The Effects of Aural Input Enhancement on L2 Acquisition

Minyoung Cho  
*The University of Hawai‘i at Mānoa*

Hayo Reinders  
*Anaheim University*

Input enhancement involves attempts to direct the learner's attention to specific linguistic forms in target language input (Sharwood Smith, 1993). One way to do this is by manipulating the input in order to attract learners' attention to the target feature, for example, by underlining or bolding it or by artificially increasing its frequency in the input (an input flood). A number of studies have investigated the effects of enriched input (e.g., Jourdenais, Ota, Stauffer, Boyson, & Doughty, 1995; Reinders & Ellis, 2009; Trahey & White, 1993; White, 1998). Although there is some evidence that enriched input can affect L2 acquisition of certain grammatical features, the results are not conclusive. Furthermore, previous studies have been limited to textual input enrichment. In this chapter we investigated the effects of aural input enhancement, a type of input enhancement that to the best of our knowledge has not been reported on before. Participants in the study were given an audiobook to listen to outside of class in which passive structures had been manipulated by 1) artificially increasing the volume slightly of the target items or by 2) slowing down the speed with which the target items were read out. A control group listened to the audiobooks in their original form. The repeated-measures ANOVA analysis showed no significant effect for the manipulated input on acquiring the target form. We discuss some possible reasons for this finding.

Introduction

There now exists a considerable body of research into the effects of different types of focus on form (FoF) or attempts to direct learners’ attention to form in an otherwise meaning-oriented context (see Ellis, 2009, for an overview). A distinction can be made between more or less obtrusive types of FoF; an example of the former would be the use of consciousness-
raising (Fotos & Ellis, 1991) or input processing (VanPatten, 1996). Unobtrusive FoF includes
the use of positive input enhancement as attempts to manipulate the input and direct
the learner's attention to a specific linguistic form (Sharwood Smith, 1993). This can be
done, for example, by underlining or bolding the target form or by artificially increasing its
frequency in the input (an input flood). A number of studies have investigated the effects
of this so-called enriched input on different aspects of language learning (Jourdenais, Ota,
Stauffer, Boyson, & Doughty, 1995; Reinders & Ellis, 2009; Trahey & White, 1993; White,
1998). Although there is some evidence that enriched input can affect L2 acquisition of
certain grammatical features, the results are not conclusive. In addition, previous studies
have been limited to textual input enrichment, and no studies exist that we are aware of that
have investigated aural manipulation. In this chapter we report on the results of a study that
adopted aural input enhancement in an extensive listening activity. Participants in the study
listened to an audiobook that was artificially enhanced in order to attract their attention to
the target structure of the study.

Effects of input enhancement

Input enhancement

The proliferation of input enhancement research reflects the recognition of the crucial role
of attention in SLA. It is clear that learners do not make use of all the input that they are
exposed to, and it has been widely argued that attention is necessary for L2 learning (Leow,
1993; Tomlin & Villa, 1994) or at least that “There is no doubt that attended learning is far
superior, and for all practical purposes, attention is necessary for all aspects of L2 learning”
(Schmidt, 2001, p. 3). As attention can be externally manipulated (Schmidt, 1990), a
number of studies have investigated ways of drawing learners’ attention to formal aspects of
the input in otherwise meaning-oriented activities, a technique referred to as focus on form.
The importance of external manipulation lies in the assumption that L2 learners, according
to VanPatten’s (1996) input processing hypothesis, tend to prioritize meaning over form
in the input. Formal aspects of the input are only processed insofar as they are crucial for
understanding meaning and insofar as the learner has the cognitive resources available to
pay attention to them. Therefore, without intentional attention to linguistic form, it is less
likely that learners will attend to or use a particular linguistic form in the input for learning.
As a result, a range of external attention-drawing techniques have been proposed, including
explicit rule presentation, input flooding, and—the focus of this study—input enhancement.

Input enhancement involves the manipulation of input salience, with an attempt to direct
learners’ attention to a specific linguistic form in the input (Sharwood Smith, 1991, 1993).
Salience refers to the ease with which learners can perceive given input. Salience of input
is determined by various learner-internal, structural, and external factors. For example,
learners’ interlanguage development (Spada & Lightbown, 1999), inherent features in
linguistic structure such as communicative value (Dulay & Burt, 1978), and external
manipulation may alter the degree of salience of certain features in the input. Although the
sources of input salience vary, it is of course learners themselves who must further process
the input for learning (Schmidt, 1990). For this reason many studies of input enhancement
have investigated whether input enhancement is effective in drawing learners’ attention to
form and its effect on subsequent acquisition of that form. Different external manipulations
have been incorporated to increase the salience of input, including manipulation of
frequency (input flood), visual salience (typographical or textual manipulation), and
corrective feedback in discourse (e.g., repetition or recast).
According to Sharwood Smith (1993), input enhancement can be categorized as either positive or negative. Positive enhancement emphasizes the correct form, such as through input flood and input enhancement, whereas negative enhancement incorporates the indication of errors in, for example, learners' production that can be enhanced by means of explicit instruction and/or corrective feedback. Both enhancement types are intended to trigger changes in input processing mechanisms of salient forms, but in this chapter we are concerned only with positive input enhancement. Examples of positive enhancement include manipulating textual or typographical aspects of the text (e.g., bolding, underlining, italicising, or changing the font type or size), visual enhancement (e.g., pictures), and technological enhancement (e.g., using a combination of keystrokes to type diacritics).

As discussed above, the use of external salience is susceptible to diverse learner-internal and external factors, and these, together with a number of methodological problems, have been speculated to explain some of the inconsistent results in input enhancement research (Han, Park, & Combs, 2008; Lee & Huang, 2008). These include the length of exposure to the target input (Leow, 1997), learners' prior knowledge of the language and the target items (Jourdenais et al., 1995; Leow, Egi, Nuevo, & Tsai, 2003), the nature of the target form (Wong, 2003), and its modality (Leow, 1993, 1995).

Types of input enhancement

Although a strict interpretation of input enhancement (Barcroft, 2003) accepts only a considerable alteration of input (through a change of or addition to the original input), a more liberal interpretation also includes other techniques such as technological enhancement and visual enhancement.

Technological enhancement refers to the use of technology as a means to draw learners' attention to the target object. Gascoigne's (2006) study, for example, involved participants listening to L2 input and then transcribing the input either on a computer or on paper. She hypothesized that the insertion of diacritics when typing on a computer, involving the pressing of additional keys, would force learners to pay conscious attention to the orthography. The results showed that keyboard transcription had a positive effect on recall of the target items compared with pen-and-paper transcription.

Another form of input enhancement is visual enhancement, which involves the inclusion of visual information (e.g., pictures) to highlight certain aspects of the text. Labrie's (2000) study compared beginning L2 learners' acquisition of vocabulary by comparing a web-based reading text enriched with images and sounds with a paper-based text without visual and aural aids. Labrie found, perhaps not surprisingly, that the inclusion of visual and aural information helped with vocabulary acquisition.

Finally, aural input enhancement involves the manipulation of listening materials, for example, by increasing the volume of target items in the text or by including a short pause before and/or after the target items. Although Gascoigne (2006) mentioned that an "oral [our emphasis] equivalent of textual enhancement" could be achieved via stress, intonation, or gestures (p. 149), we are unaware of previous studies investigating this type of input enhancement. Compared to the number of studies on written input enhancement, this lack is perhaps surprising, especially as aural enhancement may occur in natural or classroom discourse as a form of corrective feedback or recast.

Modality and aural enhancement

It is generally acknowledged that modality has a significant effect on input processing (and consequently on intake and acquisition). The separate streams hypothesis (Penny, 1980)
posits that visual and audio language input are separately and independently processed without interference. L1 studies investigating differences between reading and listening have shown that listening is more taxing than reading of the same input (Anderson, 1980; Danks, 1980; Rost, 1990), as learners do not have the same amount of control over the aural input as they do over written input. When reading, learners can more easily recognise different text elements, such as words, sentences, and paragraphs, and can re-read parts of the text (Rost, 1990), whereas in listening, segmentation of word boundaries or even boundaries between different word elements are not discrete, and learners need to rely on prosodic and intonational cues in the input to understand sequences of input (Anderson, 1980). L2 research has also shown that modality places constraints on the way input is processed; Johnson’s (1992) and V. Murphy’s (1997) studies, for instance, revealed that (adult) learners’ performance on grammaticality judgement tasks was slower and less accurate in the aural mode than in the written mode. And Wong (2001) compared the ability of learners to focus on both form and meaning in aural and written modes and found that the aural mode was more challenging than the written mode.

In terms of aural enhancement, early SLA research into teacher talk has shown frequent use of speech modifications by ESL teachers. Some studies (Dahl, 1981; Håkansson, 1986; Henzl, 1979) revealed that teachers adjusted their speech rate to the level of learners’ proficiency, and others (Chaudron, 1982; Wesche & Ready, 1985) reported teachers’ insertion of pauses around certain aspects of their speech production to make it more comprehensible to learners. Chaudron (1982), in particular, descriptively observed native teachers’ tendency to insert pauses around difficult words to make them more comprehensible to ESL learners. Other phonological, intonational, or stress characteristics have been reported to be modified by teachers (Chaudron, 1982; Henzl, 1973). Although those early studies are not generalizable due to the lack of a comparable baseline, Chaudron (1988) concludes that native teachers seem to modify their speech in certain ways to make it more comprehensible for learners. A decade later, in a study on the effects of recasts, Doughty and Varela (1998) used recasts with a rising intonation to draw learners’ attention to a particular form. Because the study did not include a baseline recast (with no intonational emphasis), it is difficult to interpret the effects of intonational emphasis in recasts; however, it seems likely that such intonational emphasis has been accepted by practitioners and researchers as one possible attention-drawing technique.

No studies, to the best of our knowledge, have investigated the effects of aural manipulation in the input on drawing learners’ attention to form in the experimental context. Leow’s (1995) study does give us some insight, as it was a replication of an earlier study (1993) into the effects of simplified written input, type of linguistic items, and L2 experience on intake, but in the aural mode, allowing for a comparison between the two. Although the studies did not reveal modality effects of simplified input on learner intake of the target form, mode of input did seem to have an effect on learners’ intake of different types of input, possibly due to the phonological salience of different morphemes. Leow emphasised the need to further investigate the effects of modality on input processing.

Study

Participants

A total of 72 Korean learners of English participated in the study. Participants were enrolled in a compulsory freshman English course, entitled “Academic English for business majors.” They were from three intact classes taught by the same instructor. Students in the classes

1It should be noted, however, that there is also an important age effect.
The effects of aural input enhancement on L2 acquisition were randomly placed into one of these classes, if they had insufficiently high scores on one of the accepted university entrance tests (e.g., TOEFL or the university entrance exam). The experiment was conducted as part of classroom activity on extensive listening, and participants received extra points for their participation. Participants’ TOEFL IBT scores ranged from 80 to 110 (out of 120) and national university entrance exam scores in English from 94 to 100 (all students had at least one of these scores), indicating that participants could be expected to be able to complete listening to an audiobook independently outside of the classroom without too much trouble. The class instructor was consulted regarding the participants’ English ability to complete the task. Participants were told the purpose of the study was to look at extensive listening practice and were given information about the length of the study and its procedures.

Participants at this level might be expected to know the basic rules pertaining to passive structures, especially considering the fact that they are introduced early in formal education (e.g., compulsory English education from Grade 3) and the fact that their formation is relatively simple. Nonetheless, many previous studies have noted challenges for students in fully mastering passive forms, even at advanced levels (Hinkel, 2002). This was confirmed for the participants in this study also as shown in their pretest scores (see below).

Design
A computer-based timed grammaticality judgement test (GJT) was administered to participants as a pretest. The GJT was preceded by instructions and a practice session. Next, the three classes were randomly assigned to either one of the two experimental groups [the Pause Group (PG, n=24), or the Reduced Speed Group (SG, n=23)], or the control group (CG, n=25). Participants in each group were given an audio version of a graded reader (see below) to listen to. They received their respective files via the university course management system so that no other individuals except the study participants could access the audio files. Students were told they were only to use the files themselves for the purposes of the study and not to share the files with others. The two experimental groups were given artificially manipulated files, which either included short pauses around the target forms (PG) or which had the speed of the audio recording reduced for the target items (SG). The control group received the original (unaltered) audio file. Participants were asked to listen to the audio files for the purpose of enjoying the story in the book as part of their course homework. They were given one week to complete listening to the approximately 90-minute recording.

After one week participants completed a posttest GJT, which included the same items as the pretest but in a different order. A survey (see Appendix A) was administered after the posttest, and this asked learners about their background in learning English, their experience in listening to the audiobook (for example, where and when they listened to it), what devices they used for listening, as well as whether they had noticed anything special about the recording (to establish whether they had noticed the input enhancement). In addition, participants were asked whether, and how often, they had listened to the book in the debriefing questionnaire. Nine participants indicated they had not listened to the audiobook or had listened to it more than once. They were removed from the dataset.

Target structure
The English passive was chosen as the target form for the study. As Hinkel (2002) points out, although the passive structure is common in English, classroom practice and instruction on passives have focused on deriving passives from active structures, making it difficult for learners to master the form on its own. Particularly in L2 production, common types of
errors in passive production include the use of the intransitive verb in passive constructions
and the use of the form in inappropriate contexts.

For example:

a. The accident happened.

b. *The accident was happened.

Previous studies on L2 learners’ passive acquisition have revealed that even advanced L2
learners tend to overpassivitize unaccusative verbs such as occur and happen (Kim, 2003;
Yip, 1995) and tend to accept ungrammatical sentences including passive structures
with unaccusative verbs. Further complicating matters for learners is the fact that the
agent of an action verb in an active structure tends to be phonologically suppressed
and that the passive structure does not follow the first noun principle. VanPatten (1996,
2004) has shown that learners, regardless of their L1, tend to perceive the first noun or
pronoun in a sentence as an agent, which is not the case in passive structures. These
factors together make it difficult for L2 learners to recognise thematic relationships in
passive sentences. Based on the level of the participants (see above), we felt the passive
form was both challenging enough, while not being beyond the participants’ grasp. In
order not to make the tests too difficult, we included mainly simple passives, similar to
those found in the audiobook. In addition, the passive structure occurred frequently in
the audiobooks we considered, thus making it a good practical choice. In general, for
the correct passive forms, grammatically correct simple sentences with a by-phrase were
presented (e.g., ‘The new rules were written by the head teacher’), and for the incorrect
forms, the be verb in the passive structure was omitted (e.g., ‘*He injured when he fell
off his bike’).

Treatment

After consulting with the course instructor, the book Frankenstein, a level three book from
the Penguin Active Reading series of graded readers was chosen. Permission was obtained
from the publisher to manipulate the audio recordings for the purpose of our study. The
audio files were digitally manipulated with the use of the sound editing software Audacity
(a freely available audio editing program) by inserting pauses of about 1.5 seconds before
and after the target forms or by slowing down the playback speed of the target items. We
slowed down the recording by 7–10% while not altering its pitch—so as not to make the
slowed down text sound unnatural. The files were then converted to mp3 files, and students
were instructed to download the files through the university course management system and
listen to the audiobook through their mobile phones or any other listening device. Further,
they were asked to listen to the book only once (so as to avoid adding length of exposure as
an additional variable). Participants were not told about the purpose of the study, other than
that it was about extensive listening. No instruction on the target structure was given during
the experimental period.

Tests

A timed grammaticality judgement test (GJT) was administered as a pretest and a posttest.
Both tests contained the same items but in a different order. There were 50 sentences of
which 20 contained the target structure and the remaining 30 were distracters. Out of the
20 target items, 10 were grammatical and 10 were ungrammatical. Items were randomly

---

2 The narrator of the story uses British English. Korean learners of English are more used to hearing
American English. However, because of the relatively low level of the text and the absence of any
clear regional expressions in the text, we considered this acceptable.
The effects of aural input enhancement on L2 acquisition presented. The items were similar to those found in the audiobook in terms of their complexity. Before the test, participants were instructed on how to complete the GJT and completed a number of practice items. Participants were asked to press the enter key on the keyboard if they thought a sentence was grammatical and the left-hand shift key if they thought it was ungrammatical. The keys were labelled with stickers indicating correct and incorrect. Since the tests aimed at measuring learners' implicit knowledge of the target feature, learners received only a few seconds to judge their grammaticality. They were encouraged to use their intuition, rather than to consciously apply grammar rules. The sentence presentation time differed, depending on the length of the sentence, and was based on Reinders (2009).

Although GJTs have some drawbacks (see, for example, Birdsong, 1989), the timed GJT seemed appropriate as the most likely effect of the extensive listening treatment would be the development of implicit knowledge. With regards to item consistency, a Kuder-Richardson 20 (KR–20) score was calculated. In the pretest the score was 0.715, and in the posttest it was 0.878, showing sufficiently high consistency.

In addition to the GJT, participants were given a comprehension test at the same time as the GJT posttest. This included six multiple-choice questions related to the main plot of the story to confirm that they had listened to the audiobook. These questions were written so as to be impossible to answer without having read the book (e.g., 'Where did Robert Walton first meet Victor Frankenstein, as he wrote in his first letter?').

Analysis

To measure performance on the pre—and post-tests, participants were given one point for each correct item and zero points for each incorrect or missing item. The score comparisons among groups included both overall GJT scores and scores on the target features. The analysis of the pretest revealed that there were no significant group differences for both scores (overall scores and passive scores), so it was assumed that the three groups were similar in terms of their knowledge of the passive structure and English grammar in general at the start of the study.

Results

In order to examine mean differences in individuals' passive knowledge improvement under different conditions of audio input, a mixed-design repeated-measures ANOVA was adopted. That is, Condition was used as a between-subject factor with three levels—the pause group (PG) vs. the reduced speed group (SG) vs. the control group (CG)—and Time was used as a within-subject factor with two levels (pretest vs. posttest). The dependent variable was learners' passive scores from the grammaticality judgement test. This analysis seemed more appropriate than a one-way ANOVA on posttest scores only because it is possible that all learners may benefit from merely listening to the audio file, regardless of the different audio input.

Table 1 shows descriptive statistics for scores on passive structures from the pretest and posttest. The mean score on the pretest for passives was 13.08 (SD=3.49) for the PG, 12.08 (SD=2.97) for the SG, and 11.88 (SD=2.773) for the CG. The posttest scores seem similar across all three groups with the average scores of PG=14.208 (SD=3.106), SG=13.217 (SD=3.147), and CG=13.120 (SD=3.059), although overall PG (pause group) performed better in both pretest and posttest, followed by SG and CG, in that order.
Table 1. Descriptive statistics for passive structures on the pretest and posttest

<table>
<thead>
<tr>
<th>group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>pretest</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PG</td>
<td>24</td>
<td>13.083</td>
<td>3.488</td>
<td>.712</td>
</tr>
<tr>
<td>SG</td>
<td>23</td>
<td>12.087</td>
<td>2.968</td>
<td>.619</td>
</tr>
<tr>
<td>CG</td>
<td>25</td>
<td>11.880</td>
<td>2.773</td>
<td>.555</td>
</tr>
<tr>
<td>posttest</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PG</td>
<td>24</td>
<td>14.208</td>
<td>3.106</td>
<td>.634</td>
</tr>
<tr>
<td>SG</td>
<td>23</td>
<td>13.217</td>
<td>3.147</td>
<td>.656</td>
</tr>
<tr>
<td>CG</td>
<td>25</td>
<td>13.120</td>
<td>3.059</td>
<td>.612</td>
</tr>
</tbody>
</table>

Note. PG=pause group, SG=reduced speed group, CG=control group (original format)

In order to examine the main effects for Time (pretest-posttest) and Condition (treatment) and the interactions of Time by Condition, the repeated-measures ANOVA was used. As can be seen in Table 2, there was a significant Time effect in the analysis ($F(1, 69)=15.821$, $p=0.00$), suggesting that learners’ passive scores significantly improved from pretest to posttest. Although the proportion of variance explained by Time in the within group design was only 18.7% (partial $\eta^2=0.187$), the significant result can be interpreted with substantial power (power=.975). However, in terms of the main effect for Condition, there was no significant difference in group scores ($F(2, 69)=0.162$, $p=0.319$, partial $\eta^2=0.033$), with Condition explaining only 3.3% of the variance in the score in the between group design. The results suggest treatment differences in the three groups did not influence the amount of improvement from pretest to posttest scores on the target form. Further, no significant interaction of Time and Condition was observed ($F(1, 69)=0.017$, $p=0.983$, partial $\eta^2=0.000$), with the interaction explaining 0.0% of the variance in the within group design. This suggests that there is no interaction effect for Time and Condition on learner scores.

Table 2. Statistics for the effects of Time, Condition, and their interaction

<table>
<thead>
<tr>
<th>source of variance</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>partial $\eta^2$</th>
<th>power</th>
</tr>
</thead>
<tbody>
<tr>
<td>within group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>48.816</td>
<td>1</td>
<td>48.816</td>
<td>15.821</td>
<td>.000</td>
<td>0.187</td>
<td>0.975</td>
</tr>
<tr>
<td>Time*Condition</td>
<td>0.103</td>
<td>2</td>
<td>0.052</td>
<td>0.017</td>
<td>.983</td>
<td>0.000</td>
<td>0.052</td>
</tr>
<tr>
<td>Condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>error</td>
<td>212.897</td>
<td>69</td>
<td>3.085</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>between group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td>37.392</td>
<td>2</td>
<td>18.696</td>
<td>1.162</td>
<td>.319</td>
<td>0.033</td>
<td>0.247</td>
</tr>
<tr>
<td>error</td>
<td>1109.914</td>
<td>69</td>
<td>16.086</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1 reveals the score changes from pretest to posttest for the three groups. As the figure suggests, all groups improved from pretest to posttest (Time 1 to Time 2), implying that all groups in general improved their knowledge of the passive forms after listening to the audiobook. However, overall there is clearly no interaction between Condition and Time (i.e., score improvement).
The effects of aural input enhancement on L2 acquisition

Figure 1. Plots for pretest and posttest scores for each group

To sum up, the results indicate that learners’ knowledge of the passive form improved from pretest to posttest by listening to the audiobook regardless of the type of aural input they received. At the same time, the results suggest no differential effect for aural enhancement on learners’ knowledge improvement of the target form. Although score improvement from pretest to posttest occurred in all three conditions with substantial power (power=.975), the findings for the effect of aural manipulation must remain tentative due to the low power (power for Condition effect=.247; power for Condition * Time effect=.052), because the lack of power may not allow us to detect any significant effects, even if such effects exist in reality.

Questionnaire

Upon completion of the posttest, participants were given a questionnaire that asked (among other questions) whether they had noticed anything special about the recordings. Approximately 70% of participants (17/24) in the pause group (PG) said they had noticed the pauses, and of those who did, about 47% (8/17) said the pauses were intended to make them concentrate on the story, two said it was to give them more time to understand the input (without mentioning the exact object of the input), and one said it was intended to encourage attention to the words that followed the pauses. Three further students said they did not know why the pauses were included. Participants in the reduced speed group (SG) did not specifically point out the slowing down of the target structure. Some also mentioned other aspects of the recording, which had not been altered by us, such as the British pronunciation, the general speed of the recording, and the length of the book. This applied to students in the control group (CG) also. None of the participants’ responses indicated they had noticed that the pauses or recording speed were intended to draw their attention to the target structure, which indicated that students’ attention to the target form had not reached the level of awareness. However, given that some participants mentioned that the reason for the pauses was to get them to pay attention to the recording (although none specified a form), it seems possible that the pauses may have increased learners’ readiness to accept or process the subsequent input. Nonetheless, the students’ responses clearly indicate the contrasts in terms of learners’ recognition of the manipulation; pauses seem to be more salient than a reduction in recording speed.
Another aspect of our study was whether extensive listening can be successfully implemented as an out-of-class activity. 60% of the learners indicated they listened to the audiobook with their own portable devices such as MP3 players and mobile phones. Only 40% of participants had listened to the book on a computer. Most participants (66%) indicated that they had listened to the book in an informal setting such as on the subway or bus when commuting to and from home or before going to bed, while 25 respondents (34%) said that they had listened to it at school. The responses point to the potential of using mobile devices for encouraging out-of-class practice.

**Comprehension test**

Students performed well on the comprehension questions they were given at the same time as the questionnaire (Mean=87.04, SD =14.91), showing that they had indeed listened to the story and the level was suitable for them.

**Discussion**

This study investigated whether external manipulation of features in L2 aural input contributes to the acquisition of the passive structure. Learners were exposed to natural language input by listening to an audiobook in which the target structure was either slowed down or in which pauses were inserted around the target structure. Participants' knowledge of the form was examined by comparing their receptive knowledge on two grammatical judgment tests (GJT), administered before and after the treatment. Although participants' test scores improved from pretest to posttest (most likely as a result of completing the extensive listening as part of the treatment), there was no effect for aural enhancement. Although it is hard to directly compare the current results to previous input enhancement studies, due to the different modality, some methodological issues can be raised in discussing possible reasons for our findings.

First, it is possible that the amount of enhanced input in the current study was not sufficient for learners to construct the rule (or at least to do so to a greater extent than participants in the control group, who simply listened to the unaltered audiobook). Similarly, the period of time over which the input was presented may have been too short. Participants listened to only one fairly short audiobook and were exposed to only 65 correct exemplars of the target structures. It is possible that they would need, if not more, than at least more repeated exposure to the target structure, for example by listening to several books over an extended period (weeks, months) of time. Schmidt (1990) emphasized the importance of frequency of input, given other things being equal, along with perceptual salience in increasing the likelihood of input to be perceived as intake and further processed for learning.

Another variable is the complexity of the target structure; simple rules are more susceptible to enhancement than complex ones. Although variables determining complexity are subject to debate (e.g., Dekeyser, 1994, 1998; Robinson, 1996), some studies (Kim, 2003; Yip, 1995) have highlighted the difficulty of acquiring the passive form. Although the passive structure adopted in the current study seems simple in terms of rule formation, adaptation of the form and use of it in a sentence is relatively complex when it comes to the incorrect use of intransitive verbs in passive and double-object constructions. Robinson's studies (1995, 1996) showed that with complex rules, learners with explicit instruction are better at acquiring the form than those in implicit and incidental learning conditions. Based on this, grammar learning in sentences beyond identifying simple rule formation may not work in incidental enhanced form learning conditions as used in the current study.
Moreover, the difficulty of processing aural input may have had an impact. A number of studies have pointed out the challenges of processing aural input compared with written input (Anderson, 1980; Danks, 1980; V. Murphy, 1997; Rost, 1990). One of the important differences between written and aural input processing is that learners have very little control over the incoming input when listening. Therefore, participants may have been unable to benefit from the input enhancements.

Although no differential effect on learning was observed for the pause and the slowed speed groups, the results from the debriefing questionnaire showed that about 70% of learners (17/24) in the pause group had noticed the pauses, whereas none in the slowed speed group (SG) had noticed the change in speed. Future studies investigating the effects of aural input enhancement may thus wish to use pauses rather than alter the playback speed. In addition, it is interesting to consider why such recognition did not lead to acquisition. Although there was no effect for the enhanced input and participants were not aware of the enhancement’s purpose to draw their attention to the target structure, it is theoretically possible that participants did in fact notice the target structure but not to the point of awareness and therefore did not engage in further processing. Williams’ study (1999) showed that, especially in meaning-oriented tasks, further processing of noticed input is necessary for the underlying rules to be formulated:

If learning distributional rules is critically dependent upon the subjects initially paying attention to relations between elements in the input, then it follows that even the simplest rules might not be learned if the subjects for some reason fail to attend to those relationships (p. 32).

In other words, participants may have noticed the form but failed to allocate attention to related grammatical segments or at least did not make the form-meaning connections necessary to establish their underlying rules (cf. Baddeley, 1990). In addition, the lack of explicit instruction and the fact that learners were not asked to focus on the target form may have prevented them from allocating attention to the target grammar (G. Murphy & Shapiro, 1994). Also, the lack of feedback may have prevented them from attempting to generate their underlying rules (Baddeley, Gathercole, & Papagno, 1998; Williams, 1999).

To reiterate, a number of methodological issues identified in previous enhancement studies may have played a role in this study also; external input saliency cannot guarantee internal focus on the form. Although learners’ attention may have been activated by the external manipulation, their attention to form may have remained at a shallow level, thus not leading to an association of the form and its underlying rule.

Nevertheless, as some students did notice the manipulation through the insertion of pauses, without having been alerted to it through instructions or examples, it seems possible to enhance learners’ readiness to accept the following input (either in terms of content or grammar) through this type of enhancement. This, nonetheless, needs more robust empirical evidence.

We feel that the methodological significance of the study lies in its attempts to incorporate authentic materials in the study of focus on form/input enhancement, as well as incorporation of input enhancement in an out-of-class learning context. That is, this study tried to examine whether L2 learners can successfully allocate their attention to salient features of grammar within (purely) meaning-oriented activities in a non-pedagogical environment. Unlike many other previous input enhancement studies where learners are placed into experimental conditions where the target form is presented more frequently than in its natural context, the extensive listening task adopted in this study more closely
resembles a natural language learning situation. Learners’ main purpose for completing
the listening activity was to understand and enjoy the story rather than to pay attention to
details of the language or text. Further studies such as this may help to better understand
the pedagogical potential of implementing form-focused instruction in informal language
learning settings.

Conclusion
This study attempted to investigate whether aural input enhancement in extensive listening
can facilitate acquisition of the passive structure. Passive structures in an audiobook
were aurally enhanced, and the effect of the enhancement on learning was measured
with a grammaticality judgement test. Analysis of the results showed no significant aural
enhancement effect on acquisition of the target structure. The results could be attributed
to the cognitive demands in processing extensive auditory input; extensive listening is a
primarily meaning-oriented type of language practice, which may have made participants
less likely to pay attention to language form.

However, a number of limitations prevent us from making definitive statements to explain
our research findings. First, a lack of a noticing measurement in the study means we were
unable to determine whether learners did not notice the form, or did notice it but did not
acquire the structure as a result. Although post-treatment self-reports were used to reveal
participants’ noticing of the enhanced target structure, self-report can only provide limited
information. Also, because participants listened to the audiobook in an uncontrolled
environment, contextual factors could have had an effect. These include the time interval
between finishing listening to the book and taking the test, as well as the conditions during
listening (e.g., listening in silence vs. listening in a noisy environment). Finally, the modality
in the treatment was different from that in the tests (listening versus reading). Our reason
for using the GJT was that it would allow us to easily include a large number of target items
in an easy-to-administer test of relatively short duration. In future studies, however, it will be
important to reconsider this issue.

Despite these limitations, we do feel it is important to investigate extensive listening in
its natural environment, and by using tools that the young adult learners in this study
are most comfortable with, we feel that the study benefits from greater ecological validity
(Van Lier, 1996). Laboratory conditions, although affording greater control in terms of data
collection, may not be suitable if the results are to be a true indication of the potential effect
of extensive listening activities. In this sense, the current study hopefully makes a small but
significant contribution to contextualizing aural input enhancement research.

References
Barcroft, J. (2003). Distinctiveness and bidirectional effects in input enhancement for
York: Springer.


Appendix

Q1. What device did you use for the listening?
   __ MP3 player
   __ cell-phone
   __ computer
   __ other, please specify:
   ______________________________________________

Q2. Where did you listen to the audiobook?
   __ on the subway
   __ on the bus
   __ at home
   __ in school
   __ other, please specify:
   ______________________________________________

Q3. How many times did you listen to the audiobook?
   __ none
   __ once
   __ twice
   __ more than three times

Q4. When did you finish listening to the audiobook?
   __ today
   __ yesterday
   __ two days ago
   __ three days ago
   __ more than four days ago

Q5. Had you read this book before (or listened to it as an audiobook)?
   __ yes
   __ no

Q6. Did you notice anything unusual about the recording (quality of recording)? If so, what?
   ______________________________________________

Q7. If you answered yes, what do you think was the purpose of the unusual recording?
   ______________________________________________