrete topics for action research.

Vandergrift, L & Goh, CM 2012, *Teaching and Learning Second Language Listening: Metacognition in Action*, Routledge, New York, NY.

This book provides a thorough examination of many dimensions of listening comprehension, from theory to application to implementation, including descriptions of concrete class activities. Woven throughout is the informing thread of metacognition.

References

American Council for the Teaching of Foreign Languages (ACTFL) 1999, 'ACTFL proficiency guidelines', viewed 22 November 2013, http://www.actfl.org/sites/default/files/pdfs/public/ACTFLProficiencyGuidelines2012_FINAL.pdf

Berne, JE 2004, 'Listening comprehension strategies: A review of the literature', *Foreign Language Annals*, vol. 37, no. 4, pp. 521–531.
Bloomfield, A, Wayland, SC, Rhoades, E, Blodgett, A, Linck, J & Ross, S 2011, *What Makes Listening Difficult? Factors Affecting Second Language Listening Comprehension*, University of Maryland, College Park, MA.
Brunfaut, T & Revesz, A 2013, 'Text characteristics of task input and difficulty in second language listening comprehension', *Studies in Second Language Acquisition*, vol. 35, no. 1, pp. 31–65.
Buck, G 2001, *Assessing Listening*, Cambridge University Press, Cambridge.
Council of Europe (COE) 2001, *A Common European Framework of Reference for Languages: Learning, Teaching and Assessment*, Cambridge University Press, Cambridge.
Cross, J 2011, 'Comprehending news videotexts: The influence of the visual content', *Language Learning & Technology*, vol. 15, no. 2, pp. 44–68.
Cross, J & Vandergrift, L 2014, 'Guidelines for designing and conducting L2 listening studies', *ELT Journal*, vol. 69, no. 1, pp. 86–89.
Cutrone, P 2005, 'A case study examining backchannels in conversations between Japanese-British dyads', *Multilingua*, vol. 24, no. 3, pp. 237–274.
Farrell, TC & Mallard, C 2006, 'The use of reception strategies by learners of French as a foreign language', *The Modern Language Journal*, vol. 90, no. 3, pp. 338–352.
Field, J 2001, 'Finding one's way in the fog: Listening strategies and second-language learners', *Modern English Teacher*, vol. 9, no. 1, pp. 29–34.
——— 2004, 'An insight into listeners' problems: Too much bottom-up or too much top-down?', *System*, vol. 32, no. 3, pp. 363–377.
——— 2008, *Listening in the Language Classroom*, Cambridge University Press, Cambridge.
Godwin-Jones, R 2012, 'Digital video revisited: Storytelling, conferencing, remixing', *Language Learning and Technology*, vol. 16, no. 1, pp. 1–9.
Goh, C 1997, 'Metacognitive awareness and second language listeners', *ELT Journal*, vol. 51, no. 4. pp. 361–369.
——— 2002, 'Exploring listening comprehension tactics and their interaction patterns', *System*, vol. 30, no. 2, pp. 185–206.
Goh, C & Taib, Y 2006, 'Metacognitive instruction in listening for young learners', *ELT Journal*, vol. 60, no. 3, pp. 222–232.
Grabe, W 2009, *Reading in a Second Language*, Cambridge University Press, Cambridge.
Graham, S 2006, 'Listening comprehension: The learners' perspective', *System*, vol. 34, no. 2, pp. 165–182.
Graham, S & Macaro, E 2008, 'Strategy instruction in listening for lower-intermediate learners of French', *Language Learning*, vol. 58, no. 4, pp. 747–783.
Graham, S, Santos, D & Vanderplank, R 2008, 'Listening comprehension and strategy use: A longitudinal exploration', *System*, vol. 36, no. 1, pp. 52–68.
Grgurović, M & Hegelheimer, V 2007, 'Help options and multimedia listening: Students' use of subtitles and the transcript', *Language Learning and Technology*, vol. 11, no. 1, pp. 45–66.
Guichon, N & McLornan, S 2008, 'The effects of multimodality on L2 learners: Implications for CALL resource design', *System*, vol. 36, no. 1, pp. 85–93.

Lynch, T 2009, *Teaching Second Language Listening*, Oxford University Press, Oxford.
Mareschal, C 2007, 'Student Perceptions of a Self-Regulatory Approach to Second Language Listening Comprehension Development', PhD thesis. University of Ottawa.
Prince, P 2012. 'Writing it down: Issues relating to the use of restitution tasks in listening comprehension', *TESOL Journal*, vol. 3, no. 1, pp. 65–86.
Robin, R 2007, 'Learner-based listening and technological authenticity', *Language Learning and Technology*, vol. 11, no. 1, pp. 109–115.
Rost, M 2006, 'Areas of research that influence L2 listening instruction', in E Uso-Juan & A Martinez-Flor (eds), *Current Trends in the Development and Teaching of the Four Language Skills*, Mouton de Gruyter, Berlin, pp. 47–74.
——— 2011, *Teaching and Researching Listening*, 2nd edn, Longman, London.
Roussel, S 2011, 'A computer assisted method to track listening strategies in second language learning', *ReCALL*, vol. 23, no. 2, pp. 98–116.
Smidt, E & Hegelheimer, V 2004, 'Effects of online academic lectures on ESL listening comprehension, incidental vocabulary acquisition and strategy use', *Computer Assisted Language Learning*, vol. 17, no. 5, pp. 517–556.
Vandergrift, L 1999, 'Facilitating second language listening comprehension: Acquiring successful strategies', *ELT Journal*, vol. 53, no. 3, pp. 168–176.
——— 2003, 'Orchestrating strategy use: Toward a model of the skilled second language listener', *Language Learning*, vol. 53, no. 3, pp. 463–496.
——— 2004, 'Learning to listen or listening to learn?', *Annual Review of Applied Linguistics*, vol. 24, pp. 3–25.
——— 2006, 'Second language listening: Listening ability or language proficiency?', *The Modern Language Journal*, vol. 90, no. 1, pp. 3–17.
——— 2007, 'Recent developments in second and foreign language listening comprehension research', *Language Teaching*, vol. 40, no. 3, pp. 191–210.
Vandergrift, L & Baker, S forthcoming, 2015, 'Learner variables in second language listening comprehension: An exploratory path analysis', *Language Learning*, vol. 65, no. 2.
Vandergrift, L & Goh, C 2012, *Teaching and Learning Second Language Listening: Metacognition in Action*, Routledge, New York, NY.
Vandergrift, L, Goh, C, Mareschal, C & Tafaghodatari, MH 2006, 'The metacognitive awareness listening questionnaire (MALQ): Development and validation'. *Language Learning*, vol. 56, no. 3, pp. 431–462.
Vandergrift, L & Tafaghodtari, MH 2010, 'Teaching students how to listen does make a difference: An empirical study', *Language Learning*, vol. 60, no. 2, pp. 470–497.