How Non-Native Speakers Perceive Listening Comprehension Problems: Implications for Adaptive Support Technologies

Xun Cao[1,2], Naomi Yamashita[2], and Toru Ishida[1]

[1] Kyoto University, Kyoto, Japan
xun@ai.soc.i.kyoto-u.ac.jp
ishida@i.kyoto-u.ac.jp
[2] NTT Communication Science Laboratories, Kyoto, Japan
naomiy@acm.org

Abstract. Previous studies have suggested many technologies to support non-native speaker comprehension in real-time communication. However, such technologies may impose an extra burden on non-native speakers (NNSs) if they do not match their current needs. To design a system that adapts to the changing needs of NNSs, we need to understand the types of problems NNSs face and how these problems are perceived by them. To explore such issues, we conducted a laboratory experiment with 40 NNSs (and 20 native speakers as a control group) who engaged in a listening task. During the task, the participants pressed a button whenever they encountered a comprehension problem. Next they explained each problem, the point at which they recognized the problem, and for how long it persisted. Our analysis identified twelve types of listening comprehension problems, which we further classified into three patterns based on their persistence and the time taken to perceive them. Our findings have implications for designing adaptive technologies to support listening comprehension of NNSs in real-time communication.

Keywords: Non-native speakers, Listening comprehension problems, Adaptive support.

1 Introduction

More and more global organizations are forming multinational teams so that people from different language backgrounds can work together to generate new ideas, solve problems, and make decisions. Even though multinational teams offer potential for gathering various creative ideas from different cultural perspectives, they also run the risk of suffering from various barriers [3]. One such barrier is caused by language [19]. To communicate and collaborate, multinational teams often adopt a common language [4]. However, a common language does not necessarily ensure effective communication [17].

Non-native speakers (NNSs) often face comprehension difficulties when listening to native speakers' (NSs') speech [2, 30]. Due to the need to process continuous streams of speech during listening, even when NNSs encounter a comprehension

problem, they cannot dedicate the time and resources to resolve such problems [25, 22]. As a result, NNSs often miss parts of the speech, and cannot comprehend the full meaning of what was said.

Previous works have proposed many technologies for supporting NNS comprehension in real-time communication, such as providing a speech translation system that translates NSs' spoken language into NNSs' language [29, 18], providing NNSs with real-time automatic speech recognition (ASR) transcripts as supplemental information for their comprehension [21, 32, 6], and adding artificial delays between NSs to provide more processing time for NNSs [31]. Some researchers even suggested providing NNSs with multiple supports, for example, automated transcripts and bilingual dictionaries, so that they can choose whichever support they wish to use that matches their needs [7]. However, providing multiple supports to a user and allowing him/her to make a choice is not necessarily the best solution – while it allows a user to deal with various kinds of problems, it often imposes extra burden to the user. Particularly in the case of a non-native user, choosing a support when encountering a listening comprehension problem could be burdensome because he/she is already overwhelmed by processing large amounts of speech information in a limited time (e.g., when listening to a lecture or in a meeting with many NSs) [25].

Our goal is to design an adaptive system, which automatically changes the type of support based on the NNSs' changing needs so that it does not impose additional burden on them. According to previous studies [5, 12], to design such a problem adaptive system, we first need to understand the types of real-time comprehension problems faced by NNSs and how these problems are perceived by them. In other words, (1) what types of listening comprehension problems emerge when NNSs are listening to native speech? (2) when do NNSs notice each problem and how long do such problems persist? In this paper, we particularly focus on NNSs' listening comprehension problems that occur during their cognitive processing of speech input. We decided to focus on the cognitive aspect of their listening comprehension problems because the accumulation of these problems leads to cognitive overload, which is the most common and fundamental problem faced by any NNS [9, 1, 2].

To answer the research questions stated above, we conducted a laboratory experiment with 40 NNSs (and 20 NSs as a control group) who engaged in a listening task followed by in-depth interviews. During the task, the participants pressed a button whenever they encountered anything about which they were unclear or did not understand: comprehension problems. In the interviews, they explained what kind of problems they faced, at what point during the listening task they recognized the problems, and for how long these problems persisted. Through an exploratory analysis of the interview data, we identified twelve types of listening comprehension problems, which we further classified into three patterns based on their persistence and the time taken to perceive them.

In the remainder of this paper, we first review previous studies and discuss how our study extends them. We then describe our study that identified the comprehension problems faced by NNSs during a listening comprehension task. We conclude with a discussion of the implications of our findings for supporting/facilitating NNS comprehension during real-time listening.

2 Background

In this section, we first review technologies that support NNS comprehension in real-time communication; we then introduce previous works that examined the listening comprehension problems of NNSs.

2.1 Technologies Supporting NNS Comprehension in Real-time Communication

Compared to asynchronous communication, NNSs face more difficulties in real-time communication. Although in asynchronous communication NNSs have more time to resolve their problems by accessing various language resources or services and considering the context [13, 22], in real-time communication they often cannot dedicate enough time and resources to resolving their problems because they are overwhelmed by processing continuous streams of speech [25].

Previous studies, which suggested technologies to support NNS comprehension in real-time communication, mainly concentrated on speech-to-speech translation and automatic speech recognition (ASR). The most direct way to support NNS comprehension is providing a speech translation system, which transcribes NSs' speech to text, translates the text into the NNSs' language, and outputs speech synthesized from the translated text [29, 18]. However, such technology remains far from satisfactory, and the combination of recognition and translation errors often disrupts comprehension.

Another widely investigated line of support uses ASR technologies. Pan et al. showed that real-time transcripts generated by ASR technologies can improve NNS comprehension when their accuracy and delay fall within a reasonable range [21]. While Pan et al. investigated the benefits of showing ASR transcripts to NNSs in a non-interactive setting (i.e., using pre-recorded speech), Gao et al. moved a step further and showed the benefits of providing ASR transcripts in an interactive setting (i.e., real-time multiparty communication) [6]. However, despite the positive effects of introducing ASR transcripts, research has also reported that NNSs are often overwhelmed when they simultaneously listen to speeches and read transcripts with errors and delays [6, 32].

Yamashita et al. provided a different perspective for supporting NNSs in real-time communication. They investigated the benefits of providing NNSs with additional processing time by adding artificial delays in NSs' speech. They found that short silent gaps produced by such delays improved the comprehension of NNSs, but more attention and effort were required to follow the speech [31].

Overall, the proposed technologies do seem to help NNSs improve their listening comprehension. However, most had some negative effects, such as placing an additional cognitive load on NNSs. We suspect that the cognitive load could be lightened if NNSs were provided with appropriate support at more propitious timing. Indeed, researchers found that NNSs themselves developed their own strategies for effectively utilizing ASR transcripts; some reviewed the transcripts only when they were not sure if they had heard a word/phrase correctly or when they had missed some words. In

most parts, they ignored the transcripts because they found it difficult to simultaneously focus on two modalities (audio and ASR transcripts) [11, 6]. This strategy implies that ASR transcripts could be useful for resolving some types of problems, but they may only impose more burdens during other parts of listening.

2.2 Listening Comprehension Problems of NNSs

To design a problem adaptive support for NNSs, we need a better understanding of the types of real-time listening comprehension problems and how NNSs perceive them.

In the second language learning field, much listening comprehension research has examined listeners' difficulties/problems while they are listening to a non-native language. Rubin conducted an extensive review of second language listening comprehension research and attributed the factors that affect listening comprehension to five characteristics: text characteristics (e.g., speech rate), interlocutor characteristics, task characteristics (e.g., task type), listener characteristics (e.g., language proficiency level, memory), and process characteristics (e.g., listening strategies) [24]. Goh offered a cognitive perspective on understanding NNSs' listening comprehension problems [9]. She used the weekly diaries of 40 students as her main data source and identified ten listening comprehension problems (Table 1).

Table 1. Listening comprehension problems identified in Goh's work [9]

	Problems
1.	Do not recognize words they know
2.	Unable to form a mental representation from words heard
3.	Cannot chunk streams of speech
4.	Neglect the next part when thinking about meaning
5.	Do not understand subsequent parts of input because of earlier problems
6.	Concentrate too hard or unable to concentrate
7.	Understand words but not the intended message
8.	Confused about the key ideas in the message
9.	Miss the beginning of texts
10.	Quickly forget what is heard

Overall, these research studies aim for improving second language learning. The findings are used for designing effective training programs or materials to improve NNSs' listening skills [28, 30, 9]. Even though these findings are also useful for our research, we still need to extend them so that they provide more detailed understanding of how each problem is perceived by NNSs (e.g., when each problem is perceived and how long it persists). We believe such detailed understanding of each comprehension problem will provide insight for designing adaptive technologies to support NNSs in real-time listening comprehension.

To gain a detailed understanding of each comprehension problem, we decided to take an approach/method that is different from previous studies. While researchers

have chosen such methods as diaries [8, 9], interviews [8, 9], questionnaires [16], and think-aloud [10] to reveal the comprehension problems faced by NNSs, they may not be suitable for our case for the following reasons: since diaries, interviews, and questionnaires are based on retrospection, we are skeptical whether they can capture the detailed process of each comprehension problem (e.g., the timing when that problem is perceived by NNSs). Furthermore, transient problems, which were tentatively confusing while listening, might not be remembered at the time of retrospection if the problem was eventually resolved. As for the think-aloud approach, even though this approach might provide a deeper understanding about some comprehension problems, participants' listening experiences during the think-aloud process could be completely different from regular listening, since the think-aloud approach requires participants to explain what they were thinking while they were listening. In our study, we use a method that allows us to record the comprehension problems faced by NNSs in real time, while keeping the listening experience as close to regular listening as possible. Using the method, we uncover the types of comprehension problems faced by NNSs and how these problems are perceived during real-time listening.

3 Current Study

In this paper, we set two research questions. First, we investigate the types of comprehension problems NNSs encounter in real-time listening. Our work builds on Goh's work, which has also focused on the cognitive aspects of NNSs' listening comprehension problems. Note that our study covers transient problems, which tentatively confused the NNSs but were eventually resolved or quickly forgotten.

RQ1 (types of real-time listening comprehension problems): What types of listening comprehension problems are identified in real-time listening?

In addition, we reveal the process of how NNSs perceive each listening comprehension problem and are burdened by them. Specifically, we posed the following question:

RQ2 (persistence and identification time of each problem): When do NNSs notice each listening comprehension problem and how long do such problems persist? Do the patterns of persistence and identification time differ among different types of problems?

This information is important when designing a problem adaptive system because providing support with inappropriate/inaccurate timing might impose an extra burden on NNSs. For example, a previous study showed that delayed transcription reduced the benefits and increased the listening effort [34].

4 Method

4.1 Overview

To explore the above research questions, we used a method that allows us (i) to record NNSs' comprehension problems in real time and (ii) to scrutinize each problem by allowing the NNSs to explain each one (iii) while keeping the listening experience close to regular listening experiences.

We developed a software tool that logs participants' listening comprehension problems in real time. During the listening task, participants pressed a button to indicate when they heard confusing language or they did not understand something: comprehension problems. Pressing the button marked specific places in the lecture transcripts, which were visited later to explain the details of the problems. We chose this "pressing a button" method because it has low-overhead, as suggested by previous work [15]. In addition, this method guarantees that we can record the problems NNSs faced in real time and simultaneously keep the task close to the actual listening experience.

4.2 Participants

We recruited 40 non-native English speakers (22 females, 18 males) as participants. 20 were native Japanese speakers and 20 were native Chinese speakers. Their mean age was 30.4 (SD = 9.97). Their English proficiency varied from intermediate to advanced, indicated by their Test of English for International Communication (TOEIC) scores, which ranged from 650 to 960 (M = 828, SD = 95.18). They did not identify themselves as fully proficient (M = 4.36, SD = 0.86, on a 7-point Likert scale; 1 = not proficient at all, 7 = very proficient). Their average overseas experience in English speaking countries was 0.3 years (SD = 0.54).

As a control group, we also recruited 20 native English speakers (13 males, 7 females) as participants whose mean age was 37.9 (SD = 11.98). Among these NSs, 14 were from the United States, three from Canada, two from New Zealand, and one from the United Kingdom.

4.3 Materials

Five audio clips from the Test of English as a Foreign Language (TOEFL) test were chosen as task materials. Two clips were conversations, and the other three were lectures. The length of the clips varied from two to five minutes. We chose such task materials to maintain consistency with Goh's setting, whose materials were collected from a second language listening course. The tasks were randomly assigned to each participant.

4.4 Procedure

Step 1 (real-time listening). The participants listened to the audio clip and pressed a button whenever they encountered anything about which they were unclear (i.e., comprehension problems). When the participants pressed a button, the software logged a timestamp.

Step 2 (retrospective listening). The participants listened to the same audio clip again. While listening, the computer automatically stopped at the place where they pressed the button during Step 1, using the timestamps logged by the software. At this point, the participants briefly explained what kind of problem they faced, at what point they recognized the problem, and for how long it persisted. This step helped participants re-experience the first step and recall their comprehension problems.

Step 3 (interviews). The participants were handed complete transcripts of the audio clip with markings that indicated their listening comprehension problems. Based on the marked-up transcripts, they further explained the problems they faced during the listening task. This step was designed to get more detailed information about the comprehension problems mentioned in Step 2. Interviews were conducted in each participant's native language.

4.5 Data Analysis

To identify each type of listening comprehension problem faced by the participants during the listening task, we first transcribed the interview data and removed any problems that were not directly related to their cognitive processing of speech input (e.g., lack of vocabulary). Then we classified the problems into ten categories based on Goh's work. We created a new problem category if a problem did not belong to any of the ten categories. All the interview data were coded independently by two coders, and discrepancies were discussed until an agreement was reached.

5 Findings

The results are presented as follows. First, we report all the types of listening comprehension problems that were identified in our experiment. We separately present the problems faced by non-native and native participants. Then we describe in further detail the two types of listening comprehension problems that were newly discovered in our study. Finally, we group the listening comprehension problems into three patterns based on the persistence and identification time of each problem.

5.1 Types of Listening Comprehension Problems

NNSs. RQ1 asked what types of listening comprehension problems emerged in real-time listening. To identify all the listening comprehension problems faced by non-native participants, we counted the number of times problems occurred based on the markups (times they pressed the button). In a few cases when participants described

two problems for one markup, the occurrences of problems were counted as two. 513 problem occurrences were initially identified by the non-native participants. Among them, 366 problem occurrences were "cognitive problems," 144 were due to "language skills" (e.g., lack of vocabulary), and the rest were due to "situational factors" (e.g., not being able to distinguish different speakers). The average number of problem occurrences identified by each non-native participant was 2.2 times per minute.

Tables 2a and 2b provide an overview of all the problems faced by non-native participants. Table 2a shows the real-time listening comprehension problems shared by Goh's work, and Table 2b shows two newly identified problems: "confused about unexpected word appearance" and "unsure about the meaning of words." The tables show the sample excerpts extracted from our interviews and the percentage of the occurrences of each problem (i.e., number of times each problem occurred/total number of occurrences).

Table 2a. Example and percentage of each listening comprehension problem faced by non-native participants: problems shared by Goh's work

Problem		**Example interview excerpt**	**Percentage**
1.	Do not recognize words they know	Since I misheard "slides" as "flive," I couldn't understand it. If I had read it, I would've understood it. (NNS 2)	27%
2.	Unable to form a mental representation from words heard	I didn't really understand "bubble gas." Although I caught both words, I couldn't form a picture of them. (NNS 6)	20%
3.	Cannot chunk streams of speech	I couldn't catch "cause you loved them too much." I couldn't divide that chunk into separate words. (NNS 9)	15%
4.	Neglect the next part when thinking about meaning	While I was wondering what "bubble gas" meant, I missed the subsequent words. They just drifted away, so I gave up. (NNS 10)	5%
5.	Do not understand subsequent parts of input because of earlier problems	I couldn't understand this part: "scientist decided that the best place to see a whole root system would be to grow it, where." Maybe the lecturer is asking a question, but since I couldn't get that part, I also couldn't understand the answer to it. (NNS 13)	5%
6.	Concentrate too hard or unable to concentrate	I couldn't concentrate. I was almost panicking. (NNS 19)	5%
7.	Understand words but not the intended message	I could understand the meaning. But I couldn't understand why he repeated the words. It seems that I didn't get the point.... (NNS 32)	4%
8.	Confused about the key ideas in the message	Until now, the lecturer has been talking about "growing stuff in water," "bubble gas through water," and "growing plants in soil." Now, she's saying that giving too much water will kill a plant… I don't understand. What on earth did	4%

		they want to say? (NNS 19)	
9.	Miss the beginning of texts	I wasn't quite ready and missed the beginning of the lecture. (NNS 10)	2%
10.	Quickly forget what is heard	When I heard "bubble gas," I thought I understood. But when the lecturer continued to the next sentence, I suddenly forgot what it was. I got confused whether it was gas or gassed water. (NNS 16)	1%

Table 2b. Example and percentage of newly identified listening comprehension problems faced by non-native participants: transient problems

Problem		Example interview excerpt	Percentage
11.	Confused about unexpected word appearance	"Commercially" came out of the blue. I got confused when I heard it because I thought they were talking about stuff happening in a lab. (NNS 1)	7%
12.	Unsure about the meaning of words	When I heard "root system," I wasn't sure what it meant. I came up with many possibilities. It could be the roots of plants, but when combined with "system," I got confused. I thought it might have something to do with a Linux file system or something related to a chart in linguistics. (NNS 10)	5%

NSs. We did the same count and categorization for the listening comprehension problems faced by native participants. Only twelve problem occurrences were identified. Among them, ten were "cognitive problems," and the other two were due to "situational factors." The average number of problem occurrences identified by each native participant was 0.27 times per minute.

Out of 20 native participants, eleven reported that the listening material was quite clear to them and they did not encounter any comprehension problem. Nine participants reported confusion, but most solved their confusion quickly and fairly easily during the listening tasks. For example, one participant mentioned:

The first time the lecturer said, "bubble water," I was like "huh?" But then she explained it (self-corrected it), I was like "ah." (NS7)

Difference between NNSs and NSs. Overall, although the NSs did encounter slight minor and infrequent listening comprehension problems, they resolved them fairly easily.

In contrast, the NNSs in our study faced many problems during the listening tasks. From the interviews with them, a snowball effect of listening comprehension problems seemed to occur during their listening, meaning that one problem triggered another problem. For example, one non-native participant reported that due to his uncertainty about the correct meaning of "root system," he couldn't understand the subsequent parts of the lecture well. Others reported that, when thinking about the meaning

of particular words, they missed subsequent speech. Some also mentioned that failing to catch some parts of the speech created concentration lapses. Such snowball effects of listening comprehension problems were only found in the NNS listening.

5.2 Transient Problems Identified by NNSs

As shown in Table 2b, since 12% of the problems did not fit into Goh's categorization, we created two new categories, each of which we describe in further details below.

Confused about unexpected word appearance. Previous research has indicated that people generally use information from a prior discourse to rapidly predict specific upcoming words as the discourse unfolds [20, 23]. However, a failed prediction hinders the processing of an unexpected word or phrase [27]. While NSs can quickly resolve problems and catch up with the current speech, NNSs tend to have difficulties recovering from such problems [25].

In our experiment, participants reported that they got confused about the appearance of a word or phrase that seemed unrelated to the current context. For example, in one lecture, the lecturer introduced an experiment of "growing plants in water to observe the root systems" but then slipped into a tangent about "how hydroponics is popular commercially." However, many non-native participants had difficulty understanding the connection between the tangent and the main topic. Some non-native participants were confused by the term "commercially." One participant explained:

I know the word "commercially," but I couldn't understand why it appeared in this context. I wondered if it had another meaning related to plant systems (NNS 15).

The non-native participant lost confidence in his ability to understand the context when he heard the word "commercially." Although this participant regained his confidence (i.e., he could follow the speech again) when the tangent was over, such problem was problematic because it confused him and sapped his confidence.

Unsure about the meaning of words. Some participants in our study got confused about the correct meaning of words/phrases that carried multiple meanings. Especially when such words/phrases were keywords that appeared repeatedly in the speech, the problem bothered them until they determined the correct meaning. Most participants in our study gradually solved their doubts using context information. When the words/phrases that confused the participants appeared only once, they tended to be easily forgotten.

Participants also reported confusion when they encountered homonyms. For example, in one listening task, the lecturer mainly discussed how big root systems of plants can be. "Root system," as one of the keywords, appeared several times during the lecture. However, the keyword "root system" confused some of the non-native participants:

At first, I couldn't tell whether this "root" meant "the root of plants" or "the route" of something. I finally realized that it meant "the root of plants" somewhere about here when I heard "the best place to see" (NNS 3).

Although they had a guess or multiple candidates in mind, they were not sure if their guess was correct, or which candidate was correct. As a result, they had to think hard to resolve the problem by listening to subsequent speech, which burdened them and sometimes triggered other problems.

5.3 Persistence and Identification Time of Problems

RQ2 asked the following two questions: (a) When do NNSs notice each listening comprehension problem and how long do such problems persist? (b) Do the patterns of persistence and identification time differ among different types of problems?

To answer these questions, we calculated the duration of each problem (T(dur)) and the response time taken to press the button (T(res)) by counting the number of words spoken in each time period. In Fig. 1, for example, T(dur) is ten words and T(res) is three words.

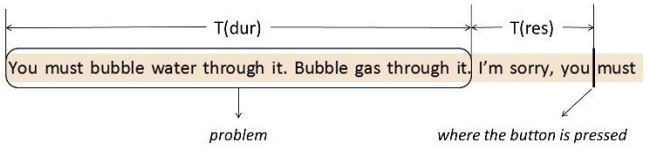

Fig. 1. Measuring "duration of each problem" (T(dur)) and "response time taken to press button" (T(res))

Fig. 2 shows how T(dur) and T(res) differed among various types of problems. Each dot represents the average T(dur) and T(res) values of each problem. To determine whether the problems can be divided into different patterns, we carried out single-linkage hierarchical clustering [14]. Based on the optimal grouping of the problems, results showed that the problems can be classified into three clusters: "immediate listening comprehension problems" (74% of all problem occurrences), "extant listening comprehension problems" (25%), and "delayed listening comprehension problems" (1%).

Pattern 1: "immediate" listening comprehension problems. The first pattern represented listening comprehension problems with short T(dur) and short T(res) values. In other words, the duration of these problems was short, and participants perceived them relatively quickly. Five types of problems fell into this pattern: "do not recognize words they know," "confused about unexpected word appearance," "unsure about the meaning of words," "cannot chunk streams of speech," and "unable to form a mental representation from words heard."

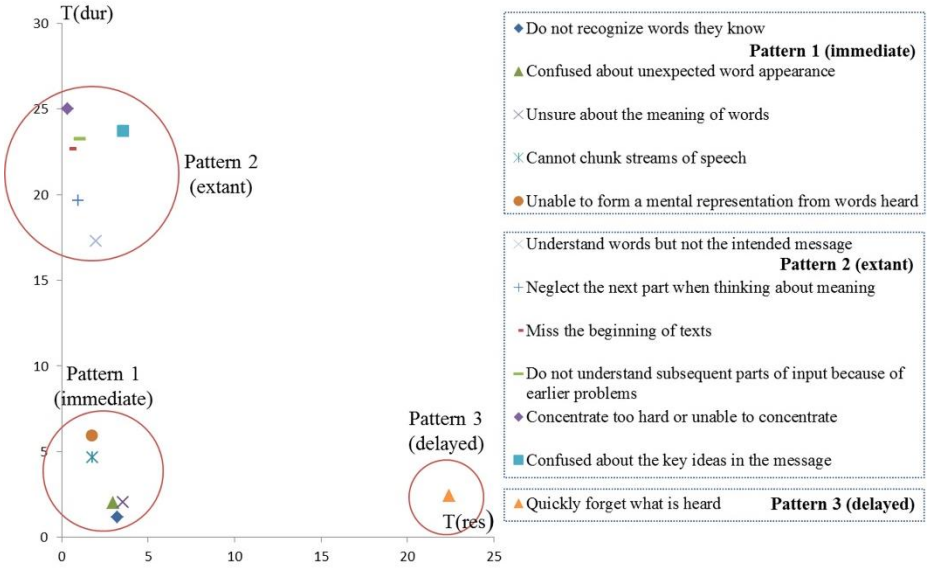

Fig. 2. Different patterns of listening comprehension problems faced by NNSs

Fig. 3 shows an example of a listening comprehension problem (in this case, "do not recognize words they know") in this pattern. In this example, the participant had a problem with the word "fertilizer," which she knew but couldn't recognize it when she heard it.

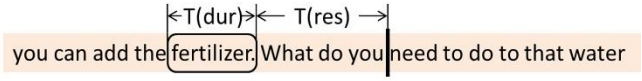

Fig. 3. Example of "immediate" listening comprehension problem

Pattern 2: "extant" listening comprehension problems. The second pattern represented listening comprehension problems with long T(dur) and short T(res) values. The duration of these problems tended to be long and they continued to burden the participants to the point at which they pressed the button. Six types of problems fell into this pattern: "understand words but not the intended message," "neglect the next part when thinking about meaning, "miss the beginning of texts," "do not understand subsequent parts of input because of earlier problems," "concentrate too hard or unable to concentrate," and "confused about the key ideas in the message."

Fig. 4 shows an example of the listening comprehension problem in this pattern. Here, the participant lost concentration and missed the entire sentence ("So there was this scientist. . . entire system got"). As shown in this example, non-native participants facing an extant listening comprehension problem had difficulty with the whole sentence, rather than just words or phrases. Compared to immediate listening comprehension problems (pattern 1), NNSs seemed to feel much more burdened when they faced problems under this pattern.

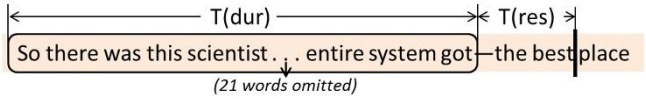

Fig. 4. Example of "extant" listening comprehension problem

Pattern 3: "delayed" listening comprehension problems. The third pattern represented listening comprehension problems with short T(dur) and long T(res) values. The duration of these problems was short, and it took participants a relatively long time to press the button. Only one type of problem fell into this pattern: "quickly forget what is heard." This problem emerged when the participants tried to recall words or phrases they had just heard a few seconds ago.

Fig. 5 shows an example of the listening comprehension problem in this pattern. Here the participant tried to recall the word "bubble gas" when he heard the lecturer's self-correction, "I'm sorry, you must bubble gas through it." According to the participant, when he tried to recall the word to understand the speech, he realized that he had already forgotten it.

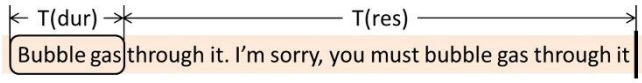

Fig. 5. Example of "delayed" listening comprehension problem

6 Discussion

In summary, we categorized twelve types of listening comprehension problems, two of which were newly identified in this study. We found that the problems can be classified into three patterns based on the persistence of each problem and the time taken to perceive it.

6.1 Interpretation of Findings

How a "pressing a button" method might have affected the results. In our experiment, participants pressed a button whenever they encountered a comprehension problem during listening tasks. Although "pressing a button" requires low over-head from the NNSs, it requires some sort of trigger or decision-making process (i.e., deciding when to press a button), which might have affected the results. For "immediate listening comprehension problems," such as "do not recognize words they know," pressing a button to indicate a problem may be easy. The word they cannot recognize would serve as a trigger to press the button. However, for such "extant listening comprehension problems" as "concentrate too hard or unable to concentrate," participants might have found it difficult to decide when to press the button. For example, one participant reported that "*While listening, I thought I needed to press the button, but I kept having problems, so I didn't know when to press it.*" Similarly, for such "delayed listening comprehension problems" as "quickly forgot what was heard," deciding to

press the button was also difficult. One participant mentioned that "*I was a little hesitant since I had a problem with the previous speech and I wonder if this was the good timing (for pressing it).*" These situations could be one reason for the unbalanced distribution of the problems identified in three patterns: "immediate listening comprehension problems" (74% of all problem occurrences), "extant listening comprehension problems" (25%), and "delayed listening comprehension problems" (1%).

6.2 Design Implications

Providing different support for different patterns of problems. Our findings show three different patterns of problems. We suggest providing different types of support for each one.

Most of the immediate listening comprehension problems are related to words or phrases. Since NNSs instantly perceive the problems, it would be best to provide support that could immediately solve their problems. For example, bilingual dictionaries, machine translation, illustrations may be helpful [13].

For extant listening comprehension problems, NNSs are already overburdened. Inappropriate support would likely to impose further burdens on them. Therefore, the support should focus on reducing their burdens and help them quickly catch up with the speech. For example, showing them keywords extracted from previous speech [26] or providing them with a small amount of time to process speech [31] may be helpful.

For delayed listening comprehension problems, NNSs notice that they forgot some words or phrases earlier in the speech. Since a possible cause of such problems is the limited capacity of the NNSs' short-term memory [1, 9], support should focus on providing memory cues for them. For example, automatically providing text summarizations of previous speech [26] or showing images that can instantly remind them what the speech was about may be of help.

Using advanced sensing technologies, we may be able to associate each problem pattern with certain NNS behaviors. For example, previous research has suggested that pupil response can be used as an indication of effortful listening [33]. If such effortful listening continues for a while, it may indicate that the NNS encounters an "extant listening comprehension problem."

7 Conclusions and Future Directions

In this study, we explored how different types of listening comprehension problems are perceived by NNSs as speech unfolds in a one-way communication setting. Through exploratory analysis of the collected data from a laboratory experiment, we identified twelve types of listening comprehension problems, which we further classified into three patterns based on their persistence and the time taken to perceive them. We believe that our findings serve as a basis for designing adaptive systems to support NNSs in real-time listening comprehension. For future studies, we plan to develop such a system. In addition, we plan to investigate if NNSs with different listening

abilities perceive comprehension problems differently. Finally, we will examine how our findings are compatible with an interactive multilingual communication setting.

Acknowledgments. This research was partially supported by a Grant-in-Aid for Scientific Research (S) (24220002, 2012-2016) from Japan Society for the Promotion of Science (JSPS).

References

1. Anderson, J.R.: Cognitive psychology and its implications (4th ed.). New York: Freeman (1995)
2. Bloomfield, A., Wayland, S.C., Rhoades, E., Blodgett, A., Linck, J., Ross, S.: What makes listening difficult? factors affecting second language listening comprehension. Tech. rep., DTIC Document (2010)
3. Cox, T.H., Blake, S.: Managing cultural diversity: Implications for organizational competitiveness. The Executive pp. 45–56 (1991)
4. Feely, A.J., Harzing, A.W.: Language management in multinational companies. Cross Cultural Management: An International Journal 10(2), 37–52 (2003)
5. Feigh, K.M., Dorneich, M.C., Hayes, C.C.: Toward a characterization of adaptive systems a framework for researchers and system designers. Human Factors: The Journal of the Human Factors and Ergonomics Society 54(6), 1008–1024 (2012)
6. Gao, G., Yamashita, N., Hautasaari, A.M., Echenique, A., Fussell, S.R.: Effects of public vs. private automated transcripts on multiparty communication between native and non-native english speakers. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 843–852. ACM (2014)
7. Gao, G., Yamashita, N., Hautasaari, A.M., Fussell, S.R.: Improving multilingual collaboration by displaying how non-native speakers use automated transcripts and bilingual dictionaries. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. pp. 3463–3472. ACM (2015)
8. Goh, C.: How much do learners know about the factors that influence their listening comprehension?. Hong Kong Journal of Applied Linguistics 4(1), 17–40 (1999)
9. Goh, C.C.: A cognitive perspective on language learners' listening comprehension problems. System 28(1), 55–75 (2000)
10. Goh, C.C.: Exploring listening comprehension tactics and their interaction patterns. System 30(2), 185–206 (2002)
11. Hautasaari, A., Yamashita, N.: Do automated transcripts help non-native speakers catch up on missed conversation in audio conferences? In: Proceedings of the 5th ACM international conference on Collaboration across boundaries: culture, distance & technology. pp. 65–72. ACM (2014)
12. Höök, K.: Steps to take before intelligent user interfaces become real. Interacting with computers 12(4), 409–426 (2000)
13. Ishida, T.: Language grid: An infrastructure for intercultural collaboration. In: Applications and the Internet, 2006. SAINT 2006. International Symposium on. pp. 5–pp. IEEE (2006)
14. Johnson, S.C.: Hierarchical clustering schemes. Psychometrika 32(3), 241–254 (1967)
15. Kalnikaitė, V., Ehlen, P., Whittaker, S.: Markup as you talk: establishing effective memory cues while still contributing to a meeting. In: Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work. pp. 349–358. ACM (2012)

16. Lotfi, G.: A questionnaire of beliefs on english language listening comprehension problems: Development and validation. World Applied Sciences Journal 16(4), 508–515 (2012)
17. Marschan, R., Welch, D., Welch, L.: Language: The forgotten factor in multinational management. European Management Journal 15(5), 591–598 (1997)
18. Nakamura, S., Markov, K., Nakaiwa, H., Kikui, G.i., Kawai, H., Jitsuhiro, T., Zhang, J.S., Yamamoto, H., Sumita, E., Yamamoto, S.: The ATR multilingual speech-to-speech translation system. IEEE Transactions on Audio, Speech, and Language Processing 14(2), 365–376 (2006)
19. Neeley, T., Hinds, P.J., Cramton, C.D.: Walking through jelly: language proficiency, emotions, and disrupted collaboration in global work. Harvard Business School Organizational Behavior Unit Working Paper (09-138) (2009)
20. Otten, M., Van Berkum, J.J.: Does working memory capacity affect the ability to predict upcoming words in discourse? Brain research 1291, 92–101 (2009)
21. Pan, Y., Jiang, D., Yao, L., Picheny, M., Qin, Y.: Effects of automated transcription quality on non-native speakers' comprehension in real-time computer-mediated communication. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1725–1734. ACM (2010)
22. Pérez, L.C.: Foreign language productivity in synchronous versus asynchronous computer-mediated communication. CALICO journal pp. 89–104 (2003)
23. Pickering, M.J., Garrod, S.: Do people use language production to make predictions during comprehension? Trends in cognitive sciences 11(3), 105–110 (2007)
24. Rubin, J.: A review of second language listening comprehension research. The modern language journal 78(2), 199–221 (1994)
25. Takano, Y., Noda, A.: A temporary decline of thinking ability during foreign language processing. Journal of Cross-Cultural Psychology 24(4), 445–462 (1993)
26. Tucker, S., Kyprianou, N., Whittaker, S.: Time-compressing speech: ASR transcripts are an effective way to support gist extraction. In: Machine Learning for Multimodal Interaction, pp. 226–235. Springer (2008)
27. Van Petten, C., Luka, B.J.: Prediction during language comprehension: Benefits, costs, and ERP components. International Journal of Psychophysiology 83(2), 176– 190 (2012)
28. Vandergrift, L.: Recent developments in second and foreign language listening comprehension research. Language teaching 40(03), 191–210 (2007)
29. Wahlster, W.: Verbmobil: foundations of speech-to-speech translation. Springer Science & Business Media (2013)
30. Yagang, F.: Listening: Problems and solutions. In: English Teaching Forum. vol. 31, pp. 16–19. Teacher Development, Making the Right Movies (1993)
31. Yamashita, N., Echenique, A., Ishida, T., Hautasaari, A.: Lost in transmittance: how transmission lag enhances and deteriorates multilingual collaboration. In: Proceedings of the 2013 conference on Computer supported cooperative work. pp. 923–934. ACM (2013)
32. Yao, L., Pan, Y.x., Jiang, D.n.: Effects of automated transcription delay on non-native speakers' comprehension in real-time computer-mediated communication. In: IFIP Conference on Human-Computer Interaction. pp. 207–214. Springer (2011).
33. Zekveld, A.A., Kramer, S.E., Festen, J.M.: Pupil response as an indication of effortful listening: The influence of sentence intelligibility. Ear and Hearing 31(4), 480–490 (2010)
34. Zekveld, A.A., Kramer, S.E., Kessens, J.M., Vlaming, M.S., Houtgast, T.: The influence of age, hearing, and working memory on the speech comprehension benefit derived from an automatic speech recognition system. Ear and hearing 30(2), 262–272 (2009)