Identification of Word Boundaries

上田真理文

Abstract

This article explores the comparison of two projects researching the identification of word boundaries in foreign language listening. It also questions whether the rate of identification of word boundaries in listening can be improved when the words are embedded in a context. It suggests that Japanese learners of English become fixated at word level, and do not have enough spare attentional capacity to draw upon cues in order to identify the word boundaries from the given context.

keywords: foreign language learning, word boundaries, reduced form, listening

Section 1 Introduction

1. Identification of Word Boundaries

James and Mullen (1973:18) claim that it is difficult for non-native speakers of English (= NNS) to acquire ‘juncture clues’ of the target language. Gimson (1989:306) also suggests that junctural oppositions are frequently neutralised in connected speech or may have such slight phonetic value as to be difficult for a listener to perceive. Nakatani and Dukes (1977:714) point out the importance of juncture recognition as well:

To understand speech, we must hear words. And to hear words, we must know where they begin and end. This is not a trivial problem in speech perception because, unlike writing which has spaces between words, speech doesn’t (sic) have silent gaps or other obvious cues that divide the sound stream into words.

Takefuta (1984) reports that the weakest points in English phonetics for Japanese learners of English (= JLE) are as follows:

<table>
<thead>
<tr>
<th>the weakest rank</th>
<th>index of signal detectability ($=d^*$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>segmentation (=juncture)</td>
<td>$d^*$=1.2</td>
</tr>
<tr>
<td>intonation</td>
<td>$d^*$=1.6</td>
</tr>
<tr>
<td>single sound</td>
<td>$d^*$=2.7</td>
</tr>
<tr>
<td>stress (=accent)</td>
<td>$d^*$=2.9</td>
</tr>
</tbody>
</table>

Takefuta (1991:229)
Firstly, before discussing these claims, it is necessary to define ‘weakest point’. Although Takefuta (1991) reports that ‘segmentation’ or ‘juncture’ is the weakest point for Japanese learners of English in his research, what he means could be called ‘identification of word boundaries’ (=IWB) in a more precise term. According to Crystal (1985), ‘segmentation’ is the identification of a discrete unit that can be identified, either physically or auditorily, in the stream of speech (Crystal, 1985: 272). Since IWB is a kind of segmentation, segmentation in general will be discussed before IWB, which is a more specific issue. There are several kinds of processes relevant to the boundary between adjacent words in English: assimilation, elision, liaison, hiatus, double consonants (Knowles, 1987: 130-132), liaison and so forth. In this research, assimilation, elision and liaison will be the main points of focus.

1. 2. Assimilation

Firstly, assimilation is the process whereby a sound is made phonetically more similar to the sounds constructing its phonetic environment (Wells, 1982; 96). When a sound anticipates some features of the pronunciation of a following sound, or continues some features of a proceeding sound, it is said to ‘assimilate’ to that sound. The alveolar consonants /t, d, n/, for instance, tend to assimilate to a following labial or velar consonant, and anticipate the place of articulation:

- *that boy* /ðat boɪ/ becomes /ðap boɪ/
- *one girl* /wɒn ɡɜːl/ becomes /wɑŋ ɡɜːl/
- *Red Cross* /red kros/ becomes /reg kros/

The sibilants, /s, z/ assimilate to the place of articulation of /ʃ, ʒ/:

- *Nice shape* /naɪs ʃeɪp/ becomes /naɪʃ ʃeɪp/
- *Boy’s shoes* /boɪz ʃuːz/ becomes /boɪʒ ʃuːz/

As a variation of assimilation, there is a sporadic assimilation of dental fricatives. It occurs when /θ, ð/ precedes or follows /s, z/:

- *both sides* /bouθ saɪdz/ becomes /bous saɪdz/
- *How’s that?* /hauθ ət/ becomes /hauz ət/ 

This rule is affected by speech rate. In a rapid conversation, assimilation is more likely to happen. The knowledge of assimilation itself would not improve the foreign language listening comprehension dramatically. However, surely JLE need to be able to decode its effects.

1. 3. Elision

Secondly, elision is the omission of sounds in connected speech and both consonants and
vowels may be affected (Crystal, 1985: 107). Crystal suggests the following categories of elision:

**i - a)** Unstressed grammatical words
such as *and* and *of* are particularly prone to elision, as when the *f* is dropped in *cup of tea* (cf. cuppa tea), or the *a* and *d* are dropped in *boys and girls* (cf. boys ‘n’ girls)

**i - b)** Polysyllabic words
the vowels and consonants in unstressed syllables regularly elide in conversational speech of normal speed:

\[
\begin{align*}
camera & \quad /\text{kam}r\text{ə}/ \\
probaby & \quad /\text{præbl}i/ \\
February & \quad /\text{fɛb}ri/
\end{align*}
\]

**ii)** Complex consonant clusters

\[
twelfths \quad /\text{twel}fs \text{ or } twelfs/
\]

Brown (1977: 66-67) argues that the most common elision is that of /t, d/:

/ **t** / elision

\[
\begin{align*}
\text{first three} & \quad /\text{fɛst} \text{ əri}/ \quad \text{becomes} \quad /\text{læs} \text{ əri}/ \\
\text{last year} & \quad /\text{læst} \text{ jɪə}/ \quad \text{becomes} \quad /\text{læs} \text{ jɪə}/ \\
\text{most recent} & \quad /məʊst \text{ rɪsənt}/ \quad \text{becomes} \quad /məʊs \text{ rɪsənt}/
\end{align*}
\]

/ **d** / elision

\[
\begin{align*}
\text{banned for life} & \quad /\text{bændən laɪf}/ \quad \text{becomes} \quad /\text{bændən laɪf}/ \\
\text{ground pressure} & \quad /\text{ɡraʊnd prɛʃə}/ \quad \text{becomes} \quad /\text{ɡraʊm prɛʃə}/
\end{align*}
\]

The elision of / **t** / and / **d** / is so common in this environment that one is surprised to hear these sounds in the stream of native English speech, even in quite slow and deliberate speech. Although / **t**, **d** / elision is the most common, it is possible for any consonant to be elided in certain circumstances. Brown gives as examples / **v**, **θ**, **l**, **r**, **n**/ and / **k** / elision:

/ **v** / elision

\[
\begin{align*}
\text{five p.m.} & \quad /\text{faɪv pɪtəm}/ \quad \text{becomes} \quad /\text{fæɪp ɪtəm}/ \\
of course & \quad /əv kɒs/ \quad \text{becomes} \quad /ə kɒs/
\end{align*}
\]

/ **θ** / elision

\[
\begin{align*}
\text{I think that was} & \quad /\text{aɪ tɪŋk ət wɛs}/ \quad \text{becomes} \quad /\text{aɪŋkətwɛs}/
\end{align*}
\]
went the way of the /wentəweɪ vandalism/ becomes /wentəweidʒə/

/1/ elision
also /ɔlsəʊ/ becomes /ɔsəʊ/
already /ɔlredi/ becomes /ɔredʒi/

/ə/ elision
a year ago /ə jɪərə ɡəʊ/ becomes /ə jɪərə ɡəʊ /
for instance /fɔr ɪnstənts/ becomes /fɪnənts/

/k/ elision
asked him /ɑsktɪm/ becomes /ɑstɪm/
excursion /ɪskɜʃən/ becomes /ɪskɜʃən/

Brown (1977: 68-71)

Not only consonants, but also vowels can be elided. Simple vowel elision unaccompanied by assimilation or syllabification can occur. Gimson (1980: 293) explains that initial /ə/ is often elided, particularly when followed by a continuant and proceed by a word final consonant (compensation for the loss of /ə/ frequently being made by the syllabicity of the continuant);

not alone /notl ˈlən/ get another /getə nədə/
run along /ˈrʌn ˈlɔŋ/ he was annoyed /hi wəznənəd/  

Or, when an appropriate vowel proceeds, word initial /ə/ may coalesce with the proceeding vowel:

go away /ɡəweɪ/ try again /trəɡən/

Or again when final /ə/ occurs with a following linking /r/ and the word initial vowel /ə/ may be elided;

after a while /əˈftraʊ wail/  as a matter of fact /əz ə jɪmətəɹə fakt/  father and son /ˈfɑɹən sɔn/  over and above /əvərəndəbəv/
Although many categories and examples have been postulated, like assimilation, the problem of elision is that the operation of the elision process is unclear even to native speakers of English (=NSE) and this causes considerable confusion in spelling, Knowles (1987: 129) provides a suggestion towards the problems of re-spelling in dictation:

Some can be solved by substituting a word beginning with a vowel, in this way avoiding the consonant cluster. Is a six-year-old child ‘age six’ or ‘aged six’? If we change six to eight it becomes clear.

This could be a useful measure when listeners are allowed to have sufficient time. In most situations, however, this would be almost impossible in a normal rapid speech stream. Listeners would be unable to have the time even to recall that such a measure exists. Elision is most likely to occur in a rapid informal speech in English, and yet as Brown (1977) points out above, JLE do expect to hear these sounds which are elided in natural spoken English. The reason derives from how and with what they are trained to listen to in English. In the tapes that they listen to, the tape transcript is normally carefully examined, written and practiced before being recorded. But this is not the case in natural conversation at a natural speed. The speech rate on these educational tapes is often made artificially slow.

What JLE need in foreign language listening comprehension is clear recognition and a precise observation of phenomena in natural spoken English. Both language teachers and JLE must be able to recognise what really happens in spoken English with a natural speed. Without recognising the target or the object, one cannot achieve it.

1. 4. Liaison

Liaison is the linking of words, in particular when a word begins with a vowel. The sequence is often spoken without any noticeable break between two words:

\[
\text{an egg} \quad /\text{ən eg}/ \quad \text{becomes} \quad /\text{əneg}/
\]

Crystal (1985: 180) gives specific examples of a major liaison phenomenon of British English:

Liaison is one type of transition between sounds, where a sound is introduced at the end of a word if a certain context follows. It may be heard in English where a ‘linking /r/’ is often found in words ending with an r in the spelling, when they occur before words beginning with a vowel.

It should be noted that there are two kinds of /r/ links in liaison: linking and intrusive. According to Gimson (1989: 302-303):
Linking

i) Retention of word final post-vocalic /ɾ/  
   RP retains word final post-vocalic /ɾ/ as a linking from when the following word begins with a vowel:

   *far off* /far of/  
   *four aces* /foɾ ɛisɪz/  
   *answer it* /ænsər it/  
   *fur inside* /foɾ insaid/  
   *near it* /nɪər it/  
   *wear out* /weər aut/  
   *secure everything* /sɪkjʊər ɛvrɪˈʃɪŋ/  

   becomes /faɾ ɾəf/  
   becomes /foɾ reisɪz/  
   becomes /ænsəɾ ɾɪt/  
   becomes /foɾ ɾɪnsaid/  
   becomes /nɪəɾ ɾɪt/  
   becomes /weəɾ ɾaut/  
   becomes /sɪkjʊər ɾevrɪˈʃɪŋ/  

   As shown above, 'linking' occurs where a word ending with 'r' in its spelling is followed by or proceeds a word which begins with a vowel. He also introduces the notion of 'insertion' as the extended usage of /ɾ/ linking to where there is no final 'r' in the spelling:

ii) /ə/ ending and /ɾ/ linking

   *Russia and China* becomes /ˈraʃər æn tʃaɪnə/  
   *drama and music* becomes /ˈdraːmər æn mjuːzɪk/  
   *idea of* becomes /aɪdɪə əd/  
   *India and Pakistan* becomes /ɪndɪə ən ˈpækɪstən/  
   *area of agreement* becomes /ˈerər əv əˈɡriːmənt/  

iii) /ɔ/ ending and /ɾ/ linking

   *Shah of Iran* becomes /ʃaɾ əv ɪrən/  
   *the spa at Bath* becomes /θæ spɑː ət ˈbæθ/  

Since the comparative rarity of the cases of such intrusive /ɾ/ s tend to make RP speakers more aware of 'correct' forms, the last two cases in /ɔ/ ending and /ɾ/ linking such as *I saw it* and *drawing* are generally disapproved as the correct pronunciation of those who are well-educated.
These characteristics such as assimilations, elision and liaison which occur at word-boundaries illustrate an important fact: there is no simple one-to-one correspondence between the segments of a speech signal and the phonemes we hear. Gimson (1989: 306-307) provides advice to foreign learners from the viewpoint of weak forms:

He must ... observe the rules concerning weak forms, should cultivate the correct variations of word rhythmic patterns according to the context, and should make a proper use of liaison forms.

In any case, whether or not he uses such forms himself, he must know of their existence, for otherwise he will find it difficult to understand much of ordinary colloquial English. This knowledge is particularly important because a second language is often learned on a basis of isolated word forms: in the speech of the native, however, the outline of these words will frequently be modified or obscured, as has been seen.

The foreign learner is recommended to aim at a careful colloquial style of English in his own speech and, at the same time, to be aware of the features which characterise the rapid colloquial (or familiar) style he is likely to hear from native speakers.

Analysing and recognising these phonetic characters of word-boundaries may provide a part of the answer to the problems of JLE.

1. 5. The Background of the Research

Having considered the phenomena which occur at word boundaries, three reviews of juncture, segmentation, and IWB experiments are introduced as the background of this research. Firstly, Nakatani and Dukes (1977) researched the locus of segmental cues for word juncture. They used 128 phrases such as 'no notion / known ocean, stay dill / stayed ill, and play taught / plate ought'. The listeners, 15 high school students, had to respond to what they had heard by choosing from four possible choices at each trial. They drew the following conclusions about juncture cues:

i) a glottal stop and / or laryngealised voicing at the onset of a word-initial vowel was a cue for juncture.

ii) for / l / and / r /, a syllable-final allophone at the end of a word and a syllable-initial allophone at the beginning of a word were cues for juncture.

iii) differences in word duration that depended on whether a word did or did not end with a voiceless consonant was not a cue for juncture.
Secondly, Christie (1974) researched cues for syllable juncture perception in English. He employed nonsense sequences derived from /asta/ synthesised on a computer driven synthesiser system. The subjects, 44 phonetically native first-year students in linguistics, were asked to indicate on their test sheets whether they perceived the stimuli as /a sta/ or as /as ta/. No other response was allowed. He deduced that silence and aspiration both affect perception of the boundary location.

Thirdly, the Japanese linguist Shimizu (1979) confirmed Christie’s research results through his own research into IWB for JLE. He reported interesting research results on IWB. He prepared 30 pairs of phrases which were phonemically identical (See Appendix 1). His subjects were 107 Japanese native speakers (=JNS) who were first-year university students. They listened to the pre-recorded phrases read by a native American English speaker. Each phrase was repeated twice per phrase and the subjects were asked to mark which of the two phrases they had heard. The percentage shown in Appendix 1 is the percentage recognised correctly. For example, in No. 1, 73.4% subjects were regarded that they had correctly identified the word boundary. He found three key factors for JLE to identify word boundary:

i) aspiration:

When there is aspiration at the word boundary, the subjects showed high identification rate.

ii) pause length:

The longer the pause length is, the higher the identification rate is. This is clearly observed in the following chart, and

iii) strength of formant:

The stronger the formant aspiration at the word boundary is, the higher the identification rate is.

For instance, in the examples of ‘I scream / ice cream’ and ‘free Danny / Freed Annie’, the portion of successful identification was 73.4% and 98.9% respectively. As Shimizu suggests it seems to be true that the longer the pause is at the word boundary the higher the identification rate is.

<table>
<thead>
<tr>
<th></th>
<th>I scream [ai] [skriːm]</th>
<th>ice cream [aːs] [klɛrm]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>at pause s k rim</td>
<td>ar s pause k rim</td>
</tr>
<tr>
<td>length[^6]</td>
<td>317 342 167 42 475</td>
<td>150 258 258 84 325</td>
</tr>
<tr>
<td>strength[^7]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[^6]: Indicates the pause length.
[^7]: Indicates the formant strength.
[^8]: Indicates the aspiration strength.
Phonemically speaking, all the pairs tested above are described as identical. For example, 'get a board' and 'get aboard' are both described as /geta:ba:rd/. Theoretically speaking, they contain the same sounds and sound superficially alike. If so, by what means do ENS distinguish these differences? To answer this question, Gimson (1989) argues that the discriminating function of phonetic features in the continuum may be said to be associated with the type of juncture (close or open) which occurs with the relevant phonetic cues:

<table>
<thead>
<tr>
<th>open juncture</th>
<th>close juncture</th>
</tr>
</thead>
<tbody>
<tr>
<td>/pit/</td>
<td>/s/</td>
</tr>
<tr>
<td>/pits/</td>
<td>/pits/</td>
</tr>
</tbody>
</table>

Shimizu (1979: 76) also concludes that ENS can sense the small difference in aspiration in order to identify the word boundary almost instinctively, while JNS do not possess this ability and that ENS can almost always identify a word boundary such as 'no notion' and 'known ocean' without context.

Section 2  Aim, Method and Data Collection

2. 1. The Purpose of the Research

As mentioned above, Shimizu (1979) made two conclusions. However, in real life, a situation where listeners are asked to identify which of the two phrases they have heard without being given a context is rare. Of course, phonetic knowledge and information could be useful, but what seems to be more important to IWB is other linguistic knowledge such as syntax and cognitive ability to facilitate as predictive ability or inferential ability. For example, on a hot day when one says that he wants to eat something cold, then 'ice cream' might be more plausible rather than 'I scream' in a certain situation. Previous research regarding juncture, segmentation and IWB such as that by Jones (1956), Lehisite (1960), Hoard (1966), Christie (1974), Nakatani and Duke (1977) seem to focus on only the phone-
tic domain itself. Other linguistic factors such as syntax and the cognitive domain have not been well researched. Therefore, in this research, the phrase which subjects are supposed to distinguish will be given with and without a proper context. More details will be explained in Method.

The main purpose of this research is to extend Shimizu's research (1979), to answer the following questions and the hypotheses.

2. 2. Questions

i) Will a significant difference be observed with and without a context? If so,

ii) How far can JLE distinguish these differences?

2. 3. Hypotheses

i) In IWB, not only phonetic cues, but also other linguistic features such as syntactic and semantic clues could be deeply involved, therefore, there could be a significant difference between aural cues with and without a context.

ii) In IWB, not only linguistic knowledge, but also cognitive knowledge is employed, therefore, where inferential clues are given with the context, the rate of IWB could be higher even when the syntax of the phrases remain identical.

iii) The expected results are as follows (in order of decreasing competence in IWB):
   1. JNS with a context
   2. JNS without a context

2. 4. Subjects

Ninety six JNS: All first-year-university students who had studied English for at least six and a half years. Their TOEFL scores lie between 400 and 450.

Though the TOEFL score of the subjects in Shimizu's research in 1979 were not shown in his research, he claims that it was in the range of between 450 and 500.

2. 5. Method

Three kinds of paper-based tests were conducted. Each test took only several minutes. In each test, from 20 to 30 pairs were printed and they were all phonemically identical. The subjects were asked to listen twice per phrase to the pre-recorded phrases and sentences read by a native American English speaker. Then, they were asked to mark which of the phrase they thought they had heard. Before Test 3, all the subjects were asked turn over the previous two tests so that they had no chance to refer to their answers in Test 2. This was because the answers in Test 2 and 3 were exactly the same; the only difference was that Test 3 was given "with" a context. In this way, it was clear that the results of Test 3 were not interfered with by Test 2 or other way round. At the end of the test, the sheets were collected. Since the subjects were divided into three classes, the identical data collection
procedure was repeated three times.

2. 6. The Design of Materials
2. 6. 1. The Status of Test 1

Firstly, they were given Test 1 (See Appendix 2). Test 1 was an exactly the same as the experimentation in the research of Shimizu in 1979. Thirty pairs of phrases which are phonemically identical were presented in Test 1. The subjects listened to the pre-recorded phrases read by a native American English speaker. The phrase was repeated twice per phrase and the subjects were asked to mark which of the two phrases they had heard. The percentage shown in Appendix 2 are the percentage recognised correctly. For example, in No. 1, 62.5% subjects were regarded to have correctly identified the word boundary.

2. 6. 2. The Status of Test 2

Next, just like Test 1, 20 pairs of phrases which are phonemically identical were presented in Test 2 (See Appendix 3). The subjects listened to the pre-recorded phrases read by the very same native American English speaker of Test 1. Each phrase was repeated twice and the subjects were asked to mark which of the two phrases they had heard. The percentage shown in Appendix 3 are the percentage recognised correctly. For example, in No. 1, 97.9% subjects were regarded to have correctly identified the word boundary.

2. 6. 3. The Status of Test 3

Finally, Test 3 was designed to observe the difference with and without a context. The face validity of Test 3 is identical with Test 2. However, in Test 3, the subjects listened to 20 pre-recorded short conversations which consist of up to five sentences. In each conversation, the selected phrase was included. The conversation was repeated twice and the subjects were asked to mark which of the two phrases they had heard in the conversation. The conversations were read by the very same native American English speaker of the previous two tests and a Japanese female who have lived in the U. K. for several years. All the phrases to be answered were intentionally put into lines read by the native American English speaker, but not by the Japanese female.

Eleven examples, Group B (Question 1, 6, 8, 9, 10, 11, 12, 14, 15, 17 and 18) was designed to prevent the listeners from choosing the correct answer from syntactic clues. This was designed to test Hypothesis ii. Group A (Question 2, 3, 4, 5, 7, 13, 16, 19 and 20) was intentionally designed so that the subject could answer by inference from the context. For example in No. 4, these words ‘recording, music, and CD’ could function as inferential clues for the response ‘grey tape’.

The percentage shown in Appendix 4 are the percentage recognised correctly. For example, in No. 1, 85.4% subjects were regarded to have correctly identified the word phrase.
Section 3  Results and Discussion

i) Statistical Analysis of the Data

Firstly, the comparative graph between Shimizu (1979) and this research is shown in Appendix 5. The t-test was employed in order to analyse the gained scores as it is useful to analyse whether two independent sets of data show a significant overall difference in the magnitude of the variable. The null hypothesis is assumed that the test Shimizu gave in 1979 and this research show the very same result. The standard deviation (=S. D.) and Mean are 22.99 and 83.24 in Shimizu's research while they are 20.37 and 87.13 respectively in this research. The calculated value for t-test is 0.69. The 5% critical value for t with 30df (degree of freedom) is 2.04. In the case of the t-test, the calculated value should be larger or equivalent to the critical value in order to be statistically significant. The result is that the null hypothesis is accepted. This means that the result cannot be said to have evidence of its significance.

Next, the comparative graph between Test 2 and Test 3 is shown in Appendix 6. The t-test was also used. The null hypothesis is assumed that Test 2 and Test 3 show the very same result. S. D. and Means are 20.91 and 80.59 in Test 2 while they 21.12 and 79.53 respectively in Test 3. The calculated value for the t-test is 0.02. The 5% critical value for t with 20df is 2.09. In the case of the t-test, the calculated value should be larger or equivalent to the critical value in order to be statistically significant. The result is that the null hypothesis is accepted. This means that the result cannot be said to have evidence of its significance.

Finally, in Test 3, Group A (Question 2, 3, 4, 5, 7, 13, 16, 19 and 20) had the questions which requested not only linguistic knowledge, but also cognitive knowledge. In these nine questions, inferential clues were given with the context while Group B (Question 1, 6, 8, 9, 10, 11, 12, 14, 15, 17 and 18) required simply linguistic knowledge. So the null hypothesis is assumed that Group A and Group B show the very same result. S. D. and Means are 23.73 and 78.49 in Group A while they are 22.24and 80.96 respectively in Group B. The calculated value for the t-test is 0.16 in Group A and 0.17 in Group B respectively. The 5% critical value for t with 7df and 10df are 2.36 and 2.23 respectively. In the case of the t-test, the calculated value should be larger or equivalent to the critical value in order to be statistically significant. The result is that the null hypothesis is accepted. This means that the results cannot be said to have evidence of its significance.

In the next section, the discussion and conclusion based on this statistical data and analysis will be discussed.

3. 2. Discussion and Conclusions

In this section, the meaning of previously presented data will be discussed. Now let us recall the hypotheses and questions again:
Identification of Word Boundaries (上田) 69

〈Questions〉
i）Will a significant difference be observed with and without a context? If so,
ii）How far can JLE distinguish these differences?

〈Hypotheses〉
i）In IWB, not only phonetic cues, but also other linguistic features such as syntactic and
semantic clues could be deeply involved, therefore, there could be a significant difference
between aural cues with and without a context.

ii）In IWB, not only linguistic knowledge, but also cognitive knowledge is employed,
therefore, where inferential clues are given with the context, the rate of IWB could be
higher even when the syntax of the phrases remain identical.

iii）The expected results are as follows (in order of decreasing competence in IWB):
1. JNS with a context
2. JNS without a context

Firstly, for Hypotheses i, iii, and Question i, from the calculated data in the previous
section, it is clear that there was no significant difference between these items with and
without a context in Test 2 and Test 3. For example, some reverse phenomena were
observed, which means that in some questions, the results were better in Test 2, without a
context such as Question 1, 2, 7, 8, 10, 12, 14, and 16. It might be an indication that the
contexts were simply obstacles rather than supportive hints. All of the subjects’ TOEFL
scores lie between 400 and 450 and no more than 450. At this level, it is assumed that the
subjects are unable to activate a context as cognitive knowledge.

Secondly, for Hypothesis ii, from the calculated data in the previous section, it is clear that
there was no significant difference between Question 1 to 11 and Questions 12 to 20 in Test
3. As mentioned above, all of the subjects’ TOEFL scores lie between 400 and 450 and no
more than 450. At this level, it might be possible to assume that the subjects are unable to
use those inferential clues in order to lead the correct answers.

In a theory of Applied Linguistics, it is said that at any level, top-down and bottom-up
processes are activated; the efficiency stems from the fact that all relevant sources of in-
formation are used at the same time to provide rapid word identification. However, through
my teaching experience mainly to JNS, it is quite doubted whether top-down and bottom-up
processes are activated simultaneously especially at the level of typical Japanese first-year
university students. As a conclusion of this research, though it has a number of limitations,
it might possibly be true that cognitive knowledge and interactive models are not well
activated regarding word identification at a lower level of learners. Eysenck and Keane (1992 :
300) also suggest the point that word recognition is only one aspect of sentence compre-
hension, and it remains to be seen whether interactive models can be applied successfully to
larger units of spoken speech.
In the next section, a number of limitation in this research will be disputed.

3. 3. Limitations of the research

The origins of this research are Shimmizu's study in 1979. In this research, however, some original and unique attempts were made. Firstly, Test 3 was designed in order to observe the effect of the existence of contexts by embedding the phrases in conversations.

Secondly, Group A (Question 2, 3, 4, 5, 7, 13, 16, 19 and 20) in Test 3 was specially designed in order to observe whether inferential clues help the immature ability to identify word boundaries. Nevertheless, there were several limitations to this research that limited its validity as an extension of Shimmizu's study.

i) Different conditions from Shimmizu's study

ii) Limited number of the subjects

iii) Limited levels of language proficiency

iv) Item Analysis was not employed.

Firstly, though this study was intended as an extension of Shimmizu's research in 1979, of course, the subjects, the level of their TOEFL scores and the material used were all different. Due to the fact that the very same tape that Shimmizu used in his study was not available at the time of this study, a newly recorded Mini Disk by an American native speaker was used. In both of the recorded materials, the very same words were read by a native American English speaker, but the speed, pitch, tone and so on were of course not identical.

Secondly, there was the limitation of the number of subjects. Only 96 were available. In order to pursue statistical validity and confident conclusions, the more the better.

Thirdly, there was a limited level of proficiency. Since the subjects were categorised between 400 and 450 in TOEFL, different results might have been observed at a higher or lower level.

Lastly, Item Analysis was not employed. Though all questions were carefully designed, it is doubtful whether every single question was actually testing what it supposed to test since in some questions, the results without the contexts or the inferential clues scored higher.

In the last section, based on these limitations, points for future research will be suggested.

3. 4. Future prospects of the research

Due to the limitations mentioned above, it should be stressed that the gained conclusions and results possess some degree of inaccuracies. As future prospects for this research, the following elements could make the conclusions a little clearer.

i) More subjects

ii) Wider variety of proficiency levels

iii) Introduction of Item Analysis
Speaking ability is always the top one, which JLE would like to gain, but listening should come before speaking. By speaking, surely one can express what s/he would like to but that is not a real communication. Without listening, one cannot communicate with others.

1) The smaller signal detectability is, the more weakly the component is learned.
2) Correction is added by M. Ueda / prəbl / (sic)
3) Correction is added by M. Ueda / febr / (sic)
4) Correction is added by M. Ueda / lar and oldr / (sic)
5) RP=Received Pronunciation.
6) length of segment
7) formant strength of segment
8) prominent fall in pitch (−)
9) prominent rise in pitch (+)
10) vowel
11) juncture
12) consonant

References


Picket, J. M. & Pollack, L. (1963). ‘Intelligibility of excerpts from fluent speech; effects of rate of


Appendix 1: A test by Shimizu (1979)

1. a) I scream  b) ice cream [3]  73.4%
2. * a) an aim  b) a name  68.1%
3. * a) no notion  b) known ocean  94.7%
4. a) you thread  * b) youth read  64.9%
5. * a) grey tape  b) great ape  24.5%
6. a) white shoes  * b) why choose  98.9%
7. * a) white shoes  b) why choose  98.9%
8. * a) freed Annie  b) free Danny  98.9%
9. * a) see zoos  b) seize ooze  92.6%
10. a) be quiet  * b) Beek Wyatt  96.8%
11. * a) be quiet  b) Beek Wyatt  95.7%
12. * a) it sprays  b) it’s praise  36.2%
13. a) up our  * b) a power  95.7%
14. a) get a board  * b) get aboard  20.2%
15. * a) pirates of Anna  b) pirate Savannah  89.2%
16. a) scrappy racer  * b) scrap eraser  96.8%
17. a) run dwarf  * b) ruined wharf  98.9%
18. * a) grasp lot  b) grass plot  91.5%
19. * a) mask ash  b) mass cash  98.5%
20. a) mash ash  * b) mass cash  100.0%
21. * a) a nice man  b) an ice man  76.6%
22. a) grade A  * b) grey day  80.9%
23. a) hoe-maker  * b) home-acre  96.8%
24. a) keeps ticking  * b) keep sticking  43.6%
25. * a) night rate  b) nitrate  93.6%
26. a) night rate  * b) nitrate  93.6%
27. * a) plum pie  b) plump eye  96.8%
28. a) seal eyeing  * b) see lying  93.6%
29. a) seem able  * b) see Mable  94.7%
30. * a) two lips  b) tulips  92.6%

[3] Correct answer
### Appendix 2: Test 1 by Ueda (2002)

<table>
<thead>
<tr>
<th></th>
<th>a) I scream</th>
<th>b) ice cream</th>
<th></th>
<th>a) an aim</th>
<th>b) a name</th>
<th></th>
<th>a) no notion</th>
<th>b) known ocean</th>
<th></th>
<th>a) you thread</th>
<th>b) youth read</th>
<th></th>
<th>a) grey tape</th>
<th>b) great ape</th>
<th></th>
<th>a) white shoes</th>
<th>b) why choose</th>
<th></th>
<th>a) freed Annie</th>
<th>b) free Danny</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>*</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>*</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>*</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>*</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>*</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>*</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>*</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>*</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>*</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>*</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>*</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>*</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>*</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>*</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>*</td>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

14) Correct answer
Appendix 3: Test 2 by Ueda (2002)

Listen to the tape. The phrase is repeated twice. Circle the phrase which you think you have heard.

1. an Asian * a nation 97.9%
2. * nitrate night-rate 57.3%
3. grey day * grade A 99.0%
4. great ape * grey tape 93.8%
5. it’s praise * it sprays 25.0%
6. * keep sticking keeps ticking 82.3%
7. I conceal it. * I can seal it. 90.6%
8. * an aim a name 100.0%
9. * see zoos sees ooze 93.8%
10. free Danny * freed Annie 99.0%
11. * roof style roof’s tile 69.8%
12. * in formal terms informal terms 52.1%
13. iced ink * I stink 89.6%
14. why choose * white shoes 97.9%
15. cow slips cowslips 47.9%
16. * iced eel I steal 91.7%
17. * that’s come that scum 93.8%
18. * ice cold I scold 70.8%
19. * no notion known ocean 86.5%
20. get aboard * get a board 72.9%

15) Correct answer
Appendix 4: Test 3 by Ueda (2002)

Listen to these short sequences which consist of up to three sentences. Then, circle the phrase which you think you have heard. Each sequence is repeated twice.

1. an Asian   *  a nation  85.4%
2. * nitrate  *  night-rate  45.8%
3. grey day   *  grade A  99.0%
4. great ape  *  grey tape  99.0%
5. it’s praise *  it sprays  46.9%
6. * keep sticking  *  keeps ticking  84.4%
7. I conceal it. *  I can seal it.  74.0%
8. * an aim  a name  95.8%
9. * see zoos  sees ooze  100.0%
10. free Danny *  freed Annie  97.9%
11. * roof style  roof’s tile  84.4%
12. * in formal terms  informal terms  28.1%
13. iced ink  *  I stink  89.6%
14. why choose  *  white shoes  70.8%
15. * cow slips  cowslips  54.2%
16. * iced eel  I steal  65.6%
17. * that’s come  that scum  99.0%
18. * ice cold  I scold  90.6%
19. * no notion  known ocean  89.6%
20. get aboard  *  get a board  90.6%

---

16) Correct answer
Appendix 5: A comparative graph of Shimizu and Ueda

Appendix 6: A comparative graph of Test 2 and Test 3
Appendix 7: Tape transcript of Test 3

Listen to these short sequences which consist of up to three sentences. Then, circle the phrase which you think you have heard. Each sequence is repeated twice.

1. A. Now, question time. What do you call a particular large group of human beings living in one area, organised in one way, and speaking a common language?
   B. A nation?

2. A. Please be careful in the chemistry laboratory.
   B. Yes, there are many dangerous things.
   A. Just be very careful when handling nitrate.

3. A. I'm so worried about my examination results.
   B. Don't worry. You know you always get grade A.

4. A. You know that I was recording music from CD to tape here this morning.
   B. Yes, what's wrong?
   A. I can't find the grey tape.

5. A. Do you know how to use this insect killer?
   B. Yes, hold like this and press here. Then it sprays automatically.

6. A. So what is your advice in this case?
   B. Well, if I were you, I'd keep sticking to the original plan.

7. A. What shall we do with that crack in the wall?
   B. Well, I can seal it with grout.

8. A. I think John’s lost his motivation.
   B. You reckon so?
   A. Yes, he said he hasn’t got an aim in life.

9. A. What do you do when you have some free time?
   B. Well, I like to see zoos because I love animals.

10. A. What happened to the hostages?
    B. They freed Annie.

---

17) The italic parts were read by the native American English speaker. The underlined words are the answers.
11. A. Do you happen to know where James lives?
    B. In Coronation street. You can't miss his house: the roof style is quite unique.

12. A. American people sometimes say 'Let's shoot some bucks!'.
    B. What does it mean?
    A. In formal terms, it means 'Let's gamble some dollars'.

13. A. You must have eaten a lot of garlic! Oh, what a smell!
    B. Really? Do I stink so badly?

14. A. I went to Alaska to visit my Eskimo friend.
    B. Did you enjoy the food?
    A. Yes except one thing: iced eel.

15. A. OK, I'm taking your cow out to pasture now.
    B. Be careful around the top of the hill. It's very slippery.
    A. Oh, yes. My cow slips there as well.

16. A. Did you know that Ann's getting married next week?
    B. Yes, but she's so casual about it that she hasn't bought anything to wear on her feet. She's gonna borrow my new white shoes.

17. A. Would you let me know if you receive a huge brown parcel for me? We're expecting a large brown parcel.
    B. That's come already.

18. A. I'm so thirsty.
    B. Drink this coke, it's ice cold.

19. A. He has never arrived on time.
    B. I know.
    A. He has no notion of time.

20. A. Oh, no! Our front window's been smashed and the rain's pouring onto the carpet!.
    B. What shall we do?
    A. We'll have to cover the hole. Get a board quickly!