What makes listening difficult?

Factors affecting second language listening comprehension

Amber Bloomfield, PhD
Sarah C. Wayland, PhD
Elizabeth Rhoades, MA
Allison Blodgett, PhD
Jared Linck, PhD
Steven Ross, PhD

PURPOSE—To establish what is currently known about factors that affect foreign language listening comprehension, with a focus on characteristics of the listener, passage, and testing conditions.

CONCLUSIONS—Research on second language (L2) listening comprehension strongly supports the importance of a number of factors, for example, a listener's working memory capacity and the number of ideas in a passage. Much of the research, however, reports weak or inconclusive results, leaving many factors and complex interactions among factors unresolved and in need of further investigation.

RELEVANCE—Identifying the factors that affect L2 listening comprehension will help Defense Language Institute Proficiency Test (DLPT) designers anticipate how qualities of selected authentic materials will impact listening comprehension.

TTO 81434 E.3.1 | CDRL A017 | DID DI-MISC 80508B | Contract No. H98230-07-D-0175

Executive Summary

PURPOSE

The U.S. Government administers the Defense Language Proficiency Test (DLPT) to military linguists and other government personnel to assess their listening and reading comprehension in a number of foreign languages, including critical languages such as Mandarin, Modern Standard Arabic, Egyptian Arabic, and Persian Farsi. The DLPT is updated every 10 to 15 years, and the most recent transition—from DLPT IV to DLPT5—included a greater emphasis on testing listening comprehension with authentic materials. In turn, this has led to a growing interest in the factors that make second language (L2) listening difficult.

To examine these factors, CASL reviewed the current scientific literature and summarized the characteristics of listeners, passages, and testing conditions. The review targeted features of particular interest to stakeholders at the Defense Language Institute (DLI). The long-term goal of the project is to support the selection of authentic listening materials that accurately reflect different proficiency levels.

CONCLUSIONS

Although the available research on L2 listening comprehension is limited, CASL's literature review identified several factors that affect listening comprehension. These factors are summarized below and in Tables 1, 2, and 3.

1 Characteristics of the listener

Understanding a foreign language taps several general cognitive abilities. For example, listeners with greater **working memory capacity**—that is, those who are most efficient at attending to, temporarily storing, and processing incoming information—understand more of what they hear when they are listening to their non-native language.[1] Further, listeners who effectively use **metacognitive strategies**—that is, those who are aware of and use effective strategies, such as avoiding mental translation—demonstrate better L2 listening comprehension.[2]

In addition to these general cognitive abilities, a number of factors pertaining to **experience** with the L2 influence listening skill. These factors include the amount of prior exposure to the language; familiarity with and an ability to understand the non-native language's phonology; vocabulary size; and background knowledge about the topic, text, structure, schema, and culture.

Familiarity with the L2 changes the extent to which the L2 listener uses top-down or bottom-up strategies in listening. For example, expert listeners use both types of strategies: They are able to accurately make sense of the speech signal (bottom-up information)[3] and integrate this information with

Table 1. Effects of listener characteristics on L2 listening comprehension

Working memory	Greater working memory capacity correlates with better comprehension.
Metacognitive strategies	The use of metacognitive strategies improves listening comprehension.
L2 proficiency and experience	As proficiency increases, the listener's ability to correctly use bottom-up information (including deciphering the L2 phonology and vocabulary) improves. Background knowledge enables the use of top-down strategies to compensate for mishearing or encountering unfamiliar words, which can improve comprehension.
Anxiety	Anxiety negatively impacts L2 listening comprehension.

background knowledge (top-down information).[4] By contrast, non-expert listeners attempt, often unsuccessfully, to use background knowledge to compensate for failure to understand speech sounds.[5,6]

Vocabulary size also impacts the extent to which L2 listeners will comprehend a spoken message, but this effect of vocabulary may be related to other more general qualities of listeners, such as their experience with the L2.[7]

Listeners' **anxiety** can also impact their ability to understand what has been said. If a listener is anxious or in some other way distracted and unable to pay attention, it will be more difficult to accurately determine what was said.

2 Characteristics of the passage

Studies directly examining the effects of passage **length** on L2 listening comprehension find little evidence that this factor alone affects comprehension difficulty.[8] However, these studies have often explored a limited range of lengths[9] or have confounded length with other factors.[10] Information density (i.e., the number of ideas in the passage) and redundancy (i.e., the extent to which passage information is repeated), which are correlated with passage length, have more consistent effects. Information density increases listening difficulty consistently across studies,[11] even when this factor is measured using different methods.[12] Redundancy improves comprehension, but the effect depends both on the proficiency of the listener and the type of redundancy (e.g., exact repetition, paraphrase).[13,14]

Passage **complexity** also may affect L2 listening comprehension. A higher number of negatives[15] and the presence of infrequent vocabulary[16] may increase difficulty. Further, simplifying the syntax of a passage does not consistently aid L2 listening comprehension.[17] Indirect passages, which include more implied information, can also be more difficult for L2 listeners to comprehend,[18] and the ability to cope with this type of information improves with L2 proficiency.[19] Concreteness, or the extent to which a passage refers to concrete objects or entities, has rarely been explored as a factor affecting L2 listening comprehension,[20] though it does affect L2 reading comprehension.[21] Finally, L2 listeners have a harder time understanding passages

Table 2. Effects of passage characteristics on L2 listening comprehension

Length	• **Overall length**—Longer length increases listening difficulty, but the effect is weak and inconsistent across studies. • **Information density**—A large number of ideas in a passage has a negative effect on listening comprehension. • **Redundancy**—Repetition of information consistently improves comprehension, but whether the listener benefits depends on the type of redundancy (e.g., exact repetition, paraphrase) and listener proficiency.
Complexity	• **Syntactic features**—Simplifying sentence structure does not consistently improve comprehension. Negatives and infrequent vocabulary have a detrimental impact. • **Directness and concreteness**—Passages with implied meaning can be more difficult to understand. Research in reading comprehension suggests that texts with more concrete objects or entities may be easier to comprehend, but little research has examined this factor in L2 listening. • **Pragmatic information**—The inclusion of L2 pragmatic constructs such as idioms and culturally specific vocabulary decreases comprehension.
Organization	• **Orality**—Passages with higher orality—that is, ones more like unscripted conversations—have greater redundancy, more disfluencies, and simpler syntax. They are easier to understand than passages with less orality. • **Coherence**—Overall coherence of a passage seems to have little effect, but only a few studies have examined its effects. Further, coherence may be difficult to define and measure objectively. • **Discourse markers**—Words and phrases that signal the relationship between adjacent propositions and the overall structure of the passage improve comprehension. However, this effect depends on the type of marker. • **Position of relevant information**—Information is most easily recalled when it occurs near the beginning or at the end of a passage.
Auditory features	• **Speaker accent**—Familiar accents are easier to understand than unfamiliar accents. • **Hesitations and pauses**—Disfluencies, like hesitations and pauses, generally aid comprehension, especially for more proficient listeners. • **Noise and distortion**—The presence of noise or distortion in the speech signal interferes with comprehension. • **Speech rate**—How quickly someone talks can hurt comprehension, but slower speech rates do not necessarily help. L2 listeners may mistakenly attribute difficulties caused by other factors to a too-fast speech rate.

that contain culturally specific words[22] and idioms.[23]

Several dimensions of passage **organization** affect comprehension, including orality—that is, the extent to which passages are similar to spoken language. Passages with higher orality have simpler syntax, greater redundancy, more hesitation markers (e.g.,"um" and "ah"), and more pauses;[24] these types of passages are easier for L2 listeners to comprehend.[25]

Coherence is an additional dimension that can be characterized as the appearance of logicality[26] in a passage or the extent to which ideas introduced at the beginning of a passage are carried through until the end.[27] Research examining the effect of coherence on L2 listening comprehension is sparse and generally inconclusive, and there are potential issues with defining this factor in a way that can be measured objectively.[28]

Discourse markers, which help to establish relationships between adjacent utterances (e.g., "yet"—a micro-marker) and the overall structure of the passage (e.g., "the first point is"—a macro-marker), improve L2 listening comprehension.[29] That said, there is some evidence that macro-markers make a passage more comprehensible, while micro-markers do not consistently help L2 listeners.[30]

Another feature of organization that affects listening comprehension is the position of the information that is necessary to answer a test question. Information toward the beginning or at the end of a passage is more easily recalled than information from the middle of the passage.[31]

Several auditory features impact L2 comprehension, including the familiarity of the speaker's accent. Accent familiarity affects passage comprehension for both first language (L1) and L2 listeners, though more so for L2 listeners.[32] The level of experience with an accent required to completely remove the detrimental effect seems to be extensive,[33] though L1 listeners show partial adaptation after very brief exposure.[34]

Disfluencies such as hesitations and pauses aid L2 listening comprehension in most studies examining these factors.[35] Pauses give L2 listeners additional processing time and act as cues about the speaker's upcoming utterances.[36] Further, some evidence shows that disfluencies that occur in the L2 must be learned before filled pauses like "um" can be useful in listening comprehension, so listeners of different proficiency levels may benefit differentially from filled pauses.[37]

Noise or distortion in the audio signal interferes with listening comprehension for L1[38] and L2[39] listeners, though the effect is larger for L2 listeners. For L2 listeners, noise that most closely resembles the signal (e.g., babble noise when listening to speech) presents the greatest challenge for listeners. Other types of distortion (e.g., white noise, filtering out high-frequency information [as often happens in telephone calls], and time-compression) differentially affect the perception and processing of speech.[40]

Faster speech rates, whether computer-manipulated or naturally produced, tend to have a negative impact on the comprehension of L2 listeners,[41] even advanced listeners.[42] However, some research suggests that L2 listeners will tend to feel the speech rate of aural materials is too fast when comprehension difficulty is caused by factors unrelated to speech rate.[43] Further, while a faster speech rate may be detrimental to L2 listening comprehension, a slower speech rate is not necessarily beneficial[44] or even preferred by listeners.[45]

3 Characteristics of the testing conditions

Research shows that imposing **time limits** makes cognitive tasks more difficult.[46] However, very little research has directly examined the effect of time limits on performance in L2 listening comprehension testing, though imposing time limits on any test is likely to affect response behavior and test-taking strategies.[47]

Performance on L2 listening tests may be unaffected by increasing time limits if the examinee has the option of pausing or replaying the passage.[48] In general, **listening to a passage multiple times** improves comprehension.[49] The improvements may be greater for lower-proficiency listeners than higher-proficiency listeners,[50] but only if they have the lexical and syntactic knowledge needed to comprehend the passage.[51] When listeners are given control over the number and timing of hearings of the passage, they will choose to replay the passage more often when the passage is difficult due to factors like rate of presentation.[52]

For L1 listeners, **note-taking** is an effortful activity that introduces time pressure due to the difference between speaking rate and writing rate,[53] and for L2 listeners, note-taking is even more cognitively effortful.[54] Note-taking can be damaging to L2 listening comprehension when listeners are urged to take notes.[55] Other factors that impact difficulty (e.g., speech rate)

Table 3. Effects of testing conditions on L2 listening comprehension

Time limits	Time pressure generally makes cognitive tasks more difficult, but time limit effects on L2 listening test performance are largely unexplored. Increasing response time may not improve comprehension unless examinees can pause or replay the passage.
Multiple hearings	Comprehension improves with additional hearings of a passage if the listener has the L2 knowledge to understand the information. Examinees are more likely to replay a passage if other factors present difficulty.
Note-taking	Note-taking is particularly effortful when listening to L2 passages. It may benefit L2 listening comprehension, but only if the participant is able to make good decisions about when to take notes.

affect whether taking notes in the L2 benefits comprehension and recall.[56] Overall, the literature indicates that if L2 listeners are able to successfully employ a metacognitive strategy for determining when to take notes and when not to take notes, note-taking can benefit listening comprehension.

The effects of item type (e.g., multiple-choice, free response) and the task associated with listening (e.g., comprehension versus transcription) were not covered in the literature review.

RELEVANCE

The current Defense Language Proficiency Test (DLPT5) includes a greater emphasis on authentic materials than prior versions of the test. This review of the scientific literature suggests that during test development and the selection of authentic spoken passages, it is possible to anticipate some of the ways in which passage, listener, and testing condition factors will influence L2 listening comprehension scores. The report provides an initial framework for assessing features of authentic spoken passages in relation to their impact on L2 listening comprehension. ■

ENDNOTES

For the full citations, see the technical details section of this report.

1 Harrington and Sawyer (1992)
2 Vandergrift et al. (2006)
3 Tsui and Fullilove (1998)
4 Field (2004)
5 Goh (2000)
6 Tyler (2001)
7 Nation (2001)
8 e.g., Rupp, Garcia, & Jamieson (2001)
9 e.g., Nissan, DeVincenzi, & Tang (1996)
10 Moyer (2006)
11 Rupp et al. (2001)
12 Rupp et al. (2001)
13 Chang & Dunkel (1992)
14 Chaudron (1983)
15 Nissan et al. (1996); Kostin (2004)
16 Kostin (2004)
17 e.g., Pica, Young, & Doughty (1987)
18 Taguchi (2008)
19 Taguchi (2005)
20 Nissan et al. (1996)
21 Freedle & Kostin (1992, 1993)
22 Sasaki (2000)
23 Kostin (2004)
24 Tannen (1982)
25 e.g., Kiany & Jalali (2006)
26 Odlin (1989)
27 Freedle & Kostin (1992)
28 Ying-hui (2006)
29 e.g., Jung (2003)
30 Chaudron & Richards (1986)
31 Freedle & Kostin (1996, 1999)
32 Major et al. (2005)
33 Weil (2003)
34 Floccia et al. (2009)
35 e.g., Blau (1991)
36 Watanabe et al. (2008)
37 Blau (1991); Watanabe et al. (2008)
38 e.g., Adank et al. (2009)
39 e.g., Lecumberri & Cooke (2006)
40 Aydelott & Bates (2004)
41 Griffiths (1990, 1992)
42 Rosenhouse, Haik, & Kishon-Rabin (2006)
43 Moore et al. (2007)
44 Griffiths (1990)
45 Derwing & Munro (2001)
46 Siemer & Reisenzen (1998)
47 Bejar (1985); Verhelst, Verstralen, & Jansen (1997)
48 Buck (2001)
49 e.g., Cervantes & Gainer (1992)
50 Field (2008)
51 Chang & Read (2006)
52 Zhao (1997)
53 Piolat, Olive, & Kellogg (2005)
54 Barbier & Piolat (2005)
55 Hale & Courtney (1994)
56 Carrell, Dunkel, & Mollaun (2002); Lin (2004)

Corresponding Author and Reprints: Sarah Wayland, PhD, University of Maryland Center for Advanced Study of Language, (301) 226-8938, swayland@casl.umd.edu, www.casl.umd.edu.

Funding/Support: This material is based upon work supported, in whole or in part, with funding from the United States Government. Any opinions, findings and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the University of Maryland, College Park and/or any agency or entity of the United States Government. Nothing in this report is intended to be and shall not be treated or construed as an endorsement or recommendation by the University of Maryland, United States Government, or the authors of the product, process, or service that is the subject of this report. No one may use any information contained or based on this report in advertisements or promotional materials related to any company product, process, or service or in support of other commercial purposes. The Contracting Officer's Representative for this project is Shannon Salyer, PhD, Senior Research Scientist—Research and Analysis Division, Defense Language Institute Foreign Language Center, (831) 242-6670, shannon.salyer@conus.army.mil.

Acknowledgment: We wish to thank the following people from the DLIFLC: Donald C. Fischer, PhD, Provost; John A. Lett Jr., PhD, Dean—Research and Analysis; Shannon Salyer, PhD, Senior Research Scientist—Research and Analysis; Mika Hoffman, PhD, Dean—Test Development; James Dirgin, Director—Test Review and Education; Ruth Mehr, Test Developer; Gerd Brendel—Test Development and Education Specialist; Nhon Le, Senior Project Manager—Test Development; and William Eilfort, Project Manager—Test Development. We also wish to thank Barbara Forsyth, PhD, Area Director for Performance & Assessment at CASL, for her calm guidance, as well as her careful and thoughtful review of this document. Catherine J. Doughty, PhD, Area Director for Second Language Acquisition at CASL, was also invaluable, providing information about issues pertaining to second language acquisition in addition to reviewing this document. Last but not least, we are grateful to Michael Bunting, PhD, Associate Research Scientist at CASL, for his insightful comments on an earlier version of this review.

UNIVERSITY OF MARYLAND
CASL
CENTER FOR ADVANCED STUDY OF LANGUAGE

Technical Details

TTO 81434 E.3.1 ■ CDRL A017 ■ DID DI-MISC 80508B ■ Contract No. H98230-07-D-0175

What makes listening difficult?

Factors affecting second language comprehension

Amber Bloomfield, PhD, Sarah C. Wayland, PhD, Elizabeth Rhoades, MA, Allison Blodgett, PhD, Jared Linck, PhD, Steven Ross, PhD

Listening is an important skill for students of a second language (L2) to master as part of learning to effectively communicate in that language (Dunkel, 1991; Lund, 1991; Rost, 2002; Rubin, 1994). It is tested directly in one component of the Defense Language Proficiency Test (DLPT), which the U.S. Government administers to military linguists and other government personnel. The current report summarizes the scientific literature on three main characteristics argued to affect L2 listening comprehension: (1) characteristics of the listener, (2) characteristics of the passage, and (3) characteristics of the test-taking conditions. For each characteristic, the review targets particular factors of interest established in discussion with stakeholders at the Defense Language Institute Foreign Language Center (DLIFLC).

In most cases, only a handful of studies have explored the role of these factors in second language listening. As a result, this review distinguishes between areas that have received enough attention to justify firm conclusions about their role in L2 listening comprehension, and those that have not. The review further highlights connections between the described factors. Ultimately, the research literature suggests that during test development and the selection of spoken passages, it is possible to anticipate some of the ways in which listener, passage, and test-taking condition factors are likely to influence L2 listening comprehension.

Main characteristics	Factors of interest
Listener	Working memory, L2 proficiency, exposure to L2, metacognitive strategies, anxiety
Passage	Authenticity, length, complexity, type and organization, auditory features
Test-taking conditions	Time limits, number and control over hearings, note-taking

TABLE OF CONTENTS

INTRODUCTION

What makes listening difficult in a second language? Of the two receptive language skills, listening to a foreign language is understudied relative to reading (Johnston & Doughty, 2006; Lund, 1991; Osada, 2004; Rubin, 1994; Shohamy & Inbar, 1991; Thompson, 1995). Most research concerned with the evaluation of second language (L2) comprehension has focused on reading rather than listening because the process of reading is more easily observed and manipulated (Osada, 2004). That said, research findings regarding reading comprehension often fail to map fully onto the processes involved in listening comprehension (Schmidt-Rinehart, 1994). For example, listeners have a worse memory for spoken information than readers do for written information, with proportionally more details recalled by readers and proportionally more main ideas by listeners (Lund, 1991). Further, characteristics of the listener influence listening performance differently than they influence reading performance (Park, 2004). In addition, there are factors that are important to listening that are not relevant for reading, like coping with a fast speech rate and disfluencies.

In short, it is important to recognize that listening is a distinct skill from reading. Listening involves real-time processing, generally without the option of going back to earlier sections of the passage the listener may have missed (Buck, 2001; Flowerdew, 1994). While slow readers can alter their reading speed without damaging comprehension, slow listeners may well miss information that cannot be recovered. Once the information is lost, it can be difficult to understand the rest of the passage (Buck, 2001). Further, while most reading involves complete control of the rate at which text is received (i.e., readers may read a text quickly or slowly, as they wish, unless some time constraint is applied), control over the speed of delivery for listeners varies much more widely (Osada, 2004). In a conversation, listeners may be able to exert some control over the speech rate of their interlocutor, while listening to a radio program provides no opportunity for control over the speed of delivery, and attending a professional lecture does so only with certain restrictions and considerable effort.

Readers typically…	Listeners typically…
…reread text as they please	…hear speech that disappears quickly
…control their own reading rate	…adapt to how quickly the speaker is talking
…encounter spaces between words	…have to figure out where words start and stop
…read text that obeys grammar and spelling rules	…encounter disfluencies, false starts, and pronunciations that may or may not resemble written forms

Aurally presented material also involves a number of phonological and lexical features that are not present in written material (Flowerdew, 1994). In some written languages, though not all, word boundaries are marked with clearly visible spaces. In spoken language these boundaries must be inferred from a variety of lexical and phonological cues (Cutler, Dahan, & van Donselaar, 1997; Leeser, 2004; Weber & Cutler, 2006). Also, written text tends to involve more planning and editing before it reaches an audience, whereas spoken information is often spontaneous (Richards, 1983). Speech, particularly in spontaneous conversations, contains irregular pauses, false starts, disfluencies such as *um*, and intonation patterns that may affect comprehension (Gilmore, 2007; Shohamy & Inbar, 1991). The pronunciation of words may also differ greatly from the way they appear in print and may be

affected by the words with which they are presented (e.g., assimilation results in the word *ten* being pronounced *tem* in the phrase *ten bikes*, Crystal, 2003; reductions result in the phrase *I'm gonna go* instead of *I am going to go*, Ito, 2001). Because of factors like reduction, spoken language may contain less lexical information than printed language, and these reduced forms have been shown to interfere with listening comprehension for non-native listeners (Ito, 2001). Differences between spoken and written texts may lead to a partial dissociation between reading skills and listening skills in the L2 learner (Lund, 1991; Song, 2008).

Given the differences between reading and listening, the current literature review focuses primarily on research in L2 listening comprehension. Separate sections cover Characteristics of the Listener, Characteristics of the Passage, and Characteristics of the Testing Conditions, in that order.

INTERPRETING THIS TARGETED LITERATURE REVIEW

While reading this literature review, there are several important issues to keep in mind. First, because the literature review is *targeted*, not all possible factors affecting L2 listening comprehension are addressed here. For example, while we understand the importance of test question construction for assessing comprehension, we do not cover that topic in this review. Decisions about the factors to review were made through discussions between the authors at CASL and DLIFLC stakeholders. The absence of any particular factor in this literature review should not be construed as a dismissal of its importance by the CASL authors or DLIFLC.

In addition, this review is intended to present a language-general and high-level discussion of the impact of various factors on L2 listening comprehension. Thus, there are several caveats to consider when interpreting the research reviewed here:

1. The studies reviewed in each section often used very different types of tasks to measure *comprehension*. That is, while some studies measured comprehension using multiple-choice questions targeting the main topic or details of the passage (analogous to the method used in the DLPT), other studies evaluate comprehension through the use of assessments like the cloze test[1], recognition of statements from the passage, analysis of the listeners' notes, in addition to many other dependent measures. A detailed investigation of the influence of task type on L2 listening comprehension and its interactions with other factors was outside the scope of this review. However, because the purpose of this literature review is to discuss factors that impact L2 listening difficulty in general, and the comprehension measures used in the studies reviewed here are diverse, the reader may well find that not all results apply to a comprehension task like that presented by the DLPT.

2. One of the most eclectic collections of measures in the studies reviewed here are those assessing L2 listening proficiency. Some studies defined proficiency by a score on an accepted test, such as the TOEFL, while others used experimenter-designed proficiency tests or simply the language course level of the student participant. Because of

[1] A cloze test presents the test-taker with a short text with blanks where some of the words should be and asks them to fill in the blanks.

these differences, it is very difficult to make generalizations about the effect of different levels of a factor (e.g., different speech rates) on the listening comprehension of listeners with varying L2 proficiency. It is also extremely difficult to make conclusions about what the literature indicates with regards to what factors will affect a listener at a particular *Interagency Language Roundtable* (ILR) level. The review notes any finding where a factor affected listening comprehension differently for listeners of differing proficiency levels, but these results must be interpreted with caution due to the aforementioned issues.

3. The level of difficulty of materials used in the studies varied, and factors affecting the difficulty level of a passage were often not fully described (e.g., the presence of discourse markers in a passage might be described in detail, while other qualities, such as vocabulary level, are not fully explicated). While this is not necessarily an issue for the investigation of the factor(s) of interest in the study, it does create some difficulty in interpreting the impact of the factor(s) on listeners with differing L2 proficiency. For instance, if introducing paraphrasing (a type of redundancy) into a passage has no effect on comprehension for lower-proficiency listeners, but does improve comprehension for higher-proficiency listeners, this may be because paraphrasing has no effect on the comprehension of lower-proficiency listeners. Alternatively, the lack of an effect of paraphrasing for lower-proficiency listeners could be due to the presence of other, uncontrolled qualities of the passage that increase its difficulty beyond the language abilities of these listeners (e.g., a fast speech rate or infrequent vocabulary). This example shows why findings indicating an interaction between proficiency level and another factor affecting listening difficulty must be interpreted carefully.

4. Finally, while the review is intended to provide a language-general discussion of how factors influence L2 listening comprehension, there are several factors reviewed here whose effects are very likely to differ across languages. In most cases, the factors reviewed here have been explored in very few languages, so there is no empirical evidence for such differences. Nonetheless, it is important to keep in mind that the predominant effect of a factor on L2 listening comprehension, as described in this review, may not hold for all languages. Further research is needed to address this issue.

CHARACTERISTICS OF THE LISTENER

While a large number of individual difference factors may affect both L2 listening comprehension and general test performance, this review covers only the subset of factors deemed by CASL and the stakeholders as relevant to the question of difficulty of listening passages on the DLPT5. The factors discussed here include working memory capacity, proficiency and experience with the L2, the use of metacognitive strategies, and anxiety.

Understanding a foreign language taps general cognitive abilities, as well as knowledge acquired as the result of dedicated study and exposure to the non-native language (L2). General cognitive abilities known to affect L2 comprehension include working memory and metacognitive strategies such as planning, prediction, monitoring, evaluation, mental translation, personal knowledge, and directed attention. Working memory is correlated with, and maybe even central to language learning aptitude (Bowles, Linck, Koeth, Mislevy, Campbell, Annis, Jackson, Doughty & Bunting, 2009); if we hope to measure language proficiency as separate from aptitude, it may be important to understand the interaction between working memory and passage difficulty. In addition, researchers generally assume that listeners use metacognitive strategies when listening to their non-native language; Vandergrift (2006) offered this explanation when he found that *native* language listening ability accounted for 14% of the variance in L2 listening ability, with L2 proficiency accounting for 25% of the variance. In addition to the metacognitive strategies identified by Vandergrift and his colleagues, Rubin and Roberts (1987) found that L2 listening comprehension scores were influenced by exposure to literary works in the L2, critical thinking, and recall abilities.

In addition to general cognitive abilities, a number of linguistic factors influence listening skill. These factors include (a) familiarity with and ability to understand the phonology of the non-native language, (b) vocabulary size, and (c) background knowledge about the topic, text, structure, schema,[2] and culture. The mental state of listeners can also have an impact on their ability to understand what has been said. That is, if a listener is anxious or in some other way distracted and unable to pay attention, it will be more difficult to process what was said. These factors are described in more detail in the sections below.

Working memory

Working memory refers to a cognitive system that is crucial to the processing, storage, and retrieval of information in memory (e.g., Baddeley & Hitch, 1974). Working memory is thought to comprise a storage component and an attentional control component known as the central executive. According to Baddeley's original model of working memory

Working Memory: A set of cognitive processes that all listeners use—with varying degrees of efficiency—as they attend to, temporarily store, and process incoming speech in L1 or L2.

[2] In this context, *schema* refers to a cognitive framework or concept that helps the listener to organize and interpret information. Schemas can be useful, because they enable the listener to interpret a vast amount of information within the context of an existing framework.

(Baddeley & Hitch, 1974), the short-term storage component is subdivided into the phonological loop (for auditory information) and the visuospatial sketchpad (for visual and spatial information), which serve as buffers where modality-specific information is maintained in a highly active and readily accessible state. Any mental manipulation or processing of information stored in these buffers is overseen by the central executive. Thus, the central executive is believed to play an important role in guiding processing within working memory.

More recent theoretical models of working memory emphasize the role of the central executive in exerting cognitive control over the contents of working memory, and also as the primary determiner of individual differences in working memory (e.g., Engle, 2002). Engle's controlled attention view of working memory places a strong emphasis on the role of the central executive in efficiently managing available attentional resources. These contemporary models are motivated in part by a body of research demonstrating that working memory is related to an individual's ability to successfully and efficiently resolve conflict or ambiguity, such as when recovering from inappropriate interpretations of ambiguous phrases in so-called *garden path* sentences (e.g., Novick, Trueswell, & Thompson-Schill, 2005).[3] As briefly reviewed below, there is a rich literature demonstrating the relationship between working memory and L1 comprehension (e.g., Daneman & Merikle, 1996).

Although reading and listening occur in different modalities, the underlying cognitive tasks required for their performance involve some of the same processes: comprehending an incoming stream of information, integrating this information online, constructing a mental representation of the syntactic structure and semantic content (i.e., meaning), resolving any ambiguities or overcoming interference from conflicting representations, etc. Because most of the research on working memory and language has been conducted with written materials, this portion of the review will describe both reading and listening research to highlight the available data on working memory's role in comprehension, while drawing particular attention to data from listening tasks where available. Note that when studying comprehension of spoken passages, one cannot return to earlier points in the input stream as is possible during reading comprehension, which may impose additional load on working memory.[4]

Further, because L1 and L2 comprehension engage an overlapping set of cognitive processes, the next section briefly reviews the literature on working memory and L1 comprehension before focusing specifically on the available research on working memory and L2 comprehension. These parallel literatures provide an overall framework for understanding the role of working memory in L2 listening comprehension.

---

[3] *John knew the answer was wrong* is an example of a garden path sentence. Readers initially misinterpret *the answer* as the direct object of *knew* (as in *John knew the answer*). They must reinterpret the noun phrase as the subject of the embedded sentence *the answer was wrong*.

[4] Of course, if the listener is engaged in a dialogue with the speaker (a.k.a., participatory listening), the listener can ask the speaker to repeat any information the listener missed.

Working memory and L1 comprehension

In a seminal study, Daneman and Carpenter (1980) introduced a new measure of working memory—the reading span test—that stressed both the storage and processing components of working memory. They found that individual differences in this processing-plus-storage measure of working memory were strongly related to measures of L1 reading and listening comprehension (correlations ranged from .42 to .90, with an average correlation of .66). Their study inspired a surge in research on working memory and L1 comprehension, much of which has supported the claim that working memory is related to L1 comprehension (e.g., Just & Carpenter, 1992; Engle, Cantor, & Carullo, 1992). In a meta-analysis of this literature, Daneman and Merikle (1996) examined the research findings from 77 studies combining data from over 6,000 participants. They found that, when using working memory measures that involved both processing and storage, the estimated population correlation between working memory and L1 comprehension ranged between .30 and .52. These correlations were notably stronger than the corresponding correlations between storage-only measures of working memory (range = .14 to .40), indicating that complex, processing-plus-storage measures of working memory are better predictors of L1 comprehension.[5]

A survey of the literature highlights the fact that there are strikingly few studies of the relationship between working memory and L1 *listening* comprehension in adults. In one study, Baddeley and Hitch (1974) imposed a working memory load on participants by visually presenting sequences of three (low load) or six (high load) digits while having them listen to a prose passage. Participants were required to recall the digit sequence before answering questions about the passage. The results indicated that taxing working memory resources by imposing the high load, but not the low load, significantly impaired comprehension of the prose passage relative to a control condition. Other evidence of the role of working memory in listening comprehension comes from an investigation of the *cocktail party phenomenon*. This phenomenon refers to a situation in which an individual selectively focuses attention on one stream of auditory input while ignoring other input (i.e., the unattended stream), but when meaningful input (e.g., the listener's name) is presented in the unattended stream, the individual's attention may be captured (Moray, 1959). In a contemporary investigation of this classic phenomenon, Conway, Cowan, and Bunting (2001) found that individuals with relatively low working memory capacities more frequently experience the cocktail party phenomenon than individuals with higher working memory capacities. They argued that this phenomenon occurs because these low-working memory individuals have difficulty blocking out or inhibiting distracting information during listening comprehension.

[5] Studies examining individual differences in working memory often refer to the measured construct as short-term memory, particularly when the task at hand measures only storage capacity (as opposed to both storage and processing capacity). However, other studies claiming to examine working memory also use tasks that measure only storage capacity (e.g., Carrell, Dunkel, & Mollaun, 2002). This review refers to both of these as working memory, and distinguishes between the use of measures of storage capacity only and those which measure both storage and processing capacity.

The available evidence from listening-specific comprehension tasks suggests that the literature demonstrating working memory effects on reading extends to listening contexts. Indeed, there are clear theoretical arguments to suggest that working memory plays a critical role in listening (e.g., Engle, 2002). Working memory has been found to be related not only to reading comprehension, but also to other higher-level cognitive processes including reasoning ability (e.g., Jonides, 1995) and multitasking performance (e.g., Konig, Buhner, & Murling, 2005), and it has been argued that working memory is a general mechanism that underlies performance on many complex tasks (e.g., Conway, Cowan, Bunting, Therriault, & Minkoff, 2002). Given that these types of cognitive processes contribute to reading comprehension, it is clear that working memory is an important factor for L1 comprehension.

Working memory and L2 comprehension

Miyake and Friedman (1998) reported one of the few examinations to date of the role of working memory in L2 listening comprehension. They described a study performed by Miyake, Friedman, and Osaka (1998; cited in Miyake & Friedman, 1998, p. 348) exploring the causal relations between working memory and L2 listening comprehension. The original study revealed both a direct and an indirect impact of working memory on syntactic processing: participants with higher working memory capacity were able to make better use of syntactic information when comprehending the L2, and demonstrated a level of sensitivity to particular syntactic cues that was near native listener levels. The authors concluded that working memory span contributes to listening comprehension, both through influencing the ability to learn what cues are important in the L2 and through influencing how well the learner can make use of these cues during listening.

McDonald (2006) used a measure of working memory presented in the L2 of the participants[6] and found that performance on this L2 working memory measure correlated significantly with the accuracy of grammaticality judgments of spoken L2 sentences. In a second experiment, McDonald imposed a high working memory load on L1 listeners by requiring participants to maintain 7-digit numbers in memory. The study found that, when L1 listeners with an induced working memory load were asked to make grammaticality judgments about L1 sentences, they showed selective impairments in their judgments. The effect directly paralleled the performance of L2 learners on the same task (in the absence of an induced working memory load), suggesting that processing in the L2 imposes a load on working memory resources.

Other investigators have employed L2 reading comprehension measures to study the relationship with measures of working memory. Harrington and Sawyer (1992) assessed participants' L2 working memory processing and storage capacity as well as L2 storage capacity alone (i.e., digit and letter span measures completed in the L2). Using

[6] Participants in this study were late L2 learners who self-rated their L2 proficiency as $M = 3.82$ on a scale of 1-5 (the range of self-ratings was 2.5-5), with 1 indicating *poor* and 5 indicating *excellent*. They also self-rated their current frequency of L2 use as nearly equivalent to their current L1 use.

TOEFL Grammar and Reading sections as well as a cloze test[7] as their measures of L2 comprehension, they found that performance on the L2 working memory test of both processing and storage capacity, but not the test of L2 storage capacity alone, significantly correlated with L2 reading comprehension scores, suggesting that the attentional control aspect of working memory is critical to L2 reading comprehension.

Working memory measures that use L1 may avoid confounds with L2 proficiency.

It is important to note that in both the McDonald (2006) and Harrington and Sawyer (1992) studies, the measures of working memory were conducted in the L2, which may confound to some extent individual differences in working memory with L2 proficiency. Although this issue is worth considering when selecting working memory measures, Osaka and colleagues (Osaka & Osaka, 1992; Osaka, Osaka, & Groner, 1993) found that L1 and L2 working memory measures in language learners were strongly correlated (*r*'s ranging from the mid-0.70s to mid-0.80s). L1 measures of working memory may provide a purer measure of working memory (particularly in the context of L2 research); nonetheless, the results of Harrington and Sawyer provide important evidence of the importance of working memory in L2 comprehension.

There also is evidence that the effect of differences in working memory resources may be moderated by features of the input. Leeser (2007) had participants read texts about familiar or unfamiliar topics. Results showed that working memory was related to performance, but this effect differed depending on the task and the familiarity of the topic. On a recall measure of comprehension, greater working memory was related to better performance for familiar, but not unfamiliar, topics. However, on a verb form recognition measure (which compared recognition for verbs that have the same future tense they had in the passage and distractors having other forms [e.g., imperfect subjunctive]), the benefits of greater working memory capacity were instead found with unfamiliar topics. Walter (2004) found that working memory correlated with the transfer of L1 reading comprehension skills to an L2 comprehension task only in the more difficult condition where a pronoun and its referent were separated by multiple clauses (versus the easy condition, where the two occurred in the same clause). These findings suggest that the difficulty of the passage (determined by factors like topic familiarity) and the difficulty of the task determine how large a role working memory capacity plays in level of performance.

[7] A cloze test is a form of assessment consisting of a portion of a text with certain words removed which the test takers are asked to replace.

It is worth noting that not all researchers are convinced working memory is critical to on-line L2 processing. In a study of reading comprehension involving native speakers of Japanese, Chinese, and Spanish, Juffs (2004) presented garden path sentences in the L2 (English) to participants using a self-paced moving window paradigm. This methodology tracks the amount of time a participant spends reading each word in a sentence. Juffs analyzed reading times at the disambiguating region, where the garden path is resolved by the presentation of a word that disambiguates the previously ambiguous word. He found that several measures of working memory, including reading span and listening span, did not predict reading times in L2. He concluded that other researchers' claims of the importance of working memory processing capacity in L2 processing should be tempered in light of his findings. However, it is critical to note that Juffs also examined L1 comprehension with these same participants and failed to find any working memory effects. Given the robustness of working memory processing capacity effects on L1 comprehension across a variety of contexts and a wide range of populations (see Daneman & Merikle, 1994), this null effect raises questions regarding the stability of Juffs's null result for working memory effects on L2 comprehension. Looking across the current literature on working memory processing capacity and L2 comprehension (both reading and listening), the pattern of results indicates that working memory is likely to impact L2 listening comprehension, and that these effects will be particularly strong in conditions that impose additional demands on working memory.

Source of characteristic	Factor of interest	Consensus in the literature
Listener	Working memory	Impacts L2 listening comprehension, particularly when other conditions, such challenging characteristics of a spoken passage, pose additional demands

Proficiency and experience with the second language

Although language proficiency is a variable in many studies examining L2 listening comprehension, it is defined in so many different ways that it is difficult to compare results across studies. One reason for this difficulty stems from the lack of standardized tests for determining proficiency level across languages. Most studies measure proficiency using age of acquisition, teacher judgment, course level, or performance on a non-standard test. In turn, the definition of *high* vs. *low proficiency* can vary from study to study, even when the same variable is measured; in one study, *beginning* language learners might be defined as students in their second year of study, whereas in another study, the same second-year students might be called *intermediate* learners. Although standardization of listening proficiency tests is still evolving, tests like those developed by the American Council of the Teaching of Foreign Languages (ACTFL) and the Defense Language Institute could serve as a way to evaluate listening comprehension levels, though these would need to be widely available to language researchers. The lack of a standard definition of *high* and/or *low proficiency* is an important caveat for interpreting findings across the L2 comprehension literature. The literature on L2 proficiency focuses on three types of knowledge—vocabulary size, phonological and grammatical information, and background knowledge.

Vocabulary size

An obvious factor that can influence comprehension of a spoken passage is the overlap between the listener's vocabulary knowledge and the vocabulary of the passage. Nation (2001) makes a compelling case that listeners must have an *adequate vocabulary* to understand a passage in another language. *Adequate vocabulary* might be estimated by the number of words a listener needs to know to understand a representative sample of texts (a.k.a. text coverage). The 5,000 most frequent words yield a coverage of 90 to 95 percent of the word tokens in an average passage in many languages, including Russian (Steinfeldt, 1965), French (Guiraud, 1954; Sciarone, 1979), English (Bongers, 1947; Carroll, Davies, & Richman, 1971; Hirsh & Nation, 1992; Johnson 1972; Nation 1993; Palmer, 1931), and Dutch (Nieuwbourg, 1992; Ostyn & Godin 1985; Sciarone, 1979; Vannes, 1952). Furthermore, Hirsh and Nation (1992) have argued that in order to understand all the main points in a text, readers need to be familiar with 95 percent of the words therein. There is no similar measure of the coverage required to understand a spoken passage; but we assume that if listeners know more than 5000 vocabulary terms, they are likely to have a good chance at understanding what has been said.

Nation (2001) also argued that vocabulary size is an indirect measure of other variables also known to influence listening comprehension ability, including world knowledge. These other factors are discussed in more detail below.

Phonological and grammatical information

Research on the impact of phonological and grammatical information in spoken language comprehension has focused on whether high-ability and low-ability listeners use top-down and bottom-up processes[8] differently. Because words are not heard in isolation, but in specific contexts, both L1 and L2 listeners will use top-down processing strategies such as inferencing and elaboration to help make sense of a passage, particularly when they do not recognize every word in the input (Goh, 1998a, 1998b). Voss (1984) had users transcribe spoken passages in their non-native language and found that many listeners were extremely dependent on top-down information. In fact, the listeners who relied most heavily on *bottom-up* information made the most errors (Voss, 1984). Voss argued that less experienced L2 listeners rely even more heavily on top-down information because they are less familiar with the non-native phonology, vocabulary, and syntax. To compensate for this lack of familiarity, they use higher-order cues to comprehend what they heard. Field (2004) describes a series of experiments by Koster (1987, cited in Field, 2004, p. 366) in which participants were asked to discriminate between words and non-words in their non-native language. Koster found that words were correctly identified more often when preceded by a closely associated word.[9] The effect was biggest when the listener had lower proficiency in the language (as opposed to advanced or native proficiency). In a second experiment, participants performed a word/non-word decision task. Non-native participants were faster to identify words that were preceded by a closely associated word, while native speakers showed the same effect, but to a lesser degree. It was as if the native listeners did not have to rely on context to help them decide whether the word was real or not, whereas non-native listeners relied more heavily on the context provided by the preceding word.

Tsui and Fullilove (1998) also observed that less-skilled L2 listeners are less able to monitor their top-down impression of a passage and modify it as necessary using incoming bottom-up information than are more-skilled L2 listeners. They analyzed the answers given by 20,000 people in Hong Kong who had to answer comprehension questions about different types of listening passages. Some of the passages were constructed to allow listeners to rely on a consistent schema introduced at the beginning of the passage as they listened, while other passages required revisions of the initial schema as more information was presented. The less-skilled listeners answered fewer questions correctly for the passages that required modification of the initial schema than did the more-skilled listeners. The authors concluded that this finding was due to less-skilled listeners relying more on top-down information (the schema activated initially) and failing to use bottom-up information (new information contradicting the initial schema) to modify the top-down information they applied in listening.

---

[8] In this review, the term *top-down processes* refers to the use of information from the highest conceptual levels (e.g., inferencing, elaboration, integration, etc.) to fill in missing details at the lower levels. The term *bottom-up processes* refers to decoding information at the lowest level (e.g., acoustic-phonetic information) and using that information to progressively build higher-level representations.

[9] A closely associated word is a word with a higher score on a "word association" test, wherein a group of participants respond to a given word with the first word that comes to mind. The more participants that generate a particular word, the more highly associated that word is to the given word.

Goh (2000) asked 40 language students to describe the processes they used to understand the spoken L2, as well as the problems they experienced. She found that the more familiar the students were with the non-native language's phonology, the more they relied on low-level (bottom-up) information. Indeed, she found that when L2 students do not know how to pronounce a word, they default to their native pronunciations. This is particularly evident with stress pattern differences in the native and non-native languages (hosTEL vs. HOStel), which, in addition to pauses, word-onsets, and perceptual salience, are used to segment the incoming auditory stream into words (Sanders, Neville, & Waldorf, 2002). This complements the findings of Voss (1984) in that the more familiar the listeners are with the non-native phonology, the more they rely on phonological cues (bottom-up information).

The results so far are consistent with Cutler's (2001) observation that L2 listeners tend to segment words in continuous speech on the basis of their usual L1 segmentation procedures. Indeed, when Weber and Cutler (2006) asked proficient German users of English as a second language to listen to nonsense sequences and respond whenever they heard an English word, they found that while the proficient L2 listeners used the permitted sound sequences of a language (the L2 phonotactics), there was still interference from the L1 phonotactics.

Interestingly, the inability of listeners to produce a phonological segment does not impair their ability to hear the phonology of their non-native language. Yamada and Tokura (1992) showed that while Japanese speakers can be trained to hear /r/ vs. /l/ many of them cannot produce this distinction.

Background knowledge about the topic, text, structure, schema, and culture

Listeners' background knowledge about a passage can have a profound impact on their ability to understand what has been said. Without a schema, understanding a passage can be extremely difficult. A classic demonstration of this phenomenon was first reported by Bartlett (1932). He asked participants to listen to passages in their native language. These passages (the most famous of which is *War of the Ghosts*) described a sequence of events that seemed logical, but were often slightly illogical, with several subtle non-sequiturs.[10] Most people found it extremely difficult to recall the story exactly even after repeated readings; where the elements of the story failed to fit into the listener's schemata, they were omitted or changed into more familiar forms. This phenomenon demonstrates that the degree to which information in the passage conforms to the listener's existing knowledge base determines how easy it is to understand.

Schema: A framework that helps the listener to organize and interpret incoming information.

[10] Here is an example from the *War of the Ghosts*. "But presently the young man heard one of the warriors say, 'Quick, let us go home: that Indian has been hit.' Now he thought: 'Oh, they are ghosts.'…." In this context, a non sequitur is a statement that does not follow logically from what preceded it.

Source of characteristic	Factor of interest	Consensus in the literature
Listener	Proficiency and experience with the second language	Most studies do not define proficiency in the same way, yet researchers agree that: • L2 listeners are more likely to rely on top-down processing when they have a weaker command of the phonology and grammar, factors that enable accurate bottom-up processing. • L2 listeners may attempt to understand spoken passages using whatever background knowledge they may have, e.g., of the topic, genre, culture, and schemas, even when their knowledge is not complete or accurate.

In the literature, authors often note the importance of shared knowledge between the speaker and listener (Churchland, 1999). Shared knowledge is important because listeners are pragmatic by nature and, whenever possible, will situate themselves in terms of the topic, the setting, the event, the speaker, and the purpose for listening (Rost, 2005; Lantoff, 1999). Vandergrift (2007) noted that L2 listeners will use prior knowledge (including topic, genre, culture, and other schemas in long term memory) to build a conceptual framework against which they interpret what they hear. Context, non-verbal information, world knowledge, cultural information, and common sense are all used to build this framework. Bodie, Worthington, Imhof, and Cooper (2008) reported that other factors specific to a particular listener (specific knowledge about the topic, world knowledge, memory span, motivation, listening capacity) interact with aspects of the speaker (interactivity, status, power, role) and the passage (objectives, purpose) to predict listening comprehension scores.

In his well-known "SPEAKING" model, Hymes (1972) identified eight situation-bound features of a message (written or spoken) that require background knowledge on the part of the L2 listener; each message has the potential to be culturally specific. Presenting material that is incongruent with the listener's expectation will result in confusion. These features include:

- the **S**etting/scene,
- the **P**articipants in the interpretive community[11] (as described in Lakoff, 2000),
- the **E**nds, or purposes of the communication (e.g., the desired outcomes and goals),
- the sequence of communication **A**cts (which can be dictated by the message's format and context),
- the **K**ey (register – formality, politeness, power relations),
- **I**nstrumentalities (channel, forms of speech),
- **N**orms, and
- **G**enre (e.g., passage type).

[11] Lakoff's term "interpretive community" refers to a group of like-minded individuals who share similar assumptions about how a text should be understood.

As described above, listeners with rich background knowledge use it to compensate for misunderstandings, unclear speech, and a lack of local or specific context from earlier parts of the passage (Goh, 2000). This use of background knowledge can be detrimental, however, because listeners that rely too heavily on prior knowledge, prior conversational units, and relational history with the speaker may be unduly biased by this information (Bodie et al., 2008). That said, the advantages likely outweigh the disadvantages; using background knowledge to help understand a passage frees up attentional resources to be devoted to other aspects of the listening task (Tyler, 2001).

Metacognitive strategies

Metacognition is both self-reflection and self-direction. Reflecting on one's thinking while listening, for example, can help listeners to adopt more appropriate ways to listen effectively. A number of studies have demonstrated the impact of metacognitive strategies on learning (Boekaerts & Niemivirta, 2000; Bolitho, Carter, Hughes, Ivanic, Masuhara, & Tomlinson, 2003; Eilam & Aharon, 2003; Mokhtari & Reichard, 2002; Palmer & Goetz, 1988; Schoonen, Hulstijn, & Boosers, 1998; Victori & Lockhart, 1995; Winne, 1995; Zimmerman & Schunk, 2001) and a few on second language reading skills. For example, Schoonen et al. (1998) described the positive effect of three types of metacognitive knowledge, including self-knowledge, task knowledge, and strategic knowledge, on the L2 reading comprehension of 6th, 8th, and 10th grade students. They found that the 8th and 10th grade students who reported using more metacognitive strategies had better L2 reading comprehension scores. Sixth grade listeners didn't use as many metacognitive strategies. This may have been because they were struggling to understand the language and didn't have the resources to effectively implement metacognitive strategies.

Vandergrift, Goh, Mareschal, and Tafaghodtari (2006) developed and validated a listening questionnaire (the Metacognitive Awareness Listening Questionnaire or MALQ) that assesses the metacognitive awareness of second language listeners as well as their (reported) use of metacognitive strategies when listening to a spoken passage. Their work revealed five distinct factors that predicted listening comprehension scores on the University of Ottawa's Placement Test (Wesche, Paribakht, & Ready, 1996):

- *Problem-solving strategies*: Strategies listeners use to make inferences and monitor them (e.g., using known words to infer the meanings of unknown words, using experience and general knowledge when interpreting the text; Rubin & Roberts, 1987).
- *Planning and evaluation strategies*: Strategies listeners use to prepare themselves for listening, and for evaluating whether their efforts are paying off (e.g., having a plan for listening, keeping a goal in mind while listening; Hinkel, 2006; Richards, 1990; Vandergrift, 2003).
- *Avoiding mental translation*: Beginning-level listeners often will mentally translate a passage as they hear it (Eastman, 1991), but they must overcome this strategy in order to become skilled L2 listeners (Vandergrift, 2003).

- *Self-knowledge*: These strategies involve understanding the difficulty of the L2 listening task and being aware of one's confidence levels and anxiety level when listening to the second language (Chemers, Hu, & Garcia, 2001; Sparks & Ganschow, 2001; Wenden, 1991; Yang, 1999; Zimmerman & Schunk, 2001).
- *Directed attention*: Strategies that listeners use to stay on task. These include recovering concentration when one's mind wanders, focusing more when understanding falters, and not giving up (Rost, 2002).

However, just being aware of these strategies is not enough; Deci and Ryan (1995) observed that a listener's motivation will affect whether he or she is able to use metacognitive strategies to help improve listening performance.

Anxiety

Source of characteristic	Factor of interest	Consensus in the literature
Listener	Metacognitive strategies	Promote comprehension during L2 listening tasks through directing attention or avoiding unsuccessful methods, like mental translation
	Anxiety	Can negatively affect comprehension, but may be ameliorated by testing conditions (e.g., the ability to take notes)

Listener anxiety can have a profound effect on comprehension abilities. When listeners are concerned that the message is too complex or that they will not be able to understand it, their ability to concentrate falters, and comprehension declines; this is true even in the native language. Listeners are more anxious when in a new situation, listening to new information, trying to sort conflicting information, listening to seemingly illogical passages (Clark; 1989; Priess & Wheeless, 1989; Wheeless, 1975), or when they think their performance reflects their abilities or intelligence (Hussey, Teubner-Rhodes, Dougherty, & Novick, 2010). Indeed Hussey et al. (2010) found that anxiety had a direct impact on listeners' abilities to resolve ambiguities in their native language in real time. Factors related to the testing conditions, such as the ability to take notes, may decrease anxiety during L2 listening comprehension (Chang & Read, 2008). Ability to take notes and other characteristics of the testing conditions are discussed below.

Summary of research findings for listener characteristics		
Factors with strong effects or convincing evidence		**Factors with sparse or inconsistent evidence**
Beneficial to listeners:	Greater **L2 proficiency**	*None in current review*
	Greater **working memory** capacity	
	Use of **meta-cognitive strategies**	
Difficult for listeners:	**Anxiety**	

CHARACTERISTICS OF THE PASSAGE

Although there are many passage-based factors that one could discuss in determining what makes L2 listening comprehension difficult, this review covers only a portion of them. The factors discussed include authenticity, passage length and related factors, passage complexity, passage type and organization, and auditory features.

Authenticity

> ***Authenticity*** has been defined in different ways. This review adopts the following definition:
>
> A piece of real language that is created by a real speaker for a real audience in order to convey a message of some sort (Gilmore, 2007; Thanajaro, 2000).

Authenticity of aural materials can be defined in many ways, involving the speaker, the listener, the context, and the message (Breen, 1985). One prevalent way of defining *authenticity* for a passage itself, and the definition that this report adopts, is that an authentic passage is a piece of real language created by a real speaker for a real audience in order to convey a message of some sort (Gilmore, 2007; Thanajaro, 2000).[12] This characterization is designed to exclude passages that are created to exemplify some aspect of the language (e.g., the use of the future tense or speech acts such as apologies) rather than to convey an actual message. The characterization does include other types of passages such as speech from a native speaker to a non-native speaker and scripted television programs (Gilmore, 2007). This definition is similar to that currently employed by the Defense Language Institute, which defines *authentic texts* as "those which are produced by users of the target language and which are intended to be read [or heard] *by other users of the target language in the target-language culture*." (Defense Language Institute Foreign Language Center, 2010, p. 23). Note that "users of the target language" indicates that the receivers of the passage may be native or non-native listeners. Despite the fact that L1 speakers delivering a spoken message to non-native listeners may alter their speech in a number of ways, there are still likely to be differences between passages that are authentic and those that are created. For instance, Flowerdew and Miller (1997) noted several important differences (e.g., use of discourse macro-markers) between an authentic lecture delivered to L2 students and a taped commercial listening passage from an *English for Academic Purposes* textbook.

The push to use authentic materials in teaching second-language listening skills began in the 1970s (Gilmore, 2007). Apart from more general concerns that using created passages rather than authentic passages robs the L2

[12] The quality of authenticity with regards to a second language listening passage may be better termed *genuine* (Widdowson, 1976, as cited in Long & Ross, in press). A passage may be a genuine example of the L2 (e.g., a recorded telephone conversation between two friends), but most activities performed with the passage (e.g., responding to comprehension items after listening) are not authentic *uses* of the passage, so the task itself (listening and responding to items) cannot be termed *authentic* (Long & Ross). Although this review acknowledges the distinction between authentic and genuine, it adopts the term *authentic* to refer to a quality of the passage rather than of the task to accord with how it is often used in the literature.

learner of experience with many elements of the L2 (e.g., lexical reductions like *can't*, a realistic speech rate, disfluencies; Breen, 1985; Cobb, 2004; Crossley, McCarthy, Louwerse, & McNamara, 2007; Flowerdew & Miller, 1997; Long & Ross, 2009; Rings, 1986; Rogers & Medley, 1988), experimental evidence shows that L2 learners benefit from experience with authentic materials (Herron & Seay, 1991; Kienbaum, Russell, & Welty, 1986). Of primary concern to the current literature review, however, is just how authentic passages might differ from created passages in terms of passage-based and context-based factors of interest, and the impact these differences are likely to have on the difficulty of the passage for an L2 listener. Throughout this section, any research uncovering differences between authentic and created (e.g., textbook) listening passages will be discussed in the section describing the relevant passage-based factor.

Passage length and related factors

One factor of concern in L2 listening comprehension is passage length and the extent to which listeners can cope with the amount of information that is presented for processing (Alderson et al., 2006; Bejar, Douglas, Jamieson, Nissan, & Turner, 2000; Carroll, 1977, cited in Dunkel, 1991, p. 440; Rost, 2006). Unlike reading, listening comprehension occurs in real time. Listeners may not have the option of going back to something they failed to comprehend (unless they can rehear the passage, see the section on Number and Control Over Hearings). Instead, the result of such a failure will be an inability to attend to new information as the listener invests additional time in attempting to understand what they missed, or an inability to comprehend later information because it relies on the understanding of earlier information (Goh, 2000; O'Malley, Chamot, & Kupper, 1989). Longer passages may be more likely to disrupt comprehension due to overwhelming listeners' working memory storage capacity (Henning, 1990). In addition, the longer a passage is, the more information listeners could miss after encountering information they do not understand.

Passage length has been defined with a number of measures, including syllables/second, duration (in minutes or seconds), and number of words or sentences.

There is also reason to predict that longer passages will have a greater impact on the listening comprehension of lower-proficiency listeners. Lower-proficiency L2 learners often try to understand a passage on a word-by-word basis; because word comprehension in the L2 is slow and effortful, this strategy is generally maladaptive for lower-proficiency L2 learners (Vandergrift & Tafaghodtari, 2010). As lower-proficiency listeners fixate on a particular word they missed, they may fail to attend to the continuing stream of information (Field, 2004; O'Malley et al., 1989; Vandergrift, 2003).

Researchers have used a number of measures to quantify passage length, including duration, number of syllables, number of words, and number of sentences, with inconsistent results at best. The lack of consistent results for passage length may be due to its relationship to other more predictive variables like redundancy (when information is presented more than once through repetition, elaboration, or other methods) and information density (the number of distinct ideas in a passage). Redundancy and density, in turn, may interact with passage authenticity and demands on working memory. In the sections that follow the review summarizes findings related to passage length, redundancy of information, and information density.

Length of passage

Studies examining passage length generally find that passage length alone does not affect L2 listening comprehension difficulty. Thompson and Rubin (1996) reported that students complained about segments longer than 2.5 minutes as being too long for them to maintain concentration. However, Thompson and Rubin found that passage length interacted with other qualities of the passage: longer dramatic segments were better tolerated than longer news reports, suggesting that number of speakers, structure of the passage, or topic may interact with length. Nissan, DeVincenzi, and Tang (1996) found no effect of overall passage length in seconds on listening comprehension item difficulty, and Kostin (2004) found no significant effect of the total number of words in the passage on comprehension difficulty. However, the passages examined in both these studies were fairly short (< 20 seconds in length). It is possible that the range of lengths was too narrow or the length of examined passages as a group too short overall to produce a reliable effect of length on item difficulty. Moyer (2006) found no significant difference between performance on comprehension questions corresponding to long passages (2–3 minutes) and those for short passages (2–4 sentences) for either non-native or native speakers. However, passage length in Moyer's study was fully confounded with passage type, since long passages were all informal dialogues and short passages were all formal news reports or announcements. The effect of passage length cannot be evaluated independently from the effect of passage type in her results.

Rupp, Garcia, and Jamieson (2001) explored the effects of passage word count and average sentence length on L2 listening comprehension. An initial analysis indicated that longer average sentence length predicted more difficult listening comprehension items, and a second analysis found that both overall word count of the passage and sentence length contributed to item difficulty. However, the results for word count must be interpreted with caution, as analyses were done for reading texts and listening passages combined together rather than each modality separately. Combining findings across the modalities may have led some factors to look as though they were important in determining both listening and reading comprehension item difficulty, when in fact they were important for only one modality.

Rupp et al. (2001) also argued that average sentence length likely increased item difficulty due to the greater syntactic complexity of longer sentences, not because longer sentences simply provided more information to be processed. The syntactic complexity of passage sentences is another factor that likely correlates with overall passage length, though the two qualities are separable: a passage containing many short, simple sentences can be equivalent in length to another passage containing a smaller number of longer and more syntactically complex sentences. Greater syntactic complexity may increase listening comprehension difficulty for L2 listeners (Chaudron, 1983), but this may not always be the case if other features, like redundancy, are present (Long & Ross, 2009; see the section discussing Redundancy below).

In summary, studies examining overall passage length, whether measured by duration, word count, or number of sentences, do not generally find that this factor predicts listening comprehension difficulty. However, materials in these studies confounded factors like passage type and length, or examined only a narrow range of passage lengths with the average length being quite short. Rupp et al. (2001) did not perform analyses separately for listening and

reading materials, leaving open the possibility that their findings obscured differences between the modalities. Passage length may also be a better predictor of difficulty when the range of considered lengths is wider, when longer passage lengths are examined, and when other important factors, such as the syntactic complexity of the passage, are controlled.

Source of characteristic	Factor of interest	Consensus in the literature
Passage	Length	Effects of this factor alone are weak and inconsistent

Working memory and passage length

There is reason to predict that working memory ability would interact with the effects of passage length on L2 listening comprehension. Demands on working memory are higher when processing is less automatic (Baddeley, 2007), and listening comprehension is less automatic for non-native listeners than for native listeners (Tyler, 2001). This should result in L2 listeners generally experiencing greater demands on working memory processing when listening to an L2 passage. Further, the more information that must be held in working memory, the greater the strain on working memory storage (Baddeley & Hitch, 1974); thus, a passage containing more information should pose a greater challenge for working memory.

The scientific literature does not currently address whether passage length interacts with ***both*** working memory storage ***and*** working memory processing capacity.

Henning (1990) examined individual differences in working memory storage capacity (measured using digit span[13]) and passage length for effects on L2 listening comprehension. He found no relationship between working memory and the effect of passage length on comprehension item difficulty; though he did find that longer passage lengths (30 words compared to 10 words) were associated with more difficult listening comprehension items. This latter result was statistically weak, however, and Henning's study confounded passage length with the number of test items corresponding to a passage and with item type, making it difficult to draw conclusions about the effect of passage length on comprehension or its potential interaction with working memory.

Carrell, Dunkel, and Mollaun (2002) examined the effects of passage length, L2 listening proficiency (measured by the Institutional TOEFL listening comprehension section), and working memory storage capacity (using the digit span test) on L2 listening comprehension. Although they found no interaction of working memory with passage length on performance for listening comprehension items, they did find an interaction between L2 listening proficiency and passage length on performance. Higher-proficiency listeners (TOEFL score $\geq$ 49) did significantly better for items corresponding to short passages (~2.5 minutes long) than those for longer passages (~5 minutes long), while lower-proficiency listeners (TOEFL score < 49), who performed worse than higher-proficiency listeners overall, showed no effect of passage length. This suggests that lower-proficiency listeners were overwhelmed by even the shorter passages.

[13] The term *digit span* refers to the number of randomly ordered digits a person can remember in the presented order.

The findings regarding the relationship between differences in working memory and passage length suggest that, to the extent that longer passage lengths hurt L2 listening comprehension, it is not likely to be due to an increased working memory load. However, as noted above, measures of working memory that focus on storage capacity alone (such as digit span) may simply not be the appropriate measure of individual differences in working memory for language comprehension (Daneman & Merikle, 1996). Research on L2 listening comprehension with measures of working memory storage *and* processing capacity is needed to address this issue.

Further, the findings of Carrell et al. (2002) point to another important consideration: passage length may only affect L2 listening comprehension once listeners have reached a particular level of proficiency. When L2 listeners are lower proficiency, they may be overwhelmed in trying to process even shorter passages, especially as the "shorter" passages examined by Carrell et al. were quite long compared to those examined in earlier studies (~2.5 minutes compared to 10 words in Henning [1990], and < 20 seconds in Kostin [2004] and Nissan et al. [1996]).

Passage length is related to other factors that can be separately manipulated, including:

- Syntactic complexity of sentences
- Length of material surrounding item-relevant information
- Redundancy
- Information density

Length of passage material surrounding item-relevant information

The degree of effort required to comprehend and integrate a particular piece of information is also likely to determine how likely it is to be comprehended. Buck and Tatsuoka (1998) found that L2 listening comprehension items were more difficult if there were more words surrounding the critical information (>8.3 words). Similar results were found by Brindley and Slatyer (2002). These results are consistent with the finding that information in the middle of a passage (which is both preceded and followed by additional words) is more difficult to comprehend and recall than information at the beginning or end of a passage (Freedle & Kostin, 1996, 1999), which will be discussed in the section on Discourse and Rhetorical Structure. However, this method of measuring the amount of information surrounding key information may have captured syntactic complexity of the sentence containing the key information rather than the amount of information, as Rupp et al. (2001) argued for their measure of sentence length. Although little research has explored the effects of length of the material surrounding key information on listening comprehension difficulty, this factor may be more predictive of difficulty for comprehending particular information in the passage than is overall passage length.

Summary: Length of passage

Studies examining the effects of passage length on L2 listening comprehension suggest that overall passage length is not a strong factor in determining passage difficulty, but this may be due to imprecise operationalization of *length*. Studies that do find a relationship find one that is weak (Henning, 1990) or qualified by interactions (Carrell et al., 2002). However, problems exist in terms of confounding other aspects of the passage with length (Moyer, 2006) or examining passages with a very limited range of lengths (e.g., Nissan et al., 1996). In addition, the range of lengths compared varied greatly across studies (e.g., 2.5 versus 5 minutes in Carrell et al., 2002, and 10 words versus 30 words in Henning, 1990). Several studies have also uncovered other factors more specific than overall passage length

that affect difficulty, such as the length of information surrounding the item-relevant information (Buck & Tatsuoka, 1998).

Redundancy

Redundancy involves repeating key information through exact repetition, paraphrase, and elaboration (Chaudron, 1983). Across a variety of passage types (e.g., both conversations and lectures), speakers will circle back to previously introduced points to check for comprehension in their listeners, or simply to reiterate information they believe is most important (Field, 2008). Perhaps counterintuitively, redundancy is often classified as a form of simplification of input because it involves re-presenting information, thus giving the listener another chance to comprehend the information, sometimes in a form that is easier to process or retain (Chaudron, 1983; Parker & Chaudron, 1987; Oh, 2001). Further, redundancy is argued to be a superior manner of simplification compared with syntactic simplification, which involves modifications like restricting the passage to canonical word order (e.g., Subject-Verb-Object is the canonical word order in English; Crystal, 2003) or using simpler syntactic forms (Cobb, 2004; Parker & Chaudron, 1987; Pica, Young, & Doughty, 1986).

***Redundancy*:** When information is re-presented through exact repetition, elaboration, or other methods

Types of redundancy and L2 listening proficiency.

Redundant information (e.g., presenting a synonym of a previously given word, elaborating on an earlier point) may be expressed in several different ways (Chaudron, 1983). These forms differ in the complexity of their syntax and the psychological salience of their semantic relationship with the initial information. Simple repetition, using the exact same word(s) as stated earlier, is the least complex and most salient form, while giving a synonym is more complex and less salient (Chaudron, 1983; Chiang & Dunkel, 1992). The complexity and saliency of the redundancy form can affect how redundancy impacts L2 listening comprehension.

Chaudron (1983) examined the effects of five types of redundancy, of differing syntactic complexity and psychological salience, on L2 passage comprehension. He found that type of redundancy determined its impact on comprehension and interacted with L2 proficiency of the listener (measured by performance on the Michigan Test of Aural Comprehension). Participants were divided into lower-, medium-, and higher-proficiency groups (MTAC scores of 30–52, 53–67, or 68–82, respectively). All participants showed the lowest level of performance on

Chaudron's (1983, pp. 441-443) five types of redundancy in increasing order of syntactic complexity	
Initial utterance: *They are selling beer at the picnic*	
1. Simple noun	*The beer tastes terrific.*
2. Repeated noun	*The beer…the beer tastes terrific*
3. Synonym	*The brew tastes terrific.*
4. Topicalizing rhetorical question	*What about the beer? It tastes terrific.*
5. If-clause	*If you afford the beer, it tastes terrific.*

comprehension items[14] when topic redundancy was presented in the least salient form (synonyms). Higher-proficiency participants showed greater comprehension than did medium or lower-proficiency participants when hearing more complex forms of redundancy (rhetorical question and *if*-clause), with lower-proficiency listeners benefitting most from repeated noun (high salience and low complexity). These results indicate that not all forms of redundancy are equally beneficial to listening comprehension for all listeners.

Later studies provided additional support to the findings of Chaudron (1983). Gainer (1997) found that dialogue passages where key information presented by the first speaker was echoed by the second speaker (Speaker 1: *He was born in 1955*. Speaker 2: *Born in 1955*.) yielded superior comprehension[15] for both higher- and lower-proficiency listeners, compared to an unmodified version of the same passage. This finding is consistent with Chaudron's (1983) finding that redundancy in the form of exact repetition (low syntactic complexity and high psychological saliency) is beneficial for both lower- and higher-proficiency listeners. Chiang and Dunkel (1992) explored the effects of redundancy presented in the form of paraphrasing ("The food of the Pennsylvania Dutch Country is very hearty and delicious. *Hearty and delicious food is nourishing and tasty*," p. 354), which is more complex than exact repetition (Chaudron, 1983), and found that this redundancy improved comprehension[16] for higher-proficiency listeners (Comprehensive English Language Test score 20–35) but not lower-proficiency listeners (CELT score 8–18). Again, this finding is consistent with Chaudron's conclusion that more complex forms of redundancy are less likely to benefit lower-proficiency listeners.

The results for type of redundancy and its interaction with proficiency indicate that more transparent types of redundancy (e.g., exact repetition) are beneficial for lower-proficiency listeners, while higher-proficiency listeners can also benefit from more complex forms of redundancy like paraphrase. Lower-proficiency listeners may experience an increased working memory processing load if redundancy is not transparent, as they try to understand this information independently from what was previously given (Blau, 1991; Field, 2008; Rubin, 1994). Similar differences in the benefit of redundancy have been found between younger versus older children who are native listeners (Sonnenschein, 1982, cited in Anderson & Lynch, 1988, p. 51). This further bolsters the idea that less

---

[14] Comprehension was measured in this study through performance on two types of items: (1) topic-related recall and (2) topic-related recognition. Recall items were verbatim sentences from the passage with a key word clozed (i.e., missing), to be filled in by the participant. Recognition items were statements for the participants to identify as having been presented by the passage or not.

[15] Gainer (1997) measured comprehension through the number of correct responses to verbatim statements from the passage with one or two key words clozed, to be filled in by the participant.

[16] Chiang and Dunkel (1992) measured comprehension with multiple-choice items. These items may have targeted the comprehension of main ideas, implied information, or details, but the authors did not specify which. It is possible that this comprehension measure was less dependent on the recall of exact details than were the measures used by Gainer (1997) and Chaudron (1983).

experienced listeners benefit less from redundancy (at least redundancy that is more complex and less salient than the simple repetition of topic nouns) than do more experienced listeners.

Interactions between redundancy and other passage factors

As will be discussed throughout this review, particular qualities of a passage, of the listener, or of the testing conditions can make the passage more or less difficult to comprehend. Some types of redundancy may be too complex for lower-proficiency listeners to benefit from (Chaudron, 1983). However, the presence of other qualities in the passage that alleviate comprehension difficulty may increase the chances that lower-proficiency listeners can more effectively use more complex types of redundancy.

Kelch (1985) investigated the effects of reduced rate of delivery and redundancy on listening comprehension. Redundancy in this study involved presenting synonyms, hypernyms (e.g., *animal* is a hypernym of *cat*), and paraphrases of important ideas. Redundancy improved recall of equivalent-meaning words (i.e., words synonymous with the exact words from the passage), but only when the speech rate of the passage had been slowed to 2/3 its unmodified rate. As synonyms and paraphrase are more complex forms of redundancy according to Chaudron (1983), participants in Kelch's study may have needed a slower speech rate to allow them to comprehend and make use of the more complex forms of redundant information.

In another investigation of redundancy effects and speech rate, Teng (2001) found a general effect of redundancy through paraphrase, but failed to find an interaction with speech rate. Paraphrase is generally described as more syntactically complex than other types of redundancy (Chaudron, 1983; Chiang & Dunkel, 1992), but the term is used somewhat loosely. Other studies have characterized paraphrase as the exact repetition of earlier-presented words in the same order (Gainer, 1997). Because Teng (2001) did not describe the method of paraphrase used, it is difficult to determine what was meant by redundancy, and Teng's failure to find an interaction between speech rate and redundancy may be due to the use of a simpler form of redundancy. Another issue in comparing Teng and Kelch's studies is that the speeds that were classified as *fast* and *normal* were not the same: Teng's (2001) normal rate passages were considerably slower than those of Kelch (1985), and thus likely to be more easily comprehended by the participants. The issue of defining *fast*, *slow*, and *normal* speech rates in L2 listening comprehension is covered in greater depth in the section on Speech Rate.

Although findings in this area are not completely consistent, more complex forms of redundancy may become accessible to lower-proficiency L2 listeners when other qualities of the passage, like speech rate, make the passage easier to comprehend. The potential interaction between redundancy and speech rate should be addressed in future studies with more consistent speech rates.

Redundancy of item-relevant information

In all the studies described above, redundancy was manipulated through repeating, in some fashion, main or key ideas in the passage. Other research has directly explored the relationship between whether a particular idea is redundant in the passage and whether a test item for which that idea is relevant is easier or more difficult. The results of this research are quite consistent. In all cases, hearing information more than once when it is important for

responding to a particular comprehension item decreases item difficulty (Buck & Tatsuoka, 1998; Freedle & Kostin, 1996, 1999; Ying-hui, 2006). Unfortunately, in these studies the forms of redundancy are not described in enough detail to determine their complexity or salience. Nonetheless, these results indicate that the comprehension of a piece of heard information, and its later recall, will be improved if the relevant information is repeated in some form in the L2 passage.

Summary: Redundancy

Overall, the research examining the effects of redundancy on listening comprehension suggests that repetition or paraphrasing of information in the passage improves comprehension for that information. This increase in comprehension is true both for item-relevant information and other information in the passage. However, it is important to take into account how forms of redundancy with different complexity and salience affect comprehension across listeners of higher and lower proficiency (e.g., Chaudron, 1983). Lynch (1988; as cited in Rubin, 1994, p. 203) noted that the usefulness of redundant information may depend on its being recognized as a repetition of previously given information (see also Blau, 1990, and Field, 2008); in lower-proficiency listeners, more complex types of redundancy may simply add to the processing load (Anderson & Lynch, 1988; Chaudron, 1983). In considering redundancy, it is also important to take into consideration its potential interaction with speech rate: redundant information in passages that are too fast for L2 listeners to comprehend may not benefit comprehension.

Source of characteristic	Factor of interest	Consensus in the literature
Passage	Redundancy	Benefits of redundancy depend on the method used to re-present information, and how this interacts with the proficiency of the L2 listener, and other characteristics of the passage

Information density

One rationale behind examining passage length as a factor in L2 listening comprehension is the belief that a greater processing load is introduced by a longer passage (Carroll, 1977, cited in Dunkel, 1991, p. 440; Henning, 1990; Rost, 2006). In service to this concern, the amount of information in a passage may be a more predictive factor for comprehension difficulty than is overall passage length. *Information* has been defined in the literature in several ways: content words (e.g., a noun, verb, adjective, or adverb, Nissan et al., 1996), the related concept of words with independent meaning (e.g., *mother*) as opposed to those without independent meaning (e.g., *a*, Gilmore, 2004), and propositions (the smallest unit of knowledge that can stand alone as a separate true-false statement, Dunkel, Henning, & Chaudron, 1993). Measures of information density involve dividing the number of pieces of information in a passage by the total number of words or the duration of the passage. Sometimes measures of density only include those pieces of information that have not been previously given in the passage in the numerator (i.e., they control for redundancy, Aiken, Thomas, & Shennum, 1975), and so these measures directly capture the density of *unique* information in a passage. While information density will tend to be highly correlated with passage length in authentic passages, such that more dense passages will tend to be longer than less dense passages, the two factors could be

independently manipulated and may vary separately in authentic speech depending on how information density is defined. Like passage length, greater information density is believed to make higher cognitive demands of L2 listeners, which may increase the effort involved in listening comprehension (Gilmore, 2004).

Information density as content word density

One way of defining a piece of information is as a *content word* (e.g., a noun, verb, adjective, or adverb; Nissan et al., 1996). Nissan et al. examined content word count alone as a measure of amount of information (i.e., they did not examine information *density*) and found no relationship between content word count and item difficulty. However, Buck and Tatsuoka (1998) found that the average number of content words per idea unit in the area surrounding item-necessary information predicted item difficulty.[17] One potential reason for the difference between the results of Nissan et al. and those of Buck and Tatsuoka is that the passages investigated by Nissan et al. were short in length overall (5–20 seconds), and so necessarily represented a limited range of content word counts. Buck and Tatsuoka reported a range of 4-20 content words per information unit; greater variation in content word counts could have increased the potential for finding an effect of this factor. In addition, Buck and Tatsuoka examined content word count in the area surrounding item-necessary information, as opposed to an overall count of content words in the passage. They argued that the average amount of information in the portion of the passage containing the item-relevant information may be more important in predicting item difficulty than the overall amount of information in the passage.

Some measures of ***information density*** include counting the number of propositions or content words and dividing this count by the total number of words in the passage.

Other measures take into account some types of redundancy (e.g., ratios of type/token or number of unique propositions/total propositions), though this may be misleading because the listener may not be aware that the information is redundant.

In addition to the number of content words/idea unit in the area around the item-necessary information, Buck and Tatsuoka (1998) also examined the proportion of content words to all words surrounding the item-necessary information (a measure of information density). They found that this factor was also a significant predictor of item difficulty, but they made no direct comparison of its predictive power and that of the content words/idea unit factor. Nonetheless, their finding for the ratio of content words to all words surrounding the item-necessary information suggests that when information is defined as content words, item difficulty increases as information density increases.

[17] Buck and Tatsuoka did not describe how they define *idea unit*, so it is difficult to determine if their count of content words involved entire sentences, clauses, or some other level of analysis. The authors cite Chafe (1985), who defined this term as "the amount of information a person can comfortably pay attention to and verbalize" (pp. 106). Because it is unclear how Buck and Tasuoka defined idea unit, it is possible that the length of the idea units, in words or duration, varied between passages. For this reason, their measure of content words/idea unit may not have been a pure measure of information density, in that it may not have controlled for length or duration.

Rupp et al. (2001) employed another measure of information density using content words: type/token ratio. This is the number of unique content words divided by the total number of words in the passage. To estimate the number of *types* in the numerator, for example, the appearance of the words *dog* and *dogs* would be counted only once. In the analysis of L2 listening comprehension by Rupp et al. (2001), type/token ratio emerged as a significant predictor of item difficulty, with test items for passages with larger type/token ratios being more difficult for the listener. This method of measuring information density provides an advantage over Buck and Tatsuoka's (1998) measure of information density, in that type/token ratio controls for one type of redundant information: exact repetition of words. Because of the findings regarding the impact of redundancy on listening comprehension (e.g., Chaudron, 1983; Gainer, 1997), it is important to consider redundancy in addition to information density. As described below, other researchers have also taken redundancy into account when measuring information density.

Type/token ratio of a passage is defined as the number of unique words that are not from the same word family (e.g., *dog* and *dogs* are from the same word family) divided by the total number of words in the passage.

Information density as propositional density

A proposition is the smallest unit of knowledge that can stand alone as a separate true-false statement (Dunkel, Henning, & Chaudron, 1993). For example, *The cat ate the meat* can be expressed in multiple forms (e.g., *The meat was eaten by the cat*) that convey the same proposition (Crystal, 2003, p. 377). Further, a single sentence can convey multiple propositions (*Those nice red apples cost a lot* expresses the propositions *the apples cost a lot, the apples are red*, and *the apples are nice*; Crystal, 2003, p. 377). Some methods of defining propositions in the literature are more specific: for instance, Rupp et al. (2001) operationalized propositional density as the number of phrases in a passage containing a noun + attributive adjective + prepositional phrase (e.g., *the fluffy cat on the table* or *that was a good suggestion you made in the meeting*). A proposition is often the information that listeners remember from a text even when they cannot recall the exact wording of the presented utterances (Eom, 2006). Propositional density can be calculated by counting the number of propositions in a passage and dividing by the total number of words or the duration of the passage (Bejar et al., 2000; Rupp et al., 2001).

In Rupp et al.'s (2001) analysis, propositional density (categorized as beginner, intermediate, and expert based on the number of propositions per 100 words) emerged as an important factor in determining item difficulty for listening passages and reading texts, with greater density indicating greater difficulty. However, Rupp et al. did not examine the effects of propositional density separately for listening and reading comprehension, so it is possible that the strength of the relationship between density and item difficulty differed for reading and listening test materials. Another interesting finding uncovered in this study was that the reading texts had significantly higher propositional density compared to the listening passages. While this finding may be due to their particular sample of passages and texts, it has been argued that speakers' cognizance of the need for their listener to process information online generally leads spoken messages to be less propositionally dense (Shohamy & Inbar, 1991). If this finding applies to spoken and written texts generally, it suggests one reason why converting written texts into listening materials may be problematic: higher propositional density in a passage increases listening difficulty, and this type of density will tend

to be higher for written texts presented aurally than passages originally delivered aurally (see also Gilmore, 2004). This possibility requires more investigation in future studies.

Other methods of calculating propositional density also refer to redundancy of information in a passage like the type/token ratio mentioned in the previous section. Bejar et al. (2000) described two ratios capturing different aspects of propositional density: the ratio of unique propositions to the count of total propositions in the passage and the ratio of total propositions to passage duration. Because of the way the first ratio is calculated, if a passage includes the sentences *The capital city of the state of Colorado is Denver. So, you see, Denver is a state capital*, the proposition containing the information about Denver being a state capital, given in both sentences, would be counted only once towards unique propositions, and twice towards total propositions.

The two ratios proposed by Bejar et al. (2000) provide far more information about a passage than its total number of propositions (as in Rupp et al., 2001). However, they are better characterized as separate measures of density and redundancy than as both describing propositional density (though they are both described as measures of propositional density by Bejar et al.). For instance, a small ratio of unique to total propositions would indicate a large amount of redundancy, a characteristic likely to make the passage easier to comprehend (e.g., Chaudron, 1983). If this were combined with a large ratio of total propositions to passage length, then there would be a large amount of redundancy in a high-density passage. The ratio of total propositions to passage length in seconds provides a global measure of how quickly information is being presented. The ability to take into account redundancy (unique propositions divided by total propositions) as a separate factor from amount of information presentation (total number of propositions divided by passage length) makes the use of these two ratios preferable to other measures of propositional density that do not take redundancy into account.

A comparison of Bejar et al.'s (2000) ratios for four passages of the same length is shown in Table 1. Though Passage A has lower propositional density than does Passage C, Passage C contains the same number of unique propositions as Passage A due to considerable redundancy. Passage C thus might not exceed Passage A in difficulty, though exactly how information density and redundancy trade off to determine passage difficulty is a topic that should be explored in future studies. Further, despite their equivalent level of density, Passage A is likely to be easier to comprehend than Passage B due to its higher level of redundancy.

Table 1. Using Bejar et al.'s (2000) measures of propositional density and redundancy, passages can have the same overall duration, but different levels of density and redundancy.

Passage label	Unique propositions	Total propositions	Duration (*seconds*)	**Density measure:** Propositions/ Duration	Relative density	**Redundancy measure:** Unique/ Total propositions	Relative redundancy
A	5	10	20 *s*	1 prop/2 *s*	Lower	1 unique /2 total	Higher
B	10	10	20 *s*	1 prop/2 *s*	Lower	1 unique /1 total	Lower
C	5	20	20 *s*	1 prop/1 *s*	Higher	1 unique /4 total	Higher
D	10	20	20 *s*	1 prop/1 *s*	Higher	1 unique /2 total	Lower

While it may seem desirable to combine the ratios described by Bejar et al. (2000) into one by dividing the number of *unique* propositions by the duration of the passage, there may be good reason to separate redundancy and information density as factors. As discussed above in the section on Redundancy, not all types of redundancy may appear to be redundant information for all listeners. In particular, more complex types of redundancy, like paraphrase, may seem redundant to higher-proficiency listeners but not lower-proficiency listeners (Chiang & Dunkel, 1992). A measure of propositional density that includes only *non-redundant* propositions should be used only in those situations where there is good reason to believe that the listeners can appreciate the redundant information as such. For example, if propositions that are *exactly repeated* are treated as redundant, dividing the number of propositions that are *not* exact repetitions by the duration of the passage may provide a better measure of propositional density than does dividing the sum total of all propositions by the duration of the passage.

Summary: Information density

Existing research on how information density impacts L2 listening comprehension indicates that greater density results in greater difficulty (e.g., Rupp et al., 2001). There are several methods of calculating information density in the literature, including some measures that account for the redundancy of information in a passage. Because redundancy tends to result in lower comprehension difficulty (e.g., Buck & Tatsuoka, 1998; Chiang & Dunkel, 1992), but can be confounded with information density in some measurements, future investigations of information density and L2 listening comprehension should be sure to use methods of measuring density that tease apart redundancy and density of information. Further, because information density, redundancy, and passage length are heavily inter-connected, any examination of one of these factors should take the others into account.

Source of characteristic	**Factor of interest**	**Consensus in the literature**
Passage	Information density	A large number of (unique) ideas in a passage has a negative effect

Differences in passage length, information density, and redundancy between authentic and created passages

One study investigated the extent to which passage length, information density, and redundancy differ for authentic and created aural materials. Gilmore (2004) examined transcripts of seven listening exercise dialogues from various language training textbooks, all involving service encounters (e.g., requesting a car from a car rental shop). To collect authentic dialogues, he took questions from the *information receiver* (e.g., the person requesting the car) in these transcripts and used them as a basis for encounters in real situations with *service providers* (e.g., the clerk at the car rental agency).

Qualities of Gilmore's (2004) spoken dialogues

Authentic	**Created**
More words	Fewer words
Higher redundancy	Lower redundancy
Lower lexical density	Higher lexical density

Gilmore (2004) found that authentic dialogues were considerably longer than textbook dialogues (M = 2,764 vs. M = 1,283 words). A principle reason for the difference in length was that the authentic passages contained a great deal more repetition of information than did the textbook passages (M = 24.14 vs. M = 1 occurrence of repetition). According to the findings reported above (e.g., Chaudron, 1983; Chiang & Dunkel, 1992; Gainer, 1997), this quality of the authentic dialogues would make them easier to comprehend than the textbook dialogues.

In addition, Gilmore (2004) found that authentic passages contained lower lexical density than textbook passages. In this case, lexical density was the ratio of words with independent meaning (e.g., *mother* has an independent meaning, whereas *a* does not; Gilmore, 2004, p.367) to the total number of words in the dialogue (similar to the measure of density of content words used by Nissan et al., 1996). The finding that textbook passages contained greater lexical density than authentic dialogues is consistent with the finding of Rupp et al. (2001) that written texts had greater information density than spoken passages, and with the finding of a corpus analysis from Flowerdew (1993) that planned passages (e.g., a radio news story) are more dense than unplanned passages (e.g., a phone conversation between friends). Thus, Gilmore's (2004) findings indicate that textbook passages may inadvertently differ from authentic passages in ways that make L2 listening comprehension more difficult for the aurally presented textbook passages.

Working memory and passage length-related factors

The impact of the listener's working memory capacity on listening comprehension is likely to be affected by the factors discussed in this section. In general, the more information that must be held in working memory during a task, the greater are the demands on working memory (Baddeley & Hitch, 1974). Greater information density should thus increase working memory demands during listening comprehension and leave less working memory processing capacity for other tasks, such as noting a particular character's name. Redundancy, because it reinforces already-given information, may decrease working memory demands, particularly if the form of redundancy has low complexity and high salience or if the listener has higher L2 listening proficiency. These possible interactions between density and redundancy and working memory demands provide another reason why these factors should be considered in conjunction with passage length.

Overall summary of passage length and related factors

Few studies present evidence that passage length itself increases L2 listening comprehension difficulty. Before discounting the role of passage length in listening difficulty, however, it is important to note that work in this area is limited in some ways. For instance, the range of passage lengths examined is often narrow, and some studies have uncovered differences in the effects of passage length on listening comprehension for lower- and higher-proficiency listeners. Some studies that do find evidence of a relationship between listening difficulty and passage length find that the amount of information surrounding the item-relevant information (in a testing context) has an effect, whereas overall passage length does not.

This section also explored two factors with which passage length is highly likely to be confounded: redundancy of information and information density. In the discussion of redundancy and its effects on L2 listening comprehension, it is clear that the effects are generally positive. More redundancy of the information in a passage decreases listening comprehension difficulty. This is true both when redundancy is examined for item-relevant information only, and when it is examined more broadly for key ideas or main points in a passage. These results suggest that, to the extent that a passage is longer because it contains more redundant information, the passage should not increase listening comprehension difficulty. However, the type of redundancy (synonyms vs. exact repetition of words), the L2 proficiency of the listener, and other factors that affect the difficulty of the passage (e.g., speech rate) should be taken into consideration, as these affect how redundancy influences difficulty.

The second factor confounded with passage length is information density. In contrast to redundancy, the findings for information density indicate that when density is higher, L2 listening comprehension is more difficult. To the extent that a passage is longer and has greater density, the passage will result in greater difficulty in listening comprehension. Further, while several measures of information density have been employed and suggested in the literature, those that take redundancy of information into account are likely to be preferable, particularly since redundancy and information density have opposing effects on difficulty.

Finally, a study examining differences between authentic and textbook dialogues shows that all the factors discussed in this section (length, information density, and redundancy of the passage) may differ depending on whether the passage is authentic or created. Care should be taken that L2 listening comprehension is not made more difficult due to decreased redundancy or increased information density in created passages.

Passage complexity

While the section above focused on factors concerning the amount of information a listener must process to comprehend a passage, the factors in this section relate to how challenging the information in a passage is to process. While two passages may share the same degree of redundancy and level of information density, they may differ a great deal in other factors that can impact L2 listening comprehension, such as their demand for pragmatic knowledge, the concreteness of the information described, and the syntactic structure of the utterances.

Syntactic complexity

Measures of ***passage complexity*** refer to such dissimilar properties as syntactic structure, concreteness, and word frequency.

Additional measures appeal to the extent to which a listener must use pragmatic knowledge (e.g., culture, context).

One way of measuring the complexity of a passage is to consider structural elements of the phrases and sentences of the passage, or its syntactic structure. The studies that have investigated the impact of the syntactic structure of a passage have done so by considering the degree of subordination (Blau, 1990; Cervantes & Gainer, 1992; Pica et al., 1987), the number of negatives (Kostin, 2004; Nissan et al., 1996; Yanagawa & Green, 2008; Ying-hui, 2006), the number of dependent clauses (Kostin, 2004; Ying-hui, 2006), or the number of references (Kostin, 2004; Ying-hui, 2006) in the passage.

Sentence structure

Blau (1990) investigated whether simplifying syntax or including surface clues for more complex sentences would affect L2 listening comprehension and found no significant effect of these manipulations. These results imply that modifying sentence structure (in terms of simplifying syntax and including cues to underlying structure) of aural passages does not impact second language learners' listening comprehension. Pica, Young, and Doughty (1987) also explored the effect of syntactic modification on listening comprehension for scripted instructions. The results revealed that participants hearing scripts with lower syntactic complexity, but without interaction with the speaker, did not perform better than those exposed to unmodified scripts and the opportunity to interact with the speaker.

Simplifying sentence structure does not consistently improve comprehension.

Some studies have found that L2 listeners benefit from syntactic simplification, however. Cervantes and Gainer (1992) found that listeners hearing a syntactically simplified version of a lecture scored significantly higher on a recall test than did listeners hearing a more complex version of the lecture. In a second study, they replicated their first findings, but found a similar improvement in comprehension when a more complex version of the lecture was played twice. While these results indicate that syntactic simplification can improve listening comprehension, they also suggest that simplifying the syntax of a passage may not be necessary if listeners can hear a passage more than once.

Negatives, dependent clauses, and referentials

Several studies have investigated the impact of additional features related to syntactic complexity on the difficulty of listening comprehension test items. These studies suggest that negatives (e.g., negative markers like *not* and negative prefixes like *un-*) may play a role in listening comprehension, but suggest less of a role for features like dependent clauses or referentials.

Nissan et al. (1996) found that the difficulty level of an item was significantly higher when the number of negatives in the corresponding passage was greater than one.[18] Kostin (2004) also explored the effect of negatives in a passage on item difficulty, in addition to the effects of referentials and dependent clauses, and distinguished between number of negatives in the first speaker's utterance and the number in the second speaker's utterance for dialogue passages.[19] She found that the presence of two or more negatives in the entire passage increased item difficulty and that a greater number of negatives in the utterance of the second speaker, but not in the utterance of the first speaker, was associated with difficulty for dialogue items. However, she included so few dialogue passages in the analysis (Kostin, 2004) that this result might be due to the particular dialogues she examined. Kostin (2004) did not find a relationship between dependent clauses or referentials in the passage and item difficulty.

Negatives may have a detrimental impact on L2 listening comprehension, but the effect is unlikely to be strong.

Other, similar analyses of test passages failed to uncover any overall relationship between negations in the passage and item difficulty (Yanagawa & Green, 2008; Ying-hui, 2006). However, Yanagawa and Green (2008) did find that negatives affected item difficulty for certain items, such that items where the *correct* answer contained many of the same words as the passage were *more difficult* when there were more negatives in the passage, and items where an *incorrect* answer contained a lot of the same words as the passage were *easier* when there were more negatives.[20] Possibly listeners understood that some information in the passage had been negated, but were unsure as to what information the negation applied to, and so avoided answers having a lot of overlap with the passage. In addition to negatives, Ying-hui (2006) also explored how dependent clauses and referentials in the passages affected item difficulty, but found no relationship.[21]

Summary: Syntactic complexity

The results from the existing literature on the effect of syntactic complexity on listening comprehension are mixed. Regarding overall syntactic complexity, Blau (1990) and Pica et al. (1987) both concluded that simplifying the syntactic structure of an aural passage does not improve second language learners' listening comprehension. Cervantes and Gainer (1992) found that learners hearing a syntactically simplified passage performed better on a recall test than those hearing an unmodified version of the passage, but also that hearing the unmodified passage a second time improved comprehension as much as hearing the simplified version. It is important to observe, however,

[18] Data for Nissan et al.'s (1996) analyses were taken from TOEFL test results, so examinees would be expected to represent a range of proficiency levels.

[19] Kostin's (2004) data were taken from several sets of post-1995 TOEFL test results, so examinees would be expected to represent a range of L2 listening proficiency levels.

[20] Participants for Yanagawa and Green's (2008) study were recruited from the Test of English As International Communication examinee pool. People from this pool should represent a range of proficiency levels.Yanagawa and Green used prior TOEIC scores as a covariate in their analyses to control for L2 listening proficiency.

[21] Data for Ying-hui's (2006) analyses were taken from National College English Test of China (CET) test results for 1000 randomly-selected examinees. These examinees likely represented a range of L2 proficiency levels.

that the studies exploring the effects of syntactic complexity on L2 listening did not explore how *increasing* the complexity of a passage (from some baseline text) impacted comprehension, but rather how *simplifying* syntax affected comprehension. Simplifying syntax may not be a good strategy to make passages more comprehensible for L2 listeners, but increasing the complexity of syntax may nonetheless hurt their comprehension.

Negatives in the passage may increase item difficulty, but the effect is unlikely to be strong. While the results from Nissan et al. (1996) and Kostin (2004) imply that the presence of two or more negatives may increase item difficulty, and those of Yanagawa and Green (2008) suggest that negatives may increase the difficulty of particular kinds of items, Ying-hui (2006) failed to find any effect of negatives on difficulty. None of the studies investigating referentials or dependent clauses found a relationship between these factors and item difficulty, so there is no evidence to date that these factors will impact L2 listening comprehension.

Concreteness

Another characteristic of a passage that contributes to its complexity and may influence item difficulty is whether a passage is concerned with concrete entities or objects (*concreteness*). In the existing literature, a passage has been considered to be concrete if there is a concrete object in the passage (Nissan et al., 1996) or if the main idea of the text and its development are concerned with concrete entities (Freedle & Kostin, 1992, 1993).

Concreteness refers to whether a passage or text is concerned with concrete entities or objects versus abstract concepts.

In reading comprehension, Freedle and Kostin (1992) found that texts in which the main idea and its development are concerned with concrete entities rather than abstract entities lead to easier inference and explicit statement comprehension items for L1 readers. In a similar study with L2 readers, Freedle and Kostin (1993) again found that greater concreteness in a text made its items easier.

However, not all conceptualizations of concreteness indicate that concreteness should make comprehension easier. Nissan et al. (1996) proposed that passages that involved references to concrete objects may be *more* difficult because L2 listeners may need to recognize the existence of the object in the setting of the passage in order to comprehend the passage. However, although they found a slight trend for passages that *did not* refer to concrete objects or entities to have easier items, there was no significant difference between those passages that referenced a concrete object in the speakers' shared environment and those that did not. Further, their method of defining *concrete* was less about concrete versus abstract objects, and more about whether the listener had to understand that there was a particular object in the speakers' context to make sense of what was being said. This study's findings thus may not bear directly on how concreteness affects L2 listening comprehension difficulty.

Although there is almost no research on the effects of concreteness in L2 listening comprehension, findings in reading comprehension for L1 and L2 readers suggest that more concrete passages and texts should be easier to comprehend. Further, other research indicates that concreteness of words might impact comprehension through demands on working memory. The translations of more concrete L2 words are easier to recover from memory because they are usually easier to define and to contextualize (e.g., De Groot & Poot, 1997). Comprehending concrete

words should thus require fewer working memory resources, and passages that contain more abstract L2 lexical items will impose additional working memory demands on the L2 listener, possibly injuring comprehension.

Directness of text

The level of directness of a passage is another factor that may impact its overall complexity and, therefore, the difficulty of the passage. Passages that are indirect tend to contain more implied information, requiring listeners to make inferences in order to comprehend the meaning of the passage. Conversational implicatures are one form of indirect communication that require the listener to infer what the speaker thinks and feels in order to arrive at the correct interpretation of the passage (Grice, 1975). There have been several studies investigating the comprehension of conversational implicatures by second language learners (Garcia, 2004; Taguchi, 2005, 2008).

Passages with implied meaning can be more difficult to understand, particularly at lower proficiency levels, as acknowledged in the ILR scale.

Ability to comprehend implied meanings seems to be a quality of advanced L2 proficiency. Garcia (2004) compared the comprehension of conversational implicatures (e.g., turning down an invitation by describing a conflicting obligation rather than outright refusal) by higher and lower proficiency L2 listeners and found that higher-proficiency learners performed significantly better than the lower-proficiency learners. Taguchi (2005) similarly found that L2 proficiency was a significant predictor of performance on a conversational implicature comprehension task. This result implies that the ability to comprehend conversational implicatures may be related to the level of the listeners' second language proficiency. In another study, Taguchi (2008) again found a significant correlation between L2 listening proficiency and accuracy in comprehending conversational implicatures. Further supporting the idea that experience with the L2 improves the ability to comprehend implicature, findings from this study revealed that participants were more accurate and faster in responding to more conventional implicatures (refusals) than less conventional implicatures (opinions, which vary more in how they are expressed). Taguchi argued that refusals, which include more common patterns of discourse within the L2, may be easier to comprehend because they require less processing effort compared to opinions.

The results from the existing literature suggest that L2 proficiency may be related to a learner's ability to comprehend the implied meaning in indirect forms of communication such as conversational implicatures. This conclusion is further supported by findings that more conventional implicatures (which language learners are more likely to have exposure to) are easier to process than are less conventional ones. Although little research has specifically examined the impact of directness of a passage on comprehension, the relationship between directness-related factors and overall L2 proficiency suggests that lower-proficiency listeners may have difficulty when a passage contains a large number of implicatures or other indirect forms of communication, particularly if they are non-conventional.

Infrequent words

The occurrence of infrequent words in a passage contributes to its complexity and may impact item difficulty. Infrequent words in a passage may impact listening comprehension item difficulty because examinees are less likely

to be familiar with low-frequency words, and so they may need to infer the meaning of any low-frequency words in a passage (for a more in-depth discussion of this process, see the section on Proficiency and Experience with the Second Language). The studies that have investigated the impact of infrequent vocabulary on item difficulty have utilized lists such as Berger's (1977) list and JACET 8000 (2003) to determine word frequency.

Infrequent words have a negative impact on L2 listening comprehension.

Nissan et al. (1996) found that the frequency of the words (determined using a list including 100,000 words compiled by Berger, 1977) in a passage was related to item difficulty: the difficulty of dialogue items corresponding to passages that contained words not on the word frequency list was greater than the difficulty of the items for passages that only contained words on the list. This result implies that infrequent vocabulary may impact the difficulty of dialogue items. Further, the authors noted that certain high-frequency words that would be used on a university campus, such as *semester* and *textbook*, were not included on Berger's list, which may have actually weakened the relationship between word frequency and item difficulty. Kostin (2004) considered both word frequency (using the same measure as Nissan et al.) and whether knowledge of the meaning of the infrequent word was necessary in order to answer the item correctly. The results of this study showed no significant relationship between the presence of an infrequent word in the passage and item difficulty, but there was a significant correlation between item difficulty and the presence of an infrequent word relevant to answering the item correctly. These results suggest that it may not be the mere presence of an infrequent word that impacts the difficulty of a dialogue item, but whether the infrequent word is relevant to responding correctly to the item.

Some evidence shows that less-frequent words *decrease* the difficulty of a passage, but this is likely due to the way *frequent* is defined. The impact of the presence of infrequent vocabulary (determined using JACET 8000, 2003[22]) in a passage on item difficulty was examined in Yanagawa and Green (2008). Their results showed that the presence of infrequent vocabulary in a text affected item difficulty, but unlike the findings from Nissan et al. (1996), it was found that less-frequent vocabulary in the passage was associated with *easier* items. This result is counterintuitive and the authors suggested that some of the words that were classified as being infrequent based on the JACET 8000 word list might have actually been more familiar to the test takers in their study than were the more-frequent words.

Research also points to a relationship between word frequency and working memory, at least for L1 listeners: listeners take longer to process low-frequency words relative to high-frequency words (e.g., Ferreira, Henderson, Anes, Weeks, & McFarlane, 1996). Thus, if a passage contains a number of low-frequency words, this may impose additional demands on L2 listeners' working memory as they attempt to recognize the low frequency word (i.e., access the lexicon). Past research shows that a greater working memory processing load will complicate listening comprehension (e.g., Baddeley & Hitch, 1974).

[22] JACET 8000 is a word frequency list for Japanese learners of English.

The bulk of research on word frequency and L2 listening comprehension suggests that infrequent vocabulary interferes with listening comprehension (Nissan et al., 1996). Whether understanding an infrequent word is important to answering a test item correctly has been found to have a significant relationship with difficulty when overall vocabulary frequency did not (Kostin, 2004). Thus, it is possible that earlier studies examining the relationship between vocabulary frequency and item difficulty were inadvertently measuring the effect of infrequent information that was necessary for answering the item rather than overall frequency of words in the passage. Yanagawa and Green (2008) uncovered conflicting results, but the authors themselves argued that this finding might have been a result of the word list they used to determine frequency. However, all the studies discussed in this section have the limitation of using word lists to determine word frequency that may not be as reliable as originally expected; this limitation should be taken into consideration when interpreting the findings.

Culturally specific vocabulary and idioms

Another characteristic of a passage that may have an impact on its level of difficulty is whether the passage contains any culturally specific vocabulary or idioms, requiring the examinee to have some pragmatic knowledge in order to understand the passage and respond correctly to the associated items. Several studies have investigated this factor, by either considering the number of culturally unfamiliar words (Nissan et al., 1996; Sasaki, 2000) or the number of idioms in a passage (Kostin, 2004; Ying-hui, 2006).

Sasaki (2000) found that L2 readers completing cloze tests containing culturally familiar words (e.g., names more common in the culture of the L1) showed correct understanding of the key terms more often, tried to solve more items, and generally understood the text better than readers completing cloze texts with culturally unfamiliar words (e.g., names more common in the culture of the L2). L2 readers completing the cloze tests with culturally familiar words also performed better on the items requiring within-sentence information (e.g., information provided by the clause or sentence in which the item appeared). This result did not seem to extend to comprehension beyond the sentence level, however (e.g., information provided by the context of the paragraph containing the item or the entire text). These results indicate that replacing culturally unfamiliar words in a text with culturally familiar words increases examinees' understanding of within-sentence information.

Research has also been conducted on the effect of culturally specific words or idioms on listening comprehension. Nissan et al. (1996) found that the presence of culturally specific vocabulary in a listening passage could not be used to predict difficulty of the associated items. However, only a few items required comprehension of culturally specific vocabulary. The authors suggest that this is because test designers are sensitive to the fact that examinees may not be familiar with certain aspects of the L2 culture, and so they often include other clues in the passages.

Kostin (2004) explored the effect of idioms in the passage on listening comprehension. The American Heritage Dictionary (2000) defines *idiom* as "an expression consisting of two or more words having a meaning that cannot be deduced from the meanings of its constituent parts." An example would be *snake in the grass*—an expression that cannot be understood even if the L2 listener is familiar with the meanings of *grass*, *in*, and *snake*. In Kostin (2004), analyses uncovered a positive correlation between whether a passage contained an idiom, the meaning of which was

central to answering the test item, and item difficulty. Again, this finding highlights that it is not simply the information in the passage that determines item difficulty, but whether understanding that information is important for answering a test item.

Although few studies have examined the effect of culturally specific vocabulary or idioms on L2 listening comprehension, results suggest that L2 listeners have a harder time comprehending a passage containing these types of features.

Source of characteristic	Factor of interest	Consensus in the literature
Passage	Complexity	This factor corresponds to several distinct features, e.g., syntax, directness, concreteness, and word frequency. Directness and word frequency have the strongest effects. Passages are harder when they require inferencing from the listener, as with indirect speech or unfamiliar vocabulary.

Passage type and organization

Passage topic

Topic of the passage is another characteristic that may affect how well L2 listeners comprehend the passage. In general, passages about familiar topics are easier for L2 listeners to comprehend than are passages about unfamiliar topics (Sadighi & Zare, 2006; Tyler, 2001). Exposure to information about a topic prior to listening to a passage about that topic improves comprehension for higher- and lower-proficiency L2 listeners (measured through the Test of English for International Communication; participants with scores ≥ 40 were classified as higher proficiency, those ≤ 39 were classified as lower proficiency; Chang & Read, 2006).

Another factor that may affect L2 listening comprehension is whether a passage is on an academic or non-academic topic. The relationship between this factor and L2 listening comprehension difficulty has been explored in two studies (Buck & Tatsuoka, 1998; Ying-hui, 2006), but only Buck and Tatsuoka found a significant relationship between topic type and difficulty (non-academic topics were associated with easier items). However, it is difficult to say from this study's findings exactly what about academic topics might make them more difficult than non-academic topics for L2 listeners. Many factors believed to constitute differences between passages covering academic and non-academic topics are covered in other sections of this review: required background knowledge, ability to distinguish between relevant and irrelevant information, amount of implied meaning, ability to cope with long passage lengths, and note-taking demands (Ferris & Tagg, 1996). The key to the effect of academic versus non-academic topic on listening comprehension may be one of these factors rather than a benefit provided simply by the non-academic topic itself.

In terms of more general differences between passages addressing different topic matter (e.g., humanities vs. mathematics lectures), differences of structure have been the focus of the relevant literature. For instance, Coulthard and Montgomery (1981) analyzed university science lectures to determine their structure and found that lectures are

composed of free-standing informative syntactic clauses, which are combined into *sequences*, indicated by prosodic phrases. These sequences are combined to make *transactions*; each lecture is composed of a number of transactions. While that structure seems to hold across a number of scientific disciplines, Rounds (1987) analyzed mathematics lectures and found a different discourse structure in which major points were named and explicitly marked by the teacher as being relevant for evaluation. Cohesion in mathematics lectures is developed using repetition and links to previous concepts, and topic changes are explicitly marked. Topics are presented in a very organized fashion using persuasion, with a question and answer format incorporated at appropriate points in the discourse. Further, the structures of scientific and mathematic lectures are quite different from those in the humanities where topics are developed from a variety of different perspectives with different interpretive frames (Strodt-Lopez, 1991).

Passage type

Lectures and recorded conversations

Different types of spoken passages have different structure, which can make them easier or more difficult for L2 listeners to understand.

Structural differences can also be found between different types of passages. The research on passage types has focused, for the most part, on lectures and conversations, which have very different structures. In a conversation, participants are allowed to ask for repetition and clarification, they follow turn-taking conventions, and they often do not distinguish between relevant and irrelevant information (Flowerdew, 1994). Furthermore, conversations do not necessarily require specialized knowledge or an understanding of implied and indirect speech acts (Flowerdew, 1994). They are by nature informal, contextualized, and involved (Biber, 1988). Furthermore, conversations have more repairs, negotiation of meaning, confirmation checks, and back-channel cues, all of which improve communication (Chaudron, 1988).

On the other hand, when attending to a lecture, listeners hear long stretches of uninterrupted speech, without the opportunity to take turns or clarify, and they must be able to distinguish between relevant and irrelevant information (Flowerdew, 1994). Furthermore, lectures generally require specialized knowledge, though they do not require that the listener be able to understand implied and indirect speech acts (Flowerdew, 1994). They are by nature formal, elaborated, decontextualized, and detached (Biber, 1988). Furthermore, lectures have more complicated syntax including *that* clauses, subordinate clauses, subordinate conditional clauses, first and second person pronouns, contractions, and the pronoun *it* (Tyler, Jeffries, & Davies, 1988), not to mention that the speakers often assume that listeners have prepared for the lecture by reading relevant material in advance.

However, even lectures can vary dramatically in style. Dudley-Evans and Johns (1981) identified four types of lectures, including (1) formal lectures, which are read from written copy (also called *reading-style lectures*); (2) less formal conversational-style lectures; (3) rhetorical lectures in which the lecturer acts more as a performer using a wide intonational range, many digressions, and shifts in key and tempo; and (4) participatory lectures (described by Benson, 1989), in which the lecturer interacts with the audience, asking questions and soliciting input. The first three

types of lectures are monologues in which the listener is not encouraged to participate; the last is processed by the speaker in real time and not composed off-line completely ahead of time.

Another finding that bears mentioning here is that comprehension of academic lectures is best when there is explicit discourse structuring indicated with terms like *First let's look at....* or *What I will do now is....* (Camiciottoli, 2004). As mentioned in the discussion of discourse markers, discourse macro-markers such as these enhance comprehension (Chaudron & Richards, 1986). Further, these types of markers are more common in college-level lectures than in college-level textbooks, likely due to a high need for organization and structure in this complex type of spoken communication (Biber et al., 2004). This may be why reading-style lectures are more difficult to understand: Coulthard and Montgomery (1981) observed that reading-style lectures have fewer of these types of markers.

Much effort has been devoted to characterizing the structure of lectures in different disciplines. Sadly, little research has examined whether one style is easier to understand than the other. Some studies, however, have examined the impact on listening of factors correlated with the different lecturing styles. This work is discussed below.

Orality

L2 listeners have less difficulty understanding passages that are more ***oral***. Such passages have simpler syntax, more disfluencies, and greater redundancy.

One quality found to differ between some passage types (e.g., dialogues versus monologues) is the degree of orality. *Orality* is the extent to which a passage contains features of spoken language as opposed to features typical of written language, with highly oral passages tending to contain more disfluencies and redundancy, and simpler syntax (Tannen, 1982). This factor differs between spoken passage types: a spontaneous dialogue between friends would be highly oral, whereas a formal lecture would be less oral (Inbar, 1988; Shohamy & Inbar, 1991).

Orality also tends to differ between authentic and created passages: authentic passages tend to have more oral qualities, in that they tend to contain more aspects of spontaneous language (e.g., ellipses, redundancy, pauses; Shohamy & Inbar, 1991). The two terms should not be understood as interchangeable, however: both a spontaneous dialogue and a formal lecture are authentic by the definition adopted here. Further, a section from an authentic text read aloud can be highly *authentic*, in the sense that the text was created by a real writer/speaker for real listeners/readers to convey a real message (Breen, 1985; Rings, 1986), and also be low in *orality*, in that it would have few characteristics of spontaneous spoken language (Brindley & Slatyer, 2002).

Greater orality in a passage seems to improve L2 listening comprehension. Shohamy and Inbar (1991; see also Inbar, 1988) explored the effect of passage orality on listening comprehension. Passages in this study included the same basic factual information, but differed on a number of factors relevant to orality: the extent to which they were planned (less planning = higher orality), potential for interaction between the speaker and the listener (more potential for interaction = higher orality), redundancy (more redundancy = higher orality), disfluency (more disfluency = higher orality), and sentence complexity (less complexity = higher orality). Degree of orality affected comprehension,

Feature	Higher orality	Lower orality
Planning	Less	More
Potential for interaction with speaker	More	Less
Redundancy	More	Less
Disfluency	More	Less
Sentence complexity	Less	More

with the passage with lowest orality (a news broadcast) producing the lowest overall performance, followed by the medium-orality passage (a lecture), and the high-orality passage (a dialogue) producing the highest overall performance, although there was no significant difference between comprehension of the dialogue and the lecture.

In contrast, Brindley and Slatyer (2002) examined performance with a monologue passage (lower orality) compared with a dialogue passage (higher orality) and found no difference between the two, although the authors stated that comprehension of the dialogue may have been complicated by a higher relative speech rate. Kiany and Jalali (2006) found significantly superior comprehension for a dialogue compared with a monologue and attributed the difference to the greater orality of the dialogue passage. It is important to note, however, that dialogues and monologues also differ in the number of speakers. Thus, it may be that the improved comprehension observed in Kiany and Jalali (2006) was due to differences in the number of speakers and not due to the differing oral qualities of the dialogues. However, Shohamy and Inbar (1991) found no significant difference in performance between participants listening to a lecture and those listening to a dialogue; if the number of speakers determined their effects, comprehension of the dialogue would have been superior to that of the lecture. Further, Brown and Yule (1983) found that when there are more speakers it is more difficult to understand a passage. It will be important for future studies to control the number of speakers when investigating the effect of orality.

In relation to the discussion of lectures in the previous section (Dudley-Evans & Johns, 1981), orality is likely to differ across different lecture types. Formal lectures delivered from written material would have lower orality than would the other types of lectures (e.g., rhetorical or participatory) because they are more planned and involve fewer qualities of spontaneous language (Shohamy & Inbar, 1991). Further, conversations will be more oral than lectures of any type, because of their greater potential for interaction between the speaker and the listener and lower formality (Biber, 1988; Flowerdew, 1994). Findings regarding the effects of orality on L2 listening comprehension suggest that non-native listeners would have greater difficulty with lectures delivered from written material than they would with lectures that were less planned, and greater difficulty with lectures than with conversations.

Rhetorical structure

Even though researchers have identified a number of different types of rhetorical structures (Meyer & Freedle, 1984), few studies have been devoted to determining whether one type is easier for listeners to comprehend than another. There are a few exceptions. Meyer and Freedle (1984) studied listening passages on the TOEFL and identified five types of rhetorical structures. The structures fell along a single dimension of *degree of organization*, with each succeeding structure incorporating qualities of the preceding ones. Listed in order of degree of organization, these are description, collection, causation, problems/solution, with the final structure, comparison, having the

potential to vary in its degree of organization (Meyer & Freedle, 1984). Meyer and Freedle (1984) found that *L1 listeners* recalled the fewest number of idea units from a passage presented as a description, the least organized of the rhetorical structures. While their findings suggest that passages with less organized rhetorical structures will be more difficult for L2 listeners to comprehend, it is possible that L1 listeners were sensitive to differences in organization to which L2 listeners would not be sensitive.

Studies examining the relationship between rhetorical structure and L2 listening comprehension do not provide strong evidence that particular structures are easier to comprehend than others for non-native listeners. Ying-hui (2006) examined passages with description, comparison, and causation structures, but did not find any relationship between rhetorical structure and difficulty. Freedle and Kostin (1996) found that items associated with less organized structures (according to Meyer and Freedle's [1984] characterization) were actually easier: items for passages with a list structure were easier than those for other structures, and items for passages with a comparison or problem/solution structure were more difficult. However, Freedle and Kostin (1996) did not directly manipulate the rhetorical structure of the passages or their associated comprehension items (they were taken from a selection of testlets), leaving open the possibility that other factors, such as type of test item, were confounded with rhetorical structure. For example, list passages may have had fewer inference test items, which have been found to be more difficult than main idea or detail items (Nissan et al., 1996). Thus, it is difficult to conclude based on Meyer and Freedle (1984), Ying-hui (2006), and Freedle and Kostin (1996) exactly how, if at all, L2 listening comprehension is affected by the rhetorical structure of the passage. However, other studies have found that certain rhetorical structures are associated with a greater number of syntactic features like negations. For example, causation passages contained more negations than did the comparison passages (Ying-hui, 2006). As there is some evidence that negations increase listening comprehension test item difficulty (see the section on Syntactic complexity), certain types of rhetorical structures which include more negations may be more difficult.

While evidence that particular rhetorical structures are easier to comprehend than others is weak, another possibility is that familiar structures are easier for L2 listeners to comprehend than are unfamiliar structures. Yang (2007) presented one of two passages to a group of L2 listeners: one with a structure common in their L1, and the other with a structure common in the L2. Analyses of the listeners' notes revealed no differences for the group hearing the familiar structure and the one hearing the less familiar structure. However,the author pointed out that requiring the participants to take notes (as this was the source of the dependent measure) may have interfered with their comprehension of both passages, preventing the discovery of an advantage for the familiar structure. This is in line with the findings of Hale and Courtney (1994), who found that participants compelled to take notes actually showed inferior listening comprehension relative to when they were not allowed to take notes at all (see section below on Note-Taking).

Position of item-relevant information

Some research supports the idea that position of the information in a passage predicts comprehension difficulty for that information. Freedle and Kostin (1996, 1999) found that when the information required by comprehension items occurred either early in the passage or in the last sentence of the passage, the item was generally easier.

Conversely, when item-relevant information was located in the middle of the passage, items tended to be more difficult. Rupp et al. (2001) failed to find the same relationship between position and difficulty, but suggested that this may have occurred because they allowed their examinees to listen to each passage as many times as they wished. Examinees could have approached additional hearings of the passage with the specific goal of attending to the information they had failed to comprehend the first time, removing any effect of position on comprehension. Yang (2007) also found a significant effect for the location of the proposition in the passage, with participants being more likely to recall a piece of information if it occurred near the end of the passage than if it occurred in the middle or at the beginning. Yang's findings are consistent with classic research showing a higher probability of recalling recent information (i.e., the recency effect) than information that occurred first (i.e., the primacy effect), though both types of information have a higher chance of being recalled than information occurring in the middle (Murdock, 1962).

Passage type, passage organization, and working memory

Some research suggests passage organization or type may impact listening comprehension because of an effect on working memory load. Presenting information in a more organized fashion makes this information easier to encode and maintain in working memory (Anderson, 2004; Baddeley, Lewis, Eldridge, & Thompson, 1984). Further, the relationship between working memory capacity and tasks involving reading comprehension or recognition differs depending on whether the topic is familiar or unfamiliar (Leeser, 2007). Findings such as these indicate that the role of working memory in listening comprehension is likely to be affected by the organization of the passage and its topic. When the passage topic is unfamiliar or its content is less organized, listening comprehension may be more difficult.

Passage organization is likely to interact with working memory in L2 listening comprehension.

Summary: Passage type and organization

The effects of passage type on L2 listening comprehension are as diverse as the conceptualizations of passage type itself. Findings regarding orality and listening comprehension show that, when there is a difference, more oral texts like dialogues are easier for L2 listeners than more literate texts like news reports. The results regarding the influence of rhetorical structure on listening comprehension are less consistent. While L1 listeners show the best recall for information from more highly organized rhetorical structures (e.g., causations rather than lists; Meyer & Freedle, 1984), L2 listeners either fail to show any difference in listening comprehension with different rhetorical structures (Yang, 2007; Ying-hui, 2006) or show better performance for less organized structures, such as lists, than more organized structures, such as comparisons (Freedle & Kostin, 1996). Other factors described below, such as the coherence of the passage, possibly including the number or type of discourse markers, may serve as better predictors of listening comprehension difficulty than the rhetorical structure of the passage. Indeed, Freedle and Kostin (1993) argued that examining the effect of the rhetorical structure of a passage on comprehension

Factors such as ***coherence*** and the use of ***discourse markers*** may be better predictors of listening comprehension difficulty than rhetorical structure.

indirectly assesses the influence of different types of discourse markers on comprehension because specific discourse markers occur for each structure (e.g., lists use *and* and *then*; comparisons use *however* and *yet*).

Coherence and relevance

Some factors have been described in the literature examining language comprehension which attempt to capture how coherent or cohesive a particular passage is, and how this factor relates to comprehension difficulty. Carroll (1986) argued that a series of meaningful sentences can nonetheless be combined in a way that makes no sense to the listener or reader if there is no overarching coherence (cited in Dunkel & Davis, 1994, p. 56). Coherence involves the appearance of logicality and relevance in a passage (Odlin, 1989). A passage will seem less coherent to the extent that it lacks strong, logical relationships between its propositions, and this may also be construed as the passage containing many propositions that seem off-topic or tangential (Odlin, 1989).

L1-L2 differences and coherence

Comprehension difficulties associated with coherence can arise from L1-L2 differences. Kaplan (2001) argued that chronic discourse organization differs markedly between English, Russian, Asian, and the Romance languages. Differences in the typical manner of organizing speech could present issues for an L2 listener who comes from a language background with a different typical discourse organization. For instance, in Japanese, texts may be presented in *ki-shoo-ten-ketsu* form, which involves a shift away from the main topic to introduce a subtopic, while English texts are generally more linear, so reading a Japanese text in *ki-shoo-ten-ketsu* form may be difficult and confusing for a reader accustomed to linear texts (Odlin, 1989). Similarly, the preferred style of speaking in Chinese is to put the topic at the end of an expository text, while English speakers tend to state the topic first (Yang, 2007). Thus, while coherence is a factor that may differ between texts or passages (Freedle & Kostin, 1992; Ying-hui, 2006), it may also arise from L1-L2 differences in discourse patterns (Odlin, 1989).

Coherence effects and L2 listening comprehension

Research examining the relationship between coherence and L2 listening comprehension is sparse, and the findings are mixed. Ying-hui (2006) examined the effects of coherence on L2 listening comprehension using a coding method developed by Freedle and Kostin (1992) for reading texts: coherence was defined as a relative rating of to what extent the elements of the first sentence of the passage were represented in the rest of the passage, as compared with the other passages in Ying-hui's sample (scored as 1 = minimal coherence, 3 = maximal coherence). In Ying-hui's study, higher coherence in a passage was associated with easier test items. These results suggest that the overall coherence of a passage might play a role in listening comprehension.

A passage is ***coherent*** when it has logical connections between its propositions. Due to differences in discourse organization norms across languages, a coherent passage could appear incoherent to L2 listeners.

Nissan et al. (1996) examined what they referred to as *local coherence*, the explicitness of the connection between the speakers' utterances. Passages containing explicit lexical links like repetition (e.g., Speaker 1: *What time are you planning on leaving for the airport?* Speaker 2: *I'm leaving for the airport at 5:30.*) or structural links like

anaphora (e.g., Speaker 1: *I hope Henry managed to catch his flight!* Speaker 2: *He just made it*.), were predicted to be easier than those with more implicit links (e.g., Speaker 1: *I heard it's going to snow tomorrow*. Speaker 2: *Oh, no! I'm supposed to fly out at 7pm*.). However, no significant relationship between coherence and item difficulty was found in their study.

The studies examining coherence do not provide strong evidence that coherence affects L2 listening comprehension difficulty. However, there are methodological issues with both studies, as well as in making generalizations across them. In Ying-hui's (2006) study, the operationalization of coherence is highly subjective, and ratings of coherence were made relative to other study passages rather than some general definition of coherence. In Nissan et al.'s (1996) study, the operationalization of coherence did not distinguish between the types of links (e.g., lexical vs. structural). This is an issue because other research in the literature has described the difficulty presented by referentials like *he* and *there* for second language readers (Freedle & Kostin, 1993; Leow, 1993). Some links may have enhanced coherence and comprehension, but others may have made the passage more complex.

Local coherence: The explicitness of the connection between adjacent utterances

Global coherence: The cohesiveness of the entire passage as a unit

Further, the studies above examined two different types of coherence: local coherence and passage-level coherence. It may not be justifiable to equate these two factors. It is easy to imagine a passage that exhibits high local coherence and little overall coherence, such as a recording of a question and answer session with a celebrity. Every individual question and its answer will be strongly connected, but neighboring question and answer pairs may be completely disparate. In future research examining the impact of coherence on L2 listening comprehension, it will be important to distinguish between these two characterizations of coherence.

Discourse markers

An alternative way of examining coherence in passages is through studying the presence of discourse markers. Discourse markers signal the rhetorical structure (e.g., comparison vs. list; Freedle & Kostin, 1993) of a passage, as well as highlight connections between adjacent propositions (Chaudron & Richards, 1986). The category of discourse markers includes macro-markers like *my first point is* and *in conclusion*, which provide clues about the overall structure of the passage, and micro-markers like *yet*, *because*, and *in fact*, which establish links between adjacent utterances (Chaudron & Richards, 1986; Dunkel & Davis, 1994; Flowerdew & Tauroza, 1995).

Certain types of discourse markers, which help establish the structure of a passage and the links between adjacent utterances, may enhance coherence.

Macro-markers are also a type of *lexical bundle*, a frequently occurring sequence of words with widespread use (e.g., occurring 40 times or more per million words) that is not idiomatic (e.g., *do you want to* is a lexical bundle, while *kick the bucket* is an idiom; Biber et al., 2004). Lexical bundles include *stance expressions* that express attitudes or assessments (e.g., *I don't know what*), *referential bundles* that directly reference a physical or abstract entity (e.g.,

one of the things), and *discourse organizers*, which reflect relationships between prior and upcoming discourse (i.e., macro-markers such as *on the other hand*; Biber et al., 2004). In a corpus analysis by Biber et al. (2004), discourse organizers (i.e., macro-markers) were found to occur more frequently in college-level teaching than in conversation or textbooks, likely due to a high need for organization and structure in this complex type of spoken communication. These findings indicate that particular types of passages may have more macro-markers than others, and spoken formal passages may tend to have more than written formal texts.

Research suggests that L2 listeners can recognize lexical bundles. Nekrasova (2009) found that, while native speakers recalled more lexical bundles verbatim from a passage than did lower-proficiency L2 listeners, higher-proficiency L2 listeners recalled even more than native speakers, likely due to the rote memorization of lexical items demanded in the L2 students' language courses (participants were assigned to proficiency groups based on their enrollment status as degree-seeking students of an Intensive English Program). Further, discourse-organizing bundles (i.e., macro-markers) were recognized more frequently than were referential bundles for all three participant groups. These findings show that macro-markers are salient to L2 listeners, and this increases with proficiency, though this salience may diminish as the listener approaches native-like proficiency.

Discourse markers have been found to improve comprehension of aural materials for L1 listeners (Hron, Kurbjuhn, Mandl, & Schnotz, 1985, cited in Jung, 2003, p. 563). These markers appear to benefit L2 listeners as well. Jung (2003) found that L2 listeners who heard a passage containing discourse markers recalled more information than participants who listened to a passage in which most of these markers were removed. However, other research finds different effects for different types of markers. Chaudron and Richards (1986) found that lower- and higher-proficiency participants who heard a version of a passage with macro-markers showed superior performance to those who heard a baseline version without added micro- or macro-markers, a version with only added micro-markers, or a version with both types of markers added; the latter three conditions did not differ from each other. They suggested that macro-markers make a passage more comprehensible, but adding micro-markers may make a passage seem less organized, increasing the listeners' cognitive load without providing useful information.

Dunkel and Davis (1994) contrasted the listening comprehension of native speakers with non-native speakers for lectures which were intact (*evident* condition) or had discourse markers like *first* and *in contrast* removed (*non-evident* condition). No difference in recall was found for native speakers or non-native speakers between those participants listening to the two versions. Although these results seem to contradict the findings of Chaudron and Richards (1986), Dunkel and Davis (1994) did not distinguish between discourse macro- and micro-markers, a distinction found to be important by Chaudron and Richards (1986).

The usefulness of micro-markers is not a closed issue, however. Flowerdew and Tauroza (1995) argued that Dunkel and Davis (1994) included too few micro-markers compared to what occur in authentic passages and too many markers associated with written rather than spoken language. They also pointed out that Chaudron and Richards (1986) specifically inserted micro-markers in such a way as to minimize the semantic information they conveyed (i.e., discourse micro-markers in their study could only be acting as *filled pauses* [a.k.a. *hesitation markers*], see

the Hesitation and Pause section for a more in-depth discussion). Flowerdew and Tauroza (1995) conducted their own study and found that recall and comprehension was lower for a passage from which discourse micro-markers had been deleted compared to one which had not been modified. Flowerdew and Tauroza's (1995) results provide evidence that micro-markers, when allowed to contain semantic information (in contrast to Chaudron and Richard's [1986] study) and when presented in appropriate number and type (in contrast to Dunkel and Davis's [1994] study), improve comprehension for L2 listeners.

Authenticity and coherence

Type and frequency of discourse markers may be one way in which textbook passages differ from authentic passages. Flowerdew and Miller (1997) analyzed a series of authentic passages and found frequent use of a variety of discourse micro- and macro-markers. They argued that textbook passages are often too short to include some of the more global discourse macro-markers, such as those that refer to segments across long sections of a passage, and that the number of discourse markers included in textbook passages can appear too dense and unnatural when an effort is made to include markers (Flowerdew & Miller, 1997). Authentic spoken materials may contain a greater variety of discourse markers and more natural use of these markers than materials attempting to emulate authentic passages.

Authentic passages may contain a greater variety and more natural use of discourse markers than created passages.

Authentic and textbook passages may differ in other ways that are related to coherence. Authentic and simplified reading texts have been found to differ in causal cohesion (the extent to which the elements of the text are connected causally), and the density of logical operators, with authentic texts having greater cohesion and more logical operators (Crossley et al., 2007). These results provide additional evidence that authentic materials may have greater coherence than created materials.

Relevance

Relevance, a factor related to coherence, can be defined as the extent to which the propositions contained in a passage are relevant to the main topic of the passage (i.e., textual relevance, van Dijk, 1978). Alternatively, relevance could be defined more narrowly as the proportion of propositions in the passage that are relevant to the test item at hand (if this item targets the main idea of the passage, these two definitions of relevance should be interchangeable). The former operationalization of this factor should be strongly related to coherence as it has been described above, while the second should be strongly related to redundancy of item-necessary information as examined by Freedle and Kostin (1996).

Relevance: An understudied area related to coherence and redundancy – It is the extent to which the propositions in a passage relate to or bear upon the main topic or particular test item

Relevance and its impact on comprehension is important to consider for listening in particular, as speakers are less explicit about connecting information to a central point or theme than are writers and often rely on the context in which the message is conveyed to provide this information (Smiley, Oakley, Worthen, Campione, & Brown, 1977).

However, a review of the literature failed to uncover any studies directly examining the role of relevance in L2 listening comprehension.

Despite a probable strong relationship between coherence and relevance, the factor of relevance is also likely to be partially independent of the established definitions of coherence. Coherence is related to whether initial topics are carried through the entire passage without regard for whether these are central topics (Ying-hui, 2006), or the degree of connectedness between adjacent propositions (Odlin, 1989). Relevance, on the other hand, takes into account the overarching theme of a passage. However, the presence of discourse markers should have an enhancing effect on both coherence and relevance, as these markers can establish the relevance of a piece of information in a passage to what has already been stated (e.g., *on the other hand* to introduce a counterpoint) and build coherence by highlight the overall structure of the passage (e.g., *my first point is*).

The effect of relevance on L2 listening comprehension is likely to be related to the listener's proficiency. In the same way that lower-proficiency listeners may have difficulty in perceiving that redundant information in a passage actually is redundant (Blau, 1990), these listeners may also have difficulty recognizing that irrelevant information is actually irrelevant. Distinguishing between irrelevant and relevant information is of high importance for listening comprehension, particularly for academic lectures (Flowerdew, 1994). For these reasons, relevance is an important factor to consider in listening comprehension research in general, and that in the L2 in particular.

Source of characteristic	Factor of interest	Consensus in the literature
Passage	Type and organization	Rhetorical structure, one measure of passage organization, does not show strong effects on listening comprehension.
		Spoken passages are easier to understand when they are more oral, more closely resembling spoken rather than written language.
		Discourse markers play a major role in establishing organization and coherence.

Summary: Coherence and relevance

Little research has directly examined the role of coherence in L2 listening comprehension. Overall coherence seems to be associated with improved comprehension, but the operationalization of this factor is a problem. A great deal of research, however, has explored the effects of discourse markers on L2 listening comprehension. These markers have been found to enhance comprehension in English, but there is some disagreement about the extent to which different markers are helpful.[23] Authentic lectures have been found to contain more discourse markers than textbook lectures. In estimating the difficulty of L2 listening passages, it may be important to take the presence of

[23] It should also be noted that Flowerdew and Tauroza (1995), Dunkel and Davis (1994), and Chaudron and Richards (1986) all explored the effect of discourse markers on comprehension for L2 learners of English; it is possible that the usefulness of discourse markers for L2 listeners will vary with the language being learned as well as with the type (micro versus macro) of marker.

discourse markers into consideration: more markers (particularly macro-markers) should make the passage more coherent and easier to comprehend and recall, though this might only be true in English.

Auditory features of the passage

Speaker accent

Comprehending spoken language involves adapting to the idiosyncrasies of a particular speaker (e.g., speaking rate or the pitch of voice). In general, listeners are quite good at this skill, but adaptation becomes considerably more challenging when the speaker has a different accent than the listener (Weil, 2003). Accented speech has been found to affect both the extent to which listeners successfully retrieve a speaker's message and the effort involved as listeners identify particular words in the message (Floccia et al., 2009). In the case of a speaker with a different accent, a listener must cope with variation arising from both the speaker's own idiosyncrasies and additional variation the speaker shares with others from the same linguistic background (Weil, 2003). Such difficulties in adaptation are further exacerbated when the spoken language is not the listener's native language, particularly when proficiency in that language is low.

Accent versus dialect

Oftentimes, the terms *accent* and *dialect* are used interchangeably in the literature to describe phonological differences in speech. For example, General American English may in the same discussion be described both as a different accent than Southern American English or as a different dialect (e.g., Clopper & Bradlow, 2008). Although the two terms are often used synonymously, Crystal (2003) defined dialect as *a regionally or socially distinctive variety of language, **identified by a particular set of words and grammatical structures*** (p. 136, emphasis added), while defining accent as ***the cumulative auditory effect of those features of pronunciation** which identify where a person is from, regionally or socially* (p. 3, emphasis added). In this section, all studies controlled for the words and grammatical structures presented to the listeners in a way that makes describing their manipulations as accent more appropriate than describing them as dialect. For this reason, the term accent will be used exclusively in this section.

Accent: Distinct from dialect, it refers to features of pronunciation that identify where a person is from regionally or socially.

Accent and L1 comprehension

Even when listening to native speakers of their own language, listeners can have difficulty if the speaker has an accent that differs from their own. Ikeno and Hansen (2006) examined the effect of native accents of varying degrees of familiarity on transcription accuracy for native speakers and found that more unfamiliar native accents led to lower accuracy. Recent data suggest that difficulty for native speakers with unfamiliar accents may be overcome through repeated exposure. Findings examining word recognition show that the first presentation of an accent to native speakers triggers a delay in word identification, indicating lower comprehensibility of the accented words (Floccia Goslin, Girard, & Konopczynski, 2006). This initial delay is followed by a subsequent adaptation across a brief series of additional trials with the same accent (Floccia, Butler, Goslin, & Ellis, 2009). However, this adaptation is not

complete: even after several trials, words with unfamiliar accents are still responded to more slowly than words in a familiar accent (Floccia et al., 2009). This suggests that, to the extent that full adaptation is possible, more extensive exposure is needed. Other results suggest that familiarity with a particular speaker is as important to determining comprehensibility as familiarity with a particular accent. Gass and Varonis (1984) found better performance on a listening comprehension measure when aural materials were presented in a familiar accent and the highest level of comprehension when the particular speaker was familiar to the listener. Their findings suggest that the less familiar the accented speech (e.g., unfamiliar non-native vs. familiar native) and the less familiar the speaker (novel vs. familiar speaker), the more difficult speech will be to comprehend.

If familiarity with an accent is the key to improving the comprehensibility of accented speech, it is possible that training listeners with a particular type of accented speech will improve comprehension in a way that generalizes to other speakers with the same accent. Weil (2003) examined the effect of training on comprehension of foreign-accented L1 speech for native listeners. Trained listeners showed superior performance relative to untrained listeners when the speaker at test was the same speaker as during training, indicating that participants did adapt to that person's speech, as found by Gass and Varonis (1984). As to whether receiving training with speech from a speaker with a particular accent improved comprehension more generally for speakers with that same accent, the results were mixed. Tasks involving single words did not show an advantage of training with a different speaker with the same accent, while those involving full sentences did show an advantage of this type of training. These results suggest that training with foreign-accented L1 can improve comprehension for other speakers with that accent for longer stretches of speech.

Recent research on the acquisition of unfamiliar phonetic contrasts suggests that the proficiency of the learner may influence whether they will be able to generalize their experience with an accent across different speakers. Lee et al. (2007) trained participants to recognize pitches from a language in which they had no proficiency. Participants completed this initial training with a high level of performance (e.g., >70% accuracy at identifying pitches) or a low level of performance (e.g., <70% accuracy in identifying pitches). These participants were then exposed to training with non-words (with the learned pitches) produced by a single speaker or produced by multiple speakers. Lee et al. found that listeners who showed a high level of performance on the initial pitch training were better able to generalize from their second round of training with non-words to a new speaker if they had been trained on multiple speakers. However, listeners with a low level of performance on the initial pitch training were better able to generalize to the new speaker if they had received their second round of training with a single speaker. Thus, it may be that listening comprehension with unfamiliar accents will be easier when passages are presented by a particular familiar speaker with that accent until listeners have attained some level of proficiency with that accent. After this point, however, exposure to multiple speakers with the accent may be the best way to improve listening comprehension.

Accent and L2 comprehension

Non-native speakers are likely to have a lower level of familiarity with any accented speech for their L2 than native speakers, but non-native speakers should still find particular L2 accents more familiar. Based on the findings for L1 comprehension of familiar and unfamiliar accents, L2 listeners should show better comprehension of the L2

when listening to speakers with familiar accents. Research supports this extension of the L1 literature. Major, Fitzmaurice, Bunta, and Balasubramanian (2005) investigated how accents of varying degrees of familiarity affected the listening comprehension of native and non-native listeners. The most familiar accent in their study was selected as such because it was the accent in which L2 listening materials are often presented, and was considered to be the most prevalent accent in the day-to-day L2 experiences of the non-native listeners. The other L2 accents were selected because of their more limited scope (they were regional, ethnic, or international [e.g., Australian English]). Results showed speaker accent significantly affected listening comprehension for non-native listeners. The more familiar the accent, the easier it was to comprehend. Further, differences in comprehension of the different accents were significant only for the non-native listeners, indicating that the effect of accent on comprehension is larger in the L2 than in the L1. Familiarity with the accented speech can explain these latter results, in that a non-native listener is less likely to have experience with any of the accents than is a native listener.

Familiarity with a particular accent or speaker is important for both native and non-native listeners.

If the effect of accent on listening comprehension is due to familiarity, it is possible that language learners may find L2 speech accented with their L1 more comprehensible, regardless of actual exposure, due to the influence of L1 phonological forms on L2 productions. Wilcox (1978) found some support for this idea, in that non-native listeners in his study understood L1-accented target language better than target language spoken by native speakers. However, two more recent studies have found only weak evidence that non-native listeners better comprehend L2 speech spoken by someone sharing their L1. Major, Fitzmaurice, Bunta, and Balasubramanian (2002) examined the comprehension of L2 listening passages presented by native speakers of the listeners' L1 with four listener groups having different L1s. Their results suggested that, while some accented L2 was difficult for all groups of listeners, only one listener group showed superior comprehension of an L2 passage produced by a speaker from their own language background. Rather, two of the three non-native listener groups comprehended speech produced by a native L2 speaker better. Munro, Derwing, and Morton (2006) found similarly weak evidence that language learners benefit from hearing the L1-accented L2: only one of their three L1 groups showed a benefit from listening to L2 produced by someone from their language background.

One reason why only two of the L1 language groups in the Munro et al. (2006) and Major et al. (2002) studies showed better comprehension for L2 speech accented with their L1 is that some listener groups experienced more L2 language training from teachers who did not share their L1 (e.g., they received language training primarily outside their native country). This experience with the L2 would lead them to be less familiar with L1-accented L2 than listeners whose training had come more from teachers sharing their same L1. This potential explanation is strengthened by the fact that the language group in Munro et al.'s (2006) study that showed better comprehension of L2 produced by a speaker from their same L1 background also reported having greater exposure to L1-accented L2 than did the other language groups. If this is the reason behind the inter-language group effects found in the two studies, however, familiarity with the speaker's accent would prove to be a more parsimonious explanation rather than a particular advantage of L1-accented L2 speech. Overall, familiarity with a particular type of accented speech is

argued to be a better explanation of how accent affects listening comprehension than is sharing the speaker's L1 (Major et al., 2002).

Accent interactions with factors influencing L2 comprehension difficulty

Other factors that influence the effort required to comprehend spoken language have been shown to interact with accent familiarity. For instance, Anderson-Hsieh and Koehler (1988) found that a faster speaking rate had a larger negative impact on listening comprehension for native listeners when the speech was produced by a non-native speaker rather than a native speaker. These results suggest that when the accent of the speaker is less familiar, L1 listening comprehension will be more affected by speech rate (Anderson-Hsieh & Koehler, 1988). The exacerbating effect of an unfamiliar accent on the negative effects of a fast speech rate may contribute to the explanation of why non-native listeners are more affected by a fast speech rate than are native listeners (e.g., Rosenhouse, Haik, & Kishon-Rabin, 2006; see the Speech Rate section for more information about the effect of speech rate on comprehension in non-native listeners).

Unfamiliar accents are likely to produce greater difficulty in listening comprehension when a fast speech rate or noise is present.

Another factor that interacts with the difficulties presented by unfamiliarly accented speech is noise. Adank, Evans, Stuart-Smith, and Scott (2009) found that, for L1 listeners, comprehension of an unfamiliar *native* accent was roughly equivalent to comprehension of a familiar native accent in quiet conditions, but adding moderate noise had a greater impact on comprehension of the unfamiliar native accent, such that it was comprehended more slowly and less accurately than the familiar native accent. By contrast, comprehension of a *non-native* accent was less accurate and slower than comprehension of a familiar native accent even in quiet conditions, and the difference between the speed and accuracy with which the two were comprehended increased when moderate noise was added. The results suggest that noise will interfere more with listening comprehension when speaker accents are unfamiliar, and the degree of interference will depend on the unfamiliarity of the accent. Similar results were found by Clopper and Bradlow (2008), who showed that the differences in the comprehension of native speakers for familiar and unfamiliar native accents were more pronounced when noise was present. The results of these studies suggest that, for non-native listeners as well as native listeners, the unfamiliarity of the speaker's accent will have a larger impact on comprehension when noise is present in the passage as well.

Summary: Speaker accent

Research examining the effect of accent on listening comprehension provides strong and largely consistent evidence that comprehension accuracy will decrease and effort (in terms of response time) will increase with the unfamiliarity of the speaker's accent. This effect has been demonstrated consistently with native speakers and with non-native speakers. Further, results indicate that the impact of accent is more extreme on non-native listeners than on native listeners and that it is difficult, within the scope of an experiment, to provide enough exposure to an unfamiliar accent to improve comprehension for that accent generally. Accent is thus an important factor to consider in choosing listening test materials for non-native speakers, as it will impact comprehension. Further, research indicates that it is

more important to consider the accent familiarity of the speaker when speech rate or noise are factors already present in the auditory materials.

Distortion and noise

Authentic recorded passages are not always recorded in ideal conditions. Recordings from telephone calls, conversations, or radio transmissions are often distorted or mixed with background noise. Cell phone signals are notoriously difficult to understand because of the loss of high-frequency information, clipping of the signal, and modulated frequency distortions. Recorded conversations are rarely held in perfectly silent places: other conversations, background noise, and modulated volume can all contribute to speech that is difficult to understand. Radio and television transmissions are subject to interference from external sources, especially static and crosstalk from other stations.

Indeed, these acoustic distortions can have a profound effect on a listener's ability to understand what they are hearing. When listening to sentences under less than ideal circumstances, even people listening to their native language struggle to understand what was said (Payton, Uchanski, & Braida, 1994; Adank et al., 2009). This is due in part to the fact that it is difficult to recognize words when the signal is degraded (Aydelott & Bates, 2004.) When words are difficult to hear due to noise and distortion in the signal, it is also difficult to build a strong semantic framework into which the listener can integrate incoming words (Aydelott, Dick, & Mills, 2006; Moll, Cardillo, & Aydelott Utman, 2001.) Indeed, all spoken language is distorted to some degree because of the phonological processes that occur in connected speech. These phonological processes include reduction, assimilation, elision, resyllabification, and cliticization, among others (Field, 2003), and all of them degrade the input from the citation form.

Distortion and noise can have profound effects on a listener's ability to understand what they are hearing, as the ILR scale acknowledges.

These difficulties are even more pronounced when a listener is trying to understand a non-native language. The ILR Listening Skill Descriptions (ILR Language Skill Level Descriptions: Listening, 1985) do not indicate that listeners should be able to understand noisy or distorted speech if they are below Level 2 (Limited Working Proficiency). Even at this level of proficiency, the ILR states that listeners can only understand *occasional words and phrases of statements made in unfavorable conditions*. Richards (2006) reports that it is extremely difficult to find authentic texts for beginning and low-proficiency learners. Even Level 4 (Advanced Professional Proficiency) listeners are reported to have *difficulty... in understanding speech in unfavorable conditions...* (ILR Language Skill Level Descriptions: Listening, 1985, paragraph 10). It is only the Level 5 listeners (with Functionally Native Proficiency) who are *able to understand fully all forms and styles of speech intelligible to the well-educated native listener, including... conversations and discourse distorted by marked interference from other noise* (ILR Language Skill Level Descriptions: Listening, 1985, paragraph 12).

Research into the influence of different types of noise on listening comprehension has typically presented speech in white noise, pink noise, and/or babble noise[24] (Bradlow & Bent, 2002; Cooke, Lecumberri, & Barker, 2008; Cutler, Lecumberri, & Cooke, 2008; Golestani, Rosen, & Scott, 2009; Lecumberri & Cooke, 2006; Rosenhouse et al., 2006; Shimizu, Makishima, Yoshida, & Yamagishi, 2002). Researchers using babble noise have investigated single-talker babble noise (see, e.g., Brungart, 2001), multi-talker babble noise (as in Simpson & Cooke, 2005), or speech played in reverse (e.g., Moll et al., 2001.) In general, these studies have shown that noise that matches the spectral and phonological characteristics of the stimulus speech interferes most with perception and noise with lower phonological and spectral similarity has a smaller impact.

Studies also have investigated the influence of audio distortion; these have focused on manipulations like reverberation (Nábělek & Donahue, 1984; Rogers, Lister, Febo, Besing, & Abrams, 2006), low-pass filtering (Aydelott et al., 2006), and time compression (Aydelott & Bates, 2004). Nábělek and Donahue (1984) found that non-native listeners identified rhyming words 10% less accurately than native listeners when they were presented with 0.8 and 1.2 second reverberation times, while Rogers et al. (2006) observed that even Spanish–English bilinguals who had learned English prior to 6 years of age and spoke English without a noticeable foreign accent had significantly poorer word recognition scores than monolingual listeners for words in noise with reverberation. These results indicate that even very skilled non-native listeners may have difficulty understanding speech in the presence of reverberation.

Both reverberation and low-pass filtering alter the intelligibility (clarity) of the signal, while time compression reduces the amount of processing time available to the listener. These manipulations do not necessarily alter the same aspects of the signal and would thus be expected to have different impacts on comprehension. Aydelott and Bates (2004) asked participants to listen to congruent and incongruent sentence contexts and make a word/nonword decision about the last word in the sentence. They found that low-pass filtering the words in the preceding sentence context reduced activation of the appropriate meaning of a word, while also reducing the inhibition of less compatible candidates. By contrast, time compression reduced the inhibition of inappropriate meanings without affecting facilitation. This led Aydelott and Bates (2004) to argue that altering intelligibility may influence the relatively early, automatic processes (reflected in the facilitation effects) involved in listening comprehension, while time compression has its primary effect on the later processes involved in selecting the most compatible word from possible candidates and inhibiting incompatible candidates. Indeed, they argue that inhibitory effects are especially vulnerable to factors that reduce processing time or increase processing demands. This is exactly what they observed in their experiments. If the preceding context was time compressed, listeners were able to make faster word/nonword

[24] White noise has equal energy across the spectrum of sound frequencies; it is the kind of noise you hear on an analog radio when the dial is not tuned to a station. Pink noise is like white noise, but it has higher amplitude frequencies in the lower frequency range, with amplitude dropping off as frequency increases. Babble noise sounds like a bunch of people talking at a cocktail party, though nothing in the "babble" is intelligible.

decisions when the target word was incongruent than they were when the target word was congruent. This *release from inhibition* was not observed when the preceding context was low-pass filtered.

Noise likely imposes an additional load on working memory as well. L1 listeners are slower and less accurate at speech processing in the presence of noise, and find understanding speech to be more effortful under noisy conditions (Larsby, Hällgren, Lyxell, & Arlinger, 2005). When noise interferes with the perception of a signal, this will be likely to increase the proportion of processing capability which a listener must devote to comprehension.

In addition to the challenges even a native listener experiences in understanding noisy speech, learners of a second language also struggle to decipher the phonology of the second language. When the competing noise shares phonological characteristics with the non-native language in the target passage, as when the passage is presented against a background of babble noise, it is harder for L2 listeners to determine which parts of the signal are from the target passage and which are from the competing noise (Carhart, Tillman, & Greetis, 1969; Brungart, 2001; Freyman, Balakrishnan, & Helfer, 2004.) Indeed, L2 learners who are *less familiar* with the language have more trouble with the sort of perceptual processing required to decipher the phonology of a passage than L2 learners who are *more familiar* with the language (Field, 2004).

As discussed above, beginning-level listeners compensate for their lack of knowledge about their non-native language's phonology by relying heavily on a top-down strategy wherein they determine the main ideas of the passage, and construct plausible contexts based on what they can understand (e.g., Lund, 1991). Field (2004) argues that listeners, and especially beginning-level listeners, do not even try to understand each and every word. Instead, they use background knowledge, co-text (information relevant to passage topic from sources like pictures, headlines, etc.), analogy, and/or knowledge about the speaker to construct a schema into which they can integrate incomplete acoustic information. Field describes this process in the context of Forster's (1989) description of *cross-word processing*, where top-down information is used to compensate for incompletely or incorrectly perceived lexical information.

Hesitation and pause

Speech, particularly spontaneous, informal speech such as conversations between friends or family members, often contains disfluencies such as pauses, hesitations, [25] or false starts (Fox Tree, 1995). For instance, in conversational speech in American English, roughly 6% of words are disfluent[26] (Fox Tree, 1995), while hesitation markers like *ano* make up about 6% of Japanese speech (Watanabe, Hirose, Den, & Minematsu, 2008)[27]. The

[25] *Filled pauses*, *hesitations*, and *hesitation markers* are used interchangeably in the literature to refer to disfluencies like *um* and *er* (e.g., Arnold, et al., 2003; Blau, 1991).

[26] Fox Tree (1995) presents this as the *non-pause disfluency* rate – disfluencies including repeated phrases or words, false starts, and hesitations like *um*, and excluding silent pauses, which are often difficult to classify as fluent or disfluent.

[27] Note that the *non-pause disfluency* rate described by Fox Tree (1995) includes hesitation markers in addition to false starts and repeated words and phrases.

usefulness of disfluencies for native listeners has been examined in several studies. One possible role for disfluencies is that they draw attention to the words around them. Collard, Corley, MacGregor, and Donaldson (2008) examined brain activity to explore the effect of disfluencies on listeners' attention to the surrounding material and found that hesitation markers orient the listener's attention, and lead to improved recall for words following a hesitation. Other research shows that hesitation markers such as *um* lead listeners to expect that new information is about to be presented by the speaker (Arnold, Fagnano, & Tanenhaus, 2003; Arnold, Tanenhaus, Altmann, & Fagnano, 2004) or provide other clues about what the speaker is about to describe (Corley & Hartsuiker, 2003). Results such as these suggest that disfluencies are not simply errors on the part of the speaker, but can present additional information for the listener.

In L1 listening, disfluencies may convey additional information about the speaker's intentions.

Despite the usefulness of disfluencies for L1 listeners, these findings do not speak to how disfluency affects comprehension for the non-native listener. A non-native listener may not yet recognize hesitations as disfluencies rather than as some other signal, such as turn-taking markers (Rubin, 1994). Hesitation markers might be misunderstood as word parts by non-native listeners (Voss, 1979). Further, L2 listeners may find pauses distracting, and experience decreased comprehension of the message as a result (Rose, 1998). For this reason, it is important to consider how disfluency affects non-native listener comprehension. This factor is particularly important to consider when predicting listening comprehension difficulty for authentic materials, as these types of materials are likely to have a greater proportion of disfluencies such as filled and silent pauses and false starts (e.g., *You said…you said that you would be there at five!*) than materials created for language learners (Gilmore, 2004).

Disfluencies and processing time in L2 listening comprehension

One reason why pauses and hesitation markers might improve listening comprehension for non-native listeners is that they allow for more processing time (Buck, 2001). Logically, a passage of a given length must contain less information with a greater amount of time to process each piece of information if the number and/or length of filled or silent pauses is increased. In fact, inserting silent pauses at clause boundaries has been used as a strategy to decrease overall speech rate (Griffiths, 1990, 1990a). It has also been specifically suggested as a way to alter authentic materials to make them more comprehensible to lower proficiency non-native listeners (Griffiths, 1990a). Blau (1990) further argued that inserting silent pauses is a preferable way to increase processing time compared to simply reducing speech rate through slowing a recording down mechanically, because pauses presented at clause boundaries do not disturb the flow of natural speech.

Research with passages containing inserted silent or filled pauses suggests that pauses improve comprehension. Blau (1990) found that non-native listeners' comprehension across a range of proficiencies (as measured by a university entrance exam) was significantly better when passages had silent pauses inserted at natural boundaries and that listeners estimated a greater degree of understanding (indicating greater confidence in their comprehension).

There was some evidence that pauses had a greater benefit for lower-proficiency listeners, but this was not consistent across participant groups. Other studies have found stronger evidence that the benefit of pauses for comprehension differ depending on listening proficiency, however. Jacobs, Chuawanlee, Hoga, Sakumoto, Saka, and Meehan (1988) found that silent pauses were related to comprehension for higher-proficiency listeners (more advanced language students), but not lower proficiency listeners (less advanced language students). Jacobs et al. (1988) suggested that lower-proficiency listeners may not be able to benefit from the additional processing time afforded by pauses.

While an earlier study found no benefit for non-native speaker listeners by increasing filled pauses in a passage (Chaudron & Richards, 1986), Blau (1991) examined the impact of filled pauses on listening comprehension and found that these had even greater benefits for comprehension than silent pauses for a group of lower-proficiency listeners (classified as lower proficiency due to less experience in the L2 compared to another group from a different L1 background). She argued that, for filled or silent pauses to benefit listening comprehension, they must be obviously without meaning (e.g., the listener must understand that filled pauses are hesitation markers rather than words with some independent meaning). To the extent that listeners realize that filled pauses are semantically empty, they may benefit from the additional processing time offered by their presence (Blau, 1991).

Disfluencies as cues in L2 listening comprehension

The studies described above approach the advantage of silent and filled pauses for listening comprehension from the viewpoint that these disfluencies may allow for more processing time for the listener in those cases where the listener might need it (e.g., when the speech rate of the passage is fast). For native listeners, however, disfluencies like silent pauses and hesitation markers like *um* (i.e., filled pauses) seem to create an expectation on the listener's part for particular types of information, such as objects not previously mentioned in the current interaction (Arnold et al., 2003, 2004). If disfluencies can be cues for L1 listeners, it is possible they may serve this same purpose for L2 listeners.

Disfluencies can provide additional processing time for L2 listeners, or they can act as cues about upcoming information in the message.

Watanabe et al. (2008) explored whether filled pauses and silent pauses served as cues for complex information for non-native listeners. Results from this study showed that response times were faster to requests for a more complex object (an object that would be more difficult for a speaker to describe) when the portion of the request regarding the complex attribute was preceded by a filled or silent pause than when there was no disfluency. However, this effect differed depending on the proficiency of the listener. For lower-proficiency listeners (less than 1.5 years in the L2-speaking country), neither filled nor silent pauses had an effect on response time to the complex object. For intermediate-proficiency listeners (1.5–2.5 years in the L2-speaking country), filled pauses led to shorter response times for the complex object as compared to no pauses, but silent pauses did not significantly decrease response times compared to no pauses. These participants also showed a negative effect of filled pauses when the object was simple, suggesting that their expectation of a complex object following the pause in the request slowed down their choice of the simple object. In the expert condition, choice of the complex object was faster when silent or filled pauses were present, but these participants did not suffer when a silent or filled pause preceded the information about the simple object, showing a pattern of responses similar to that of native listeners.

The effects of Watanabe et al.'s (2008) study suggest a developmental trajectory for the comprehension of disfluencies as cues to the speaker's upcoming message. Watanabe et al. concluded that, while lower-proficiency listeners had not yet learned to use filled or silent pauses as cues in L2 listening, and experts had attained near-native ability to use this information, intermediate learners had only partial ability to use pause information to anticipate a speaker's message. Intermediate listeners experienced difficulty when their expectations were violated (a filled pause preceded a simple object attribute) and had difficulty recovering.

Pause phenomena may present another part of a language that must be learned, making ability to recognize and use this information another aspect of proficiency.

Difficulties in L2 listening comprehension introduced by disfluencies

Most studies find a benefit of disfluencies such as silent or filled pauses for L2 listening comprehension, depending on the proficiency of the listener. However, two studies uncovered in our review of the literature found negative effects of disfluencies. Freedle and Kostin (1996, 1999) found that having a greater number of pauses (filled and silent) in a passage was associated with greater difficulty for corresponding comprehension items. However, this factor was not a significant predictor of item difficulty in regression analyses, suggesting that it did not contribute a large degree of explanatory power over other factors that Freedle and Kostin (1996) considered. Further, Freedle and Kostin mention that the incidence of pauses (filled or silent) was low in the sample of passages examined in their study. The association between number of pauses and item difficulty may have simply arisen because the passages that happened to contain pauses had other, idiosyncratic qualities which made them more difficult to comprehend.

Other evidence suggests that the benefit of pauses on L2 listening comprehension may depend on the passage topic. Leeser (2004) examined how long silent pauses (3 seconds long) at the end of each sentence in listening passages affected comprehension items specifically examining the recognition of verb tense as well as more general comprehension. While silent pauses were found to be useful for general comprehension (i.e., recall of idea units in a free recall task) when the topic of the passage was unfamiliar, they had a detrimental effect when the topic was familiar. There was no effect of pauses on the recognition of verb tense or on direct translation. These findings indicate that pauses do not necessarily provide a general advantage for listeners, but that they may alleviate difficulties caused by unfamiliar topic matter. However, these results also highlight the potential for pauses to be distracting, which would explain the inferior performance of listeners receiving the familiar passage with pauses compared to the familiar passage without pauses. Leeser (2004) also pointed out that pauses are likely to alleviate time pressure caused by a normal-to-fast speech rate, allowing for better listening comprehension; in those cases where the speech rate is slow, pauses may have no effect.

Other types of disfluency

One point important to make is that, while the bulk of the research evidence with L2 listeners suggests that disfluencies such as silent or filled pauses may improve L2 listening comprehension, different types of disfluencies exist which may have a different impact on comprehension. One example of this would be repairs, where a filler like

uh appears in a mid-word interruption to indicate that the speaker has misstated a word (e.g., *Move to the yel-uh, purple circle*, Brennan & Schober, 2001, p. 278). In cases such as these, a non-native listener, particularly a lower-proficiency listener, might be much more likely to misunderstand the disfluency as part of a word. In the example above, listeners might believe that the speaker was referring to a particular shade of purple or another color term with which they are unfamiliar. Until studies are conducted with non-native listeners, it cannot be concluded that all types of disfluency are undisruptive to listening comprehension. Further research is needed in this area.

Summary: Hesitation and pause

The majority of research on disfluencies in L2 listening comprehension has indicated that pauses can be helpful to non-native listeners, though this effect may depend on the listener's proficiency level and whether the pause is filled or silent. Some evidence suggests that knowledge of filled pauses such as *um* must be learned like other features of a language, and until this knowledge is complete, these disfluencies may be misinterpreted (Voss, 1979; Watanabe et al., 2008). Thus, when listeners are lower-proficiency, the avoidance of disfluencies in the form of filled pauses in listening passages is likely to be desirable. However, Rose (1998) pointed out that knowledge of an L2's hesitation and pause phenomena is important to L2 listening proficiency, particularly given their prevalence in authentic speech.

Speech rate

Think about two speakers—one fast, one slow—producing the same spoken passage. The speaker with the faster speech rate takes less time, conveying the given content more quickly, than the slower speaker. Now imagine that these speakers speak for the same amount of time. Given enough material, the faster speaker produces more speech, conveying more of the passage overall, than the slower speaker. The ILR scale explicitly refers to the ability of L2 listeners to handle different *speech rates*, including a *slower than normal rate* at Level 1, a *normal rate* or *normal speed* at Levels 2 and 3, and an *increased ability to understand native speakers talking quickly* at Level 3+. The current section explores the literature on speech rate and L2 listener comprehension, including its interactions with other factors, its numerous definitions, and factors that contribute to differences in rate.

Speech rate and L2 listening comprehension

Speech rate can make it harder for L2 listeners to understand a spoken passage, as the ILR scale acknowledges.

Results of several studies suggest that speech rate can negatively affect L2 comprehension. Critically, Griffiths (1990, 1992) used an experimental design in which a given text appeared with different speech rates (rotated across listeners so that individuals never heard the same content twice). This design controls for text-based characteristics, such as topic or vocabulary, and isolates the influence of speech rate. In the first study, Griffiths (1990) observed higher comprehension scores for passages presented at 1.93 or 2.85 syllables per second (syll/sec) relative to passages presented at 3.8 syll/sec. In the second study, Griffiths (1992) observed higher comprehension scores for passages presented at 2.5 syll/sec relative to passages presented at 3.75 or 5 syll/sec. Although Griffiths (1990, 1992) described the listeners as *lower intermediate learners*, more recent evidence suggests that speech rate also influences listening comprehension among relatively advanced L2 users. Specifically, Rosenhouse et al. (2006) investigated the effects of increased rate and noise on a group of proficient Arabic-Hebrew

bilinguals. All spoke Arabic at home, but were studying at a Hebrew-speaking university and had been exposed to Hebrew for more than 10 years. Listeners performed more poorly in their L1 and L2 when speaking rates increased from 3 syll/sec to 4 syll/sec (or when background noise was introduced), but they exhibited a significantly greater drop in performance when listening in their L2. Although these studies used tasks that defined comprehension in different ways (Griffiths tested passage understanding using a set of true/false questions and Rosenhouse et al. tested recall rates for individually presented sentences) these studies as a group are consistent in demonstrating that speech rate by itself is a factor that can result in decreased comprehension among L2 listeners.

Given the experimental findings that faster speech rates can lead to lower comprehension, it is not surprising that L2 learners sometimes explicitly point to speech rate as a source of difficulty. For example, Flowerdew and Miller (1992) interviewed a small group of eight language learners taking a university class taught exclusively in their L2. When the researchers asked the learners whether the lecturer spoke too fast, all but one responded in the affirmative. In addition, diary entries from the larger group of 30 learners mentioned lecturer speed as an issue. Consistent with these self-reports Zhao (1997) found that L2 learners took advantage of the opportunity to adjust the speaking rate of a passage as part of a listening experiment: 14 of the 15 listeners reduced the preset rate of 194 words per minute, and none of them increased the speed.

Listeners do not perceive speech to be fast or slow purely on the basis of objective quantitative measurements.

Listeners are more likely to think that speech is fast when other features of the passage challenge comprehension (e.g., low redundancy, unfamiliar accents).

None of the studies on speech rate and L2 comprehension, however, pinpointed a critical turning point at which speech rates become unmanageable for L2 listeners—of any proficiency level. One reason for this is the fact that listeners do not perceive speech to be fast or slow—in either their L1 or L2—purely on the basis of objective measurements of speech rate. Other factors influence their judgments. For example, Moore, Adams, Dagenais, and Caffee (2007) found that native speakers judged reverberated speech to be faster than filtered or unfiltered speech despite a constant speech rate. Griffiths (1990) observed that native speakers perceived non-existent differences in rate while pre-testing materials that varied in text length and difficulty. Native speakers in a study by Anderson-Hsieh and Koehler (1988) perceived heavily accented speech as faster than less accented speech. L2 listeners exhibit comparable effects, according to Cheung (1994) and Dahl (1981) (as cited in Tauroza, 2001, p. 146), and consistent with this, Derwing and Munro (2001) found that L2 listeners tended to prefer slower rates for speech from other non-native speakers, particularly if those speakers came from a different language background than the listener.

Because the perception of speech rate interacts with other factors, it is not always the case that slower is better among L2 learners. Derwing and Munro (2001) found that a group of non-native speakers (described as high proficiency) preferred an original speaking rate of 4.9 syll/sec to a reduced rate of 3.4 syll/sec when asked to judge native speaker passages on a nine-point scale ranging from *too fast* to *just right* to *too slow*. Similarly, Griffiths (1990) observed no difference in comprehension for passages presented at 2.85 or 1.93 syll/sec.

Additional studies have demonstrated interactions between speech rate and other factors. For example, Lin (2006) demonstrated increased comprehension for note-taking on a standardized listening comprehension exam (Taiwan's *General English Proficiency Test*), but only when it was delivered at its usual speaking rate of 180 wpm. Note-taking had no effect at a reduced rate of 120 wpm.

Buck and Tatsuoka (1998) attempted to identify a (non-exhaustive) set of person and item characteristics contributing to listener difficulty as measured by test scores. They included speech rate not as a global factor of the entire passage, but as a local factor of the *necessary information* contained within the text, where listeners must understand such information in order to answer comprehension questions correctly. The final analysis included two item characteristics related to speech rate: (1) the speech rate of the necessary information exceeded 180 wpm and (2) the speech rate of the text surrounding the necessary information exceeded 160 wpm. Ultimately, speech rate by itself was not a predictor, but it interacted with other features of the text, such as whether the necessary information was expected or appeared in longer units.

Brindley and Slatyer (2002) also considered speech rate as a local factor within a text. In their study, they examined some of the ways in which test materials might vary while still meeting pre-defined criteria for selecting materials that feed into a national system of assessing foreign language proficiency in English in Australia. Their study materials varied in speech rate, delivery (recording vs. live speaker), number of speakers (monologue vs. dialogue), response type (sentence completion, short answer, or table fill-in), and number of hearings (one vs. two). Not surprisingly, the large number of variables involved made it difficult for the researchers to examine all possible interactions, even with 284 participants. Regardless, the analysis did suggest that the speech rate of passages could influence the outcome of proficiency assessments for borderline candidates, i.e., assuming other features are held constant, an examinee would be more likely to pass if he or she listens to a test passage of 180 wpm than a comparable one of 200 wpm.

The many variables in the Brindley and Slatyer (2000) design, of course, highlighted that other features are typically not held constant across authentic passages. One of the easiest items for listeners corresponded to a slow speech rate (157 wpm) on the necessary information and surrounding text, and a highly constraining single-word sentence completion format. The corresponding short answer version was harder, indicating that type of test item, rather than the slower speech rate of the corresponding passage, may have led to its lower difficulty for examinees.

Although this review did not find any studies that directly explored the interaction between speech rate and informational density in L2 listening comprehension, it is reasonable to expect such an interaction because both factors (faster speech rate and greater density) have been used to increase speech processing load for L1 listeners (Wingfield, 2000), and increased propositional density has a greater negative impact on L1 listening comprehension at faster speech rates (Stine, Wingfield, & Leonard, 1986). The impact of propositional density will also depend on redundancy. Across authentic passages, different speakers produce different content, and that content may be more or less redundant. As a result, a faster speaker who repeats or paraphrases previously stated text may be easier to

understand than a slower speaker who conveys *more new information* to a listener using speech that has little redundancy. (See the Information Density section for a discussion of this passage-based characteristic.)

In short, speech rate can negatively affect L2 listener comprehension, but its direct effects are most visible in experiments that artificially isolate and manipulate the speech rate of a given passage. Across authentic passages, listeners are more likely to perceive speech as fast when other features, such as audio quality, speaker accent, text length, text difficulty, and test format, challenge listener comprehension.

Speech rate definitions

Speech rate measures vary in their units of analysis (e.g., words vs. syllables) and in their treatment of pause duration.

Speech rate may appear to be a basic concept that distinguishes between faster and slower speakers, but in actuality, it corresponds to multiple measures across the research literature. One common measure, particularly in the literature on L2 listening comprehension, is *words per minute* (e.g., Blau, 1990; Brindley & Slatyer, 2002; Griffiths, 1990, 1992; Jacobs et al., 1988; Zhao, 1997). This measure is also referred to as *speaking rate* (e.g., Robb, Maclagan, & Chen, 2004) and it critically includes silent intervals (which include pauses) in its duration calculation. As Griffiths noted, it is important to interpret measures of words per minute (wpm) with respect to the ratio of syllables to words within a passage. Although two passages could have similar wpm rates, one could contain largely monosyllabic words and the other could contain multisyllabic words.

To help control for gross variation in word length, some studies express speaking rate in terms of syllables per second (e.g., Derwing & Munro, 2001; Kang, Rubin, & Pickering, in press). Any measure at the syllable level, however, necessarily requires closer inspection. Measures of syllables per second or seconds per syllable (also known as average syllable duration, e.g., Quené, 2007, 2008; Rosenhouse et al., 2006) may indicate *articulatory rate*. Whereas speaking rate includes silent intervals in its duration calculation, articulatory rate excludes silent intervals that exceed a given threshold. The aim of these thresholds is to eliminate silent intervals that correspond to pauses or disfluencies and to preserve silent intervals that correspond to articulatory gestures, in order to obtain a purer measure of the speed of motor movements related to the pronunciation of speech sounds. For example, a threshold is intended to preserve silent intervals related to the articulation of stop consonants (e.g., the *b* and *p* sounds both temporarily block the flow of air through the vocal tract, creating silence) and to exclude pauses related to speech planning or hesitation. Threshold values vary in the literature; Robb et al. (2004) rejected a previously reported threshold of 150–250 ms and adopted a more conservative 50 ms cut-off in their analysis of average speech rates for speakers of American or New Zealand English.

Speaking rate and articulatory rate are both dependent on how quickly a speaker produces speech sounds, but speaking rate provides a more global measure of content over time that incorporates silent durations related to pause phenomena. For studies that use sentence-length or shorter materials (e.g., Barreto & Ortiz, 2008; Moore et al., 2007), utterances containing disfluencies, or dissimilar patterns of intonation phrasing (which may involve differences in the location and duration of pauses) are likely excluded prior to use as listening comprehension materials, and any

contrast between speaking rate and articulatory rate is likely to be minimal for such short, controlled utterances. Corpus studies that examine naturally occurring sentence-length or shorter utterances may stand a greater chance of using materials that result in distinct speaking and articulatory rates, but they too have selected utterances without pauses (e.g., Quené, 2007, 2008; Yuan, Liberman, & Cieri, 2006). For studies that use discourse-length materials (e.g., Anderson-Hsieh & Koehler, 1988; Lin, 2006), calculations of speaking rate are the norm, even though such rates provide no information regarding the frequency, duration, location, or type of pause phenomena. Given that studies by Jacobs et al. (1988) and Blau (1990) suggest that pausing improves L2 listener comprehension (See the Hesitation and Pause section for a thorough discussion of pausing), future studies of speech rate should consider incorporating careful descriptions of articulatory rate and pause phenomena. One advantage to including articulatory rate as a measure—in conjunction with separate measures of pausing—is that it can apply over portions of a passage, thereby capturing dynamic changes in rate that may interact with other properties, such as the location of information that listeners must understand in order to answer a test question or the use of unusual or unfamiliar vocabulary.

Measuring speech rate independently from pausing would help to capture the unique contributions of these two factors on L2 listening comprehension.

Additional speech rate measures in the literature include *realized phones per second* (or the actual consonant and vowel segments in the utterance) and *intended phones per second* (or the canonical or unreduced form which the utterance could maximally take; Koreman, 2006). For example, a speaker who says *wanna* for *want to* deletes the *t* sound, resulting in one fewer realized phones than intended phones. The ratio of these two measures further provides a measure of articulatory precision. Hirai (1999) adopted a *standard words per minute* measure of speaking rate based on the written form of a text in which a *standard word* was equivalent to six character spaces. Lastly, Kang et al. (in press) examined *mean length of run*, where a run was a stretch of speech bounded by silent intervals of no less than 100 ms, as they occurred in 60-second speech samples; mean length was established as total number of syllables within runs divided by total number of runs.

In short, investigations of speech rate involve multiple measures that are not always similarly defined within or across studies. For example, *words per minute* may or may not include pauses, even within a given study (Yuan et al., 2006), and *speaking rate* may indicate a specific mathematical calculation that includes silent intervals or a more general term for how quickly someone speaks that subsumes speaking rate and articulatory rate (e.g., Robb et al., 2004). Studies that intend to assess the speech rate of authentic materials should define measures clearly and use them consistently across passages.

Assessments of authentic materials should define measures clearly and apply them consistently across passages.

Factors contributing to variations in speech rate

Many factors drive naturally occurring variation in speech rate. At the level of the individual, studies such as Tsao, Weismer, and Iqbal (2006) suggest that differences in the articulatory rates of people who are habitually fast or habitually slow speakers may have a biological basis rooted in differences in neuromuscular control. At the level of languages, studies such as Robb et al. (2004) raise the possibility that differences in vowel inventories and language

contact (i.e., the other languages with which speakers interact) may play a role in explaining differences between dialects. More generally, sociophonetic studies (e.g., Jacewicz, Fox, & O'Neill, 2009; Verhoeven, De Pauw, & Kloots, 2004) examine articulatory rate as one of the many manifestations of social identity within speech communities and demonstrate the interaction of speech rate with such factors as region, age, gender, and situation (e.g., reading vs. informal conversation). For example, both Robb et al. and Jacewicz et al. document faster average articulation rates for tasks that involve informal conversation relative to tasks that involve reading out loud.

Phrase length is a major factor in articulatory rate: longer spoken phrases tend to have faster articulatory rates than shorter phrases.

A recent study by Quené (2008), however, demonstrates the important role that phrase length plays in determining articulatory rate and its possible confounding with other factors. *Phrase length* in this context refers to chunks of speech that are set off by pauses. This is because calculations of articulatory rate apply over stretches of speech demarcated by silent intervals that exceed a particular threshold, e.g., 50 ms (Robb et al., 2004).Using a spoken corpus of Dutch, Verhoeven, De Pauw, and Kloots (2004) found significant effects of age, gender, country, and region on articulatory rate. Using materials from the same corpus, Quené modeled their factors alongside phrase length and several other novel factors. In this expanded analysis, significant effects were routinely mediated by phrase length, which Quené attributed to well-known effects of *anticipatory shortening* (Nooteboom, 1972; Lindblom & Rapp, 1973; De Rooij, 1979; Nakatani, O'Connor, & Aston, 1981, as cited in Quené, 2008, p. 1111), that is, *speakers shorten their syllables if they anticipate more syllables within a phrase.* Although the speakers' countries of origin continued to explain significant variance in articulatory rate, such that speakers from the Netherlands produced faster and more varied rates than those from Flanders, phrase length was also significantly shorter in the Netherlands than in Flanders. Similarly, the previously reported effect of gender remained significant, with faster articulatory rates for males than females, but the magnitude of the effect was reduced, falling near the *just noticeable difference* for articulatory rate (Quene, 2007). Furthermore, the previously reported effect of age was solely explained by differences in phrase length. Older speakers produced shorter phrases than younger speakers, as well as greater variation in phrase length.

Other factors affecting speech rate, some of which may actually correspond to differences in phrase length, include dialect, situation or task (e.g., informal conversation vs. read speech), gender, age, emotional content, and predictability of content.

Additional factors may reflect aspects of the individual or the spoken text. For example, Murray and Arnott's (1993) review article on human vocal emotion suggests that higher speech rates are associated with anger and fear, and slower rates, with sadness and disgust. In a different line of work, using analyses of spoken and written corpora from a medical domain, Pan, McKeown, and Hirschberg (2001) demonstrated that spoken phrases containing unexpected words (i.e., words with a low frequency relative to the given corpora) tended to exhibit faster articulatory

rates than those containing expected words.[28] Not surprisingly, given the many contributing factors, articulatory rate is a dynamic property, and speakers may speed up or slow down throughout a passage.

Speech rate is a dynamic property—a speaker may speed up or slow down through a passage.

Speech rate is of interest in part because of the effects it may have on the speech signal. First, imagine holding articulatory rate constant. In order for speaking rate to increase, pause durations must necessarily decrease. Then, imagine increasing articulatory rate. Segments and their acoustic cues must arrive more quickly. In turn, increased articulatory rate is associated with, for example, decreases in vowel duration, second formant (F2) vowel onsets, and stop closure intervals, as well as with increases in coarticulation, deletion of consonant or vowel segments, and reduction in pitch range and pitch resets (e.g., Guion, Flege, Akahane-Yamada, & Pruitt, 2000; Koreman, 2006; Lindblom, Sussman, & Agwuele, 2009). As segment durations decrease (relative to some slower rate), speakers may undershoot or hypoarticulate segments given that they will have less time to reach the intended motor targets (e.g., Byrd & Tan, 1996; Lindblom et al., 2009). Despite this, articulatory rate is distinct from speech clarity, and speakers may accompany faster articulatory rates with greater effort thereby reducing hypoarticulation. Indeed, Koreman (2006) was able to extract naturally occurring utterances from a corpus of spoken German that independently varied in articulatory rate (fast vs. slow) and articulatory precision (clear vs. sloppy). In clear speech, the realized phone rate paralleled the intended phone rate. In sloppy speech, the realized phone rate was lower than the intended rate.

Faster speech is often less clear than slower speech, although speech rate and auditory clarity are distinct properties.

In short, multiple factors are likely to affect the speech rate of authentic passages. However, it is important to note that the literature cited here fails to address the extent to which the individual findings apply cross-linguistically. For example, although Quené (2008) found that males tended to speak more quickly than females in Dutch, it does not follow that males speak more quickly than females cross-linguistically. The research can only point to variables that are likely to drive variations in speech rate. One factor that may have a universal component is the role of phrase length in articulatory rate. To the extent that speakers have a limited amount of air in which to produce a spoken utterance (set off by pauses), longer utterances will necessarily require a faster rate.

Perhaps in part because speakers vary the length of their spoken phrases throughout a passage, speakers may speed up or slow down. While it is possible to calculate average articulatory or speaking rates for a given passage to obtain a gross measure of how quickly someone is speaking, it is important to remember that the values are just that, averages. As a result, such global measures of speech rate may miss important interactions of rate with other features

[28] Pan et al. (2001) do not explain why unexpected words might have faster articulatory rates, and the finding is somewhat unexpected given that less predictable words also tend to have longer durations than more predictable words (e.g., Bell, Brenier, Gregory, Girand, & Jurasky, 2009). Other aspects, such as phrase length, or intonation or discourse structure, might be driving the effect.

of the passage. Finally, it is important to remember that although faster speech has a tendency to be less clear than slower speech, rate and clarity are unique properties that correspond to distinct measures.

Summary: Speech rate

The research literature provides evidence that speech rate can negatively affect L2 listener comprehension. However, its direct effects are most visible in experiments that artificially isolate and manipulate the speech rate of a given passage. In the real world, L2 listeners move from passage to passage and encounter different speakers and different content. Because listeners are more likely to perceive speech as fast when other features challenge comprehension, speech rate must be considered in conjunction with other aspects of the listener, passage, and environment.

Source of characteristic	Factor of interest	Consensus in the literature
Passage	Auditory features	Familiarity with a speaker's accent and the degree of distortion and noise can have profound effects on L2 listening comprehension. Learning to use disfluencies, such as filled pauses, and silent pauses represent another aspect of L2 proficiency. Listeners are more likely to perceive speech as fast when other aspects of the passage challenge comprehension.

Any assessment of the speech rate of authentic materials will need to define measures clearly and use them consistently across passages. The selection of particular measures will depend in part on other measures in use. More specifically, it may be useful to examine articulatory rate (which focuses on speech sounds and excludes pauses) in conjunction with separate descriptions of pause phenomena. Such an approach would allow for separate influences of speech rate and pausing, and capture dynamic changes in rate that are likely to interact with other properties of the passage (e.g., the location of new vs. redundant information).

Variation in speech rate stems from multiple sources, but two factors—dialect and the length of spoken phrases—may provide some guidance to those who are charged with selecting authentic passages.

Summary of research findings for passage-based factors			
Factors with strong effects or convincing evidence		**Factors with sparse or inconsistent evidence**	
Beneficial to listeners:	Greater **redundancy**	Possibly beneficial to listeners:	Simplifying **syntactic structure**
	Use of **discourse markers**		Greater reference to **concrete** entities/objects
	Positioning item-relevant information near the beginning or at the end of a passage decreases difficulty		Greater overall **coherence**
	Disfluencies such as filled and silent pauses		
Difficult for listeners:	Greater **information density**	Possibly difficult for listeners:	Overall **length**
	Need for **pragmatic information**, including for example, understanding implied meanings, indirect text, idioms, and culturally specific information		**Syntactic features,** such as negatives
	Unfamiliar **accents**		
	Noise or distortion		
	Faster **speech rates**		

CHARACTERISTICS OF THE TESTING CONDITIONS

For an L2 listening comprehension examinee, there are many challenges. Some of the factors related to comprehension difficulty are aspects of the passage that may map onto proficiency in the language (e.g., ability to recognize disfluencies in the L2 or cope with a faster speech rate). Other factors involve personal characteristics of the examinee that lessen or heighten the difficulty of the listening comprehension exam which cannot be controlled by the test developers. However, some factors in L2 listening comprehension testing are completely under the control of the test developers. When determining the specifications for how the test will be designed and administered, the role these factors play in the difficulty of the test should be considered.

One topic that could appropriately be discussed in this section is the type of task (e.g., specific comprehension items versus free recall) or test item type (e.g., multiple-choice versus open-ended). The purpose or goal for listening will differ depending on whether the examinee is asked to create a transcript of a passage, engage a free recall task or answer specific comprehension questions about the passage; the purpose for listening can also affect the way various factors impact difficulty (Dunkel, 1991). Performance on listening comprehension tests can be affected by whether the examinee knows test items before they hear the passage, possibly by affecting the listening strategies they adopt (Yanagawa & Green, 2008). However, a discussion of the effect of test item type or task type on L2 listening comprehension and the resulting purpose for listening, or interactions between item type and task type and other factors discussed in this review, is beyond the current scope of this review.

As alluded to above, a number of factors in the testing conditions can contribute to the outcome of the exam. The review only discusses time limits, number of hearings, and note-taking.

Time limits

Discussions of time limits in testing often focus on how introducing a time limit changes response patterns. Bejar (1985) defined a speeded test as one where some portion of the test-takers does not have enough time to attempt every item within the test in the allotted amount of time. For multiple-choice or true/false tests, one possible effect of speededness is that some of the test-takers answer some of the test items in a more or less random manner, having run out of time to attempt to answer the items thoughtfully (Bejar, 1985; Yamamoto & Everson, 1997). The propensity for this type of response increases for more difficult items, so the effect of a time limit will interact with the difficulty of the test item (Bejar, 1985). Scoring of items on the exam must also take into account time limits. If items answered incorrectly are treated the same as unanswered items, the test may fail to discriminate between those who are working slowly, but accurately, and those who are answering incorrectly (Verhelst, Verstralen, & Jansen, 1997).

Little research has explored how time limits affect performance on listening comprehension items.

Time pressure decreases a listener's available working memory processing capacity.

In addition to changing how examinees respond, speeded tests should also be generally more difficult for test-takers than are tests without time limits due to the effects of time pressure. Research in

performance of cognitive tasks demonstrates that time pressure reduces available working memory processing capacity (Siemer & Reisenzen, 1998). These results suggest that any test measuring cognitive skills will be affected by whether the examinee experiences time pressure. Whenever other qualities of the testing environment (e.g., distractions in the testing room), materials (e.g., a faster speech rate in a listening test passage), or the examinee (e.g., a low working memory capacity) affect working memory processing capacity, speededness may have an even greater impact on performance.

Time limits and L2 listening comprehension

Research on language tests which require production, such as tests of second language oral proficiency, have shown benefits from increased response time allocation for both lower- and higher-proficiency test-takers, though these effects are arguably due to the additional planning time examinees are granted (Ellis & Yuan, 2004; Hale, 1992; Powers & Fowles, 1996). Almost no research has examined the effects of response time on L2 listening comprehension test performance. Coniam (1996) examined factors affecting performance on an L2 dictation task and found that a 5 second/word time limit for responses was acceptable for the participants, as no participants reported running out of time, though he did not explore the effect of different time limits. The time required to address other types of tasks (e.g., multiple-choice comprehension items) is also likely to differ from that required for a dictation task, but the review of the literature uncovered no studies exploring the effect of time limits on other types of L2 listening tasks.

An important consideration in predicting the effects of speededness on listening comprehension tests is that time pressure is an intrinsic quality of listening, regardless of time limits imposed on the duration of the exam. This is because listening occurs in real time, such that information must be processed on-line while it is being delivered (Buck, 2001). Passage-based factors that affect processing demands or processing time available are described in the section Characteristics of the Passage. In considering the effect of time limits on the comprehension of the passages themselves, rather than time limits for responding to test items or for the exam overall, these factors will be most important.

Time pressure is an intrinsic quality of listening comprehension testing because listening occurs in real time.

Buck (2001) argued that time limits for responding to test items were unlikely to impact L2 listening comprehension test performance. He claimed that test-takers who spend extra time thinking about a response to a listening comprehension item have not understood the passage well enough to respond, and unless the examinee has the option of re-hearing the passage, additional response time would be unlikely to affect performance. As an alternative, Buck (2001) suggested pre-testing items to determine a reasonable time limit for each item on a listening comprehension test.

Summary: Time limits

It was difficult to locate any research on the effects of time limits on L2 listening comprehension performance. However, other research on how time pressure affects cognitive tasks, test strategies or response behavior is relevant (e.g., Bejar, 1985; Siemer & Reisenzen, 1998). Research on the effect of time limits on cognitive tasks suggests that

time pressure will have a negative impact on performance because it limits working memory processing capacity (Siemer & Reisenzen, 1998). It has also been argued that providing additional response time will not change performance in listening comprehension tasks, because examinees cannot go back to the material if they did not comprehend it the first time (Buck, 2001). This latter point is relevant only if the test does not allow the examinee to listen more than once to the test passages.

Number of and control over hearings

As the above discussion of speededness or time pressure in the evaluation of listening comprehension suggests, the examinee's ability to pause or replay a passage may have an impact on comprehension of the passage and performance on corresponding test items. Similarly to response time, the number of hearings examinees may have of a test passage, and the extent to which they can choose to pause the passage or go back to particular segments of the passage, is under the control of the test designer. Thus, this factor can be introduced in the testing situation if it is desirable to do so.

Repetition and multiple hearings

Several studies have examined the impact of multiple hearings on comprehension of aurally presented information. However, some of these studies classified this manipulation as exact repetition, of the sort investigated with respect to redundancy (e.g., Cervantes & Gainer, 1992). For the purposes of this review, there is a distinction between the repetition that occurs through the re-presentation of words, phrases, or information units within a passage and that which occurs through multiple hearings of a passage. There are several reasons why this is an important distinction to make. Repetition, as a form of redundancy, may be introduced by a speaker in an attempt to ease the comprehension of the message. L2 listeners may encounter this type of repetition in their experiences with the L2, and there is even evidence that they will be more likely to encounter it when experiencing authentic samples of the L2 than in the language classroom (Gilmore, 2004). Repetition of this type is a quality of the passage itself. In contrast, repetition of the passage in the form of multiple hearings is a quality of the testing conditions, not of the passage itself, and will be described as such here.

Multiple hearings provide a type of repetition that is categorically distinct from repetition described with respect to redundancy.

Impact of multiple hearings on testing

Allowing examinees to listen to test passages multiple times is a decision for the test developer. There are several potential benefits for testing to accompany multiple hearings of a passage. First, allowing a listener to hear a passage more than once is believed to counteract the difficulties encountered by lower-proficiency listeners who are preoccupied with decoding and finding links between adjacent utterances: listeners can use the first hearing to gather bits of information (e.g., picking out propositions) from the passage, and use the second and additional hearings to discover the structure of

Playing a passage multiple times can correct for idiosyncratic problems in the test environment (e.g., a sudden loud noise) and reduce the effects of factors such as test anxiety.

the passage as whole, such as the central theme (Field, 2008). Further, allowing more than one hearing of the passage may overcome idiosyncrasies of the test environment, such as experiencing an interruption from a noise in the testing room while listening to the passage (Buck, 2001).

There is also some evidence, however, that multiple hearings are not useful for L2 listening comprehension tests. Henning (1990) found a trend for multiple hearings to decrease item difficulty, but found greater convergent and discriminant validity for items associated with passages that were *not* repeated. Based on this, he argued that repeating the listening passages failed to improve the validity of the test items.[29] Although a number of issues with Henning's study may have led to his results (the test passages used differed in length and the passages of different lengths had different numbers of corresponding test items), this study nonetheless raises validity as an important issue to be considered when evaluating the desirability of replaying a listening passage for testing purposes. Further, it has been argued that allowing multiple hearings presents an inauthentic listening situation because real-life listening, when the speaker is not recorded, rarely offers the opportunity for an exact repetition of what was said (Buck, 2001). However, because many authentic listening scenarios allow for multiple hearings (e.g., replaying a podcast or a voicemail message), this is not necessarily a serious issue.

Comprehension across hearings

Studies exploring the effect of multiple hearings on listening comprehension in an L2 have often used a within-subjects methodology. In this methodology, how much participants recall after a first hearing is compared to their recall after a second hearing. Lund (1991) allowed participants to listen to the passage once, then engage in a free recall test, then listen to the same passage again after their first recall efforts had been taken away, and again recall as much as possible from the passage after the second hearing. Recall performance on the second recall occasion was superior to that on the first occasion. Using a parallel methodology, Sakai (2009) reported similar results: recall was superior following a second hearing compared to a first hearing. Berne (1995) presented a multiple-choice listening comprehension test to participants after the first presentation of a video with audio and again after a second viewing (removing their first responses prior to the second viewing). In agreement with Lund's findings, performance on the comprehension test improved from the first to the second viewing of the video.

Though studies using the test-retest design find evidence that multiple hearings of the passage aid listening comprehension, their results should be interpreted with caution. Participants in all three studies both heard passages twice and responded twice to the dependent measure. Additional familiarity with the dependent measure likely led to some improvement of performance (i.e., a practice effect) regardless of the benefits of hearing the test passage multiple times.

---

[29] If multiple hearings help overcome issues like noise in the testing room, allowing more than one hearing would be expected to improve the validity of test items because performance would no longer be affected by idiosyncracies of the testing session.

Multiple hearings and listener proficiency

Fortunately, other studies have manipulated the number of hearings of a passage independently of other factors. The studies have uncovered results pointing to the general benefit to listeners of multiple hearings as well as differential benefits for listeners with higher listening proficiency.

Cervantes and Gainer (1992) found better comprehension performance following two hearings of a passage compared to only one for both higher and lower-proficiency listeners (proficiency determined through an experimenter-provided pre-test). Other studies tend to find an interaction with listener proficiency, however. Chang and Read (2006) found that higher-proficiency listeners showed a more pronounced benefit from hearing the passage three times (compared to other methods of improving listening comprehension) than did lower-proficiency listeners (measured through the Test of English for International Communication; participants with scores ≥ 40 were classified as higher proficiency, those ≤ 39 were classified as lower proficiency). Similarly, Chang (1999; as cited in Chang & Read, 2006, pp. 379–380) found that higher-proficiency listeners benefited from a single replay of the passage, but lower-proficiency listeners failed to show improvement even after several hearings (method of measuring language proficiency was not specified). Chang and Read (2006) concluded that lower-proficiency listeners may lack the language knowledge to comprehend certain aspects of a passage (e.g., specific vocabulary knowledge) even after hearing it multiple times.It should be noted that the passages used in Chang and Read's (1996) study were rated by other language learners prior to the study in terms of the suitability of their topic, the familiarity of the vocabulary, speech rate and other criteria; because of this, it is somewhat unlikely that the passages far exceeded the listening ability of the lower-proficiency participants, who came from the same class level as the raters. That said, it is still possible that the listening comprehension of lower-proficiency listeners will improve over repeated hearings for simpler passages. Without more information about the objective level of difficulty of Chang and Read's (1996) passages in terms of factors like vocabulary level, speech rate, and syntactic complexity, the potential for an interaction between multiple hearings and listener proficiency remains an open issue.

Higher-proficiency listeners benefit more from hearing a passage multiple times than do lower-proficiency listeners.

Multiple hearings versus redundancy

Gainer (1997) compared the benefits of multiple hearings of a dialogue to the benefits of redundancy (which included, in this study, the exact repetition of some of the words within the passage itself). Participants performed better on a comprehension task if they heard the unmodified (without redundancy) passage played twice than if they heard the unmodified passage only once, but the former group did not differ from participants who heard the passage with added redundancy once. These results were similar for both higher (first and second-year English Communication majors) and lower-proficiency listeners (first-year students from departments other than English Communication; difference in L2 proficiency was confirmed through an experimenter-designed cloze task). This result suggests that allowing multiple hearings of a passage presents the same level of benefit to listening comprehension as adding redundant information to the passage for both high and low-proficiency listeners. Further, performance in the condition where participants heard the redundant version twice was superior to both the condition

where they heard the unmodified passage twice and where they heard the redundant passage once, again for both proficiency levels.

Gainer's (1997) findings have important implications for the use of authentic materials in testing. Even though redundancy tends to be greater in authentic than in created passages (Gilmore, 2004) a given authentic passage may not include much redundancy. If it is desirable to include redundant information to improve comprehension without modifying the authentic passage, playing the passage more than once may compensate for a lack of naturally occurring redundancy.

Control over hearings

Another issue concerning multiple hearings of a listening passage is whether the listener controls when and if the passage replays. Zhao (1997) assigned L2 listeners to one of four conditions that differed in their ability to replay and change the rate of delivery for passages. Zhao found a significant effect of control over rate of delivery on comprehension, but simply allowing participants to replay the sentences of the passages did not significantly improve performance over a condition where participants could not replay the passage or control the rate of delivery. In interpreting these results, however, it is important to note that the test passages in the one-hearing/no rate control condition were individual sentences, while those in the replay/no rate control condition were passages 15–20 sentences in length. The materials in the one-hearing/no control condition may have been easy for the participants to comprehend without multiple hearings.

Zhao's (1997) study also provided some evidence that, the more difficult a passage is to comprehend for other reasons, the more likely listeners will choose to rehear the passage if they are allowed to do so. Participants in one of the conditions could both control the speech rate of each sentence in the passage and replay each sentence as they wished, while participants in the replay/no rate control condition could only replay each sentence. The proportion of participants in the former condition who chose to replay the passages was considerably smaller than the proportion that chose to replay the passages when they could not control the speech rate. These findings suggest that listeners will choose to rehear a passage more often if other factors make the passage difficult to comprehend.

Multiple hearings and other factors

Playing a passage multiple times may counteract the effects of other factors. For instance, Freedle and Kostin (1996; 1999) found that the position of the item-relevant information in a passage affected the difficulty of the item. However, Rupp et al. (2001) failed to find a significant effect of location of the item-relevant information, which they attributed to allowing participants in their study to listen to the passages as many times as they wished. Allowing multiple hearings of a passage may remove the effect of where in a passage item-relevant information occurs, which may be desirable. Further, allowing multiple hearings has been shown to decrease test anxiety (Chang & Read, 2008) and improve participants' self-ratings of comprehension (Brown, 2007). If there is evidence that anxiety is interfering with listening comprehension, allowing for multiple hearings may partially alleviate this issue.

Summary: Number and control over hearings

Overall, studies examining the effect of multiple hearings of the passage on L2 listening comprehension tend to show that replaying a passage improves comprehension. However, Henning's (1990) finding that allowing multiple hearings did not improve validity of test items must be considered when this factor is manipulated for test materials. Further, though the interaction between multiple hearings and listener proficiency is not completely consistent across studies, there is some evidence that this factor impacts listeners of different proficiency in different ways. Offering multiple hearings to examinees has also been criticized based on authenticity grounds. However, the capacity of multiple hearings to help the examinee overcome idiosyncrasies of the test environment and possibly lessen the effects of other factors which are not directly relevant to L2 listening proficiency (e.g., test anxiety) may make it a desirable factor to introduce in L2 listening comprehension testing.

Source of characteristic	Factor of interest	Consensus in the literature
Passage	Testing Conditions	Multiple hearings decrease difficulty, and note-taking may have similar effects when listeners can decide when to take notes.

Note taking

Compared with other test condition factors described above, the effects of note-taking on L2 listening comprehension are not as easy to predict. Because writing operates at a fraction of the speed of speaking (writing speed is 0.2 to 0.3 words/second and speaking is 2–3 words/second), note-taking introduces considerable time pressure (Piolat, Olive, & Kellogg, 2005). Also, note-takers must monitor incoming information to create and update representations in their working memory at the same time that they are making decisions about how to express the information they write in their notes (Piolat et al., 2005). The need to switch attention between multiple, concurrent activities puts demand on the central executive of working memory over and above that imposed by the activities themselves (Baddeley, 2003). Barbier and Piolat (2005; described in Piolat et al., 2005) examined the cognitive effort involved in note-taking from an L2 passage and found that taking notes required more cognitive effort than taking notes on a passage presented in the L1. These findings show that note-taking is particularly effortful for L2 listening. Because of the effort involved, taking notes in the L2 may actually hurt listening comprehension.

While L1 research suggests that having notes to review is more important than taking notes, results in L2 research are less clear-cut.

Alternatively, note-taking should benefit listening comprehension (particularly when a passage is available only once) because it allows the listener to capture the ephemeral bits of information in the passage (Chaudron, Loschky & Cook, 1994; Lin, 2006). The bulk of research on note-taking in the L1 indicates that having notes while answering test items is the most important aspect of note-taking, as opposed to encoding benefits provided by the act of taking notes alone (Dunkel et al., 1989; see Hartley, 1983, and Kiewra, 1985 for reviews of research examining note-taking in the L1). Taking notes while listening in the L2 may improve

Note-taking is an effortful activity for listeners, particularly those listening in their L2.

performance on comprehension items because L2 listeners can off-load information from the passage into their notes rather than having to retain all information mentally.

Note taking in the L2

Most research examining the effects of note-taking on L2 listening comprehension has found that note-taking can be beneficial for some types of passages or under certain circumstances. Lin (2006) examined note-taking for passages presented at faster or slower speech rates. Note-taking did not significantly affect overall performance on comprehension test items. However, analyses performed separately for the two speech rates used in the study found that participants who took notes outperformed those who were not allowed to take notes when speech rate was faster. Lin's (2006) finding that note-taking was more beneficial for faster passages than slower passages seems to contradict findings that note-taking in the L2 requires a large amount of cognitive effort (Piolat et al., 2005). Generally, faster speech rates hurt L2 listening comprehension (e.g., Rosenhouse et al., 2006), so the combination of note-taking and a faster speech rate would be expected to hurt comprehension more than taking notes while listening to slower passages. However, a potential reason for this inconsistency is that participants who listened to the faster lectures sometimes chose not to take notes even though they could, perhaps feeling that they could not manage both comprehending the lecture and taking notes on the material (Lin, 2006). The decision to forgo note-taking in order to cope with listening and comprehending may have been more clear-cut for the faster rate lectures than the slower rate lectures, leading fewer participants to attempt to take notes when their comprehension would have been aided more by focusing on listening.

Carrell et al. (2002) allowed their participants to take notes while listening to half of the passages in the study, but not while listening to the other half. Participants performed significantly better on test items when allowed to take notes during the corresponding passage. Both higher and lower-proficiency listeners (proficiency measured by Institutional TOEFL listening comprehension section, higher scores ≥ 49; lower scores < 49) showed the same benefit from note-taking, but note-taking had a larger impact on performance for the short (~2.5 minutes) lectures than for the long (~5 minutes) lectures and for passages with less familiar topics. These results demonstrate that note-taking can be beneficial to performance in listening comprehension tests, but may be less helpful for longer passages or those involving more familiar topics, though in neither case is note-taking likely to be detrimental to performance. It is also worth pointing out that, due in part to the findings of Carrell et al., the 2006 version of the TOEFL allowed note-taking on the listening section of the test (Zareva, 2005).

Note-taking may be most useful when the listener can make good decisions about when to take on the additional effort of note-taking.

Hale and Courtney (1994) examined how insisting that L2 listeners take notes while listening affected the impact of note-taking. For half of the passages, each participant was allowed to take notes while listening. In one condition, participants were merely permitted to take notes for this half of the passages; in the other condition they were *urged* to take notes for half the passages. Note-taking had no significant effect on performance on comprehension items when it was simply allowed; however, when note-taking was urged, performance was actually worse than when no notes were taken. Participants in both the urged and optional note-taking conditions did report feeling more at ease when they were able to take notes than when they were not, and they

believed taking notes helped them remember more information. This finding suggests that most participants in both conditions believed note-taking helped their performance on the test items, even though it did not do so significantly in either case, and actually hurt performance when participants were urged to take notes. Hale and Courtney (1994) explained the negative effects of urged note-taking through the speech rate of the passages, which may have been too fast to allow participants to listen and comprehend and write notes when they were urged to do so. In light of Lin's (2006) findings, another possible explanation for Hale and Courtney's findings is that insisting on note-taking prevented the L2 listeners from feeling that they could decide when to take notes and when not to take notes. The authors explained the lack of a positive effect of note-taking on performance for participants in the optional note-taking condition as arising from the brevity of the passages, which may not have provided enough information to require note-taking (< 250 words presented at 145 words/minute). However, this seems to contradict the findings of Carrell et al. (2002), whose shorter passages were roughly the same in length. As always, however, it is difficult to interpret a null effect. What can be concluded from Hale and Courtney's (1994) study is that urging L2 listeners to take notes can hurt their comprehension test performance.

The findings of Lin (2006) and Hale and Courtney (1994) suggest that note-taking must be appropriately used by the L2 listener in order to be helpful or, at the least, not detrimental, to L2 listening comprehension. The decision about whether to take notes qualifies as a metacognitive strategy for listeners: in order to make good decisions about note-taking, listeners must be able to monitor their own level of comprehension while listening and how their comprehension is affected by note-taking, as well as their likelihood of forgetting information if they do not take notes. A more in-depth discussion of these types of strategies can be found in the section on Metacognitive Strategies.

Summary: Note taking

Overall, the literature suggests that the ability to take notes can be advantageous for L2 listening comprehension under certain circumstances. Although Hale and Courtney (1994) found that participants performed worse when they took notes than when they did not, this was only in the case where they were urged to do so. When participants are allowed to choose whether to take notes, there is evidence that they can make the decision that best fits their needs (Lin, 2006). Other studies have found that, at worst, there is no effect of taking and having notes available on test performance for some participants or for certain types of passages (Carrell et al., 2002). Overall, the research exploring the effects of optional note-taking in L2 listening comprehension is limited but suggests that simply permitting note-taking is not detrimental to the examinees. However, the potential benefit to examinees must be weighed against the risk that they will feel compelled to take notes because it is possible to do so, which may damage their performance (Hale & Courtney, 1994).

Summary of research findings for test condition factors			
Factors with strong effects or convincing evidence		**Factors with sparse or inconsistent evidence**	
Beneficial to listeners:	**Multiple hearings**	Possibly difficult for listeners:	Shorter **time limits**
		Possibly beneficial to listeners:	**Note-taking,** but only when listeners can decide when to take notes

CONCLUSION

In general, the availability of research examining how certain characteristics of the listener, passage, and test-taking conditions affect L2 listening comprehension is limited, and the literature is plagued by inconsistencies in how factors of interest within these characteristics, such as L2 listening proficiency, coherence, and, note-taking, are described. In spite of these issues, the research does point to some conclusions, some more tempered than others, about what factors will impact L2 listening and the direction (positive or negative) of their effects.

The scientific literature points to several characteristics of listeners that play a role in L2 listening comprehension. If capable and motivated, listeners can apply particular metacognitive strategies while listening, such as preventing themselves from fixating on a particular word they missed, which can improve comprehension. Further, L2 listening proficiency will obviously impact listening comprehension and interact with many of the passage-based factors described in this review. For example, the usefulness of discourse markers seems to depend on the listener's ability to recognize these markers, which necessitates more experience with listening in the L2. However, the variety of characterizations of proficiency used in the literature (e.g., years of formal instruction, experimenter-developed pretests) make the results for this factor particularly difficult to summarize. Better research is needed to speak to how proficiency with a language interacts with passage-based and testing condition-based factors. In addition to differing in L2 listening proficiency, a listener may also have more or less working memory capacity. Research suggests that this quality of listeners will impact how they cope with larger amounts of spoken information (e.g., greater information density), low-frequency words, and noise. Working memory capacity is likely to influence the effects of other passage-based factors as well, but research has not yet addressed all the possibilities. Further, differences in working memory capacity may affect how readily listeners can take advantage of testing conditions, such as their ability to take notes while listening, and how much they will be affected by the time pressure imposed by an exam.

In terms of passage characteristics there is little research to suggest that longer passages necessarily increase the difficulty of L2 listening comprehension. Nonetheless, passage length is a predictive factor for listening difficulty because it is correlated with redundancy of information and information density. Because more redundancy decreases the difficulty of a passage, and greater information density increases difficulty, both factors must be taken into account in order to predict the impact of passage length on L2 listening comprehension. Research on passage complexity suggests that negatives and infrequent vocabulary increase difficulty, as does more culturally specific or implied information. Improving comprehensibility for L2 listeners through simplifying the sentences in a passage is not consistently successful, however. Increasing the number of discourse markers in a passage may improve comprehension, as well as including additional pauses. Familiar topics reduce L2 listening difficulty, possibly by enabling listeners to apply top-down knowledge more readily while listening. Factors that make listening difficult for L1 listeners, such as noise or distortion and unfamiliar speaker accents, have a similar impact on L2 listening comprehension, though to a larger degree than for native listeners. Research on the comprehension of passages with relatively fast speech rates suggests that this factor increases difficulty for the L2 listener, but inconsistency in how

speech rate is defined and what rate is determined to be fast makes specifying the particular rate at which L2 listeners will struggle challenging.

Another quality of passages which relates to several of the passage-based factors described here is its authenticity. Authentic passages differ from created passages in terms of their orality, a feature that captures factors such as redundancy, disfluencies, and syntactic complexity. Research on orality suggests that more oral passages are easier for L2 listeners to comprehend. Authentic dialogues have also been found to differ from created (textbook) dialogues in terms of information density and length. Though the amount of research addressing how authentic passages differ from created passages is very small, this combined with findings concerning orality suggest that authentic passages may be easier for listeners than created passages. However, other factors likely to be more prevalent in authentic passages, such as unfamiliar speaker accents, varying speech rates, culturally specific information, distortion, and noise, are likely to increase the difficulty of the passage. The authenticity of a passage is thus likely to have a mix of negative and positive impacts on L2 listening comprehension.

While passage-based factors can be partially controlled through the selection of passages, factors arising from the testing conditions can be fully controlled by test developers. Research shows that playing a passage multiple times improves listening comprehension if it affects comprehension at all. In contrast, allowing the listener to take notes has the capacity to both harm and help: if listeners attempts (or are urged) to take notes when their proficiency is low or the passage is challenging, note-taking may actually detract from comprehension; conversely, when voluntary and used judiciously, note-taking can help comprehension and later recall. Finally, findings addressing response time and its effects on the performance of cognitive tasks, though not specifically listening comprehension, indicate that increasing time pressure reduces working memory processing capacity, interfering with performance. Despite this general finding, it is likely impractical to have no time constraints in testing and there is reason to believe that simply increasing the time allotment for responding to test items (without permitting multiple hearings) will not help listening comprehension test performance. Pre-testing will help determine an approximate amount of time for responding to test items.

A number of the factors reviewed here have been found to interact with each other. While pauses generally improve L2 listening comprehension, this may only be when speech rate is not fast. Playing a passage multiple times, which generally results in lower L2 listening difficulty, may only have a positive effect when the passage exceeds a certain length. The speech rate of a passage may affect decisions about note-taking, thereby influencing the effect of allowing examinees to do so. Interactions such as these indicate that it is important to consider factors in conjunction when predicting the difficulty posed by a passage. Further, the prevalence of interactions that have been examined in only one or two studies points to the need for further research to examine how factors combine to decrease or increase L2 listening comprehension difficulty.

This review of the scientific literature suggests that during test development and the selection of authentic spoken passages, it is possible to anticipate some of the ways in which passage, examinee, and testing condition factors will influence L2 listening comprehension. The report provides an initial framework for assessing features of authentic

spoken passages in relation to proficiency levels. Further, areas where more research effort is needed in order to fully understand the impact of a factor can be clearly seen. In the next stage of the project, a taxonomy fully describing the demonstrated and potential relationships between the factors discussed in this review and their impact on L2 listening comprehension will be developed. This taxonomy will provide guidance for the development of future research projects exploring how passage, examinee and testing condition factors impact L2 listening comprehension.

REFERENCES

Adank, P., Evans, B. G., Stuart-Smith, J., Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 520–529.

Aiken, E. G., Thomas, G. S., & Shennum, W. A. (1975). Memory for a lecture: Effects of notes, lecture rate and informational density. *Journal of Educational Psychology, 67*(3), 439–444.

Alderson, J. C., Figueras, N., Kuijper, H., Nold, G., Takala, S., Tardieu, C. (2006). Analysing tests of reading and listening in relation to the common European framework of reference: The experience of the Dutch CEFR Construct Project. *Language Assessment Quarterly*, *3*(1), 3–30.

American Heritage Dictionary of the English Language (4th ed.). (2000). Boston: Houghton Mifflin Co.

Anderson, J. R. (2004). *Cognitive psychology and its implications* (6th ed.). New York: Worth.

Anderson, A., & Lynch, T. (1988). *Listening*. Oxford: Oxford University Press.

Anderson-Hsieh, J., & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning, 38*(4), 561–613.

Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal theee, um, new information. *Journal of Psycholinguistic Research, 32*(1), 25–36.

Arnold, J. E., Tanenhaus, M. K., Altmann, R. J., & Fagnano, M. (2004). The old and thee, uh, new: Disfluency and reference resolution. *Psychological Science, 15*(9), 578–582.

Aydelott, J., & Bates, E. (2004). Effects of acoustic distortion and semantic context on lexical access. *Language and Cognitive Processes,19*(1), 29–56.

Aydelott, J., Dick, F., & Mills, D. (2006). Effects of acoustic distortion and semantic context on event-related potentials to spoken words. *Psychophysiology, 43*, 454–464.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*, 829–839.

Baddeley, A. (2007). *Working memory, thought and action*. Oxford: Oxford University Press.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.

Baddeley, A. D, Eldridge, M., Lewis, Y, & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General*, *113*, 518–540.

Barbier, M. L., & Piolat, A. (2005). L1 and L2 cognitive effort of note taking and writing. In L. Allal, & B. Schneuwly (Eds.). *Proceedings of the Special Interest Group on Writing*. Geneva, Switzerland.

Barnett, J., Di Vesta, F., & Rogozinski, J. (1981). What is learned in note taking? *Journal of Educational Psychology, 73*(2), 181–192.

Bartlett, F.C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, England: Cambridge University Press.

Barreto S., & Ortiz, K. (2008). Influence of speech rate and loudness on speech intelligibility. *Pró-Fono Revista de Atualização Científica, 20*(2), 87–92.

Bejar, I. I. (1985). *Test speededness under number-right scoring: An analysis of the Test of English as a Foreign Language*. (Report No. RR-85-11). Princeton, NJ: Educational Testing Service.

Bejar, I., Douglas, D., Jamieson, J., Nissan, S., & Turner, J. (2000). *TOEFL 2000 listening framework: A working paper*. (TOEFL Monograph Series No. MA-19). Princeton, NJ: Educational Testing Service.

Bell, A., Brenier, J., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability Effects on Durations of Content and Function Words in Conversational English. *Journal of Memory and Language, 60*(1), 92–111.

Berger, K.W. (1977). *The most common 100,000 words used in conversations*. Kent, Ohio: Herald Publishing House.

Berne, J. E. (1995). How does varying pre-listening activities affect second language listening comprehension? *Hispania, 78*(2), 316–329.

Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25*(3), 371–405.

Blau, E. K. (1990). The effect of syntax, speed and pauses on listening comprehension. *TESOL Quarterly, 24*(4), 746–753.

Blau, E. K. (1991). More on comprehensible input: The effect of pauses and hesitation markers on listening comprehension. Unpublished paper presented at the Annual Meeting of the Puerto Rico Teachers of English to Speakers of Other Languages, San Juan, Puerto Rico.

Bodie, G., Worthington, D., Imhof, M., & Cooper, L. (2008). What would a unified field of listening look like? A proposal linking past perspectives and future endeavors. The International Journal of Listening, 22, 103–122.

Boekaerts, M., & Niemivirta, M. (2000). Self-regulated learning: Finding a balance between learning goals and ego-protective goals. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation: Theory, research, and applications* (pp. 417–450). San Diego, CA:Academic Press

Boekaerts, M., Pintrich, P., & Zeidner, M. (2000). *Handbook of self regulation*. San Diego: Academic Press.

Bolitho, R., Carter, R., Hughes, R., Ivanic, R., Masuhara, H., & Tomlinson, B. (2003). Ten questions about language awareness. *ELT Journal, 57*(3), 251–260.

Bongers, H. (1947). The *History and Principles of Vocabulary Control*. Woerden, Netherlands: Wocopi.

Bowles, A., Linck, J., Koeth, J., Mislevy, M., Campbell, S., Annis, R., Jackson, S., Doughty, C. & Bunting, M. (2009). Predicting success at DLIFLC with the DLAB 2: Interim results from field testing. *Technical Report*. College Park: University of Maryland Center for Advanced Study of Language.

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *The Journal of the Acoustical Society of America, 112*(1), 272–284.

Breen, M. P. (1985). Authenticity in the language classroom. *Applied Linguistics, 6*(1), 60–70.

Brennan, S. E., & Schober, M. F. (2001) How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language, 44*, 274–296.

Brindley, G., & Slatyer, H. (2002). Exploring task difficulty in ESL listening assessment. *Language Testing, 19*(4), 369–394.

Brown, G., & Yule, G. (1983). *Teaching the spoken language: An approach based on the analysis of conversational English*. Cambridge: Cambridge University Press.

Brown, R. A. (2007). Enhancing motivation through repeated listening. *Information and Communication Studies, 37*, 9–15.

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *Journal of the Acoustical Society of America, 109*, 1101–1109.

Buck, G. (2001). *Assessing Listening*. Cambridge: Cambridge University Press.

Buck, G., & Tatsuoka, K. (1998). Application of the rule-space procedure to language testing: Examining attributes of a free response listening test. *Language Testing, 15*(2), 119–157.

Byrd, D., & Tan, C. (1996). Saying consonant clusters quickly. *Journal of Phonetics, 24*, 263–282.

Camiciottoli, B. C. (2004). Intereactive discourse structuring in L2 guest lectures: Some insights from a comparative corpus-based study. *Journal of English for Academic Purposes*, *3*, 39–54.

Carhart, R., Tillman, T. W., & Greetis, E. S. (1969). Perceptual masking in multiple sound backgrounds. *Journal of the Acoustical Society of America, 45*, 694–703.

Carrell, P. L., Dunkel, P. A., & Mollaun, P. (2002). *The effects of notetaking, lecture length and topic on the listening component of the TOEFL 2000*. (TOEFL Monograph Series No. MS-23). Princeton, NJ: Educational Testing Service.

Carroll, D. W. (1986). *Psychology of Language*. Monterey, CA: Brooks/Cole Publishing.

Carroll, J. B. (1977). On learning from being told. In M. C. Wittrock (Ed.), *Learning & Instruction* (2nd ed., pp. 496–512). Berkeley, CA: McCutchan.

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American Heritage Word Frequency Book*. Boston, MA: Houghton Mifflin.

Cervantes, R., & Gainer, G. (1992). The effects of syntactic simplification and repetition on listening comprehension. *TESOL Quarterly*, *26*(4), 767–774.

Chafe, W.L. (1985). Linguistic differences produced by differences between speaking and writing. In D.R. Olsen, N. Torrance, and A. Hilyard (Eds.), *Literacy, language and learning: the nature and consequences of reading and writing* (pp. 105–147). Cambridge: Cambridge University Press.

Chang, A. C., & Read, J. (2006). The effects of listening support on the listening performance of EFL learners. *TESOL Quarterly, 40*(2), 375–420.

Chang, A. C., & Read, J. (2008). Reducing listening text anxiety through various forms of listening support. *TESL-EJ, 12*(1), 1–25.

Chang, C. S. (1999). The effect of repeated listening on different levels of ESL learners. In *Proceedings of the 1st Conference on Applied English Teaching* (pp. 72–80). Taoyuan, Republic of China: Ming Chuan University, Department of Applied English.

Chaudron, C. (1983). Simplification of input: Topic restatements and their effects on L2 learners' recognition and recall. *TESOL Quarterly, 17*(3), 437–458.

Chaudron, C. (1988). *Second language classrooms: Research on teaching and learning*. New York: Cambridge University Press.

Chaudron, C., & Richards, J. C. (1986). The effect of discourse markers on the comprehension of lectures. *Applied Linguistics, 7*(2), 113–127.

Chaudron, C., Loschky, L., & Cook, J. (1994). Second language listening comprehension and note-taking. In J. Flowerdew (Ed.) *Academic Listening: Research Perspectives* (pp. 75–92). Cambridge: Cambridge University Press.

Chemers, M., Hu, L., & Garcia, B. (2001). Academic self-efficacy and first-year college student performance and adjustment. *Journal of Educational Psychology, 93,* 55–64.

Cheung, H. (1994). *The perceived and actual effects of speech rates on listening comprehension*. Unpublished master's thesis, City University of Hong Kong.

Chiang, C. S., & Dunkel, P. (1992). The effect of speech modification, prior knowledge, and listening proficiency on EFL lecture learning. *TESOL Quarterly, 26*(2), 345–374.

Churchland, P. (1999). Learning and conceptual change: The view from the neurons. In A. Clark and P. Millican (Eds.), *Connectionism, concepts, and folk psychology* (pp. 7–43). Oxford, UK: Oxford University Press.

Clark. A. J. (1989). Communication confidence and listening competence: An investigation of the relationships of willingness to communicate, communication apprehension, and receiver apprehension to comprehension of content and emotional meaning in spoken messages. *Communication Education, 38*, 237–248.

Clopper, C. G., & Bradlow, A. R. (2008). Perception of dialect variation in noise: Intelligibility and classification. *Language and Speech, 51*(3), 175–198.

Cobb, M. (2004). Input elaboration in second and foreign language teaching. *Dialog on Language Instruction, 16*(1), 13–23.

Collard, P., Corley, M., MacGregor, L. J., Donaldson, D. I., (2008). Attention orienting effects of hesitations in speech: Evidence from ERPs. *Journal of Experimental Psychology: Learning, Memory and Cognition, 34*, 696–702.

Coniam, D. (1996). Computerized dictation for assessing listening proficiency. *CALICO Journal, 13*(2 & 3), 73–85.

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: the importance of working memory capacity. *Psychonomic Bulletin & Review, 8*(2), 331–335.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*, 163–183.

Cooke, M., Lecumberri, M. L. G., & Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *The Journal of the Acoustical Society of America, 123*(1), 414–427.

Coulthard, M. and Montgomery,M. (1981) . The structure of monologue. In M. Coulthard and M. Montgomery (Eds.) *Studies in Discourse Analysis* (pp.31–39). London: Routledge and Kegan Paul.

Corley, M., & Hartsuiker, R. J. (2003). Hesitation in speech can...um...help a listener understand. *Proceedings of the Twenty-Fifth Meeting of the Cognitive Science Society* (pp. 276–281). Mahwah, NJ, USA: Lawrence Erlbaum.

Crossley, S. A., McCarthy, P. M., Louwerse, M. M., & McNamara, D. S. (2007). A linguistic analysis of simplified and authentic texts. *The Modern Language Journal, 91*, 15–30.

Crystal, D. (2003). *A Dictionary of Linguistics and Phonetics*. (5th ed.). Malden, MA: Blackwell Publishing.

Cutler, A. (2001). Listening to a second language through the ears of a first. *Interpreting, 5*, 1–23.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*(2), 141–201.

Cutler, A., Lecumberri, M. L. G., & Cooke, M. (2008). Consonant identification in noise by native and non-native listeners: Effects of local context. *Journal of the Acoustical Society of America, 124*(2), 1264–1268.

Dahl, D. (1981). The role of experience in speech modifications for second language learners. *Minnesota Papers in Linguistics and Philosophy of Language, 7*, 78–93.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*(4), 422–433.

De Rooij, J. (1979). *Speech punctuation: An acoustic and perceptual study of some aspects of speech prosody in Dutch*. Unpublished doctoral dissertation, Rijksuniversiteit Utrecht, Utrecht, The Netherlands.

De Groot, A. M. B., & Poot, R. (1997). Word translation at three levels of proficiency in a second language: The ubiquitous involvement of conceptual memory. *Language Learning, 47*(2), 215–264.

Deci, E. L., & Ryan, R. M. (1995).Human autonomy: The basis for true self-esteem. In M. H. Kernis (Ed.), *Efficacy, agency, and self-esteem* (pp. 31–49). New York: Plenum.

Defense Language Institute Foreign Language Center. (2010). *Defense Language Proficiency Testing System 5 Framework*. Monterey, CA.

Derwing, T., & Munro, M. (2001). What speaking rates do non-native listeners prefer? *Applied Linguistics, 22*(3), 324–337.

Dudley-Evans, A., & Johns, T. (1981) . A team teaching approach to lecture comprehension for overseas students . In *The Teaching of Listening Comprehension*: *ELT Documents Special* (pp. 30–46). London: The British Council.

Dunkel, P. (1991). Listening in the native and second/foreign language: Toward an integration of research and practice. *TESOL Quarterly, 25*(3), 431–457.

Dunkel, P. A., and Davis, J. N. (1994). The effects of rhetorical signaling cues on the recall of English lecture information by speakers of English as a native or second language. In J. Flowerdew (Ed.) *Academic Listening: Research Perspectives*. (pp. 55–74). Cambridge, UK: Cambridge University Press.

Dunkel, P., Henning, G., & Chaudron, C. (1993). The assessment of an L2 listening comprehension construct: A tentative model for test specification and development. *The Modern Language Journal, 77*(2), 180–191.

Dunkel, P., Mishra, S., & Berliner, D. (1989). Effects of note taking, memory, and language proficiency on lecture learning for native and nonnative speakers of English. *TESOL Quarterly, 23*(3), 543–549.

Eastman, J. K. (1991).Learning to listen and comprehend: The beginning stages. *System, 19,* 179–187.

Eilam, B., & Aharon, I. (2003). Students' planning in the process of self-regulated learning. *Contemporary Educational Psychology, 28*(3), 304–334.

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition, 26*, 59–84.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19–23.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 972–992.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.) *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). New York: Cambridge University Press.

Eom, M. (2006). *An investigation of operationalized constructs of second language listening tests*. Unpublished doctoral thesis, University of Iowa.

Ferreira, F., Henderson, J. M., Anes, M. D., Weeks, P. A., & McFarlane, D. K. (1996). Effects of lexical frequency and syntactic complexity in spoken-language comprehension: Evidence from the auditory moving-window technique. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(2), 324–355.

Ferris, D., & Tagg, T. (1996). Academic listening/speaking tasks for ESL students: Problems, suggestions, and implications. TESOL *Quarterly, 30*(2), 297–317.

Field, J. (2003). Promoting perception: lexical segmentation in L2 listening. *ELT Journal, 57*(4), 325–334.

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

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

Floccia, C., Butler, J., Goslin, J., & Ellis, L. (2009). Regional and Foreign Accent Processing in English: Can Listeners Adapt? *Journal of Psycholinguistic Research, 38*(4), 379–412

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? Journal of Experimental Psychology: *Human Perception and Performance, 32*, 1276–1293.

Flowerdew, J., & Miller, L. (1997). The teaching of academic listening comprehension and the question of authenticity. *English for Specific Purposes, 16*(1), 27–46.

Flowerdew, J. (1994). *Academic Listening: Research Perspectives*. Cambridge: Cambridge University Press.

Flowerdew, J., & Miller, L. (1992). Student perceptions, problems and strategies in second language lecture comprehension. *RELC Journal*, 23(2), 60–80.

Flowerdew, J. and Miller, L. (1997). The teaching of academic listening comprehension and the question of authenticity. *English for Specific Purposes, 16*(1), 27–46.

Flowerdew, J., & Tauroza, S. (1995). The effect of discourse markers on second language lecture comprehension. *Studies in Second Language Acquisition, 17*(4), 435–458.

Forster, K. I. (1989). Basic issues in lexical processing. In: Marslen-Wilson, W. (Ed.), *Lexical Representation and Process*. MIT, Cambridge, MA.

Fox Tree, J. E. (1995). The effect of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language*, *34*, 709–738.

Freedle, R., & Kostin, I. (1992). *The prediction of GRE reading comprehension item difficulty for expository prose passages for each of three item types: Main ideas, inferences and explicit statements*. (ETS Research Report RR-91-59). Princeton, NJ: Educational Testing Service.

Freedle, R., & Kostin, I. (1993). The prediction of TOEFL reading item difficulty: implications for construct validity. *Language Testing, 10*(2), 133–170.

Freedle, R., & Kostin, I. (1996). *The prediction of TOEFL listening comprehension item difficulty for minitalk passages: Implications for construct validity*. (TOEFL Research Report No. 56). Princeton, NJ: Educational Testing Service.

Freedle, R., & Kostin, I. (1999). Does the text matter in a multiple-choice test of comprehension? The case for the construct validity of TOEFL's minitalks. *Language Testing, 16*(1), 2–32.

Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *Journal of the Acoustical Society of America, 115*, 2246–2256.

Fuster, J. M. (2002). Frontal lobe and cognitive development. *Journal of Neurocytology, 31*, 373–385.

Gainer, G. (1997). *A comparison of the effects of pure repetition and embedded elaboration on the listening comprehension of EFL students*. Unpublished manuscript.

Garcia, P. (2004). Pragmatic comprehension of high and low level language learners. *TESL-EJ, 8*. Retrieved January 25, 2010, from http://www-writing.berkeley.edu/tesl-ej/ej30/a1.html

Gass, S., & Varonis, E. M. (1984).The Effect of Familiarity on the Comprehensibility of Nonnative Speech. *Language Learning, 34*(1), 65–89

Gilmore, A. (2004). A comparison of textbook and authentic interactions. *ELT Journal, 58*(4), 363–374.

Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching, 40*, 97–118.

Goh, C. (1998a). Strategic Processing and Metacognition in Second Language Listening. PhD thesis, Lancaster University, Lancaster, UK. Cited in Goh, C. C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. *System, 28*, 55–75.

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

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

Golestani, N., Rosen, S., & Scott, S. K. (2009). Native-language benefit for understanding speech-in-noise: The contribution of semantics. *Bilingualism: Language and Cognition, 12*, 385–392.

Grice, P. (1975). Logic and Conversation. In P. Cole and J. Morgan (Eds.), *Syntax and Semantics* (Vol. 3, pp.41–58). New York: Academic Press.

Griffiths, R. (1990). Speech rate and NNS comprehension: A preliminary study in time-benefit analysis. *Language Learning, 40*(3), 311–336.

Griffiths, R. (1990a). Facilitating listening comprehension through rate-control. *RELC Journal, 21*(1), 55–65.

Griffiths, R. (1992). Speech rate and listening comprehension: Further evidence of the relationship. *TESOL Quarterly, 26*(2), 385–390.

Guion, S., Flege, J., Akahane-Yamada, R., & Pruitt, J. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America, 107*(5), 2711–2724.

Guiraud, H. (1954). *Les Caractères Statistiques du Vocabulaire*. Paris: Presses Universitaires de France.

Hale, G. A. (1992). *Effects of amount of time allowed on the test of written English*. (TOEFL Research Reports RR-39). Princeton, NJ: Educational Testing Service.

Hale, G. A., & Courtney, R. (1994). The effects of note-taking on listening comprehension in the Test of English as a Foreign Language. *Language Testing, 11*(1), 29–47.

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition, 14*(1), 25–38.

Hartley, J. (1983). Notetaking research: Resetting the scoreboard. *Bulletin of the British Psychological Society, 36*, 13–14.

Henning, G. (1990). *A study of the effects of variation of short-term memory load, reading response length, and processing hierarchy on TOEFL listening comprehension item performance*. (TOEFL Research Reports RR-33). Princeton, NJ: Educational Testing Service.

Herron, C. A., & Seay, I. (1991). The effect of authentic oral texts on student listening comprehension in the foreign language classroom. *Foreign Language Annals, 24*(6), 487–495.

Hinkel, E. (2006). Current perspectives on teaching the four skills. *TESOL Quarterly, 40*(1), 109–131.

Hirai, A. (1999). The relationship between listening and reading rates of Japanese EFL learners. *The Modern Language Journal, 83*, 367–384.

Hirsh, D., & Nation, I. S. P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language, 8*, 689–696

Hron, A., Kurbjuhn, I., Mandl, H., & Schnotz, W. L. (1985). Structural inferences in reading and listening. In G. Richheit & H. Strohner (Eds.), *Inferences in Text Processing* (pp. 221–245). Amsterdam: North-Holland.

Hymes, D. (1972). *Towards communicative competence*. Philadelphia: University of Pennsylvania Press.

Hussey, E. K., Teubner-Rhodes, S. E., Dougherty, M. R., & Novick, J. M. (2010). Parsing under pressure: The role of performance pressure in cognitive control and syntactic ambiguity resolution. Poster presented at *CUNY 2010: Conference on Human Sentence Processing*, New York City, NY.

Ikeno, A., & Hansen, J. H. L. (2006). Perceptual recognition cues in native English accent variation: Listener accent perceived accent, and comprehension. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing* (pp. I-401–I–404), Toulouse, France, May 2006.

Inbar, O. (1988). The effect of text and question-type on achievement in EFL listening comprehension. English Teachers Journal, 37, 45–48.

Interagency Language Roundtable Language Skill Level Descriptions: Listening. (1985). Retrieved from http://www.govtilr.org/Skills/ILRscale3.htm. Retrieved 12 February 2010.

Ito, Y. (2001). Effect of reduced forms on ESL learners' input-intake process. *Second Language Studies, 20*(1), 99–124.

JACET (Japan Association of College English Teachers). (2003). *JACET list of 8000 basic words*. JACET, Tokyo.

Jacewicz, E., Fox, R., & O'Neill, C. (2009). Articulation rate across dialect, age, and gender. *Language Variation and Change, 21*, 233–256.

Jacobs, George., Chuawanlee, W., Hoga, B.K ., Sakumoto, D., Saka, S., & Meehan,K. (1988). The effect of pausing on listening comprehension. Paper presented at the *8th Annual Second Language Research Forum*.Honolulu,Hawaii (March 3- 6) .

Johnson, D.B. (1972). Computer frequency control of vocabulary in language learning reading materials. *Instructional Science, 1*, 121-131.

Johnston, J. and Doughty, C. J. (2006, September 01). *Final Technical Report M.2: Technology-Mediated Language Training*. (TTO 101: Deliverable). College Park: University of Maryland Center for Advanced Study of Language.

Jonides, J. (1995). Working memory and thinking. In E. E. Smith & D. N. Osherson (Eds.), *An invitation to cognitive science: Thinking* (pp.215–265). Cambridge, MA: Bradford.

Juffs, A. (2004). Representation, processing and working memory in a second language. *Transactions of the Philological Society, 102*(2), 199–225.

Jung, E. H. (2003). The role of discourse signaling cues in second language listening comprehension. *The Modern Language Journal, 87*, 562–577.

Just & Carpenter (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122–149.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122–149.

Kang, O., Rubin, D., & Pickering, L. (in press). Suprasegmental measures of accentedness and judgments of English language learner proficiency in oral English. *The Modern Language Journal*.

Kaplan, R. B. (2001). Cultural thought patterns in inter-cultural education. In T. Silva & P. K. Matsuda (Eds.) *Landmark Essays on ESL Writing* (pp. 11–25). Mawah, NJ: Lawrence Erlbaum Associates, Inc.

Kelch, K. (1985). Modified input as an aid to comprehension. *Studies in Second Language Acquisition, 7*(1), 81–90.

Kiany, G. R., & Jalali, S. (2006). Relationship between modality, types of passages, and performance of advanced EFL learners on listening comprehension tests. *Iranian Journal of Applied Linguistics, 9*(2), 79–99.

Kienbaum, B. E., Russell, A. J., & Welty, S. (1986). *Communicative competence in foreign language learning with authentic materials: Final project report*. (ERIC Document Reproduction Service No. ED 275200).

Konig, C. J., Buhner, M., & Murling, G. (2005). Working memory, fluid intelligence, and attention are predictors of multitasking performance, but polychronicity and extraversion are not. *Human Performance, 18*, 243–266.

Koreman, J. (2006). Perceived speech rate: The effects of articulation rate and speaking style in spontaneous speech. *Journal of the Acoustical Society of America, 119*(1), 582–596.

Koster, C. J. (1987). *Word Recognition in Foreign and Native Language: Effects of Context and Assimilation*. Dordrecht: Foris Publications.

Kostin, I. (2004). *Exploring item characteristics that are related to the difficulty of TOEFL dialogue items*. (TOEFL Research Report RR-79). Princeton, NJ: Educational Testing Service.

Lakoff, R. (2000). *The Language War*. Berkeley: University of California Press.

Lantoff, J. (1999). Second culture acquisition: cognitive considerations. In E. Hinkel (Ed.), *Culture in second language teaching and learning*. Oxford, UK: Oxford University Press.

Larsby, B., Hällgren, M., Lyxell, B., & Arlinger, S. (2005). Cognitive performance and perceived effort in speech processing tasks: Effects of different noise backgrounds in normal-hearing and hearing-impaired subjects. *International Journal of Audiology, 44*(3), 131–143.

Lecumberri, M. L. G., & Cooke, M. (2006). Effect of masker type on native and non-native consonant perception in noise. *Journal of the Acoustical Society of America, 119*(4), 2445–2454.

Leeser, M. J. (2004). The effects of topic familiarity, mode, and pausing on second language learners' comprehension and focus on form. *Studies in Second Language Acquisition, 26*, 587–615.

Leeser, M. J. (2007). Learner-based factors in L2 reading comprehension and processing grammatical form: Topic familiarity and working memory. *Language Learning, 57*(2), 229–270.

Leow, R. P. (1993). To simplify or not to simplify. *Studies in Second Language Acquisition, 15*, 333–355.

Lin, M. (2006). *The effects of note-taking, memory and rate of presentation on EFL learners' listening comprehension*. Unpublished doctoral dissertation, La Sierra University, California.

Lindblom, B., & Rapp, K. (1973). Some temporal regularities of spoken Swedish. *Papers in Linguistics from the University of Stockholm, 21*, 1–59.

Lindblom, B., Sussman, H., & Agwuele, A. (2009). A duration-dependent account of coarticulation for hyper- and hypoarticulation. *Phonetica, 66*, 188–195.

Long, M. H. and Ross, S. J. (2009). Input elaboration: a viable alternative to “authentic” and simplified texts. In K. Nakai (Ed.) *Festschrift for Yasutaka Yano* (pp. 307-323). Tokyo: Waseda University Press.

Lund, R. J. (1991). A taxonomy for teaching second language listening. *Foreign Language Annals, 23*(2), 105–115.

Lynch, A. J. (1988). *Grading foreign language listening comprehension materials: The use of naturally modified interaction*. Unpublished doctoral dissertation, University of Edinburgh.

Major, R., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2002). The effects of non-native accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly, 36*(2), 173–190.

Major, R., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2005). Testing the effects of regional, ethnic and international dialects of English on listening comprehension. *Language Learning, 55*, 37–69.

McInnes, A., Humphries, T., Hogg-Johnson, S., & Tannock, R. (2003). Listening comprehension and working memory are impaired in attention-deficit hyperactivity disorder irrespective of language impairment. *Journal of Abnormal Child Psychology, 31*(4), 427–443.

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language, 55*(3), 381–401.

Meyer, B. J. F., & Freedle, R. O. (1984). Effects of discourse type on recall. *American Educational Research Journal, 21*(1), 121–143.

Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. F. Healy, & L. E. Bourne (Eds.), *Foreign Language Learning* (pp. 339–364). London: Lawrence Erlbaum Associates.

Mokhtari, K., & Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology, 94*(2), 249–259.

Moll, K., Cardillo, E., & Aydelott Utman, J. (2001). Effects of competing speech on sentence-word priming: Semantic, perceptual, and attentional factors. In Moore, J. D., & Stenning, K. (Eds.), *Proceedings of the Twenty-third Annual Conference of the Cognitive Science Society* (pp. 651–656). Mahwah, NJ, USA: Lawrence Erlbaum Associates, Inc.

Moore, R., Adams, E., Dagenais, P., & Caffee, C. (2007). Effects of reverberation and filtering on speech rate judgment. *International Journal of Audiology, 46*, 154–160.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *The Quarterly Journal of Experimental Psychology, 11*, 56–60.

Moyer, A. (2006). Language contact and confidence in second language listening comprehension: A pilot study of advanced learners of German. *Foreign Language Annals, 39*(2), 255–275.

Munro, M. J., Derwing, T. M., and Morton, S. L. (2006). The Mutual Intelligibility of Speech. *Studies in Second Language Acquisition 28*, 111–131.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64*(5), 482–488.

Murray, I., & Arnott, J. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America, 93*(2), 1097–1108.

Nábělek, A. K., & Donahue, A. M. (1984). Perception of consonants in reverberation by native and non-native listeners. *Journal of the Acoustical Society of America, 75*(2), 632–634.

Nakatani, L., O'Connor, K., & Aston, C. (1981). Prosodic aspects of American English speech rhythm. *Phonetica, 38*, 84–105.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, England: Cambridge University Press.

Nation, I. S. P. (1993). Using dictionaries to estimate vocabulary size: Essential, but rarely followed, procedures. *Language Testing*, 10(1), 27–40.

Nekrasova, T. M. (2009). English L1 and L2 speakers' knowledge of lexical bundles. *Language Learning, 59*(3), 647–686.

Nieuwborg, E. 1992 "Tekstdekking en tekstbegrip: Een experimenteel onderzoek" {Text coverage and reading comprehension: An experimental investigation} in A Halbo (Ed.) *Evaluation and Language Teaching Liber Amicorum Frans van Passel.* Bern: Peter Lang

Nissan, S., DeVincenzi, F., & Tang, K. L. (1996). *An analysis of factors affecting the difficulty of dialogue items in TOEFL listening comprehension*. (ETS Research Report 95-37). Princeton, NJ: Educational Testing Service.

Nooteboom, S. (1972). *Production and perception of vowel duration: A study of durational properties of vowels in Dutch*. Unpublished doctoral dissertation, Rijksuniversiteit Utrecht, Utrecht, The Netherlands.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Nueroscience, 5*(3), 263–281.

O'Malley, J. M., Chamot, A. U., & Küpper, L. (1989). Listening comprehension strategies in second language acquisition. *Applied Linguistics, 10*(4), 418–437.

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.

Oh, S. Y. (2001). Two types of input modification and EFL reading comprehension: Simplification versus elaboration. *TESOL Quarterly, 35*(1), 69–96.

Osada, N. (2004). Listening comprehension research: A brief review of the past thirty years. *Dialogue, 3*, 53–66.

Osaka, M., & Osaka, N. (1992). Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society, 30*, 287–289.

Osaka, M., Osaka, N., & Groner, R. (1993). Language-independent working memory: evidence from German and French reading span tests. *Bulletin of the Psychonomic Society, 31*, 117–118.

Ostyn, P. & Godin, P. (1985). An alternative approach to language teaching. *The Modern Language Journal, 69*, 346-355.

Palmer, H. (1931). *The Second Interim Report on Vocabulary*. Tokyo: I.R.E.T.

Palmer, D. J., & Goetz, E. T. (1988). Selection and use of study strategies: the role of the students' beliefs about self and strategies. In Weinstein, C. E., Goetz, E. T., & Alexander, P. A. (Eds.) *Learning and Study Strategies: Issues in Assessment, Instruction, and Evaluation* (pp. 41–61). San Diego, CA: Academic Press.

Pan, S., McKeown, K., & Hirschberg, J. (2001). Semantic abnormality and its realization in spoken language. *Proceedings of Eurospeech 2001*, 333–336.

Park, G. P. (2004). Comparison of L2 listening and reading comprehension by university students learning English in Korea. *Foreign Language Annals, 37*(3), 448–458.

Parker, K., & Chaudron, C. (1987). The effects of linguistic simplification and elaborative modifications on L2 comprehension. *University of Hawaii Working Papers in ESL, 6*, 107–133.

Payton, K. L., Uchanski, R. M., & Braida, L. D. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *Journal of the Acoustical Society of America, 95*(3), 1581–1592.

Pica, T., Young, R., & Doughty, C. (1987). The impact of interaction on comprehension. *TESOL Quarterly, 21*(4), 737–758.

Piolat, A., Olive, T., & Kellogg, R. T. (2005). Cognitive effort during note taking. *Applied Cognitive Psychology, 19*(3), 291–312.

Powers, D. E., & Fowles, M. E. (1996). Effects of applying different time limits to a proposed GRE writing test. *Journal of Educational Measurement, 33*(4), 433–452.

Preiss. R. W., & Wheeless, L. R. (1989). Affective responses in listening: A meta-analysis of receiver apprehension outcomes. *The International Journal of Listening, 3*, 72–102.

Quené, H. (2007). On the just noticeable difference for tempo in speech. *Journal of Phonetics, 35*, 353–362.

Quené, H. (2008). Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo. *Journal of the Acoustical Society of America, 123*(2), 1104–1113.

Richards, J. (1990). *The language teaching matrix*. Cambridge, UK: Cambridge University Press.

Richards, J. (2006). Materials development and research – Making the connection. *RELC Journal, 37*(1), 5–26. Cited in Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching, 40*(3), 191–210.

Richards, J. C. (1983). Listening comprehension: Approach, design, procedure. *TESOL Quarterly, 17*(2), 219–240.

Rings, L. (1986). Authentic language and authentic conversational texts. *Foreign Language Annals, 19*(3), 203–208.

Robb, M., Maclagan, M., & Chen, Y. (2004). Speaking rates of American and New Zealand varieties of English. *Clinical Linguistics and Phonetics, 18*(1), 1–15.

Rogers, C. V., & Medley, F. W. (1988). Language with a purpose: Using authentic materials in the foreign language classroom. *Foreign Language Annals, 21*(5), 467–476.

Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics, 27*, 465–485.

Rose, R. L. (1998). T*he communicative value of filled pauses in spontaneous speech*. Unpublished Masters Thesis, University of Birmingham, Birmingham, U.K.

Rosenhouse, J., Haik, L., & Kishon-Rabin, L. (2006). Speech perception in adverse listening conditions in Arabic-Hebrew bilinguals. *International Journal of Bilingualism, 10*(2), 119–135.

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

Rost, M. (2005). L2 Listening. In E. Hinkel (Ed.), *Handbook of Research on Second Language Teaching and Learning*. (pp. 503–528). Mahwah, NJ: Lawrence Erlbaum.

Rost, M. (2006). Areas of research that influence L2 listening instruction. In E. Uso-Juan and A. Martinez-Flor (Eds.) *Current Trends in the Development and Teaching of the Four Language Skills*. (pp. 47–74). New York: Mouton de Gruyter.

Rounds, P. (1987). Characterizing successful classroom discourse for NNS teaching assistants. *TESOL Quarterly, 21*(4), 643–671.

Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal, 78*(2), 199–221.

Rubin, R., & Roberts, C. V. (1987). A comparative examination and analysis of three listening tests. *Communication Education, 36*, 142–153.

Rupp, A. A., Garcia, P, & Jamieson, J. (2001). Combining multiple regression and CART to understand difficulty in second language reading and listening comprehension test items. *International Journal of Testing, 1*(3 & 4), 185–216.

Sadighi, F., & Zare, S. (2006). Is listening comprehension influenced by the background knowledge of the learners? A case study of Iranian EFL learners. *The Linguistics Journal, 1*(3), 110–126.

Sakai, H. (2009). Effect of repetition of exposure and proficiency level in L2 listening tests. *TESOL Quarterly, 43*(2), 360–371.

Sanders, L. D., Neville, H. J., & Woldorff, M. G. (2002). Speech segmentation by native and non-native speakers: The use of lexical, syntactic, and stress-pattern cues. *Journal of Speech, Language and Hearing Research, 45*, 519–530.

Sasaki, M. (2000). Effects of cultural schemata on students' test-taking processes for cloze tests: A multiple data source approach. *Language Testing, 17*(1), 85–114.

Schmidt-Rinehart, B. C. (1994). The effects of topic familiarity on second language listening comprehension. *The Modern Language Journal, 78*(2), 179–189.

Schoonen, R., Hulstijn, J., & Boosers, B. (1998). Metacognitive and language-specific knowledge in native and foreign language reading comprehension: An empirical study among Dutch students in grades 6, 8 and 10. *Language Learning, 48,* 71–106.

Sciarone, A. G. (1979). *Woordjes Leren in het Vreemde-talenonderwijs*. Muiderberg, Netherlands: Coutinho.

Shimizu, T., Makishima, K., Yoshida, M., & Yamagishi, H. (2002). Effect of background noise on perception of English speech for Japanese listeners. *Auris Nasus Larynx, 29*, 121–125.

Shohamy, E., & Inbar, O. (1991). Validation of listening comprehension tests: The effect of text and question type. *Language Testing 8*(1), 23–40.

Siemer, M., & Reisenzein, R. (1998). Effects of mood on evaluative judgements: Influence of reduced processing capacity and mood salience. *Cognition and Emotion, 12*(6), 783–805.

Simpson, S., & Cooke, M. P. (2005). Consonant identification in N-talker babble is a non-monotonic function of N. *Journal of the Acoustical Society of America, 118*, 2775–2778.

Smiley, S. S., Oakley, D. D., Worthen, D., Campione, J. C., & Brown, A. L. (1977). Recall of thematically relevant material by adolescent good and poor readers as a function of written versus oral presentation. *Journal of Educational Psychology, 69*(4), 381–387.

Song, M. Y. (2008). Do divisible subskills exist in second language (L2) comprehension? A structural equation modeling approach. *Language Testing, 25*(4), 435, 464.

Sonnenschein, S. (1982). The effects of redundant communication on listeners: When more is less. *Child Development, 53*, 717–729.

Sparks, R., & Ganschow, L. (2001). Aptitude for learning a foreign language. *Annual Review of Applied Linguistics, 21,* 90–112.

Steinfeldt, E A. (1965). *Russian Word Count*. Translated by V. Korotky. Moscow: Progress Publishers.

Stine, E. L., Wingfield, A., & Poon, L. W. (1986). How much and how fast: Rapid processing of spoken language in later adulthood. *Psychology and Aging, 1*(4), 303–311.

Strodt-Lopez, B. (1991). Tying it all in: Asides in university lectures. *Applied Linguistics, 12*, 117–140.

Taguchi, N. (2005). Comprehending implied meaning as a foreign language. *The Modern Language Journal, 89*, 543–562.

Taguchi, N. (2008). The effect of working memory, semantic access, and listening abilities on the comprehension of conversational implicatures in L2 English. *Pragmatics & Cognition, 16*(3), 517–539.

Tannen, D. (1982). Oral and literate strategies in spoken and written narratives. *Language, 58*(1), 1–21.

Tauroza, S. (2001). Second language lecture comprehension in naturalistic controlled conditions. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for academic purposes* (pp. 360–374). Cambridge: Cambridge University Press.

Teng, H. C. (2001). Impact of speech modification on EFL listening. *Proceedings of the PAC3 at JALT 2001*, Nov. 22–25, 2001.

Thanajaro, M. (2000). *Using authentic materials to develop listening comprehension in the English as a Second Language classroom*. Unpublished dissertation, Virginia Polytechnic Institute and State University.

Thompson, I. (1995). Testing listening comprehension. *AATSEEL Newsletter, 37*, 24–31.

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

Tsao, Y., Weismer, G., & Iqbal, K. (2006). Interspeaker variation in habitual speaking rate: Additional evidence. *Journal of Speech, Language, and Hearing Research, 49*, 1156–1164.

Tyler, A.E., Jefferies, A.A. & Davies, C.E. (1988) . The effect of discourse structuring devices on listener perceptions of coherence in non-native university teacher's spoken discourse. World Englishes, 7(2), 101-110.

Tyler, M. D. (2001). Resource consumption as a function of topic knowledge in nonnative and native comprehension. *Language Learning, 51*(2), 257–280.

Van Dijk, T. A. (1979). Relevance assignment in discourse comprehension. *Discourse Processes, 2*, 113–126.

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

Vandergrift, L. (2006). Second language listening: Listening ability or language proficiency? *The Modern Language Journal, 90*, 6–18.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching, 40*(3), 191–210.

Vandergrift, L., & Tafaghodtari, M. H. (2010). Teaching L2 learners how to listen does make a difference: An empirical study. *Language Learning*, doi: 10.1111/j.1467-9922.2009.00559.x

Vannes, G. (1952). *Votabulaire de base du Neerlandais*. Antwerp, Belgium: De Sikkel.

Victori, M., & Lockhart, W. (1995). Enhancing metacognition in self-directed language learning. *System, 23*(2), 223–234.

Voss, B. (1984). *Slips of the ear*. Gunter Narr Verlag, T€ubingen.

Vandergrift, L., Goh, C. C. M., Mareschal, C. J., & Tafaghodtari, M. H. (2006). The metacognitive awareness listening questionnaire: Development and validation. *Language Learning, 56*(3), 431–462.

Verhelst, N. D., Verstralen, H. H. F. M., & Jansen, M. G. H. (1997). A logistic model for time-limit tests. In W. J. van der Linden & R. K. Hambleton (Eds.) *Handbook of Modern Item Response Theory*. (pp. 169–185). New York: Springer-Verlag.

Verhoeven, J., De Pauw, G., & Kloots, H. (2004). Speech rate in a pluricentric language: A comparison between Dutch in Belgium and the Netherlands. *Language and Speech, 47*(3), 297–308.

Walter, C. (2004). Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. *Applied Linguistics, 25*(3), 315–339.

Watanabe, M., Hirose, K., Den, Y., & Minematsu, N. (2008). Filled pauses as cues to the complexity of upcoming phrases for native and non-native listeners. *Speech Communication, 50*, 81–94.

Weber, A., & Cutler, A. (2006). First-language phonotactics in second-language listening. *The Journal of the Acoustical Society of America, 119*, 597–607.

Weil, S. A. (2003). *Foreign accented speech: Adaptation and generalization*. Unpublished master's thesis, The Ohio State University.

Wenden, A. L. (1991). *Learner strategies for learner autonomy.* Englewood Cliffs, NJ: Prentice-Hall.

Wesche, M., Paribakht, T., & Ready, D. (1996). A comparative study of four ESL placement instruments. In *Performance testing, cognition and assessment: Selected papers from the 15th Language Testing Research Colloquium* (pp. 199–210). Cambridge, UK: Cambridge University Press.

Wheeless, L. (1975). An investigation of receiver apprehension and social context dimensions of communication apprehension. *The Speech Teacher, 24*, 261–268.

Wilcox, G. K. (1978). The Effect of Accent on Listening Comprehension-A Singapore Study. *English Language Teaching Journal, 32*(2), 118–127

Winne, P. (1995). Inherent details in self-regulated learning. *Educational Psychologist, 30,* 173–187.

Wingfield, A. (2000). Speech perception and the comprehension of spoken language in adult aging. In D. Park & N. Schwarz (Eds.) *Cognitive Aging: A Primer*. (pp. 175–195). Philadelphia, PA: Psychology Press.

Yamada, R., & Tokura, Y. (1992). Perception of American English /r/ and /l/ by native speakers of Japanese. In E. Tohkura, E. Vatikiotis-Bateson, & Y. Sagisaka (Eds.), *Speech perception, production, and linguistic structure* (pp. 155–74). Tokyo: Ohmsha.

Yamamoto, K., & Everson, H. (1997). Modeling the effects of test length and test time on parameter estimation using the HYBRID model. In J. Rost & R. Langeheine (Eds.), *Applications of latent trait and latent class models in the social sciences*. New York: Waxmann. Retrieved January 4, 2009, from http://www.ipn.uni-kiel.de/aktuell/buecher/rostbuch/c07.pdf.

Yanagawa, K., & Green, A. (2008). To show or not to show: The effects of item stems and answer options on performance on a multiple-choice listening comprehension test. *System, 36*, 107–122.

Yang, N. D. (1999). The relationship between EFL learners' beliefs and learning strategy use. *System, 27*, 515–535.

Yang, X. Q. (2007). The influence of discourse organizational patterns on Chinese EFL learners' listening comprehension. *US-China Foreign Language, 5*(3), 22–31.

Yano,Y., Long, M. H., & Ross, S. (1994). The effects of simplified and elaborated texts on foreign language reading comprehension. *Language Learning, 44*, 189–219.

Ying-hui, H. (2006). An investigation into the task features affecting EFL listening comprehension test performance. *The Asian EFL Journal Quarterly, 8*(2), 33–54.

Yuan, J., Liberman, M., & Cieri, C. (2006). Towards an integrated understanding of speaking rate in conversation. *Proceedings of Interspeech 2006*, 541–544.

Zareva, A. (2005). What is new in the new TOEFL-iBT 2006 test format? *Electronic Journal of Foreign Language Teaching, 2*(2), 45–57.

Zimmerman, B. J., & Schunk, D. H. (2001). *Self-regulated learning and academic achievement.* Mahwah, NJ: Erlbaum.

Zhao, Y. (1997). The effects of listeners' control of speech rate on second language comprehension. *Applied Linguistics, 18*(1), 49–68.

Corresponding Author and Reprints:
Sarah C. Wayland, PhD, University of Maryland Center for Advanced Study of Language, (301) 226-8938, swayland@casl.umd.edu, www.casl.umd.edu.

Funding/Support: This material is based upon work supported, in whole or in part, with funding from the United States Government. Any opinions, findings and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the University of Maryland, College Park and/or any agency or entity of the United States Government. Nothing in this report is intended to be and shall not be treated or construed as an endorsement or recommendation by the University of Maryland, United States Government, or the authors of the product, process, or service that is the subject of this report. No one may use any information contained or based on this report in advertisements or promotional materials related to any company product, process, or service or in support of other commercial purposes. The Contracting Officer's Representative for this project is Shannon Salyer, PhD, Senior Research Scientist - Research and Analysis Division, Defense Language Institute Foreign Language Center, (831) 242-6670, shannon.salyer@conus.army.mil.

Acknowledgment: We wish to thank the following people from the DLIFLC: Donald C. Fischer, PhD, Provost; John A. Lett Jr., PhD, Dean - Research and Analysis; Shannon Salyer, PhD, Senior Research Scientist - Research and Analysis; Mika Hoffman, PhD, Dean - Test Development; James Dirgin, Director - Test Review and Education; Ruth Mehr, Test Developer; Gerd Brendel - Test Development and Education Specialist; Nhon Le, Senior Project Manager - Test Development; and William Eilfort, Project Manager - Test Development.

We also wish to thank Barbara Forsyth, PhD, Area Director for Performance & Assessment at CASL, for her calm guidance, as well as her careful and thoughtful review of this document. Catherine J. Doughty, PhD, Area Director for Second Language Acquisition at CASL, was also invaluable, providing information about issues pertaining to second language acquisition in addition to reviewing this document. Last but not least, we are grateful to Michael Bunting, PhD, Associate Research Scientist at CASL, for his insightful comments on an earlier version of this review.