영어어문교육, 15권 4호 2009년 겨울

An Evaluation of Listening Studies concerning Discourse Signaling Cues: Focus on Research Designs

Euen Hyuk Jung
(Yonsei University)

Jung, Euen Hyuk (2009). An evaluation of listening studies concerning discourse signaling cues: Focus on research designs. *English Language & Literature Teaching, 15*(4), 55-74.

Although a considerable amount of research on discourse signaling cues has been conducted in reading, little attention has been paid to such cues in the area of listening Moreover, despite the solid evidence showing that cues have beneficial effects for reading comprehension, L2 listening research has produced mixed findings about the role of cues Such discrepancies among these findings might be due in part to inadequate research methodologies as well as the idiosyncratic features of their experimental designs However, no study, to date, has thoroughly examined the research designs of listening comprehension studies on cues Consequently, this study critically evaluates the present state of research designs and reporting practices of studies investigating the role of cues in listening comprehension The present study aims to provide insights into areas that require empirical attention and systematic investigation It also seeks to encourage improved and refined research practices for future studies This paper is organized as follows It will first critically review the empirical findings regarding cues in the area of L1 listening comprehension Second, it will present a critical evaluation of L2 listening studies on cues Finally, it will address the major research design issues of currently available listening studies and provide suggestions for improvement in future research

[research design/critical evaluation/listening/discourse signaling cues]

I. INTRODUCTION

Text comprehension has been shown to be an interactive process, which is affected by both comprehender (reader/listener) variables and text variables (Brindley, 1998, Carrier, 1999, Dunkel, 1991, Im, 2000, Lynch, 1998, Mendelson, 1998, Morley, 1995, Park, 2008, Rubin, 1994, Thompson, 1993, Weir, 1993) Much research has been devoted to

investigating comprehender variables, such as background knowledge, language proficiency, motivation, memory, and anxiety Some studies have examined text variables, such as text type, text structure, text readability/listenability, and discourse signaling cues Of interest here are discourse signaling cues that may affect the reader's/listener's comprehension of the text Discourse signaling cues (from here on, for simplicity's sake, *cues*) are metadiscursive markers that indicate the structure of a text and explicitly show the relations of ideas within it These cues include previews (e g , "Today, we will discuss air pollution"), exemplifiers (e g , "For instance"), summarizers (e g , "To sum up so far"), and logical connectives (e g , "first," "finally, and "or"), among others

A considerable amount of research on cues has been conducted in the area of reading (Chung, 2000, Kintsch & Yarbrough, 1982, Lorch, 1989, Lorch & Lorch, 1995, 1996, Lorch, Lorch, & Inman, 1993, Meyer, 1987, Spyridakis & Standal, 1987) However, little attention has been paid to such cues in the area of listening Moreover, despite the solid evidence found for the beneficial effects of cues on reading comprehension, L2 listening research has produced mixed findings about the effect of cues on understanding spoken text (Chaudron & Richards, 1986, Dunkel & Davis, 1994, Flowerdew & Tauroza, 1995, Jung 2003a, 2003b) Such discrepancies among these findings might be due in part to inadequate research methodologies or idiosyncratic features of the experimental designs used However, no study, to date, has thoroughly examined the research designs of listening studies on cues Consequently, this study intends to closely examine the research designs of existing research that has investigated the effects of cues on listening comprehension Achieving this objective may provide insights into the current state of knowledge about cues and help future research employ sound and refined research methodology

The organization of this paper is as follows The first section critically reviews the empirical findings on cues in the area of L1 listening comprehension The second section presents a critical evaluation of L2 listening research on cues The L1/L2 listening studies reviewed in this paper were selected on the basis of the following criteria (a) the study was either experimental or quasi-experimental in research design with cues as one of the independent variables The research domain was confined to experimental or quasi-experimental studies due to the main research question about the effectiveness or usefulness of cues in understanding a spoken text, (b) the dependent variable of the study was a measure of language behavior related to the effect of cues, (c) the study employed a statistical analysis to examine the effects of cues on listening comprehension, (d) the study was available in its published form at the time of current synthesis, (e) the study was published in a referred international journal [indexed in social science citation index (SSCI)] or as a book chapter published by an international book publisher between the years of 1980-2009 (see Appendix for the list of publication sources), and (f) the study was

published in English In the third section, major research design issues of the currently available research are presented, and suggestions for improvement in future research are outlined The final section offers conclusions

II. EMPIRICAL RESEARCH ON CUES IN L1 LISTENING

As mentioned earlier, despite considerable attention that cues have received in the area of L1 and L2 reading, their role in the domain of L1 listening has not been actively researched Hron, Kurbjuhn, Mandl, and Scnotz (1985) compared the effects of cues (e g , previews, summarizers, pointer words, and logical connectives) on spoken text comprehension with those on written text comprehension In Hron et al 's study, 75 native speakers of German were randomly assigned to one of four groups of signaling variation (a) a signaled written text, (b) a non-signaled written text, (c) a signaled spoken text, and (d) a non-signaled spoken text Comprehension measures included a written recall protocol and nine open-ended comprehension questions The facilitating effects of cues were found in both written and spoken modalities, although these cues appeared to play a more important role in reading than in listening comprehension Compared to those receiving the non-signaled versions, both readers and listeners receiving the signaled versions recalled significantly more main ideas and performed better on open-ended questions immediately after the exposure However, on tests administered one week later, only the readers' performance scores continued to reflect the facilitating effects of cues on text comprehension

The results of Hron et al 's (1985) study were rather unexpected, especially in light of the additional cognitive demand potentially placed on listeners during the act of listening Whereas listeners need to build a coherent mental representation of the text from the continuous flow of information, readers have more control over what they read simply by having the text at hand and by having the opportunity to review it if necessary Their findings should be interpreted with caution, however, because of a problem with the material used for the spoken modality In order to test the effects of cues on listening comprehension, the passage originally designed for reading was simply read aloud That is, the material was not a naturally-occurring spoken text The experimental task was merely a reading of a written text, rather than authentic spoken discourse

Another problem with this study is its method for measuring text comprehension First, the study failed to control for test bias The same nine open-ended questions were used for both the immediate and delayed testing sessions Hence, there may have been a practice effect In order to control for test bias, two different versions of comparable open-ended questions (versions A and B) should have been prepared, the order of these versions should

have been counterbalanced in the experimental procedures Second, no interrater reliability was reported in the recall scoring Finally, a time limit was imposed for the completion of the recall task Participants should have been provided with ample time to complete their tasks to allow them to demonstrate their skills as readers and listeners, respectively, and to allow researchers to clearly distinguish the reading and listening abilities (i e , comprehension abilities) from the writing abilities (i e , production abilities) involved in the comprehension tasks

Another L1 listening study on cues (e g , previews, summarizers, pointer words, and evaluators) was conducted by Rickards, Fajen, Sullivan, and Gillespie (1997) In their study, forty native English-speaking college students participated The effects of cues on L1 listening comprehension was measured in terms of both the quantity and quality of students' notes and recall protocols The non-signaled text version was developed by deleting cues from a journal article (a signaled text) and by rewriting parts of the text to keep the meaning consistent with its signaled counterpart Rickards et al tested two hypotheses (a) The presence of signaling devices which explicitly cue listeners to the important ideas and their relationship would provide a solid basis for improving the quantity and quality of both listeners' note-taking and their recall, and (b) the absence of signaling devices would make it difficult for listeners to identify important ideas and thus require them to make inferences about the appropriate logical relationships between ideas, this effort to infer would then strain their cognitive resources, possibly resulting in poor quantity and quality of note-taking and recall Their findings supported both hypotheses The presence of signaling devices in the text led to a significant increase in the amount of notes taken overall and notes taken on the main ideas Furthermore, cues enabled listeners to understand the text in terms of both quantity and quality In other words, compared to those receiving the non-signaled text, those receiving the signaled text comprehended not only significantly more information, but also more important information from the text

As with Hron et al 's (1985) study, Rickards et al 's (1997) findings should be interpreted with caution It seems that they failed to examine how cues actually operate in L1 spoken discourse comprehension In their study, participants listened to an oral reading of a written passage excerpted from a magazine instead of an authentic spoken discourse In addition, these researchers inserted cues in such a way that the signaled texts were syntactically different from the non-signaled texts, which created an uncontrolled variable The passage was also relatively short (400 to 600 words long) Furthermore, the study did not provide a control for the participants' background knowledge of the text topic

III. EMPIRICAL RESEARCH ON CUES IN L2 LISTENING

1 Critical Evaluation of L2 Listening Studies on Cues

1) Chaudron and Richards's Research (1986)

Chaudron and Richards's (1986) study was the first empirical investigation of the effects of cues on L2 listening comprehension Their study differentiated macro-markers from micro-markers According to their definitions, macro-markers indicate how the major ideas are interrelated in the text, while micro-markers signal how the minor ideas are interrelated In their study, four text versions were constructed on the basis of a transcript of an American history lecture (a) a baseline lecture with no cues, (b) a lecture with macro-markers, (c) a lecture with micro-markers, and (d) a lecture with a combination of both macro- and micro-markers Each of these lecture versions was randomly assigned to English as a Second Language (ESL) learner groups at two different proficiency levels from various L1 backgrounds (a) a pre-university ESL group (lower proficiency level group, n = 71), and (b) a university ESL group (higher proficiency level group, n = 81) Participants had to stop every 60 to 90 seconds to perform a cloze listening test during the lecture, and thus did not listen to the whole lecture without interruptions Participants were not allowed to take notes during the lecture After completing the cloze listening test, participants took a multiple-choice question test and afterward took a true-false question test The scores on the multiple-choice and true-false question tests were discarded from data analysis because of their low reliability, only the cloze listening test scores were included in the data analysis Findings showed that only macro-markers had positive effects, neither micro-markers nor a combination of macro- and micro-markers had positive effects

These results might have been the product of the following research design flaws in their study First, the lecture materials used in Chaudron and Richards's (1986) study were questionable because the lecture materials did not seem to reflect the appropriate use and placement of micro-markers as they occur in natural, spoken English discourse (Flowerdew & Tauroza, 1995) Spoken micro-markers were incorporated into the written passage without retaining the naturalness of the text Special care was also taken to redundantly insert these micro-markers into places in the baseline lecture version where such insertions would add no value to comprehension Furthermore, when constructing a lecture version with both macro- and micro-markers, the micro-markers were inserted in such a way that they might have merely increased the participants' cognitive demands without providing any useful information, thus detracting from the effect of the macro-

markers alone [1] Because the inserted micro-markers in the lecture were unnatural and redundant, it should not be surprising that both micro-markers and a combination of macro- and micro-markers were not found to enhance participants' understanding of the lecture

Second, the division of cues into macro- and micro-markers according to word length in their study should be reconsidered All the macro-markers were phrases or sentences that included three to nine words In contrast, more than two-thirds of the micro-markers were single words Such classification of these markers according to word length is problematic, since the function of cues, rather than their length seems to distinguish macro- from micro-markers According to Chaudron and Richards's definition, micro-markers (e g , *now,* and *so*) should operate only at the micro-level simply to link sentences and to fill pauses However, as several researchers (DeCarrico & Nattinger, 1992, Flowerdew & Tauroza, 1995, Schiffrin, 1987, Tyler, 1992) have demonstrated, these cues do not merely link adjoining sentences, but signal major discourse boundaries and mark the main transition points in discourse For example, the cue *Now* serves as a topic shifter in the following lecture excerpt The cue indicates the major transition point and signals the introduction of a new topic (DeCarrico & Nattinger, 1992)

> I think it is much easier to discuss the abiotic components of an ecosystem first ***Now***, I'll spend quite a bit of time on the availability of water in the soil One of the weaknesses of this first edition is that there is very little discussion of the ways in which light and temperature vary (p 156)

In the above lecture excerpt, shifting away from the previous topic (*the abiotic components of an ecosystem*), the lecturer used the cue *Now* to mark a major transition point and to signal the introduction of a new topic (*the availability of water in the soil*) According to Chaudron and Richards's definition, the cue *Now* should have been classified as a macro-marker (i e , a high-order cue, marking major discourse segments)

Third, the measurement task used in Chaudron and Richards's (1986) study is questionable Instead of using integrative tasks, they employed only discrete-point tasks (i e , multiple-choice question, true or false, and cloze listening tests) to assess the impact of cues on L2 learners' listening comprehension As Riley and Lee (1996) argued, discrete-point tasks often fail to measure accurately the listener's genuine understanding of a text It

[1] This particular account as to why beneficial effects were not found for a combination of macro- and micro-markers was provided in Chaudron and Richards's (1986) study

is necessary to employ integrative tasks to make a valid assessment of listeners' comprehension, which can reflect a complex interplay between the text and the listener

Chaudron and Richards analyzed scores only on the cloze listening test For this test, participants were required to stop every 60 to 90 seconds to answer questions while listening to the lecture As a result, they may not have had the opportunity to construct a coherent or elaborate mental representation of the lecture Spoken text comprehension involves both top-down and bottom-up processes that function interactively and whose ultimate goal is to construct a coherent and complete mental representation of the text (Brindley, 1998, Dunkel, 1991, Long, 1989, Lynch, 1998, Mendelsohn, 1998, Morley, 1995) Because the lecture was frequently interrupted, participants did not have the chance to listen to the full-length lecture, if they had listened to the entire lecture without interruption, they might have been able to construct a coherent mental representation of the lecture As such, the task might have simply taxed the participants' cognitive resources as they completed the cloze listening test during the lecture In other words, participants were not allowed to engage in the normal processing of an extended spoken text

The cloze listening test might not have been a valid measurement of comprehension because the process involved in completing it did not reflect the natural listening process and thus might have failed to tap into authentic spoken text processing The experimental procedure used in their study might have failed to capture the facilitating effects of cues on authentic spoken discourse comprehension, such as helping listeners synthesize the incoming information, build a coherent mental representation of the text, and generate elaborate inferences on the basis of the given information An invalid comprehension measurement can undermine both the internal validity (i e , cause-and-effect relationship) and external validity (i e , generalizability of the findings) of the study (Oxford, 1986)

Finally, in their study, participants were not allowed to take notes during the lecture Considering the real academic listening situation in which it is customary for students to take notes during lectures, it is ecologically not valid for participants to be prevented from taking notes while listening to the lecture Moreover, from the experimental perspective, there is no reason to suppose that note-taking and reference to notes would favor the group listening to the lecture with cues over the group listening to the lecture without cues, or vice versa

2) Dunkel and Davis's (1994) study

Similar to Chaudron and Richards's (1986) findings concerning the effects of macro- and micro-markers overall, Dunkel and Davis's (1994) study failed to find positive effects for cues overall for both L1 and L2 listeners' lecture comprehension In their study, 29 native English-speaking college students and 26 ESL learners listened to a scripted lecture on the

sinking of the two ocean liners, the *Titanic* and the *Andrea Doria* Half of each group was randomly assigned to either a baseline lecture version with cues or an experimental lecture version without them [2] During the lecture, participants were asked to take notes (ESL learners were allowed to take notes in their L1) After listening to the lecture, participants were asked to write down everything they remembered from the lecture, and they were allowed to refer to the notes taken during the lecture (ESL learners were also allowed to use their L1 to write recall protocols) Participants' comprehension of the lecture was measured in terms of their notes (i e , the quantity of notes taken during the lecture) and recall protocols (i e , the quantity of information units and the total number of words used in the recall protocols) Cues were not found to have any facilitating effects in any of these three comprehension measurements [3]

Dunkel and Davis's (1994) findings that cues had no significant effects on lecture comprehension might have been due to the lack of sensitivity of the comprehension measurements used in the study Dunkel and Davis examined the effects of cues on listening comprehension only in a gross quantitative manner That is, to assess the effects of cues on listening comprehension, they used only the quantity of notes, information units, and words used in the recall protocols Qualitative analysis of the effects of cues on the learners' comprehension (i e , learners' comprehension of high- and low-level information in the lecture) might have accurately detected the role of cues in lecture comprehension

Other methodological weaknesses included problems with the lecture materials, which were scripted In preparing for the lecture material, Dunkel and Davis also failed to take into account the listener variables such as topic appropriateness, relevancy, and the extent to which listeners were interested in the lecture In other words, they did not provide any rationale for the lecture selection, they did not give any explanation about why Speech Communication majors had to listen to a scripted lecture about *Titanic* and *Andrea Doria* As several researchers (Kirschner, Wexler, & Spector-Cohen, 1992, Long & Macian, 1994) maintain, one should consider the comprehender variables in devising materials for a study to preserve the ecological validity of the experimental study

---

[2] Report of the lecture material preparations of Dunkel and Davis's (1994) research was mischaracterized in Flowerdew and Tauroza s discussion (1995) of this study Flowerdew and Tauroza claimed that similar to the material preparations in Chaudron and Richards's study (1986), Dunkel and Davis had inserted cues into the baseline lecture However, these cues were actually removed from the baseline lecture (i e , signaled lecture) in Dunkel and Davis's study

[3] There is a discrepancy in Dunkel and Davis's (1994) report in the amount of notes taken by the participants during the lecture In the abstract of the study, Dunkel and Davis claimed that participants who listened to the signaled lecture took significantly more notes than those who listened to the non-signaled lecture However, the results of the study indicated that cues had no beneficial effects on the amount of notes taken during the lecture

Moreover, the participants' prior knowledge of the text topic and L2 listening proficiency level were not controlled for, which undermines the internal and external validity of the study Dunkel and Davis also failed to provide sufficient information about the study participants The only information included in the study was that there were two groups of participants, one native English speaker group majoring in Speech Communication and the other nonnative speaker group studying ESL at a university No information was provided about the characteristics of the participants, such as their age, gender, educational background, years of L2 learning, and L2 proficiency level

3) Flowerdew and Tauroza's (1995) Study

Flowerdew and Tauroza (1995) examined the impact of authentic micro-markers (as defined by Chaudron and Richards (1986)) (e g , *and, well, now,* and *so*) on English as a Foreign Language (EFL) learners' listening comprehension Sixty-three Cantonese-speaking EFL college students in Hong Kong participated in the study In this study, two versions of an authentic video-recorded science lecture were prepared (a) a baseline lecture version with naturally-occurring micro-markers, and (b) an experimental lecture version with these micro-markers deleted Participants were randomly assigned to either of these lecture versions While viewing the lecture, participants were allowed to take notes using their L1, L2, or a mixture of both Three comprehension measures were employed to assess the effects of micro-markers on the EFL learners' lecture comprehension (a) a self-assessment test, (b) a short answer test, and (c) a summary task using their L1, L2, or a mixture of both, in which they had access to the notes they had taken during the lecture The results indicated that the presence of micro-markers facilitated L2 learners' listening comprehension of the lecture Across all the three measures, learners who viewed the lecture with micro-markers outperformed those who viewed the lecture without them Group differences were statistically significant for the summary and short answer test scores, whereas group differences in the self-assessment scores were marginally below statistical significance

Although beneficial effects of cues were found in Flowerdew and Tauroza's (1995) study, the study is not without problems Flowerdew and Tauroza failed to ensure that the participants in both control and experimental groups had similar L2 listening proficiency prior to the treatment, thereby bringing in a compound variable They also failed to provide detailed information about the study participants, no information about the participants' age, gender, and years of L2 study was provided Furthermore, participants were given only 10 minutes to complete the summary task in their study Learners should be provided sufficient time to complete the comprehension task, and no time limit should be imposed for the completion of the comprehension assessment tasks The absence of a time limit is

considered appropriate on the assumption that listening comprehension measurement tasks should require learners to fully demonstrate their listening skills and should differentiate receptive skills from production skills involved in comprehension assessment tasks

4) Jung's (2003a) Research

Jung (2003a) examined the overall effects of cues (i e , the combined effects of macro- and micro-markers) on oral text comprehension by Korean ESL learners in terms of their perceived text comprehension and recall of text information at both high and low levels (i e , main ideas vs supporting ideas) She compared two matched groups of learners according to both their L2 listening proficiency and their L2 general proficiency (a) the signaled lecture group that listened to the lecture on culture with cues, and (b) the non-signaled lecture group that listened to the lecture without them During the lecture, participants were allowed to take notes in Korean, English, or a combination of both After listening to the lecture, participants were asked to rate their perceived comprehension of the lecture on the assessment scale ranging from 0% (nothing comprehended) to 100% (everything comprehended) Participants were then asked to write down what they thought was important in the lecture using Korean, English, or a combination of both, in which they were allowed to refer to the notes they had taken during the lecture The summaries were scored according to the presence of each of the 17 high-level information units (main ideas) and 17 low-level information units (supporting ideas) in the lecture The signaled lecture group outperformed the non-signaled lecture group in recall of both high-level and low-level information units The signaled lecture group also performed better on self-assessment tests than the other group [4] Unlike the signaled lecture group, the non-signaled lecture group was confused about the high-level and low-level information units

Jung's (2003a) study is meaningful because it was the first L2 empirical listening study to demonstrate the positive effects of cues overall on learners' text comprehension However, the study suffers from several methodological weaknesses, such as the use of a scripted text, a small sample size, and a lack of a report about the interrater reliability for the summary task

5) Jung's (2003b) Research

Following up on Jung (2003a), Jung (2003b) conducted a large-scale study with an improved research design to examine the impact of cues on EFL learners' listening

[4] Statistically significant group differences were not found for the recall of low-level information units and self assessment, perhaps because of the small sample size of this study

comprehension of high- and low-level information in a naturally-occurring academic lecture The study involved 80 Korean EFL learners, whose L2 listening proficiency level and background knowledge were controlled for Of the 80 learners, half of them listened to the lecture with cues (i e , signaled lecture group), and the other half listened to the lecture without such cues (i e , non-signaled lecture group) Half of the learners in each group performed summary tasks, the other half performed recall tasks Each group performed its task in a language laboratory using headphones During the lecture, they were allowed to take notes using Korean, English, or a combination of both They were also allowed to use Korean, English, or a combination of both to complete the summary or recall task The findings showed that cues play an important role in L2 listening comprehension The signaled lecture group performed significantly better than the non-signaled lecture group in recall of both high- and low-level information In contrast to the signaled lecture group, the non-signaled lecture group tended to misinterpret the main ideas and to become confused about the relationships among topic points Furthermore, participants in the non-signaled lecture group were confused about supporting ideas

2 Overall Picture of L2 Listening Research on Cues

To create an overall picture of L2 listening studies on cues as they are represented in the currently available literature, methodological and reporting features were analyzed and compared across the existing empirical findings (see Tables 1-3) Table 1 summarizes the characteristics of participants in the L2 listening studies on cues These studies have been conducted with widely differing populations and in widely varying instructional contexts (e g, ESL vs EFL, classroom vs language laboratory)

TABLE 1
The Characteristics of Participants in L2 Listening Studies on Cues

	Chaudron & Richards (1986)	Dunkel & Davis (1994)	Flowerdew & Tauroza (1995)	Jung (2003a)	Jung (2003b)
Participants	Pre-university ESL learners & university ESL learners (N = 152)	Native English university speakers & ESL learners (N = 55)	Cantonese EFL university learners (N = 63)	Korean ESL university learners (N = 16)	Korean EFL university learners (N = 80)
L1	Various	Various	Cantonese	Korean	Korean
L2	English	English	English	English	English
Control for L2	Controlled	Uncontrolled	Uncontrolled	Controlled	Controlled

proficiency level					
Educational context	ESL (USA)	ESL (USA)	EFL (Hong Kong)	ESL (USA)	EFL (South Korea)
Age	NR	NR	NR	22-25	21-26
Gender	NR	NR	NR	F 9, M 7	F 14, M 66
Years of L2 learning (average)	NR	NR	NR	9 years	7 years & 9 months
Educational background	NR	Native English speaker speech communication majors & ESL learners NR	First years of a 3-year degree program	High intermediate & advanced ESL learners	Mostly juniors from various academic backgrounds
Research setting	Language lab & classroom	NR	Classroom	Language lab	Language lab

Note NR = Not Reported

Table 2 summarizes the characteristics of lecture materials used in the L2 listening studies on cues These studies employed a variety of text types and text topics, with differences in lecture duration, lecture source, lecture mode, speech rate and so forth

TABLE 2
The Characteristics of Lecture Materials in L2 Listening Studies on Cues

	Chaudron & Richards (1986)	Dunkel & Davis (1994)	Flowerdew & Tauroza (1995)	Jung (2003a)	Jung (2003b)
Cue types	Macro-markers vs micro-markers vs combination of macro- and micro-markers	Combination of macro- and micro-markers (cues overall)	Micro-markers	Combination of macro- and micro-markers (cues overall)	Combination of macro-and micro-markers (cues overall)
Cue manipulation	Cue insertion	Cue deletion	Cue deletion	Cue deletion	Cue deletion
Lecture source	Scripted	Scripted	Authentic	Scripted	Authentic
Lecture topic	American history	*Titanic* and *Andrea Doria*	Computer	Culture	Culture
Lecture mode	Audio	Audio	Video	Audio	Audio

Lecture duration (min)	23 vs 27 (8 min of pauses included)	7 vs 10	23 vs 26	8 vs 10	14 vs 17
Number of words	1760 vs 2064	878 vs 1263	3283 vs 3502	788 vs 1180	1765 vs 2155
Speech rate (wpm)	117 3 vs 107 1	NR	219 vs 207	NR	122 7 vs 120 6
Text type	Narrative	Narrative & comparison-and-contrast	Expository	Expository	Expository
Consideration for listener variables (topic relevancy, appropriateness, & interestingness)	Not considered	Not considered	Considered	Not considered	Considered
Control for background knowledge	Uncontrolled	Uncontrolled	Controlled	Controlled	Controlled

Note NR = Not Reported

Table 3 summarizes the features of comprehension measurement tasks employed in the L2 listening studies on cues A variety of comprehension measures have been used within the domain to test the impact of cues on L2 listening comprehension Dependent measurement tasks have varied the form of such tasks has included self-assessments, multiple-choice question tests, short answer tests, True or False (T or F) tests, cloze listening tests, and summary and recall tasks

TABLE 3
The Characteristics of Measurement Tasks in L2 Listening Studies on Cues

	Chaudron & Richards (1986)	Dunkel & Davis (1994)	Flowerdew & Tauroza (1995)	Jung (2003a)	Jung (2003b)
Discrete vs integrative measures	Discrete cloze listening test, T or F question test, multiple-choice question test	Integrative recall task (L1) & notes (L1)	Integrative & discrete self-assessment, short answer test & summary task (L1, L2, or	Integrative & discrete self-assessment & summary task (L1, L2, or mixture of	Integrative recall & summary task (L1, L2, or mixture of both)

			mixture of both)	both)	
Immediate vs delayed measures	Immediate task	Immediate task	Immediate task	Immediate task	Immediate task
Note-taking	Not allowed	Allowed (L1)	Not allowed	Allowed (L1, or L2, a mixture of both)	Allowed (L1, or L2, a mixture of both)
Statistical analysis	MANOVA	ANOVA	ANOVA	T-test	ANOVA
Interrater reliability	NA	NR	R	NR	R

Notes
NR = Not Reported, NA = Not Applicable, R Reported

IV. MAJOR RESEARCH DESIGN ISSUES AND SUGGESTIONS

This section overviews the major research design issues of existing L2 listening studies on cues and makes suggestions for improvement in future research The studies are discussed below under headings that reflect major areas of concern in this area of research

1 Use of Naturally-Occurring Materials

Most previous L2 listening studies (Chaudron & Richards, 1986, Dunkel & Davis, 1994, Jung, 2003a) used scripted written texts for lecture materials, only Flowerdew and Tauroza (1995) and Jung (2003b) used a naturally-occurring lecture A scripted written lecture and a naturally-occurring spoken lecture differ in numerous ways For instance, a scripted written text is characterized by "a higher lexical density (largely due to heavier use of nominalization) [A scripted written text] contains longer information units with more complex relations of coordination and subordination It features a wider range and choice of vocabulary, has longer average word length and features more frequent use of the passive voice" (Flowerdew & Tauroza, 1995, p 442) In contrast, a naturally-occurring spoken text is characterized by "hesitations, false starts, filled pauses, phonological contractions and assimilations [A naturally-occurring spoken text] feature[s] a higher proportion of sentence fragments, as opposed to complete sentences [It is] structured according to tone units, as opposed to clauses, and use[s] discourse [signaling cues] to mark the beginning or end of tone groups" (Flowerdew & Tauroza, 1995, p 442) It is

important to use authentic/natural spoken texts to accurately examine the role of cues in comprehending L2 oral texts

2 Consideration of Listener Variables

Flowerdew and Tauroza (1995) and Jung (2003a, 2003b), but neither Dunkel and Davis (1994) nor Chaudron and Richards (1986), provided a rationale for the particular lecture materials used for the study Because of the interactive nature of listening comprehension, which involves both text and listener variables, in preparing lecture material, it is necessary not only to employ a naturally-occurring text, but also to take into account listener variables such as lecture topic relevancy, appropriateness, and interestingness (i e , the topic should be interesting enough to hold L2 learners' attention during the lecture) It is necessary to consider these listener variables to preserve the ecological validity of the experimental research

3 Control for Extraneous Variables

Flowerdew and Tauroza (1995) and Jung (2003a, 2003b), but neither Dunkel and Davis (1994) nor Chaudron and Richards (1986), controlled for the learners' background knowledge of the text topic Research has documented a significant role that background knowledge plays in spoken text comprehension (Chiang & Dunkel, 1992, Long, 1990, Markham & Latham, 1988, Nam, 1997, Schmidt-Reinhardt, 1988) In order to avoid the potentially confounding effect of topic familiarity, it is necessary to control for the learners' background knowledge of the text when examining the of role of cues The use of texts with which participants are unfamiliar would seem to make it inevitable that listeners will depend heavily on cues instead of their background knowledge in an attempt to understand an L2 spoken text

Chaudron and Richards (1986) and Jung (2003a, 2003b), but neither Dunkel and Davis (1994) nor Flowerdew and Tauroza (1995), ensured that the learners in both experimental and control groups had similar L2 listening proficiency via an independent comprehension measurement Future studies should be careful to control for the participants' L2 listening proficiency, failing to control for this factor may jeopardize the internal and external validity of research findings Therefore, future research must ensure a similar L2 listening proficiency in the experimental and control groups so that results are valid and accurate

4 Information about Study Participants

The majority of studies (Chaudron & Richards, 1986, Dunkel & Davis, 1994, Flowerdew

& Tauroza, 1995), except for Jung (2003a, 2003b) did not report sufficient information about study participants, such as their age, gender, years of L2 learning, L2 proficiency level, and educational background Researchers should report important characteristics of participants, especially those that might potentially affect results As Norris and Ortega (2006, as cited in Lee, 2006, p 182) argued, "it is, quite simply, impossible to take full advantage of research synthesis, when the primary studies that define a domain provide impoverished accounts of their research methods and study findings" Therefore, future studies should provide sufficient and accurate information about participants, which would enable research findings to be generalized more widely and have greater explanatory power Another noticeable feature of existing L2 listening studies on cues involves the fact that all the participants in these studies were adult learners, no study addressed the effects of cues on young L2 learners' understanding of spoken text, which indicates an area for future research

5 Comprehension Measurement

Measurement tasks in the previous L2 listening studies on cues have ranged from discrete-point to integrative tasks Only Chaudron and Richards (1986) employed discrete-point tasks and relied primarily on the scores of the cloze listening test to assess the impact of cues on L2 listening comprehension As pointed out earlier in the discussion of this particular study, the reliability and validity of the cloze listening test are questionable

Researchers need to put efforts into developing and using valid and reliable comprehension measurement tasks Measurement devices affect the internal and external validity of research findings "If a particular measurement is not valid, then the internal and external validity of the study in which it is used are suspect Any lack of reliability can also detrimentally affect the internal and external validity of a piece of research In other words[,] if the instrument or set of instruments employed in a study is questionable, then the results and interpretation are also questionable" (Oxford, 1986, p 166)

To measure the effect of cues on listening comprehensibility accurately, researchers in future studies should use integrative tasks instead of discrete-point tasks, and participants should be provided with sufficient time to complete the task so that they can fully demonstrate their comprehension abilities For the integrative tasks such as recall and summary tasks, it is particularly important to report the interrater reliability All the previous studies on cues have focused on the short-term effects of cues on learners' comprehension of the text That is, they used only immediate comprehension assessment tasks Future research should incorporate the use of delayed comprehension assessment tasks to examine the long-term effect of cues on listeners' understanding of a text To do so

will enable us to gain a fuller understanding of how these cues affect memory representation and cognitive encoding processes

V. CONCLUSION

The aim of this paper was to critically examine the present state of empirical research methods and the reporting practices of studies investigating the role of cues in listening comprehension The objective of this study was to provide insights into areas that need empirical attention and systematic investigation Another goal of this study was to point out improved and refined research practices for use in future research The current study attempted to make contributions to the field of English language teaching by critically evaluating listening research on cues, which has not been done so far

The limitations of the studies on cues pointed out in this paper may pose difficulties for and/or constraints on the interpretation of their findings Readers of these studies need to keep such limitations in mind Moreover, future researchers should avoid these research design drawbacks Researchers will need to use more careful, rigorous experimental and quasi-experimental designs to enhance our knowledge of the role of cues in comprehending spoken text

REFERENCES

Brindley, G (1998) Assessing listening abilities *Annual Review of Applied Linguistics, 18,* 171-191

Carrier, K (1999) The social environment of second language learning Does status play a role in comprehension? *Modern Language Journal, 83,* 65-79

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

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

Chung, J S (2000) Signals and reading comprehension--theory and practice *System, 28,* 247-259

DeCarrico, J , & Nattinger J R (1992) *Lexical phrases and language teaching* Oxford, England Oxford University Press

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

Dunkel, P A , & Davis, J N (1994) The effects of rhetorical signaling cues on the recall

of English lecture information by speakers of English as a native and second language In J Flowerdew (Ed), *Academic listening Research perspectives* (pp 55-74) New York Cambridge University Press

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

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, The Netherlands Elsevier Science Publisher

Im, B B (2000) Effective learning tasks and activities to improve EFL listening comprehension *English Language & Literature Teaching, 6*, 1-24

Jung, E H (2003a) The effects of organization markers on ESL learners' text understanding *TESOL Quarterly, 30,* 31-58

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

Kintsch, W , & Yarbrough, J C (1982) Role of rhetorical structure in text comprehension *Journal of Educational Psychology, 74*, 828-834

Kirschner, M , Wexler, C , & Spector-Cohen, E (1992) Avoiding obstacles to student comprehension of test questions *TESOL Quarterly, 26,* 537-556

Lee, S K (2006) A synthesis of research on the effect of typographical visual input enhancement *English Teaching, 61*(3), 169-187

Long, D R (1989) Second language listening comprehension A schema-theoretic perspective *Modern Language Journal, 73,* 32-40

Long, D R (1990) What you don't know can't help you An exploratory study of background knowledge and second language listening comprehension *Studies in Second Language Acquisition, 12,* 65-80

Long, D R , & Macian, J (1994) Listening skills Acquisition and assessment In R Charles & S Brooks-Brown (Eds), *Northeast conference on the teaching of foreign languages Teaching, testing and assessment making the connection* (pp 111-138) Lincolnwood, IL National Textbook Company

Lorch, R F (1989) Text signaling devices and their effects on reading and memory processes *Educational Psychology Review, 1*, 209-234

Lorch, R F , & Lorch, E P (1995) Effects of organizational signals on text-processing strategies *Journal of Educational Psychology, 87,* 537-544

Lorch, R F , & Lorch, E P (1996) Effects of organizational signals on free recall of expository text *Journal of Educational Psychology, 88,* 38-48

Lorch, R F , Lorch, E P , & Inman, W E (1993) Effects of signaling topic structure on text recall *Journal of Educational Psychology, 85*, 281-290

Lynch, T (1998) Theoretical perspectives on listening *Annual Review of Applied*

Linguistics, 18, 3-19

Markham P, & Latham, M (1988) The influence of religion-specific background knowledge on the listening comprehension of adult second language students *Language Learning, 37,* 157-170

Mendelsohn, D (1998) Teaching listening *Annual Review of Applied Linguistics, 18,* 81-101

Meyer, B J F (1987) Following the author's top-level organization An important skill for reading comprehension In R J Tierney, P L Anderson & J N Mitchell (Eds), *Understanding readers' understanding Theory and practice* (pp 59-76) Hillsdale, NJ Lawrence Erlbaum Associates

Morley, J (1995) Academic listening comprehension instruction Models, principles, and practices In D Mendelsohn & J Rubin (Eds), *A guide for the teaching of second language listening* (pp 151-165) San Diego, CA Dominie Press

Nam, J Y (1997) The effects of topic familiarity on English listening comprehension in relation to language proficiency *Journal of the Applied Linguistics Association of Korea, 13,* 61-88

Oxford, R (1986) Research on language loss A review with implications for foreign language teaching *Modern Language Journal, 66,* 160-169

Park, M (2008) Out-of-class listening activities *English Language & Literature Teaching, 14*(1), 59-74

Rickards, J P, Fajen, B R, Sullivan, J F, & Gillespie, G (1997) Signaling, notetaking, and field independence-dependence in text comprehension and recall *Journal of Educational Psychology, 89,* 508-517

Riley, G, & Lee, J (1996) A comparison of recall and summary protocols as measures of second language reading comprehension *Language Testing, 13,* 175-187

Roller, C M (1990) The interaction of knowledge and structure variables in the processing of expository prose *Reading Research Quarterly, 25, 2,* 79-89

Rounds, P (1987) Characterizing successful classroom discourse for NNS teaching assistant training *TESOL Quarterly, 21,* 643-672

Rubin, J (1994) A review of second language listening comprehension research *Modern Language Journal, 78,* 196-204

Schiffrin, D (1987) *Discourse markers* New York Cambridge University Press

Schmidt-Rinehart, B C (1994) The effects of topic familiarity on second language listening comprehension *Modern Language Journal, 78,* 179-188

Spyridakis, J H , & Standal, T C (1987) Signals in expository prose Effects on reading comprehension *Reading Research Quarterly, 22,* 285-310

Thompson, I (1993) An investigation of the effects of texts and tasks on listening comprehension Some evidence from Russian In J E Alatis (Ed), *Georgetown*

university round table on languages and linguistics (pp 294-305) Washington, DC Georgetown University Press

Tyler, A (1992) Discourse structure and the perception of incoherence in International Teaching Assistants' spoken discourse *TESOL Quarterly, 26,* 713-729

Weir, C J (1993) *Understanding and developing language tests* London Longman

APPENDIX

List of Publication Sources

Source	Study report frequency
Applied Linguistics (L2 listening comprehension)	1
Journal of Educational Psychology (L1 listening comprehension)	1
Modern Language Journal (L2 listening comprehension)	1
Studies in Second Language Acquisition (L2 listening comprehension)	1
TESOL Quarterly (L2 listening comprehension)	1
Book chapter (L1 & L2 listening comprehension)	2
- Cambridge University Press (L2 listening comprehension)	
- Elsevier Science Publisher (L1 listening comprehension)	

Examples in: English
Applicable Languages: English
Applicable Levels: College

Euen Hyuk Jung
Department of English Language and Literature
Yonsei University
134 Sinchon-dong Seodaemun-gu Seoul, South Korea
Tel 02-2123-2326
Email junge@yonsei ac kr

Received in October, 2009
Reviewed in November, 2009
Revised version received in December, 2009