A Review of Second Language Listening Comprehension Research

JOAN RUBIN
2011 Hermitage Avenue
Wheaton, MD 20902

AS TEACHERS AND SCHOLARS GROW TO understand the unique characteristics of the listening skill and the significant role it plays in language learning and communication, they recognize more and more the importance of teaching listening comprehension in the second language classroom. This recognition has resulted in an increase in the number of listening activities in student textbooks and even in methodology texts designed specifically for listening (2; 99; 120; 121).

In the past decade, scholars have reviewed the factors that may influence second language (L2) listening comprehension (19; 38; 39; 44; 68; 69; 84; 98; 99), referring, in some cases, to factors isolated for first language (L1) learning (2; 29; 36; 37; 42; 43; 70; 72). Surprisingly few of the myriad of factors listed in these reviews have been researched and, of those that have, even fewer have been subjected to adequate experimental controls. In the main, many factors are cited as relevant either because they are suspected on logical grounds to affect listening or because they are thought to be relevant based on parallels found in reading research. Of those factors that have been researched, most have been the object of only a limited number of research studies, often only on one of the more commonly taught languages. Finally, most of the research results are based on listening comprehension measures that have not been standardized, making it difficult to compare results.

There is, however, a small—but growing—body of research on listening comprehension that forms a lively, ongoing dialogue about how learners interact with oral input. How and when do learners rely on top-down factors (background knowledge, semantics) and on bottom-up factors (acoustic features, stress and rhythm, syntax)? When do they use parallel processing? What kind of learner-interlocutor interaction affects listening comprehension? What kind of learner strategies improve L2 listening comprehension? How can we best train learners to use strategies?

Emerging from this dialogue are five major factors that researchers believe affect listening comprehension: 1) text characteristics (variation in a listening passage/text or associated visual support); 2) interlocutor characteristics (variation in the speaker’s personal characteristics); 3) task characteristics (variation in the purpose for listening and associated responses); 4) listener characteristics (variation in the listener’s personal characteristics); and 5) process characteristics (variation in the listener’s cognitive activities and in the nature of the interaction between speaker and listener). This review of listening comprehension research will consider each of these five factors in turn.

It is hoped that this review of current scholarship will provide some direction for teaching and facilitate future inquiry by identifying some of the most potentially productive research questions. Indeed, recent dissertation research (7; 27; 107) has begun to explore relationships among these factors. Much cooperation and rigorous scholarship is needed to push forward the young field of L2 listening comprehension studies.

TEXT CHARACTERISTICS: ACOUSTIC—TEMPORAL VARIABLES

Griffiths (58) suggests that three temporal variables—speech rate, pause phenomena, and hesitation, studied within the science of prosody—are directly relevant to L2 research and English language teaching methodology. Speech rate includes a "normal rate" (the rate at which speakers can usually easily comprehend
a text) and a "threshold rate" (the rate at which comprehension begins to decrease rapidly). Pause phenomena include pause duration, pause distribution, and pause frequency. Hesitation phenomena comprise filled pauses (for example, English "er," "um," "schwa," etc.) as well as repeats and false starts.

**Speech Rate.** Anderson-Hsieh and Koehler, Blau (9), and Griffiths (58) all note that it is difficult to compare rate studies. All the studies reviewed by these authors use subjects from different language and culture backgrounds and different age groups. Each also uses different mechanical compressor-expanders, different texts of different lengths, and different measurement techniques. In addition, they usually use different definitions of normal rate and different points where the threshold rate occurs.

Griffiths (58) suggests that different languages have different "normal" rates and the rates defined in studies using English cannot be applied exactly to studies of other languages. Tauroza and Allison also note that normal rates vary among text types and that the range of what is considered normal may vary from language to language.

A study by Rader considered the effect of slowing down speech for third-quarter university students of Spanish. Rader notes that the nonsignificant results may be related to: a) the control of background knowledge, b) the high level of difficulty of the texts, c) the lack of student exposure to authentic extended aural discourse, and d) the fact that recall was not done immediately after exposure. This study further illustrates how speed can interact with experimental procedural issues as well as with both declarative and procedural knowledge.

Despite these criticisms, most research quotes a normal speech rate of 165 to 180 words per minute for native speakers of English (e.g., 116). On the other hand, while Foulke reports a threshold level between 250–275 w.p.m., others (116) state that comprehension decreases as a function of mental aptitude and difficulty level. Carver concurs, suggesting that each student has his/her own level of comprehension. Tauroza and Allison compare normal speed of British speakers for four types of speech. They found that while the mean for radio and interview speech events lies within the range of 160 to 190 w.p.m., the means for conversation and lecture categories are outside this range. The mean rate for conversation in words per minute was 210; for lectures, 140. They note further that thirty-three percent of their lecture data was slower than 130 w.p.m. and twenty-three percent of the conversation data was faster than 220 w.p.m. The issue of normal speech rate is one that still needs a great deal more research that takes into account all of the variables mentioned above.

For L2 listeners of English, there is conflicting evidence about how speech rate affects comprehension. Using a text that assumed little background knowledge, Griffiths (59) found potential evidence that speech faster than two-hundred w.p.m. is hard for lower-intermediate learners to understand. He found that this level of student performed best at 127 w.p.m. Working with intermediate-level students, Kelch found significant comprehension effects for slowed speech (124 w.p.m.). On the other hand, Blau (8; 9) found that speech ranging from 145 to 185 w.p.m. did not significantly affect comprehension of intermediate- and advanced-level L2 students. This conflicting evidence about the effect of speech rate on comprehension may relate to a number of variables such as how proficiency is measured, type of text used, and amount of background knowledge required.

Yet another factor in measuring the impact of rate of speech is the kind of listening required. King and Behnke studied the interaction of task with listening type for L1 subjects. They identified three kinds of listening: comprehensive listening (understanding a message and remembering it for future use); interpretive listening (detecting inferred meaning); and short-term listening (receiving, processing, and recalling limited amounts of information over short time periods). King and Behnke found that comprehensive listening performance deteriorated significantly as speech compression levels increased (i.e., faster speech), while interpretive and short-term listening comprehension remained stable until a high degree of time compression (60%) was reached.

From another perspective, there are proponents for using syllables per second as a more reliable unit of measure than words per minute, since the latter vary considerably among speech categories (i.e., radio broadcasts, lectures, interviews, conversations) (57; 117). Vanderplank (125) argues that neither words per minute nor syllables per second are wholly satisfactory in determining difficulty of understanding speech because both ignore the influence of stress and rhythmic patterning. Certainly, research is needed to determine which variables are more critical.
Hesitation and Pause Phenomena (usually grouped together). Several researchers have examined hesitation phenomena for L2 learners (8; 9; 45; 67; 129). Voss (129) is the only one who regards hesitation phenomena as detracting from comprehension. The others find that hesitation phenomena are beneficial to comprehension. However, Voss is the only one of these researchers to use real, spontaneous speech. All of the other studies use a written text read aloud. Clearly, notice must be taken of the role of hesitation within a language, including the expectations listeners have concerning hesitation phenomena.

Voss (129) studied repeats, false starts, filled pauses, and unfilled pauses in spontaneous speech. He found that all types of hesitation phenomena cause perceptual problems and thus comprehension errors for non-native speakers (NNs). Voss suggests that native speakers (NSs) are usually not bothered by hesitation phenomena since they possess strategies to recognize and process such phenomena while focusing on meaning. NSs get stuck in bottom-up processing of phonetic utterances that do not affect meaning, while NSs discard these utterances in favor of top-down processing.

In a second study, Voss (127) analyzed the kinds of perceptual errors that native German speakers made while listening first to German and then to English texts. Both texts had a high proportion of hesitation phenomena. Voss found that “successful speech perception depends on an active reconstruction process applying top-down strategies to the acoustic input, i.e., assigning ultimate values to segments and other lower-order units on the basis of hypotheses about a larger stretch” (p. 148). He argues further that language “is not primarily processed through a sequential identification of segments and units of increasingly bigger size” (p. 148).

Jacobs et al. considered the effect of longer than normal pauses at clause and sentence boundaries on comprehension of expository lecture notes read aloud at both slow and normal speeds. Jacobs et al. found that pauses related significantly to comprehension for more advanced students, but not for less advanced students. Jacobs et al. suggest that students may need to be at a certain ability level in order to benefit from the processing time that pauses afford and also that input with overly long pauses can bore learners.

Dunkel (45) worked with both low and high listening proficient groups and used taped versions of a lecture. She found that use of hesitation phenomena gave students time to take a greater quantity of notes.

In the first of two studies, Blau (9) compared the effect of speed and blank pauses on comprehension. At most language proficiency levels she found that blank pauses facilitated listening comprehension more than speed. In a second study, Blau (8) found that blank pauses did not help comprehension while filled pauses did. Blau speculates that this difference in findings may relate to different conversational norms among NSs. (The subjects in the first study were Puerto Rican and Polish, and those in the second were Japanese).

TEXT CHARACTERISTICS: ACOUSTIC—OTHER VARIABLES

Level of Perception. Hieke raises an important question: what do listeners perceive as they listen? Whereas it is generally assumed that L1 listeners can and do perceive individual words, Hieke suggests that in fact it might be “syllables and the segments that make them up” (p. 127) that help listeners decompose the stream of speech. As Hieke notes, more research is needed about how L2 listeners break up speech.

Dejean de la Bâtie notes that beginning language learners have a difficult task listening since they have limited exposure to connected speech, inadequate phonological competence, and inefficient processing strategies. Bradley and Dejean de la Bâtie argue that word boundary association is negotiated under the simultaneous constraint of lexical and prosodic analysis.

Sandhi. Some researchers suggest that one problem in dividing the stream of speech is the effects of sandhi. Sandhi is defined as “the phonological modification of grammatical forms which have been juxtaposed” (62: p. 311) and includes assimilation, mutation, contraction, liaison, and elision. Henrichsen, considered the effect of sandhi-variation on the comprehension of input by English as a Second Language (ESL) students. This study included high- and low-level students and NSs. Henrichsen reached three conclusions. 1) Listeners with considerable knowledge of the system and elements of a language can use that knowledge to compensate for the loss or reduced saliency of portions of the input created by sandhi-variation; when sandhi was present, NSs' comprehension was significantly better than that of high-level ESL learners. 2) Language learners with only a limited knowledge of the linguistic code are not
able to compensate for the missing signal information when input is characterized by sandhi; students of both ESL levels showed comprehension levels that were significantly similar to each other and significantly different for those of NSs when sandhi was present. 3) When sandhi was not present, comprehension of high-level ESL students was closer to that of NSs and was significantly different from that of low-level ESL students.

Bradley and Dejean de la Bâtie looked at perception of /t/ at word boundaries (a form of sandhi) in French by three groups: NSs, first-year, and second-year university students of French. The data suggest strongly that /t/ detection is lexically guided and that recognition of spoken words is negotiated through the simultaneous constraint of lexical and prosodic analyses. The perception of /t/ tended to be less for NSs than for second-year students, which is in line with presumed differences in processing efficiency.

Ahn considered the effect of Sandhi on listening comprehension of Korean university students learning English in Korea and Korean graduate students at an American university. Ahn found that sandhi significantly hampered listening comprehension. Although this author has seen only an abstract of this research, it appears that subjects were presented with words in sentences. However, no comparison is made of proficiency levels.

Since Henrichsen and Bradley and Dejean de la Bâtie both used invented, isolated sentences, much more research using other more natural contexts is in order. Further research is needed to delineate the role of segmentation in listening, especially in languages with reduced syllables, extensive sandhi, and/or complex morphological transformations. Finally, as Dejean de la Bâtie suggests, segmentation needs to be researched in the context of perceptual saliency, that is, stressed syllables, high intensity segments, and the way phonemes are articulated.

Stress and Rhythmic Patterning Perception. Vanderplank (124) argues that perception of stress is an important factor in rapid and efficient "tuning in." Vanderplank (125) found interesting differences in NSs' and NNSs' ability to perceive stress. "NSs were in general agreement as to sentence stress location, while only a small number of NNSs agreed with NS judgments as to stress location, and the ability to perceive stress location was not significantly linked to level of English ability in learners" (p. 38). Vanderplank (123) argues that stress and rhythmic patterning comprise an important factor in determining level of difficulty. Bond and Garnes agree that speech perception is also subject to what they call "heuristic strategies" such as paying attention to stress, intonation patterns, and stressed vowels.

Given the contribution that stress and rhythm make in assigning meaning in English and the fact that they differ substantially from language to language, this would appear to be a critical area for future research. We need to know which students use this acoustic variable, how they use it, and the conditions under which it is productive to train students to rely on stress and rhythm to enhance listening comprehension.

L1/L2 Differences. Although it is clear that more than acoustic differences distinguish L1 and L2, this category is included here because there is research in this area.

Two researchers have considered whether the kinds of errors made in listening are the same for both native and non-native listeners. Fishman compared NSs of English with NNSs who were at a fairly high level of ESL competence. She identified ten categories of error (largely phonological) and found that although native speakers made 2.5 times fewer errors than non-natives, the same error categories turned up in L1 and L2 listeners. Voss (128) compared errors in German and in English of German students of English who were in their second or third year of study and had had between seven and nine years of English before taking university courses. He found that all types of slips made by NNSs were also made by NSs. He concluded that, in principle, the perception strategies used in the L1 and L2 follow similar lines.

TEXT CHARACTERISTICS: MORPHOLOGICAL AND SYNTACTIC MODIFICATIONS (INCLUDING RESTATMENTS)

Krashen proposed that learners need "comprehensible input" to learn a language. Although research in the acoustic section above was conducted independently of the Krashen hypothesis, much of it relates to making language comprehensible. The section below reports on research that considers the role of morphological and syntactic modifications in making input more comprehensible.

There is a great deal of controversy over what kinds of syntactic modifications improve comprehension. The major issue is whether syntactic modifications made by NSs are helpful or
whether they are best made in response to a learner's requests for clarification. In this section, we will consider only those modifications initiated by the speaker.

**Syntactic Modifications.** Chaudron (23) looked at NNSs of low, medium, and high proficiency and evaluated the effect of syntactic modifications on recognition and recall for all three proficiency levels. He found redundancy in the form of a repeated noun to be much more helpful in both recognition and recall than any of the other devices tested.

On the other hand, Kelch tested the effect of modifications (synonymy, hyperonymy, parallel syntactic structures, and paraphrase) with intermediate students and found no significant main effect for NS syntactic modifications alone. He found that there was an effect only for those passages with both modifications and a slowed rate of speech or for those passages with a slowed rate of speech alone.

Another study by Cervantes and Gainer, comparing the relative effectiveness of syntactic simplification with repetition, found "no significant difference between scores for groups hearing the syntactically simplified version and the complex version with repetition" (p. 769). The authors worked with first-year students and seniors at a Japanese university; proficiency level was not stated, but the levels might be intermediate and advanced, respectively. They found that proficiency level did not affect their findings.

Chiang and Dunkel investigated the listening comprehension of high-intermediate and low-intermediate students (based on scores on the Comprehensive English Language Test [CELT]) on lectures that were modified and others that were not. The modified discourse contained information redundancies and elaborations. They found that high-intermediate students benefited from speech modification while low-intermediate students did not.

These four studies point to several variables that interact to influence comprehensibility: student proficiency level, type of modification, type of passage, and amount of background knowledge. Since comparability for proficiency level has not been established, it is hard to verify comparability of results.

**Redundancy.** Both Chaudron (23) and Chiang and Dunkel report that redundancy improved comprehension. Chiang and Dunkel found that modification (repetition of constituents, paraphrase, and use of synonyms) works best with high-intermediate students as compared to low-intermediate students. Pica, Young, and Dougherty, working with low intermediate ESL students, also reported improved comprehension when content words were repeated. Hence, redundancy in input appears to be important in comprehension.

Glisan also found that advanced students understood longer, modified sentences more readily and more successfully than shorter, unmodified sentences. She concludes that providing advanced students with explanatory and/or redundant information helps them understand. Derwing concludes that “simple redundancy” of the type used by Chaudron (23) with low-intermediate students facilitates understanding while redundancy of another type—the increased use of background detail—detracts from comprehension.

However, other researchers (84; 106) have suggested that, at the lowest level of proficiency, some kinds of redundancy add to the listening load in that they provide more language to process. As Lynch (84) notes, providing rephrased and reformulated expressions may go unused because they are perceived as alternative not additional information.

Hence, two variables appear to affect the utility of redundancy for enhancing listening comprehension: proficiency level and kind of redundancy.

**Morphological Complexity.** To our knowledge, there is no research that studies languages in which the complexity of the morphology changes with the degree of formality (e.g., Indonesian).

**Word Order.** Research by Glisan examines the effect of Spanish word order patterns on aural comprehension of English-speaking students learning Spanish. Working with advanced students, Glisan reports a hierarchy of comprehension difficulty associated with word order. The hierarchy was: 1) subject-verb-object; 2) verb-subject-object; and 3) object-verb-subject. Glisan also found a significant effect for the position of the key sentence in the text. Mean comprehension scores were lower for initial sentences than for medial and final sentences. Glisan notes that these results are appropriate for advanced students of Spanish as less competent students may not pay as much attention to syntax.

**Discourse Markers.** Chaudron and Richards consider two kinds of discourse signals in a lecture: macromarkers—which signal the macro structure of a lecture through highlighting the major information in the lecture and the se-
quencing or importance of that information—and micromarkers—which indicate links between sentences within the lecture, or function as fillers. Working with a pre-university group of ESL students (level was not specified) and a university group (at a "higher" level of proficiency), the authors found a consistent result across groups: macromarkers signaling major transitions and emphasis in the lecture are more conducive to successful recall of the lectures than micromarkers. They suggest that the micromarkers may make the entire lecture appear less well organized [see Voss (129) above and the role of hesitation in reducing comprehension].

Hron et al. compared the listening comprehension of German science students of two versions of a lengthy English text with and without discourse markers. Discourse markers included signaling where 1) logical relations are emphasized (e.g., first, second); 2) a specific summary statement containing the most important ideas precedes text segments; 3) summary statements are put at the end of text segments; and 4) pointer words are used to represent the author’s viewpoint (e.g., the most important”). Hron et al. conclude that listeners who heard the text with the additional discourse markers reproduced significantly more macropropositions than those who heard the nonmarked version. However, they note that use of signaling with listening is more limited and has different perception prerequisites from that of reading.

TEXT CHARACTERISTICS: TEXT TYPE

Text type has been noted as a factor affecting listening comprehension. It is assumed that since most written texts are more syntactically complex, less redundant, denser, and use fewer pauses than spoken texts, they are potentially more difficult to understand (i.e., cognitively taxing). Conversely, because conversational texts are less complex syntactically, more redundant, and have more pauses, they are potentially easier to understand.

Shohamy and Inbar considered the relative comprehensibility of three text types: a news broadcast using a pre-written, edited monologue, a lecturette consisting of a monologue based on written notes, and a consultative dialogue involving constant interaction. Their subjects were twelfth-grade Israeli students of English as a Foreign Language (EFL). For each text type there were two topics. The text types in both topics followed the same hierarchical order of difficulty: news was the most difficult, followed by the lecturette, with the dialogue being the least difficult type text.

This finding parallels those of Brown et al. (15) who worked with L1 English-speaking students (approximately age 15) to provide empirical evidence that narrative texts are easier to listen to and recall than expository texts. Furthermore, they report data to support their view that “events described in chronological order are easier to recall than narratives with disrupted sequences or flashbacks” (p. 36).

On the other hand, a recent dissertation by Berne considered whether the listening comprehension of FL learners varied according to text type and assessment task. Berne worked with English NSs studying beginning and advanced-intermediate Spanish at the university level. She found no main effect for text type (lecture vs. interview). In a secondary series of analyses, however, she found that, although text type was not a significant factor in testing for main ideas, it was significant in the comprehension of details. However, this difference emerged only on the multiple choice test type; neither the open-ended task nor the cloze task showed a difference. Suggesting a major limitation of her study, Berne noted that since the texts were constructed to control for background information, the discourse style of the lecture and the interview may not have been different enough. She noted also the length of the texts were two to three times longer than the recommended length of two to three minutes, causing her to wonder if memory demands may have been more critical than text differences. She suggests that text type might have been significant had less difficult and/or shorter texts been used.

These studies indicate the need for further research that rates texts by a variety of criteria such as kinds of discourse markers, word order, redundancy, morphological complexity, length, and difficulty, and that measures student proficiency through questions keyed to the proficiency level of the text.

Visual Support for Texts. Studies have suggested visual support can enhance listening comprehension. Rubin (102) found that the listening comprehension of high-beginning Spanish students who watched dramas on video improved significantly over students who received no video support for their listening training. She argues that “video can serve as a haven to enhance listening comprehension if it is selected so that it provides sufficient clues for in-
formation processing. It is the selection that is critical, not just the use of video alone" (p. 315).

Secules et al. considered the impact of video on listening comprehension of second-semester university French students. The control group used a “direct method” text and did pattern practice, pronunciation exercises, reading activities with a focus on grammar, and communication activities (question/answer and role play). The experimental classes used the French in Action video series. The research design controlled for both teacher and class composition. The treatment included viewing the dramatic section of the tape, followed by comprehension questions and guided discussion. Workbooks, audiotapes, structure exercises, and communicative activities were also used. The experimental group scored higher overall in listening comprehension than the control group. In particular, the experimental group outscored the control group on questions involving main ideas, details, and inferences. Similarly, Herron et al. (64) reported that for first-year university French students listening comprehension improved more after one year’s exposure to a video-based curriculum than after the same length of exposure to a text and audio-based curriculum.

Further clarification of when and how video can help is provided by Mueller’s study of English-speaking students of German. Mueller showed that when the text is an interview, with a single, visual line drawing, the helpfulness of visual clues is inversely related to proficiency levels—in the more proficient the learner, the less crucial the visual aids are for comprehension.

Another variable in the use of visual support is the perceived difficulty of the text. Wolff’s study of German ESL learners between ages twelve and eighteen found that the more difficult the text was, the more subjects used the illustration. Wolff suggests that with an easy text, subjects ignored the illustration because bottom-up processing was not impeded, but with a harder text, subjects needed a more schema-based approach, which the illustration permitted.

The amount of support gained from video can vary both within and between text types. For example, a drama or soap opera may consist of “talking heads” that provide little support, or the images may correlate fairly closely with the conversation, at times even to the point of negating the need to listen at all! The visual aspects of many interviews and most newscasts do not correlate closely with the text. However, as Phillips observes, the use of visuals to support particular text types can vary from language to language.

We need more research to demonstrate Phillips’s assertion that there is no strict hierarchy of video texts. The difficulty of the same type of text depends on whether a) the text is supported or not with visuals, b) the visuals are consistent with the speech, c) the background music distracts the NNS, d) there is a variety of voices or speech styles, and e) how all of these variables interact with the learner’s proficiency level.

INTERLOCUTOR CHARACTERISTICS

Markham considered sex bias and perceived speaker expertise on ESL student listening recall. Working with intermediate and advanced university level ESL students, he reports that: 1) both groups recalled more from the nonexpert male speaker than from the female nonexpert; 2) the advanced group recalled more from the male expert than from the female expert (the same was true for the intermediate group but the means did not differ as markedly); and 3) both groups performed at a noticeably higher level when the presentation was given by a female expert than by a female nonexpert. He concludes that “gender bias is a pervasive factor that exerts an influence on ESL students’ recall of orally presented material” (p. 404). However, this result may be more true for some cultures than for others (over 51% of their subjects were North Asian). As he notes, “Perhaps female listeners are gradually conditioned to be more attentive to male speakers as a result of gender-related status divisions in the speech community” (p. 404).

TASK CHARACTERISTICS

No research has directly related the many variables already discussed above to task type. However, Lynch (84: p. 198) suggests that grouping students to work on listening may lead to more negotiation and hence greater clarification of text. Buck (16) suggests that more research is needed on task motivation and test results.

Task Type. There is little research comparing task types for L2. Brown et al. (15) considered ways to assess the relative difficulty of a task for native English listeners. They note that output tasks should not rely heavily, or exclusively, on memory or writing abilities. They found that
informative titles facilitated comprehension if listeners were directed to pay attention to them. They also found that students varied greatly in their ability to summarize texts from one listening to another.

Eyken studied the impact of four tasks on the recall protocol of novice high school French learners while watching authentic video material. The tasks were multiple choice, choose-a-picture, French to English vocabulary lists, and WH-questions. Results indicated that learners produced protocols with more idea units when working with multiple choice than with WH-questions. Eyken found considerable variability with the picture guide and vocabulary lists. The author does not provide student responses for each task independent of the recall protocol. Further, the results on the WH-task need reexamination because, as the author notes, students may not have clearly distinguished among the WH-questions.

Shohamy and Inbar considered how type of question influenced success in L2 listening tasks. They found that subjects performed better on questions referring to local cues in the text than on those referring to global cues. They concluded that it is apparently more difficult to generalize, infer, and synthesize information than to look for data-specific information. They further noted that “low level test-takers were more affected by the type of information demanded, because their performance on local and global question types differed to a greater extent than that of learners of the high group” (p. 35). Finally, they reported that most students who responded to global questions were also able to respond to local questions but not vice versa.

Lund (81) also looked at how task type affects learners’ ability to remember more main ideas or details and to give more inferences and elaborations. The study was conducted with university students in the first three semesters of German. Lund found significant main effects for task and semester in the total number of propositions recalled to conclude that tasks affect 1) recall of task-specific propositions, 2) the proportion of macropropositions to micropropositions recalled, and 3) the number of distortions. He found no significant effects either for inferences or elaborations.

As Lund notes, “though the effects of tasks are clear the underlying causes and their dynamics are less so” (p. 12). Further, Lund notes “the results of the study do not necessarily indicate that one task is superior to another. It all depends on the objectives for listening” (p. 14).

The study shows that learners do modify their behavior depending on the task they are given. Lund suggests that further research should consider the method of responding, learner traits, and topics as well as task sequences and combinations.

Pedagogical Research. Postovsky, working with beginning students of Russian at the Defense Language Institute, considered the effect of delayed oral practice on all four skills. An experimental group was deliberately kept from oral production, being required instead to respond only in writing. After four weeks, the experimental group was merged with students in the regular Russian program. Tested after twelve weeks, the experimental group performed significantly better in listening comprehension than the control group. In addition, the experimental group outperformed the control group on the other three skills. Since this is a unique experiment, it deserves repetition for its policy implications.

LISTENER CHARACTERISTICS

Listener characteristics appear to have considerable impact on an individual’s listening comprehension. Among these characteristics are: language proficiency level, memory, affect, age, gender, learning disabilities in L1, and background knowledge as well as aptitude, processing skills, background biases, motivation, and confidence level. Only some of these have been subject to research.

Language Proficiency Level. Language proficiency is a major variable in almost all of the studies in this review. A major problem in comparing studies is that there are few standardized tests to determine proficiency level. Most studies use either teacher judgment, course level, or performance on a nonstandard test. Still, most researchers suggest that cognitive processing will vary depending on learners’ knowledge of the language. It is not clear what role grammar, vocabulary, background knowledge of the culture, and knowledge of discourse processes play at different proficiency levels. In addition, all tests are dependent on task, context, and lesson purpose as well as on learner purpose, interest, and motivation. Although standardization of listening proficiency tests is still evolving, tests such as the American Council of the Teaching of Foreign Languages (ACTFL) and The Interagency Language Roundtable (ILR) are working toward a standard; their tests may serve as a needed standard
in future listening comprehension research. Language proficiency is clearly a variable that needs to be considered in every study. (For detailed discussions of the many variables that must be taken into account in listening comprehension assessment, see 46; 118.)

Memory. The relationship between memory and listening comprehension is complex. Our understanding of this relationship is hindered because our knowledge about the nature of memory is still evolving [see Stevick's reviews (114; 115) of memory research for second language learning teachers] and because we do not yet clearly understand how language is processed during listening.

Nonetheless, there are several compelling hypotheses that deserve further investigation, as suggested by Call:

-memory span for target language input is shorter than for native language input, the amount of target language input that can be successfully processed seems to increase as proficiency in the language increases. A corollary of this finding is that length of memory span for linguistic input is a good indicator of overall language proficiency. Knowledge of target language syntax seems to be an important factor in increasing the amount of linguistic material that can be retained in short-term memory (p. 769)

Call considered whether listening comprehension was related to short-term memory for five types of auditory input: sentences in context, isolated sentences, random words, random digits, and musical tones. She found that listening comprehension was correlated with these five types in the following order: isolated sentences, sentences in context, musical tones, random words, and random digits; with musical tones being closer to natural speech than expected. Questioning Call's conclusion that short-term memory is important in listening comprehension, Buck (17) argues that although forty-two percent of the difference in the listening test was explained by a correlation of all five short-term memory tests, "... the correlation would surely be due to the fact the tests of sentence recall and the test of listening comprehension are likely to be dependent on some degree of general language proficiency, and this could be the cause of the correlations observed" (p. 90). Nonetheless, Buck agrees that Call's research indicates at least a weak relationship between short-term memory and listening comprehension. Greenberg and Roscoe worked with German students in the first six weeks of study to determine if echoic memory differs between students with "weak" or "strong" listening comprehension. Considering how using a suffix or tone control interfered with comprehension, they found that the weaker group was more affected by the interference than the stronger group.

Dunkel et al. (46) considered the influence of short-term memory on encoding lecture material in English. The authors studied students taking college freshman English classes. The students were NSs and NNSs who were presumably at an advanced level in English. They found that subjects (not divided by NS and NNSs) who had high short-term memory correctly recognized significantly more content information and detail information than subjects who had low short-term memory. In addition, they found that NSs recognized significantly more of the lecture concepts and detail than did NNSs of English.

These studies may need to be reconceptualized in the light of new thinking about short term memory as a state rather than as a stage [cf. Stevick's discussions, 114; 115]. This line of thought suggests that memory—with every act of recall, and even every act of recognition transforming it—is always in transition. Studies may need to focus more on levels of activation at any point in time. This transitory and interactive quality of short-term memory requires a more dynamic research model to characterize listeners and its impact on listening.

Conrad (30) compared memory of NNSs at high- and medium-skill levels for recorded sentences at different speaking rates. She found that NNSs tended to ignore information in the middle of sentences; rather, they tried to duplicate beginnings or ends of sentences. (Although Conrad was not studying strategy use, she provides evidence of the strategy of ignoring information used by listeners.)

Attention. Another area that logically must affect memory and in some way listening comprehension is attention [Butler (18) discusses attention and memory for L1]. There is little research on the role of attention in L2 listening comprehension, perhaps because of the difficulty of observing degrees of attention or asking learners to report them.

However, O'Malley et al. (91) provide anecdotal evidence from their research on listening strategies that "... effective listeners seemed to be aware when they stopped attending and made an effort to redirect their attention to the task" (p. 428). They found further that "ineffective listeners reported that when they encoun-
tered an unknown word or phrase in a listening text, they usually just stopped listening or failed to be aware of their inattention." (p. 428).

Affect. The role of affect in listening appears to be related to attention and the functioning of memory. Using verbal protocols and think aloud reports, a dissertation by Fujita examined the listening strategies of beginning students of Japanese at the US Air Force Academy. Students were divided into successful and unsuccessful listeners respectively in the top third and bottom third of the class on six listening comprehension exercises administered over an academic year. He found that successful students reported self-confidence as one of three major factors affecting their listening.

A dissertation by Anciro considered the relationship between apprehension in a second language, listening comprehension, and language competency. She worked with college students in Puerto Rico (proficiency is not given) and found that: 1) high receiver apprehension was significantly related to lower listening comprehension and L2 language proficiency, 2) high exposure to English was significantly related to lower receiver apprehension, and 3) dyadic communication created the greatest amount of apprehension, followed by receiving information, communication in a group, and watching television.

There appear to be very few studies comparing affect and listening comprehension. Research in this area should take into account the interaction between language competency, self-confidence, empathy with the listener, listening proficiency, and the type of listening tasks.

Apprehension or self-confidence is only one kind of affect. There can be others such as affinity for a language, the teacher, the country, and the oral tradition which impinge on listening comprehension. None of these appears to have been researched.

Age. Seright considered the relationship between age and L2 achievement of adults in an instructional setting. Working with members of the Canadian Armed Forces undergoing English language training in Quebec, she used seventy-one learners who ranged in age from seventeen to forty-one years of age with a mean age of 22.76 years. Seright divided the students into two groups: an older group, age twenty-five and older, and a younger group, age twenty-four or younger. She found that the mean gain in listening comprehension made by the younger subjects exceeded that made by the older subjects. Seright notes that this study investigated only the rate of development but made no contrast in differences in ultimate attainment.

Since most of the subjects were not really old enough to experience loss of auditory acuity and reaction time, Seright suggests other variables to explain the differences, such as years away from school and associated difficulties in adapting to formal instruction, and perhaps degree of inhibition or confidence. Seright notes that this study is consistent with an earlier study by Halladay that showed that younger adults acquired L2 skills more quickly than older adults.

Seright suggests that there are many variables that may affect listening comprehension: first language, educational background, occupation, purpose of instruction, initial L2 proficiency, type of instructional program, and previous L2 exposure. Her results are thus only a beginning to our understanding of how age affects listening comprehension.

Gender. Three studies consider how gender may relate to differences in listening comprehension. Boyle researched Chinese students of English between eighteen and twenty years old (proficiency level was not given, but part of the group was reported to be at the lower 40% of a placement test). Using a battery of tests—some published and some not—including a vocabulary recall, a listening passage, a listening conversation, two dictations, vocabulary identification given orally, as well as other items not related directly to listening, he found that males did significantly better on two tests of listening vocabulary, and women did significantly better on all other tests. It is difficult to agree with Boyle’s conclusion that females do better on “general language ability,” and males do better on “listening vocabulary,” especially since the females did do well on vocabulary recall.

Feyten looked at university students of French and Spanish and failed to find a significant relationship between gender and any foreign language proficiency measure. Bacon (6) looked at university students of Spanish and also failed to find a significant relationship between gender and listening comprehension. Thus, the small amount of research on gender and listening comprehension is inconclusive.

Learning Disabilities in L1. Sparks and Gan- chow (112) suggest that students who experience difficulties learning a FL may have L1 problems. The authors provide evidence from an initial study by Sparks et al. (113) of twenty-two college students who were granted a waiver from the FL graduation requirement because of
their inability to pass a FL course. They found that of thirteen students with phonological deficits in listening, seven failed the FL in the first semester, and six failed in the second semester. In this and their later article, Sparks and Ganschow suggest that L2 learning difficulties are a direct result of a deficit in L1 phonological processing. The authors open up a new variable to explain individual differences in listening comprehension; further research is in order.

Background Knowledge. Current views on listening comprehension agree that background knowledge or prior knowledge (also called declarative knowledge) can affect listening comprehension. To make sense of the rapid-fire "noise" that comes from oral speech, learners often try to find an overall schema. Even at the word, phrase, or sentence level students attempt to associate prior knowledge of the language with the incoming "noise."

Several studies have considered the role of background knowledge in listening comprehension. These studies parallel experiments carried out for L1 listening (see 13 for a classic experiment of this type.)

Long (80) considered whether Spanish FL listeners comprehend better when they possess schemata relevant to the listening topic. Subjects were students enrolled in a third-quarter university Spanish course. Before beginning the experiment, students completed a survey of their background knowledge of the two subjects used, namely, gold rushes and rock groups. Two measures of comprehension were used: a recall protocol and a recognition test consisting of paraphrased statements in English about the texts.

According to the survey results, students in Long’s study possessed less information about gold rushes than about rock groups. Long found that recall protocols for the rock passage revealed a significantly higher number of correct idea units than protocols for the gold rush passage, suggesting that background knowledge can relate to listening comprehension. Long further found that schemata can also have dysfunctional effects on listening comprehension. Learners overextended their gold rush schema onto a set of data that were clearly incongruent. Similar misapplications of background knowledge are reported by O’Malley et al. (91). The question of how students monitor their use of background knowledge and how they come to recognize incongruities is a very important issue that needs research.

Schmidt-Rinehart (105) also considered how topic familiarity and proficiency level affect recall measures of listening comprehension. Students from three intact sections of first-, second-, and third-quarter university Spanish classes listened to two passages, one about a familiar topic and another about a novel topic. Comprehension was assessed through a recall protocol. Results showed that both topic familiarity and course level affect measures of listening comprehension. All subjects recalled significantly more information from the familiar topic and the improvement in comprehension scores ascended with the quarter level.

Finally, Chiang and Dunkel studied male undergraduate students at an intermediate level of English at the Chinese Naval Academy in the Republic of China. They used a 650-word lecture on “The Amish People and the Pennsylvania Dutch Country” (a topic considered unfamiliar to students) and a 680-word lecture on “Confucius and Confucianism” (a topic considered familiar to students). Analysis found significant main effects for prior knowledge. However, when a within subjects variable of test type (passage independent versus passage dependent) was considered, the main effect was nullified. The authors explain that the significant effect appeared only on the passage independent items. Subjects’ performance on the passage dependent items did not differ significantly between the passage containing a familiar topic and the passage containing an unfamiliar topic.

In all three studies, background knowledge is shown to improve listening comprehension. However, as Chiang and Dunkel show, measures of background knowledge can vary and, as a result, may confound results. Further research is needed on the effect of the many different kinds of background knowledge and their interaction with proficiency level. It would also be helpful to consider the interaction between text type and types of background knowledge. In addition, as mentioned above, we lack very important information about how students monitor their use of background knowledge.

Another slant on the role of background knowledge is provided by Markham and Latham who investigated whether religious background influences listening comprehension. The study worked with advanced, university level ESL students who reported being either religion-neutral, practicing Moslems, or practicing Christians. Two expository prose passages were played: a 613-word description of Islamic prayer rituals and a 642-word description of Christian
prayer rituals. On the Christian passage, the Christian units recalled a greater number of idea units and had more elaborations and fewer distortions than Moslems, and vice versa. With regard to the neutral group, they recalled more major idea units in the Moslem passage. The authors suggest that the Moslem passage may have been slightly easier to follow, as the major concepts were numerically labeled. The study points to another way in which background knowledge can influence listening comprehension.

**PROCESS CHARACTERISTICS**

The process of listening is probably more difficult to research than the four factors discussed above (text, interlocutor, task, listener characteristics) because process consists of internal operations or behaviors and hence is not easy to measure directly. Still, the process of listening comprehension has been emerging as a major research focus.

Process refers to how listeners interpret input in terms of what they know or identify what they don’t know. There is ample evidence that this is what L1 learners do (e.g., 52; 15). Process also refers to the way in which listeners use different kinds of signals to interpret what is said (e.g., 87). The following processes are currently being examined in L2 contexts.

*Top-down, Bottom-up, and Parallel Processing.* For L1 listeners, Marslen-Wilson and Tyler provide extensive evidence for real time interactive language processing theory in which lexical, structural (syntactic), and interpretive knowledge sources communicate and interact in an optimally efficient and accurate manner during processing. Bond and Garnes argue further that, in addition to using grammatical, phonological, lexical, and sentence level information listeners employ heuristic strategies (e.g., find a phrase or word).

There is an ongoing discussion in L2 research as to whether listeners use their knowledge of the world, situations, and roles of human interaction to focus on meaning (top-down) and then use their knowledge of words, syntax, and grammar to work on form (bottom-up) or vice versa, and as to when and how these two interact (i.e., when and how does parallel processing take place). Lund (82) provides evidence for top-down processing in a study comparing listening and reading of first-, second-, and third-semester university students of German. He found that listeners made misinterpretations on their recall protocols and were able to invent plausible contexts for conversations, which lead him to conclude that they relied considerably on top-down processing. Looking at meaning from a discourse perspective, Anderson and Lynch point out that, in the early stages of language learning, the meaning level may consist of merely recognizing the topic of a conversation or being able to make predictions about likely developments of the topic.

VanPatten considered how top-down and bottom-up processing relate to proficiency level. He gave narrative passages to university students of Spanish in first-semester, fourth-semester, and third-year university conversation courses. Task one involved listening for content only; Task two involved listening for content and a word-final morpheme; Task three involved listening for content and a separate morpheme; and Task four involved listening for content and a key lexical item. The study yielded a significant effect for the interaction of level and task. VanPatten found that when learners had trouble directing attention toward both content and form, doing a task that required paying attention to form interfered with their comprehension of content. He suggests that when students have difficulty they should focus on meaning first.

Wolff worked with twelve- to eighteen-year-old German students of English; their proficiency level is not given, but it was probably beginning- to low-intermediate. Wolff found while students appeared to make a “harmonious use of bottom-up and top-down processes” with an easy text, they used more top-down processing strategies for the more difficult texts.

In contrast to the findings of the above two studies, Conrad’s study (31, 32) of university students of English (scores ranged between 83.7% and 96% on the Michigan State University English Language Exam) showed that as L2 listeners decrease in proficiency they rely more on syntax than on contextual semantic cues. Differences in procedures and scoring may explain some of the differences in results. Conrad used a cloze test while Wolff and VanPatten used recall protocols. Conrad created her own systems of categories to score responses as syntactic or semantic. Both Wolff and VanPatten studied idea units. However, their methods of analysis differed; Wolff looked at how similar these idea units were to the original text, while VanPatten scored for notation of lexical or grammatical category.

Bacon (5) tried but was unable to verify VanPatten’s findings that listeners process meaning
before they process form. Her conclusions are closer to those of Conrad discussed above. She found that students of Spanish "in the first Spanish course beyond the Arts and Science foreign language requirement" used more top-down strategies with the more familiar passage (about mobile homes) than with the less familiar passage (about electric converters). She suggested that listeners employ more "text-based" or bottom-up strategies on more difficult input.

In a study of secondary school intermediate ESL students, O'Malley et al. (91) found that "effective listeners seemed to be listening for larger chunks, shifting their attention to individual words only when there was a breakdown in comprehension" (p. 429). Thus, at some level, there is a delicate interaction between top-down and bottom-up processing. The authors suggest that this interaction occurs especially when there is a communication breakdown.

Despite the obvious contrast among their conclusions, these studies illustrate the complexities of the listening processing and how difficult it is to observe, especially since most of the measures are indirect. It seems clear that many factors influence how and when listeners use and integrate syntactic, phonological, lexical, pragmatic, and discourse information. Some of those that merit attention include the difficulty of the text (relating to its genre, its interactive/transactional nature, and to the listener's background knowledge of the topic); the amount of text provided at any one time; and the learner's language proficiency.

**Listening Strategies.** Current views of listening comprehension propose that listeners actively process language input. Two types of processing have been discerned: cognitive strategies and metacognitive strategies. Cognitive strategies involve solving learning problems by considering how to store and retrieve information. Metacognitive strategies involve planning, monitoring, and evaluating comprehension. In addition to metacognitive strategies, listeners use metacognitive knowledge about themselves and how they listen best (person knowledge); about how much text to listen to, how difficult a text is, etc. (task knowledge); and about strategies and their relation to texts and tasks (strategy knowledge).

Research on listening strategies includes: work on several languages (ESL most of all, but also French, Italian, Russian, and Spanish); work contrasting strategy use at several proficiency levels; work with audio or video texts; work with interactive or transactional listening; work with cognitive and metacognitive strategies; work considering the relation of strategy use to test, task, and setting.

A major approach to studying strategy use is to consider the types of strategies used by FL students in order to answer the following questions: 1) Are strategies used in different ways?; 2) With different frequencies?; 3) In combinations of strategies or strategy types, such as cognitive or metacognitive? These studies do not lead to firm conclusions because the measures of student differences are quite subjective.

Murphy worked with intermediate-level ESL university students. He found that "more proficient" listeners placed greater emphasis on "personalizing" (on elaborating from their own knowledge) and also inferred, drew conclusions, self-described, and anticipated more often than less proficient listeners. Murphy analyzed cognitive and metacognitive strategies together, identifying four patterns of strategy use to find that the more proficient listeners tended to use a strategy called "wide distribution," and the less proficient listener most frequently used the "text heavy" strategy. "Wide distribution" refers to an open and flexible use of strategies while "text heavy" refers to a dependence on the text and a consistent use of paraphrase.

Identifying different strategy patterns and associating them with different proficiency levels is potentially very useful. However, since Murphy's definition of "more proficient" appears to be circular—proficiency was distinguished by the frequency of strategies used and the sequential patterns they followed—further research is needed to validate the findings.

DeFillipis worked with second-semester elementary French students to find that the listening strategies of both skillful and unskillful listeners were more similar than dissimilar. Both groups reported using the same list of strategies, and the total number of strategies employed by each group was nearly equal. However, rank ordered in terms of total frequency, the three major listening strategies of skillful listeners were key word strategy, contextual inferencing, and grammar strategy. The three major strategies of unskillful listeners were key word strategy, translation strategy, and contextual inferencing. In addition, skillful listeners used five times more instances of visualization, three times more French-English cognates, and two times more role identification than unskilled listeners. To differentiate between skillful and unskillful listeners, DeFillipis used raw
test scores of the listening section of the MLA Cooperative Foreign Language Test. The top thirteen students were identified as skillful (range 43-37), and the bottom thirteen students were identified as unskillful (range 35-17). Although DeFillipis used a standard instrument, the rationale for selecting the point where she divided the group is not clear.

O’Malley et al. (91), working with intermediate-level ESL students at the secondary level, found significant differences between effective and ineffective listeners. Effective listeners used more self-monitoring (checking their comprehension or production as it was taking place), elaboration (relating new information to prior knowledge or to other ideas in the new information), and inferencing (using information in the text to guess meaning or to complete missing ideas). Their sample consisted of five effective and three ineffective learners. Their definition of student effectiveness was subjective, namely, teacher designations of effective and ineffective students.

Laviosa (76; 77) considered strategies that five advanced students of Italian used to solve listening comprehension problems. Each of her subjects tended to apply listening strategies in distinct ways. These distinct ways were influenced by previous knowledge about the topic, ability to perceive and decode words, identification of important meanings, and flexibility in establishing the relationship between word meaning and main topic. Hence, for example, one subject had little background knowledge about commercials and unless she heard the key word that indicated the product advertised, she would miss the whole point. On the other hand, another subject listened repeatedly until the sounds separated themselves into meaningful connections. The subject often tried to visualize a word to get the meaning. Laviosa concluded that listener’s strategies are closely tied to each other and are interrelated. She also found that there was a statistically significant correlation between listening problems experienced by the subjects and the strategies they selected to solve the problems. Unlike Murphy, perhaps due to her limited sample, she did not attempt to categorize the patterns used.

Rost and Ross considered how strategies varied by proficiency level to find that use of certain strategies correlates with proficiency. They used EFL students at three different colleges in Japan. Using a dictation test, they identified two proficiency ranges: elementary and intermediate/advanced proficiency. Beginning-level students were found to have a persistent pattern of global queries (asking for repetition, rephrasing, or simplification) across settings, speakers, and topics, whereas more advanced students used forward inference (asking a question using information already given in the story) and continuation signals (backchannel communication).

Finally, Vandergrift looked at differences in strategy use by proficiency level of successful and less successful learners. Vandergrift used students of French in their first, second, and fifth years of language study (called novice) and students in their eighth year of study (called intermediate). He found that novice listeners rely heavily on elaboration, inferencing, and transfer to build up meaning and that they overcome their limited knowledge of words by using what they know (cognates). This finding led him to suggest that the cognitive constraints of processing at the novice level are so great that there is little room for metacognitive processing strategies such as monitoring. He argues that the data indicate that listeners focus mostly on semantic cues—cognates, contextual and extralinguistic cues such as type of text, background noise, tone of voice, kinesics, and relationships between speakers.

At the intermediate level, Vandergrift found greater use of metacognitive strategies than at the novice level, but still a predominant use of cognitive strategies—especially elaboration and inferencing—to construct meaning. Students were also able to process larger chunks of information as indicated by their summarization skills. These conclusions must be verified with much larger numbers since the intermediate group consisted of seven people, and the novice group consisted of fourteen.

Vandergrift argues that successful and less successful learners use different patterns of strategies. Although Vandergrift initially hoped to distinguish strategy use by proficiency level, given the small group sizes, Vandergrift divided his subjects into those that were successful and those that were not. He found that successful learners appear to use more metacognitive strategies than unsuccessful learners. However, there is a circularity in Vandergrift’s definition of success: for him successful learners are those who report the greatest frequency, variety, and sophistication of language learning strategies. Further studies should use an independent measure of success.
Notetaking. Another approach to listening strategies is to consider whether a particular strategy is effective for a particular task. Along these lines, notetaking for academic listening has been the focus of research. Di Vesta and Gray's early studies in 1972 and 1973, considered the value of note-taking for L1 listeners. In 1988, Chaudron et al. described the complexity of determining how this strategy can be useful, especially for L2 listeners. They note that factors such as pre-training, test expectancy, quality of notes, and retrieval mode can affect the value of notetaking. In addition, they suggest that notetaking can affect the encoding process or the review/rehearsal process. For the encoding process, notetaking can increase meaningful chunking and thereby encoding of information, level of attention, effort, and assimilation of new and old information. For the review/rehearsal process, notetaking can provide mnemonics and information for reconstruction.

In a 1992 study, Chaudron et al. considered the effect of notetaking on L2 student comprehension and the extent to which quality of notes correlates with student comprehension. Working with adult ESL students who had a mean TOEFL listening score of 52.2, they found the more accurate notetaking was, the more effective the notes were in facilitating further study. Information had to be noted unambiguously to be usable. Finally, they found that how well learners understand the content of a lecture is related to the clarity of structure and presentation of that lecture. Although they found a lack of overall effect for notetaking on overall comprehension, Chaudron et al. suggested that certain information in the lectures was better retrieved with the aid of notes.

At this point, there is a fairly extensive list of the strategies L2 listeners use. However, more research is needed to clarify the relationship among proficiency level, skillfulness, and combinations of strategies used.

Strategy Training. Now that the research detailing the range of listening strategies is relatively extensive, researchers have begun to investigate how best to promote more successful listening comprehension. In the main, there are two foci: what kind of strategies should be taught and how to teach strategies.

1. What Kind of Strategies to Teach: Cognitive, Metacognitive, or Both? Several research studies in cognitive psychology report that in order to improve learning it is most effective to teach both cognitive and metacognitive strategies [see, Rubin (101) for a review of this literature]. Two studies consider the differential impact of teaching only cognitive or cognitive combined with metacognitive strategies for L2 listening comprehension.

O'Malley worked with high school ESL students at an intermediate level. There were two experimental groups (one was taught a metacognitive and a cognitive strategy and the other a cognitive strategy only) and a control group. In an analysis of daily tests, the metacognitive group outperformed the cognitive group on three of the days. However, in daily test two, the cognitive strategy group outperformed the metacognitive group. According to O'Malley, "analyses of posttest scores on listening approached but failed to reach significance . . ." (p. 141). O'Malley suggests that these inconclusive findings might have been caused by the amount of strategy training students received or by the fact that students needed the opportunity to select which strategies they would use.

Schwartz investigated the effects of strategy training on listening comprehension with intermediate-level university Spanish students. Cognitive strategies, metacognitive strategies, and attributional retraining that emphasized the link between effort, strategy use, and listening comprehension were studied. Three treatments were additively combined: Group one received only cognitive training; Group two received cognitive and metacognitive training; and Group three received cognitive, metacognitive, and attributional training. There were no statistical differences found between treatment groups in listening comprehension or in causal attributions. Schwartz suggests the lack of differences could relate to any of several conditions of the study: 1) the study was conducted outside the classroom so students did not have the opportunity to use the strategies and internalize them, 2) a standardized Spanish listening comprehension test was not available, 3) the sample size was small, and 4) too many treatment groups made it difficult to distinguish training effects.

A second area for strategy training involves teaching students to use the metacognitive strategy of self-questioning to enhance comprehension. Self-questioning functions as a form of self-testing that helps students keep a continuous check on their understanding. Exemplary of research on self-questioning is King's study (73) of lecture comprehension in which the subjects were L1 English-speaking ninth graders.
King tested whether doing cooperative self-questioning was more effective than independent use of the self-questioning strategy. Students were divided into four treatment groups: 1) self-questioning with reciprocal peer questioning, 2) self-questioning by oneself, 3) group discussion, and 4) control. Results showed highest performance for self-questioning with reciprocal peer questioning, followed first by self-questioning, then by the discussion group, and finally by the control group. The research points to the importance of both self-questioning and cooperative learning for L1 lecture comprehension.

Paulauskas' dissertation considers the effect of type of strategy training on listening comprehension: reciprocal questioning vs. teacher-centered in an university intensive ESL course with high beginner/low intermediate students. Paulauskas used a variety of audiotape texts (interviews, news broadcasts, lectures, conversations, and stories) to find that both treatment groups—the reciprocal questioning and the teacher-centered—benefited significantly from training relative to the control group. Although no significant differences were found between the reciprocal questioning and teacher-centered strategy training groups on the two post-test measures, group means revealed that the teacher-centered group benefited initially more than the reciprocal questioning group, but that over time the reciprocal questioning group marginally outperformed the control group.

Since only the dissertation proposal and a few letters were available as of this date, it is difficult to determine what other variables might have influenced these results. From these sources, Paulauskas reports no convincing evidence that group work in questioning improved listening comprehension at this level.

Thus, while the cognitive psychology literature is relatively clear on the importance of teaching both cognitive and metacognitive strategies, specific experiments on their effects in L2 listening comprehension are still lacking.

2. How to Teach Strategies. Brown and Palincsar found that performance increased as students moved from blind to controlled training.

Herron and Seay (63) worked with third-semester university students of French. Both the control and experimental groups used the French in Action video series. The experimental group—also listening to the audiotapes Champs-Élysées that resemble a radio program—was trained in listening strategies: generating and refining hypotheses, eliciting pertinent vocabulary, searching for main ideas, and key vocabulary. It is not clear from the description of the experiment whether the strategies were consistently labeled or whether students were told why strategies are useful. One part of the training was clearly blind: students were asked to fill in missing words in a song. So the study may be either blind or informed. On the final tests of listening comprehension with both the video and the audio, the experimental group did significantly better than the control group in which no strategies were taught. What needs to be sorted out is what contributed more to this improved performance: strategy training, additional exposure to audio input, or both.

Rubin et al. (103) considered the impact of three types of teaching strategies—modeling, corrective and/or supportive feedback; explanation; and assessment—on student performance. Type one (T1) was a Blind Treatment Group in which the teaching strategies (modeling and corrective/supportive feedback) were neither labeled nor the value of the strategies discussed. Type two (T2) was the Informed Treatment Group in which the teaching strategies not only included corrective and/or supportive feedback but also explanation. This latter teaching strategy developed awareness of when and how the strategy would be useful. Type three (T3), the Self-Control Treatment Group, not only included explanation and corrective and/or supportive feedback teaching strategies but also assessment. Assessment is a teaching strategy that encourages students to consider the usefulness of strategies in particular circumstances. It encourages students to monitor their comprehension in order to gain control over their use of strategies.

The experiment was conducted with second-year Spanish high school students. The research hypothesis predicted three outcomes: 1) the Self-Control Treatment Group (T3) would outperform the Blind Treatment Group (T1) and the Informed Treatment Group (T2);
2) the Informed Treatment Group (T2) would outperform the Blind Treatment Group (T1); and 3) all three treatment groups would outperform both control groups (one that watched video but had no strategy training and a second that neither watched the video nor had strategy training). The hypotheses that T3 would outperform T2 and T1 and that T2 would outperform T1 received no confirmation. This lack of confirmation may be due, at least in part, to uneven length of the lesson design and insufficient teacher training. Rubin et al. note: “just as it takes many demonstrations and a lot of practice time for students to integrate learning strategies into their learning repertoire, teachers need many opportunities to observe, practice, and integrate new teaching strategies” (p. 23).

Rubin and Thompson and Thompson and Rubin considered techniques that encourage student control in applying their learner strategies. The experiment was conducted with third-year university students of Russian. The focus was on noninteractive listening, namely, listening to a video segment, rather than on interactive listening where meaning between the participating parties can be negotiated. The teacher gave students labels for strategies, told them why the strategies were useful, and, once students were familiar with the strategies, offered them choices of which strategies to use, asking them to provide a rationale for their choices. Choices included: listening with the audio on or off; listening with the video on or off; determining if additional listenings were needed; and, if so, the goal of these listenings; and considering which strategies were most appropriate for particular kinds of texts. The study showed that training in the use of strategies improves listening proficiency and that students can be taught to control their use of listening strategies. Students regularly made choices and were able to provide a rationale for their decision: “It’s too distracting with the sound on,” or “It doesn’t help to have the sound off,” if the segment consists of talking heads (119: p. 34).

Negotiation of Comprehensible Input. Long’s research (78; 79) asserts that modification of interaction between a speaker and listener (by conversation checks, expansions, repetition, and clarification) will facilitate comprehension more than a speaker’s modification of his/her language independently of the listener. Long (78) called for further research to tease out the relative contributions of different kinds and combinations of adjustments both in cooperative interaction and speaker monologue.

Pica (94) investigated the important role of negotiation in comprehension and learning. For Pica, negotiation begins as one interlocutor signals to another that the other’s message is not clear. There is now a considerable body of literature (48; 79; 84; 95) that underscores the importance of this kind of interaction in improving comprehension. As Pica notes, the process involves two steps: 1) partners call each other’s attention to individual words or phrases that impede their access to the meaning of a message (a cognitive process of attention); and 2) partners respond to each other’s signals by repeating and reformulating all or part of their utterances (linguistic processes). Pica suggests that negotiation unites linguistic processes with cognitive processes and argues that repetitions, rephrasings, and other linguistic elaborations are crucial to the comprehension process. Obviously, these are effective only when made at the request of the listener. As discussed above in the section on redundancy, repetitions and other elaborations may make understanding difficult for students at the lowest proficiency levels.

Using NNSs in a low-intermediate ESL class, Pica et al. compared the effects of two conditions: 1) where the input is modified in some way before the learner sees or hears it, and 2) where both parties modify and restructure the interaction in order to arrive at a mutual understanding. Under condition one, subjects heard a linguistically adjusted script read by a female NS. Under condition two, the female NS read directions to the NNS and subjects were encouraged to ask for verbal help if they had any difficulty. The NS also asked whether the directions were understood or needed repeating. The authors found that overall comprehension was significantly higher for condition two than for condition one and concluded that interactional modifications of input led to significantly greater comprehension than at prior manipulations of text.

Lynch (83; 84) took this research a step further and considered two questions: 1) whether NSs modify their input and interaction more as they talk to L2 listeners at decreasing levels of proficiency, and 2) whether secondary listeners (i.e., third parties to a conversation) find it easier to understand discourse in which the NS primarily modifies interaction rather than input. Subjects for the first question included native Edinburgh ESL teachers working with four
levels of proficiency: native, advanced, intermediate, and elementary. Lynch found that NSs do modify their speech differently according to their assessment of the proficiency level of the listener. They make three types of adjustment: input, interaction, and information choice.

To investigate the second question concerning secondary listeners, Lynch used elementary and post-elementary ESL adults. He found that there were two kinds of native speakers: a block type, where speakers produced a monologue having modified their own output without consulting the listener, and a string type, where speakers took short turns followed by comments or questions from the listener. The narrators appeared to produce spontaneously the two types of modification experimentally controlled by Pica et al. Lynch found that the strongest influence on comprehensibility was the extent to which speakers and listeners engaged in cooperative modification of interaction.

However, Ehrlich et al. questioned whether the negotiation of meaning necessarily guarantees that input is comprehensible. They argued that success or failure of negotiation is directly related to the point in the discourse where the negotiation occurred. Ehrlich et al. worked with adult high-intermediate and low-intermediate students of ESL who talked with native English speakers. They found, first of all, that there were differences among NSs in the amount of detail they provided. The authors divided native speakers into two categories: skeletonizers (speakers who provide only the bare events of a narrative) and embroiderers (speakers who use a great deal of expansion and embellishment in a narrative). They found that when communication breakdowns occur, skeletonizers tended to abandon negotiation of meaning more readily than embroiderers. They note that the embroidering strategy can lead to confused problematic discourse in NS-NNS interactions. On the other hand, the skeletonizing strategy may result in more comprehensible input. The authors note that "this strategy is maintained by negotiation abandonments at deeply embedded points in the discourse and not by persistently engaging in negotiations of meaning" (emphasis mine) (p. 411).

This is obviously a very productive area of research that needs more studies to check reliability and to define more clearly the kinds of negotiation that are most productive.

SUMMARY

This paper has reviewed over 130 studies of which over 115 report directly on research in listening comprehension. There are some promising research directions:

1. For both bottom-up processing and top-down processing, we need to prioritize the elements that affect listening most for each proficiency level.

2. For strategy training, we need more work identifying problems and strategies used in languages with varying kinds of transformations, and on the effect of these transformations on the listening process for each proficiency level. We also need more research on what kinds of training are most effective for which kind of students.

3. For the definition of comprehensible input, we need more research on the kinds of negotiations that enhance understanding.

Within each of these areas, research that takes the following factors into account will prove most useful for teaching: the research should be carried out over fairly long periods of time and with a variety of languages of differing degrees of difficulty. The research should use authentic materials; provide a clear definition of language proficiency; test for proficiency, not achievement (so that transferability is clear); and bear in mind differences related to proficiency level. Much more research in listening comprehension is sorely needed; these are some clear directions upon which it can be built.

ACKNOWLEDGEMENTS

This review was prepared while the author was Senior Research Scholar (1990-93) at George Washington University under a grant from the National Security Agency (Project No. MDA904-91-H-0058). This manuscript is published with the understanding that the United States Government is authorized to reproduce and distribute reprints. I wish to express my thanks to research assistants, Vanessa Bittner and Michael Wilson, for their assistance and dedication. Finally, I wish to acknowledge the valuable comments and suggestions of Sally Sieloff Magnan, Eileen K. Blau, Anna U. Chamot, Bernadette Dejean de la Bâtie, Tony Lynch, Randall J. Lund, Flavia Laviosa, and Robert Vanderplank. Any errors are my responsibility alone.


79. ———, “Native Speaker/Non-native Speaker Com-
106. ———. Personal communication, 1993.
Major focus: a refereed publication devoted to questions and concerns about the learning and teaching of foreign and second languages; publishes articles, research studies, editorials, reports, book reviews, and professional news and announcements pertaining to modern languages, including TESL.

Circulation: approximately 7000.

Frequency of Publication: four issues per year (approximately 150 pages each).

Evaluation: manuscripts refereed by at least two specialist readers.

Time from Submission to Decision: normally within ninety days; time from acceptance to publication usually no longer than one year.

Acceptance Rate: ten to thirteen percent of manuscripts submitted are accepted.

Articles Published per Volume: approximately twenty-five to thirty articles plus editorials and comments in Readers' Forum.

Book Reviews per Volume: approximately 200.

Ownership: The Modern Language Journal is owned and copyrighted by the National Federation of Modern Language Teachers Associations and published by it at the University of Wisconsin Press. The MLJ has no legal relationship to the Modern Language Association of America (MLA).

Payment: authors of articles receive two complimentary copies of the issue in which their work is published; authors of book reviews receive one copy.

Offprints: available at cost, but must be purchased directly from the reprint service at the time author reads galley proofs.

Language of Publications: English.

Manuscript Format: typewritten, double-spaced throughout (including bibliography, any notes, citations, figures, and tables); leave a two-inch left margin, one-inch elsewhere; submit original and three clear copies (only one will be returned).

Manuscript Length: twenty pages preferred (excluding bibliography, tables, notes); longer articles acceptable, depending on merit.


Multiple Submissions: manuscripts submitted simultaneously for publication elsewhere not considered; author(s) must inform editor at time of submission of similar/related versions of manuscript that have appeared or are being considered elsewhere.

Special Requirements: 1) indicate full name of institution where (each) author is employed; 2) include 4 copies of your submission, a self-addressed 9 x 12 manila envelope, and US $5.00 in loose postage (non-US submissions $10.00 in a check drawn on a bank with US affiliation or twenty-five international postal coupons; 3) prepare a cover sheet containing the title of the manuscript, separate short title [see 4 below] for identification, name, address, office and home telephone numbers, Fax number, and Email address of author(s); 4) select one or two key words from the title of the manuscript to function as identification markers; type those works at the top right of all pages of the manuscript followed by the appropriate page number; 5) refer to your own previous publications in the third person—not "as I said in . . .", but "as Jane Smith noted . . ."; 6) refer to an institution where research was conducted or where the author teaches as "institution X" during the refereeing process; 7) number consecutively each entry in the bibliography, indenting the second and subsequent line of entry four spaces.