JAPANESE ESL STUDENTS' USE AND PERCEPTION OF ENGLISH LINKING AND INTRUSIVE SOUNDS.

A thesis presented in partial fulfilment of the requirements for the degree of

Masters of Arts

in

Second Language Teaching

at Massey University, Albany, New Zealand

GLENN WICKINS
2016

Abstract

This study investigates the use of linking and intrusive sounds in English by Japanese students studying university preparation English as a Second Language (ESL) courses at a tertiary institution in New Zealand. Such students covet a native-like accent, which in part comes from the ability of native English speakers to resolve the interruptions in the constant stream of language that they are producing. Producing a constant stream of English sounds more "natural" and "fluent". Native English speakers have options to help them do this, which are called *sandhi*. This study focuses on six such features of connected speech: linking /j/, /w/ and /r/ sounds which are only pronounced when the following word begins with a vowel; and the intrusive versions of the same three sounds used to remove hiatus (interruptions in the flow of speech) between two vowel sounds. The purpose of this study is to investigate potential links between the usage of sandhi techniques and the current proficiency level of learners to see how teachers can best encourage the use of these "natural-sounding" features. By collecting speech samples from 38 participants and comparing them to the linking patterns of native English speakers as reported by other studies, this study endeavours to draw some conclusions about the usage of sandhi in Japanese ESL students.

Contents

Table of Figures

Chapter One:

Introduction

The term *boundary* was first used to define the limits of spoken words in the 1930s by Trubetzkoy (1969). Since then, academics and teachers have had increased appreciation for how different the written form of English (with spaces separating the text into lexical units) is from the stream of continuous language that is spoken English. It is easy to understand how learners of English have trouble listening to English when they cannot tell where one word ends and another begins.

The way in which English is written onto the page can mislead learners. Wong (1987) states that learners of English often learn with their eyes rather than their ears and this leads to learners believing that words should be pronounced as they appear on the page. They believe that words should be separated by blank spaces. This, however, is not the way that native speakers converse.

Some argue that the difference between the way that a native English speaker converses and the way in which a learner converses could potentially

define the gap which a learner needs to close in order to be considered *fluent*. This introduction will examine one particular challenge that learners face in mimicking the speaking style of a native speaker: the connection of words into fluent speech.

As seen in the literature review, there is some research into the area of sandhi usage by ESL learners (Alameen, 2007; Hieke, 1984), but I believe that more research in this area can help ESL learners to more confidently produce English and understand the English of the native speakers around them.

1.1 Definitions

The terms used in this study have been defined many different ways by many different studies. Therefore, in order that the meanings that I intend to assign to these terms are clear, it is best to start by defining the terms *hiatus, sandhi, linkers and intrusion* for the purposes of this study. In the second chapter, sandhi and rhoticity will be examined further by means of existing studies.

1.1.1 Hiatus defined

English speech consists of a string of syllables which are quite flexible in their nature. An English syllable must have a vowel sound as a nucleus (the core of the syllable). However, a syllable may or may not have an onset (a consonant or consonant cluster sound before the vowel) and it may or may not have a coda (a consonant or consonant cluster sound after the vowel). Having the option to omit onsets and codas creates the possibility that two vowel sounds may exist in succession.

When two vowel sounds exist in succession, the speaker needs to pause in order to create an audible break between the two vowels and/or transition the

articulating organs (for example the tongue) to the new position. This break, known as *hiatus*, interrupts the flow of speech and is seen by some as a hindrance to the flow of natural speech. Native English speakers often use techniques called *Sandhi* in order to resolve such issues when they speak. By introducing consonant sounds between two vowels (intrusion) or by pronouncing a consonant sound that is usually silent (linking), speakers of English can avoid the need to interrupt their stream of language.

Hiatus comes from Latin and means "an opening, crevice" (Bussmann, 1996; p. 206). According to Crystal (2003), two vowels are said to be *in hiatus* when two adjacent vowels belong to different syllables. This can occur both within a word (internal hiatus) such as in the word "hiatus" itself (an intrusive /j/ sound is inserted between the /aɪ/ and the /eɪ/ sounds). Hiatus can also occur across word boundaries (external hiatus) when the terminal syllable of the preceding word ends with a vowel sound and the initial syllable of the following word begins with a vowel sound (such as the /j/ sound that often appears between the words in the utterance "see it").

Bussman's (1996) definition of hiatus focuses on the fact that the two vowels are *heterosyllabic* (each vowel belongs to a different syllable) *monophthongs* and that the gap between them can be perceived by the human ear. This study has no need for the distinction between monophthongs (single sound vowel sounds such as /e/) and diphthongs (combinations of two vowel sounds such as /eɪ/). It is enough to take from this definition that the focus is on the perceivable gap between two adjacent vowel sounds from two different syllables. Bussmann goes on to state that in the case of hiatus, a semivowel (/j/ or /w/) may be inserted to resolve the hiatus.

R. L. Trask (1996) concurs with the definitions above that *hiatus* occurs when two consecutive vowels form two separate syllables. So, for the purposes of this report, *hiatus* shall be defined as the gap created when two consecutive vowels form two separate syllables. The term *in hiatus* will refer to the situation of the two adjacent vowels. Trask goes on to state that such vocalic combinations tend to be "unstable" (p. 170). This refers to the fact that not even native English speakers produce sandhi consistently.

1.1.2 Sandhi defined

The word *sandhi* is derived from Sanskrit and loosely translated, it means, "put together". It is the way that native speakers connect their speech in order to create the constant stream of language that is spoken English. In order to develop a deeper understanding of sandhi, this report will now examine several definitions of sandhi given by various authors.

Bloomfield (1933) defines *sandhi* to include the following: assimilation, mutation, contractions, liaison and elision. This is a definition broader than is required for the purposes of the current study. In order to examine two features of pronunciation in more detail, this study does not include mutation, elision, or assimilation. Contractions will be briefly mentioned during the analysis of the data so it is important that contractions remain a part of the definition of *sandhi* for this study.

Brown and Miller (1988) define sandhi as "a general term ... applied to phonological modifications that occur between juxtaposed forms." The Routledge Dictionary of Language and Linguistics gives a similarly vague definition: "... merging of two words or word forms and the resulting systematic

phonological changes" (Bussman 1996, p.413). The commonality between these definitions is that sandhi is a feature of phonetics; it is a modification of sounds in certain conditions and it covers a broad range of features that we will later look at more specifically.

The most simplistic yet precise definition is from Henrichson (1984), who states that sandhi is simply the difference between written and spoken text. Sandhi is one way in which native speakers subconsciously judge the abstract concept of *naturalness* of the speech of learners.

In this work, sandhi will be the umbrella term covering both intrusives and linkers. Before delving into definitions of intrusives and linkers, we shall first look at the six phonological features that these terms define.

1.1.3 Six phonological features

This study will look at six phonological features which are categorised as three linking sounds and three intrusive sounds. The six phonological features are as follows.

- /j/ sandhi   (linking /j/ and intrusive /j/)
- /w/ sandhi   (linking /w/ and intrusive /w/)
- /r/ sandhi   (linking /r/ and intrusive /r/)

The linking /j/ is the 'Y' sound in "say a", the intrusive /j/ is the 'Y' sound that often occurs in "he eats". The linking /w/ is the 'W' sound in "how a", while the intrusive /w/ is the 'W' sound that often occurs in "go out". Finally, the linking /r/ is the 'R' sound in "far and" and the intrusive /r/ is the 'R' sound that often occurs in "saw a". The main difference between a linker and an intrusive is

that a linker is represented in the written form of the utterance. You can see that "say a" has the letter 'Y', "how a" has the letter 'W' and "far and" has the letter 'R'.

1.1.4 Linkers defined

The next terms to define are *linking, linking sounds* and *linkers*, which are synonymous and are therefore used interchangeably throughout this study. A definition of *linking sounds*, provided by David Crystal's A dictionary of Linguistics and Phonetics (2003) is, "...a sound which is introduced between linguistic units, usually for ease of pronunciation" (p. 274). According to this definition, intrusive sounds are actually a subset of the larger set called *linkers* since intrusive sounds also fall within the bounds of this definition. Intrusive sounds are sounds which are introduced between linguistic units (words or syllables) and are introduced for the purpose of making pronunciation easier.

For the purposes of this project, however, the terms *linkers* and *intrusives* will be defined as two exclusive sets. This is to ensure clarity by eliminating the overlap. Linkers will herein be defined as sounds that are usually or often dropped unless the following word begins with a vowel. Intrusives are therefore sounds that are not present in the individual words and do not appear in the written form of the utterance. Therefore, the utterance "be in" may contain a /j/ sound which, according to Crystal's definitions is an intrusive sound and, by definition therefore a linker as well. However, by the definitions of this project, linkers and intrusives are two separate sets and sounds will be labelled by intrusives if possible and linkers if not. So the /j/ in "be in" is an intrusive but not a linker. *Linkers*, *linking sounds* and *linking* shall be used to refer to the /j/, /w/, and /r/ sounds that are written but only vocalised before a vowel.

R. L. Trask (1996) gives a definition similar to Crystal's stating that a *linking sound* is a sound which is present in specific environments in connected speech, but linking sounds are absent in the words when pronounced in isolation. This definition gives the condition of specific environments which is worth a quick mention here. Linking sounds occur *intervocalically*; this means that they occur between two vowel sounds. The purpose of recovering the sound of the silent consonant is to resolve the issue of *hiatus* which is discussed below. To illustrate, the letter 'Y' in the word "they" is effectively silent. The utterance "They eat" would be phonetically transcribed as /ðeɪ/ and /iːt/ in isolation. Notice that "they" ends in a diphthong. However, when spoken in connected speech, the utterance becomes /ðeɪjiːt/. A /j/ sound has been inserted between two vowel sounds in order to make the utterance smoother.

The term *liaison* could be seen as a synonym for *linker*. According to the Routledge Dictionary of Language and Linguistics by Bussman (1996), liaison is a pronunciation rule in French whereby a "normally silent" (p. 281) final consonant is articulated when the following word starts with a vowel. In the context of this project, that would incorporate the linking /r/ in "here I…" / hɪə (r)aɪ/ but not the intrusive /r/ in "saw a" /sɔ:(r)ə/ as *normally silent* implies that the sound is represented with a letter but that letter is not normally pronounced (i.e. it is orthographically represented). To avoid confusion, the term liaison will be avoided in this study.

1.1.5 Intrusion defined

Crystal (1992) defines intrusion as, "The addition of sounds in connected speech which are not heard when words or syllables are said in isolation"

12

(p.194). An example of this is the /w/ sound that appears in the utterance, "to a" (often pronounced /tuːwə/). This definition is not complex enough to distinguish between linkers and intrusives. In his fifth edition (Crystal, 2003), he goes on to give the example of intrusive /r/ saying that there is no historical justification for the /r/ sound to appear in utterances such as "law and order" (often pronounced /lɔː rænd ɔːdə/). There is not, nor has there ever been a letter 'R' in these words. So even though Crystal's definition fails to distinguish between intrusive and linking sounds, he does so in the examples he gives.

R. L. Trask (1996) gives a more precise definition of *intrusion*. Trask states clearly that intrusion is where a sound is added to an utterance without *etymological justification*. This means that not only the present spelling of the words in question but also the historical spelling of the words need to be examined to ensure that the sound which is injected did not exist in the past and has been dropped in more recent spellings of the terms.

This project will adopt Trask's definition of *intrusion* but will perhaps use it slightly more specifically. Intrusion is where a sound has been added between two vowels in order to make pronunciation smoother and the etymology of the terms does not justify its presence. The example above of the /j/ sound in "be in" and the /w/ sound in "to a" are both examples of this. *Intrusives* and *intrusive sounds* shall be used interchangeably throughout this report to refer to the /j/, /w/, and /r/ sounds.

Trask (1996) also provides a definition for the classic example of an intrusive sound which is the intrusive /r/. Intrusive /r/ is the /r/ sound that often occurs in the utterances "law and order" (Crystal's example from above) and "saw a" (often pronounced /sɔːrə/). He states that the intrusive /r/ is

13

automatically inserted after any of the following vocalic sounds /ɑː/ (as in "star"), /ɔː/ (as in "store"), /ɜː/ (as in "sir"), /ə/ (as in "teacher"). He adds that intrusive /r/ also appears after centring diphthongs which include /ɪə/ (as in "here"), /eə/ (as in "hair"), /ʊə/ (as in "tour"), and /ɔɪ/ (as in "boy").

1.1.6 Definitions summarised

In short, this study will look at how Japanese learners use the following features. First is *linking sounds or linkers*. These are the /j/, /w/ or /r/ sounds that are included in the written form of the utterance, but they are usually silent unless the next sound is a vowel sound, in which case, they are pronounced. Next is *intrusives*. These are the /j/, /w/ or /r/ sounds that are not included in the written form of the utterance but are added in order to resolve the hiatus.

Sandhi is a collection of many different such techniques to smooth out the language so that it requires less effort on the part of the speaker. For this study, many of those techniques are beyond the scope of this study, so sandhi will be used as a collective term to include linking and intrusion.

The verb *link* will be used with linking sounds and the verb *connect* will be used for intrusion.

1.2 Japanese learners' issues with sandhi

One factor that impedes learners' ability to comprehend and produce native sounding speech is the group of techniques used by native speakers to smooth the words into a continuous flow of speech, called *sandhi*.

The occurrence of sandhi in native speech is by no means consistent. According to Wells (1982) the occurrence of /r/ sandhi across word boundaries depends not only on speaking style but also other "random" factors. These factors are discussed later in the study. Brown (1988) concurs, adding that intrusive /r/ is more frequently realised in "fluent colloquial style" (p.145). Using the frequency of linking in native speech as a baseline, this study will investigate the usage of linking and intrusive sounds in speech samples from Japanese ESL students studying degree preparation ESL courses at a tertiary institution in New Zealand.

Sandhi could be seen as a distortion of the vernacular to a learner who hears sounds that are not meant to exist in a particular word combination. In addition, the inconsistency of the usage of sandhi may leave the learners wondering if they heard an error or misheard.

Resolving hiatus is assumed to be something that learners naturally "pick up" as they gain confidence and familiarity with the language. However, this simply may not be the case. There are many possible reasons why a learner might have difficulties in acquiring the correct usage of sandhi. Below is a brief look at just a handful of the most influential issues in developing the habit of sandhi.

According to Swan and Smith (2001), errors in pronunciation occur when the mother tongue has no equivalent or when the equivalent in the mother

tongue is similar but not alike. When looking to explain why the Japanese learners fail to produce sandhi like their native speaking counterparts, the differences between the Japanese language and English must be investigated. Japanese uses a syllabary not an alphabet, which means that words are made up from a group of prescribed syllables. With only one exception, consonants do not exist alone in the Japanese language; they are always coupled with a vowel (for example, KA, TE or SU).

Ohata (2004) demonstrates with a comparison of syllable structures where C is a consonant and V is a vowel. In English syllable structures can be varied (see-CV, sit-CVC, spit-CCVC, spits-CCVCC, sprint-CCCVCC). In Japanese, with the one exception of the character ん (/n/ sometimes pronounced /m/), all words are made in the following structure (ke-CV, kare-CVCV, kakureru-CVCVCVCV). YA や, YU ゆ and YO よ exist, but YI and YE do not. This could lead to difficulties with Japanese learners applying a /j/ to resolve a hiatus where the following sound is /ɪ/ or /e/. If this is a real complication for Japanese learners, one would expect learners to correctly liaise in the case of "play a" more often than they would "play in". Similarly, わ WA, ゐ WI, ゑ WE, を WO exist, but WU does not. This may hinder the use of linking /w/ before /uː/. In summary, students are more likely to produce combinations of sounds that exist in their own language than those that do not.

A similar concern that was discussed in Heselwood (2009) was the relationship between the lexical glides (/j/ and /w/ sounds used within words) and the epenthetic glides (/j/ and /w/ sounds that are added to resolve hiatus). Because the sounds employed to create speech differ from language to language,

not all sounds that exist in English exist in other languages. Uffman (2007) suggests that if a language does not employ a /j/ sound to make words, then it is unlikely that same language would employ the /j/ sound as a hiatus resolution strategy. Ladgefoged and Maddieson (1996) reveal that 15 percent of languages do not employ a lexical /j/. However, Heselwood uses research by Mahootian (1997) and Bijankhan (2005) to demonstrate that Persian is one language that uses an epenthetic sound without employing its lexical counterpart. This suggests that although it is possibly more difficult, Japanese leaners may be able to adopt epenthetic glides (intrusive /y/ and /w/) before a vowel sound even if the do not have that consonant and vowel combination in their own syllabary.

Another issue that might complicate learners' development of intrusive sandhi is not knowing which sound to insert. Hay and Sudbury (2005) concede that it might not be as obvious as one would assume which sound needs to be inserted when there is no orthographic clue to follow. So, while it seems perfectly obvious that the hiatus in "play a" would be resolved using /j/, it is not so obvious that the hiatus in "draw a" would be resolved with an /r/.

The intrusive /r/ is a rather different case to the other types of sandhi mentioned above because of its elusive origins and indeed it may not be immediately apparent that an /r/ sound is the sound required to resolve the hiatus. Take for example the utterance "law of" a native English speaker may insert an /r/ sound to connect the /ɔː/ and /ɒ/ sounds. However, there are no orthographic clues to suggest an /r/ sound, so in the absence of articulatory clues, how is a learner of the language to know which sound they need to insert?

Japanese students could have issues producing a particular sound like a native speaker. The likely candidate here is /r/ as Japanese ESL learners struggle

with the distinction between /l/ and /r/. In the Japanese section of his book, Learner English, Michael Swan (2001) states that Japanese ESL students have difficulties differentiating between /l/ and /r/ and instead produce the Japanese /r/ (a flap like a short /d/). The Japanese /r/ sound has a range of allophones (variations of the sound) which range from something like an /l/ sound to something like an /r/ sound depending on the class of the vowel that it precedes (Collins and Mees, 2013). Swan goes on to mention that they also have difficulties with linking /r/ because of their habit of adding a *rounding off* vowel to words that end in a consonant sound. So "there are" may be pronounced /ðɪəru: ɑ:/ instead of /ðɪə rɑ:/. This is likely to be a result of interference as the Japanese write, and to some extend speak, in syllables as opposed to creating syllables with letters as in English.

One final complication that learners may face in acquiring the correct use of linking is an affective one. It is possible that there is a stigma in using linking as learners may potentially see it as "lazy" or "incorrect". It is known that there is a stigma amongst native English speakers concerning the intrusive /r/ (Hay & Maclagan, 2012). Intrusive /r/ is often regarded negatively in cases such as "saw a" or "drawing". In the same way, it is possible that learners see the intrusion of /w/ in "do it" as a sound that should not be present. So the perspective of the learners may influence the token rate of the linking.

In summary, the issues that any learner may have in acquiring the skill of sandhi could be categorised as pronunciation, fluency, affective factors and knowledge gap. Pronunciation issues could stem from the alphabet or syllabary of their mother tongue not containing the required sound or sound combination, or perhaps the sound exists as a lexical sound but not as an epenthetic sound in

their language. Lack of fluency may be an issue given that pausing to consider content may or may not interrupt word connection. Affective factors mean that the learner may view sandhi as incorrect or too informal for the situation. Finally, the learner may not know which consonant sound is required.

This section has examined the issues that a Japanese speaker faces in learning to speak English with fluency. It should be noted that ESL classes often include students from a variety of backgrounds. Each of these backgrounds brings with it different challenges.

It is believed that by understanding more about learners' usage of sandhi, the approaches employed by teachers in helping their students to speak more naturally can be improved. In particular, the question of whether linking should be directly taught at all is an interesting query, but it is unfortunately beyond the scope of this project. However, it is hoped that the third research question (see below) goes a little way toward determining whether sandhi increases as a direct result of formal education or not.

1.3 Why speakers resolve hiatus

Why do speakers need to resolve hiatus? Speakers of English link to make speech more fluid and less complex to articulate (Hieke, 1984). As mentioned earlier, the use of sandhi is not consistent and indeed Hieke (1984) goes on to state that other factors also affect the usage of linkers in speech and that they include the formality of register and the pacing of speech. So the use of sandhi is more common in informal spoken English (Pennington & Richards, 1986; Richards, 1983).

A series of experiments carried out in 1998 by Derwing and Munroe suggest that native English speakers' comprehension of learners' utterances often improves when learners increase the rate of their speech. It is suggested that the slow rate of production leads to issues such as over-enunciation (and therefore non-standard word stress) and broken, not connected, speech (leading to non-standard sentence stress). A distortion in the stress patterns of a sentence may lead to a disruption in the word connection in the participants of the current study.

1.4 Glottal stops

Hiatus, when not resolved leads to the use of a glottal stop. A glottal stop is a momentary blockage of the glottis, which is an aperture between the vocal cords (Crystal, 2003). A glottal stop is seen by phonologists as an unnecessary interruption in the flow of speech. Resolving hiatus leads to faster, smoother production. However, Hieke (1984) suggests that the glottal stop is retained by native speakers under the following conditions:

1. In deliberate speech for reasons of extra clarity

2. Where special stress assignment overrides absorption phenomena

3. Where plus juncture is phonemic and thus obligatory

4. After silence (with no prior syllable to draw on)

The first point here is fairly self-explanatory, when people speak slowly and enunciate purposefully, they tend to leave a gap between the syllables for clarity. The second point allows for extra stress on words to alter the meaning of a sentence. An example would be, "I said he IS going." A glottal stop before "IS"

further isolates the stressed word to clarify with a listener who may think the speaker meant that he was not planning to go.

The third point allows for obligatory pauses in utterances that help to divide sentences into clauses, nestled clauses and interjected segments. An example of an interjected segment is "He, I think, is a genius." The pauses before and after, "I think" distinguish it from the main clause in order to avoid the confusion of running two pronouns together. Finally, point four suggests that a silence should disrupt word connection and prevent the intrusive consonants from occurring.

1.5 Epenthetic consonants and their preceding vowels

The vowel environment dictates which sound will appear in order to resolve a hiatus. A /j/, /w/ or /r/ sound will be required depending on the class of the vowel (high-front, high-back, mid or low class) that precedes it. In the case of linkers, an additional clue exists to which sound will be required to resolve the hiatus. The written form of the utterance contains a consonant which is silent. This silent consonant is usually pronounced as a linker. However, the class of the consonant overrules the written form of the utterance in cases such as "saw a" where the class of the preceding vowel dictates an /r/ sound is to be used but the written form of the utterance contains a silent 'W'.

If the preceding vowel is a high-front vowel (such as /i:/), then a /j/ sound will be required to resolve the hiatus. A high-back vowel needs to be followed by a /w/ sound to resolve the hiatus and other vowels require an /r/ sound. It is important to keep in mind that the intrusive sounds are a result of the intrinsic relationship between the vowels (Alameen, 2007).

When consideration is given to the vowel sounds that precede the glides, it becomes apparent that there is a connection between the preceding vowel sound and the sound that is chosen to link or intrude. If the vowel sound is a mid- or high-tense vowel such as the /uː/ in "do it", it is produced with a rounding of the lips. When the lips are unrounded to produce the next sound /ɪ/, the lips transition from the rounded to unrounded position. Now, consider the production of /w/. Air is expelled whilst the lips transition from a rounded to an unrounded position. This means that the only action necessary to produce a /w/ between a mid- or high-tense vowel is simply to keep the air flowing. Since all sounds are produced with flowing air, the airflow is a characteristic of continuous, flowing speech.

Similarly, the /j/ sound is characterised by a raising of the middle of the tongue. The tongue then transitions down to begin the next sound. High-front vowel sounds such as the /iː/ in see it /siː(j)ɪt/end with the middle of the tongue raised. Therefore, in order to transition from a high-front vowel to another vowel, the speaker needs to make the same transition. Again, the only extra element is the flowing of air. It can therefore be concluded that the production of the glides is therefore simply a by-product of the articulation of the vowel sounds that surround it. The flow of air is simply not halted between the two vowel sounds for the sake of the ease of continuing the momentum of speech. Newton and Wells (2002) aptly explain this by saying that /j/ and /w/ are simply *low-level articulatory transitional phenomena* and then quickly point out that /r/ sandhi is not. Throughout this report it is accepted that the /r/ sound is distinguished from the glides because the glides appear to be a bi-product of the transition of the articulating organs (the lips and tongue).

In summary, it is the class of the preceding vowels that decides which sound is required to resolve the hiatus. The production of the glides (/j/ and /w/) appears to be a result of the transition of the articulators from the preceding vowel sound to the following vowel sound. However, this does not appear to be the case with /r/ sandhi.

1.6 Research Questions and hypotheses

In order to learn more about sandhi and the way in which Japanese ESL students learn to use it, this study will pose the following research questions.

1.   Do Japanese ESL students use sandhi when speaking in conversation or reading?   Is the extent to which they use them comparable with native English speakers?

2.   Can Japanese ESL students identify intrusive sounds when spoken by a native English speaker?

3.   Does the frequency of sandhi use correspond with the participants' current level of study?

4.   Does the frequency of sandhi use correspond with the speed of speech production?

The final question is to establish whether students will naturally develop their ability to use sandhi as they are exposed to the language. It may be logical to assume that ESL students naturally begin to incorporate sandhi into their speech, but it is also possible that it is not done without either direct instruction or at least increased awareness.

The hypothesis for this study is that the linking and intrusive /j/ and /w/ sounds will naturally increase as the production speed of the learners is artificially increased suggesting that the learners will increase their usage of sandhi as they develop their fluency. Also, it is hypothesised that the intrusive /r/ sound will not be more frequently used just because the speaking fluency of the learner is improving. The author believes that in order to develop the intrusive /r/, an advanced level of awareness of phonemes is required, but whether that is developed exclusively in the higher levels of fluency is not known.

It is also hypothesised that the rate of sandhi usage will correlate with the students' current study level. A correlation here is only expected because the current study level is a weak indicator of the amount of English to which they have been exposed and the amount that they use. It is believed that the amount of exposure to English builds the learners' linking abilities not the more advanced material that they are studying. However, any correlation between study level and sandhi usage is suspected to be quite weak in order to allow for individual experiences, differing profiles amongst the students (in terms of their various abilities) and varying levels of exposure to English environments.

1.7 The researcher's perspective

I have always found the debate between descriptive and prescriptive perspectives to be quite interesting. The prescriptive perspective on language states that there are certain rules, grammars and ways of pronouncing a language that are correct and that people should adhere to these rules if they want to be respected as a well-educated individual or a competent second

language user. The descriptive perspective disagrees that the rules dictate the language of its speakers and believe instead that the language is constantly evolving and we use grammar as a tool to describe the language that is spoken.

There is an inherit belief in ESL teaching that students speak "correctly" or "incorrectly" depending on the likeness of their speech with that of a native speaker. I presume that educators believe it is necessary to define speech that is different to that of native speakers as "incorrect" in order to identify the areas in which a student can improve. Therefore, by its very nature, ESL instruction is, at least to some extent, prescriptive in nature.

This prescriptive perspective in my occupation is contradictory to the descriptivist approach taken by this study. Throughout this report, the terms *correct* and *incorrect* will be avoided. Instead, I will use the terms *standard* and *non-standard*. For further discussion on this issue of "correctness", see section 5.5.

Chapter Two:

Literature Review

This study will now examine the research that has previously been conducted on the subject of sandhi and in particular, linking sounds and intrusion. The issue of rhoticity will then be discussed.

2.1 Sandhi

Being quite a specific field, the area of the pedagogy of sandhi is one in which there has been little previous research. There have been some quantitative studies on the use of sandhi by adults (Kennedy & Blanchett, 2014 and Mompea & Mompea, 2009). Mompea and Mompea (2009) demonstrated that BBC newsreaders resolved hiatus in cases of potential /r/ sandhi about 60% of the time. This research was performed on archives of BBC news broadcasts and that means that they are dealing with a specific accent known as Received Pronunciation (RP) English. RP pronunciation (informally known as *the Queen's English*) is known for being very clearly enunciated. This means that the levels of sandhi usage could be expected to be low compared to the population of English

speakers in the population of Britain. In other words, it is potentially a conservatively skewed sample.

Kennedy and Blanchett (2014) revealed that learners of French underwent improvement in comprehension of language that included sandhi (with a focus on liaison) after a course, which included explicit instruction. This study was conducted with a different target language (French learners not ESL learners), but if these results are transferrable, it provides support for direct teaching of sandhi.

There have been some studies, which were more qualitative such as Gick (1999) and Broadbent (1991). These tend to focus more on the theory behind the emergence of sandhi rather than collecting data to make assertions about how ESL learners use it.

There have been some studies on the development of sandhi techniques in young children (Chevrot et al., 2009; Newton & Wells, 2000). Newton and Wells (2000) presents an interesting point in the development of linking in children. They state that children as young as eighteen months start to produce multi-word utterances and that by the age of three and a half the process of developing connected speech is all but complete. He states that linking does not appear to be learned but acquired by the children through listening and verbally experimenting. The researchers do highlight one exception and that is /r/ sandhi which is developed later.

The implications of this research to the current study are interesting. It suggests that word connection using the glides (/j/ and /w/) are developed as a child learns to speak. This could suggest that the ESL learners also develop sandhi techniques not only in the latter stages of English development but as

they speak, they connect. This would suggest that there would be no significant correlation between the current level of study of the participants and their performance in the sessions in terms of word connection with the glides. We will see later that this is indeed the case.

Counter-intuitive as it may sound, the participants who have achieved a higher level of study do not use more sandhi. It would be logical to assume that the level of study would correlate with the amount of time that the participant has invested in studying English and therefore higher-level students have had more exposure to English and would use sandhi more. However, in Japan, most elementary schools offer English lessons with native speakers throughout and these usually continue through to high school. Therefore, the level of study does not correlate with the length of exposure to English: it more likely correlates to the language learning abilities of the student. Some students just learn faster than others due to motivation, memory, a sense of belonging and other social and psychological factors.

Instead of sandhi developing slowly over time as does the students' grammar and vocabulary knowledge, it is possible that the habit of linking, at least in terms of the glides, is done in the very early stages and as soon as the learner can string two word utterances together, they are able to link them. Perhaps the only way to witness the development of linking is in a classroom of true beginners. /r/ sandhi is the exception here and as discussed earlier, /r/ sandhi may require some direct teaching in order for learners to develop.

In the specific area of sandhi usage by adult learners of ESL, there have been few attempts to quantify and investigate. Some examples of the attempts that have been made are Zahedi, Sahragard and Nasirizadeh (2007), Henrichson

(1984), Hieke (1984) and Alameen (2007). The first two studies focused on the participants' ability to listen and identify word boundaries in utterances, which is only part of the scope of this study. This study looks at the effect of sandhi on both receptive skills (listening) and productive (speaking).

Zahedi, Sahragard and Nasirizadeh (2007) demonstrated that participants, who were studying at a tertiary institution, were generally weak at perceiving material which included different forms of sandhi including elision[1], liaison, assimilation[2], juncture[3], transition[4] and palatalization[5]. The study also demonstrated that there is no interaction effect between proficiency level and the perception of phonological features. If the present study concurs with this study, Part Three will show no correlation between the participants' scores and their current level of study.

Zahedi, Sahragard and Nasirizadeh (2007) also demonstrated that there is no interaction between gender and the perception of phonological features and that participants had more problems with elision and assimilation (except for the elementary learners who had more issues with palatalization and liaison). Transition and juncture appeared to be the least troublesome features for the Iranian learners. The study concluded that the presence of phonological features in materials used in the EFL classroom was a major factor in the comprehension of the learners.

---

[1] *Elision* is the dropping of sounds like the missing vowel in "comfortable" – usually pronounced /kʌmftəb(ə)l/

[2] *Assimilation* is a sound that changes to be more like the sounds around them like the /n/ to /m/ transformation in /hæmbæg/)

[3] *Juncture* is a pause, lengthening or strengthening of a sound in order to mark the break between words. It is the difference between "an aim" and "a name"

[4] *Transition* is the action of moving the articulators from the position of one consonant sound to the position required to produce a different consonant sound.

[5] *Palatalization* refers to the production of consonants when the tongue is moved close to the hard palate. This makes the non-palatalised /k/ in "cut" sound slightly different to the palatalised /k/ in "key". Palatalization also includes the transformation of /t/ into /ʧ/ as in 'tune' and /s/ into /∗/ as in 'assume' before high vowels. In Japanese, palatalization includes the transformation of /t/ into /ʧ/ before /i/.

Lynn Henrichson (1984) conducted a study that also demonstrated the effect that phonological features such as sandhi had on the comprehension of the learners. She stated that learners must have comprehensible input in order to acquire language and that such phonological features can act as a filter in the input-intake process.

Hieke (1984) compared the frequency of linkers and intrusives in native English speakers to non-native English speakers. He discovered a measurable difference and concluded that sandhi could potentially be used as a factor in measuring fluency in non-native speakers.

Alameen (2007) followed on from Hieke (1984) and produced the study most similar to the present study. Alameen demonstrated a significant difference in the linking of native and non-native English speakers. In addition, she showed that native speakers link function words more than content words. This confirms the observations of Selkirk (1995) that function words maintain a close phonological connection with the words around them. Unlike the current study, however, both Henrichson (1984) and Zahedi et al. (2007) focused on testing learners' comprehension of sandhi where the current study focuses on the production of sandhi.

So the research above suggests that even though native speakers use sandhi inconsistently, it has implications for the learners' comprehension in the classroom and outside. There is however, a severe lack of information about to what extent learners produce these features of natural language when they are engaged in conversation with each other. That is the gap, which this project will help to fill. It is clear that use of sandhi by the teacher is a factor in the learners'

comprehension, but this investigation aims to confirm that learners are using sandhi in their speech and investigate to what extent they acquire these features.

In terms of pedagogy, phonetics research tends to neglect linking. Phonetic transcriptions tend to omit linking even though they include other types of sandhi such as elision (Alameen, 2007). Failing to recognise linking in transcriptions means that it is effectively ignored by researchers and teachers alike.

It may seem surprising that this area of ESL learners' education is not researched more. Any area which provides the potential for assisting students in achieving a much coveted *native accent* surely deserves the attention of researchers. However, various studies have reported that pronunciation instruction has limited impact in terms of altering learners' accents to sound more like native speakers or may hinder the acquisition of grammar and/or vocabulary (Krashen & Terrell, 1983; Leather, 1983).

It is clear that due to the commercial pressure, more attention is paid to the lower levels of the English learning spectrum (which due to the attrition of students tend to be the most heavily populated). The author believes that research on how all levels develop sandhi techniques will help ESL teachers to prioritise the development of these techniques. The word prioritise is used here because as Skehan (2009) discussed, there is a trade-off between fluency, accuracy and complexity. Teachers need to make informed decisions about where class time is best utilised.

2.2 Rhoticity

As mentioned above, most of the studies on linking tend to focus more on /r/ sandhi than on the glides (/j/ and /w/). This is because of the history surrounding its origins and its connection with rhoticity, which is a comparatively well-researched area. English accents are described as being either *rhotic* or *non-rhotic* depending on whether words that end with a letter 'R' have an /r/ sound. Many American accents pronounce the word "car" as /kɑr/ whereas many British accents terminate the word with a vowel sound /kɑː/. The existence of the /r/ sound in rhotic languages eliminates the possibility of a linking /r/ and intrusive /r/s are rare in rhotic languages.

The history in terms of New Zealand English is explained in Hay and Sudbury (2005). "Pronunciation of non-prevocalic /r/ in non-boundary positions decreased and eventually vanished, ... -resulting in a non-rhotic dialect, with high rates of linking /r/ at word boundaries." Because New Zealand English is (with the exception of some southern accents) a non-rhotic accent of English, it opens up the possibility of linking /r/ by New Zealand English speakers.

There is a plethora of debate surrounding the relationship between rhotic and non-rhotic accents of English. One perspective states that non-rhoticity is a result of */r/ dropping* (Vennemann 1972, Johansson 1973, Wells 1982, McMahon et al. 1994). Proponents of this theory argue that non-rhotic languages stem from a rhotic language but the speakers drop the /r/ sound. It is preserved in cases where the following sound is vocalic. The opposing theory states that the /r/ at the end of words is always underlyingly present. It is then argued that the /r/ sound in saw comes as a result of the listener making the analogy with the /r/ sound in soar (Hay & Sudbury, 2005).

This research does not intend to enter the debate of whether the intrusive sounds are *inserted* or *preserved* (given that they have a historical context) so throughout this report, the term *inserted* is used without the intention of entering this debate.

Hay and Sudbury (2005) state that "most dialects that exhibit linking /r/ also exhibit intrusive /r/" (p. 801). Many researchers have discovered that intrusive /r/ occurs less commonly than linking /r/, and the present study concurs. It is usually implied that this is a result of the stigma involved in using intrusive /r/ (Lewis 1975, 1977, Pring 1976, Fox 1978, Brown 1988). Although as this research demonstrates, the rarity of intrusive /r/ is more likely to be a result of the fact that few word combinations necessitate the intrusion with /r/. Some examples are "law and order" and "saw it". The requirement for an intrusive /r/ is an /ɔː/ sound such as in "law", "saw", and "raw", but without the letter 'R' which would make it a linking /r/. So words such as "door", and "floor" end with the same vowel sound, but existence of the letter 'R' in the spelling of the word means that an /r/ sound would be a linker not an intrusive. Therefore, there are few word combinations that require an intrusive /r/ for word connection.

So rhoticity means that the background of the participant and the nationality and/or background of each of his or her teachers becomes a major factor in whether a participant will produce a linking /r/ or just a terminal /r/. Take for example the segment "car is". If the participant was mostly influenced by teachers from rhotic areas within the United States, he or she may say them in isolation as /kɑːr/ and /ɪz/. A participant who learned with a British teacher from the non-rhotic areas of the U.K. may pronounce them /kɑː/ and /ɪz/. So the

participant with the rhotic teachers is not producing a linking sound because as our definitions above state, linkers need to voice letters, which are not pronounced when the words are pronounced in isolation. Yet the /r/ sound would not be an intrusive either as the letter 'R' is present in the spelling of the word. Hence, without a complete history of each participant and all of their previous teachers, it is not possible to tell whether each /r/ is really a linking /r/ or not.

2.3 Summary

There is a paucity of studies that shed light on whether ESL students in general use sandhi and to what extent. Mompea and Mompea (2009) demonstrated that native English speakers do not consistently use linkers and intrusion. In a survey of BBC newsreaders, only about 60% of possible cases were connected or linked. Zahedi et al. (2007) and Henrichson (1984) suggested that sandhi decreases a learner's ability to comprehend spoken English. While many researchers believe that direct instruction does not improve pronunciation, Kennedy and Blanchett (2014) achieved some success with direct instruction.

The literature tends to debate whether non-rhotic languages are a variation of rhotic languages which do not pronounce the /r/ sound or whether the /r/ sound is intrinsic in every hiatus which consists of certain vowels. Although this study does not intend to enter this debate, it is important that we recognise the underlying relationship between the vowels that create hiatuses.

My search of the literature did not uncover whether the learners in more advanced English classes use sandhi more frequently than those in less advanced

classes. The literature also does not reveal how the speaking speed of the learners affects the rate of success with sandhi. With this study, I hope to discover whether the production of the lateral glides (/j/ and /w/) as intrusives increases as speaking speed is increased and whether the intrusive /r/ increases in a similar way.

Chapter Three:

Methodology

This report examines speech samples, which were collected during *sessions* with participants. These *sessions* involved the participants performing three tasks while responses to the tasks were recorded. Throughout this report, the term *session* will refer to the sample collection session where the researcher met with participants to collect speech samples using these three tasks. The methodology section details information about the participants, how the speech sample collection sessions were conducted, and how the data were extracted from the sessions.

3.1 Definitions used in analysis of data

Potential token will be used to refer to a combination of two words that create a hiatus, which could potentially be resolved using one of the six techniques discussed in this study. A *token* will refer to the standard resolution of sandhi by using either a linker or an intrusive sound. For example, the words "say" and "anything" when recorded in succession create a potential token. If the

learner links these two words with a /j/ sound, then that occurrence will be referred to as a *token*.

A *nonstandard token* is a word combination which was linked with an unexpected consonant. To illustrate, if the potential token "saw a" occurs and the speaker used a /w/ sound to link whereas a native speaker would more likely have used an /r/ sound to connect these words, this is deemed to be a nonstandard token. The terms correct and incorrect are avoided in this study. See the Discussion section for an explanation on why this terminology was avoided.

The term *token rate* will refer to the proportion of potential tokens that are realised as standard tokens. In other words, the proportion of cases in which sandhi was used to resolve a hiatus as a native speaker would resolve it.

3.2 The participants

The institution at which this study was conducted is a tertiary institution located in New Zealand. It has a sister school in Japan where it is compulsory for students studying particular courses to study one year in New Zealand. Because of this large number of incoming students from Japan, the student body is predominantly Japanese students, many of who are studying English courses to prepare them for future study in English.

The participants in this study come from various places around Japan. Some of them will return to Japan where they will study to be English teachers, airline pilots and other English-intensive careers. Others will stay in New Zealand, having achieved an appropriate standard of English to continue on to further training or start their careers in New Zealand.

Most of the students at this institution live on campus in dormitories provided by the college. For some, because of the geographical isolation of the college, the experience of living in New Zealand is limited to the teachers they encounter, but the more adventurous student may become completely immersed in the language and culture of New Zealand. Some of the students become immersed because they take on the challenge of living as a homestay student with local families. So even the experience of studying in New Zealand varies from student to student.

Thirty-eight participants volunteered in response to a call for participation in this research. All were Japanese ESL students studying at this institution. All of the participants had been studying in New Zealand for approximately 6 months prior to participating. The average age of the population of students from which this sample was taken was 19 years old at the time of completing the session. The sample consisted of 22 males and 16 females. An equivalent sample size of both genders was obtained even though Mompea and Mompea (2009) demonstrated in an empirical study of native English speakers that gender was not a significant variable in the production of /r/ sandhi.

Before the session, each participant was asked to complete a background survey. The survey is included as Appendix B. This survey was designed to quantify the study experience of the participants. Questions included how long the participant regularly studies (both inside and outside class) and how long he/she has studied English to date. The amount of study experience to date was then broken down into the amount of time studying, the amount of time studying at a cram school and the amount of time immersed in an English speaking country. This information would have been important had there been any

anomalous performances in the sessions. However, since there were no mysteriously anomalous performances in the sessions, the data collected with this survey were simply used to define the sample of participants.

The other questions in the survey investigated the participants' first language, nationality and motivation for studying English. The first two questions were to check that no participant was raised in a household where English was the first language. It was important to screen out any person who had experienced much more exposure to English than the other participants and would therefore skew the results.

The background survey gave some insight to the background of the students, which was fairly varied, but all students were at least false beginners. This means that they had all studied English in Japan before coming to New Zealand to the point that they were familiar with English letters. They could read written words even if they could not understand the meaning of the words. The average number of years that the participants claimed to have studied in Japan ranged from one to twelve, and the average was 6.7 years. All participants used Japanese as their first language at home.

Of the 36 participants who completed the background survey, 26 had studied English in a "cram school" or other private institution. A *juku* or *cram school* is a common way to supplement a school child's formal education in Japan with almost all university-bound high school students attending a cram school. The average attendance at a cram school for those who attended was 4.4 hours per week (ranging from one to ten hours per week). The average duration of study at a cram school was 4.1 years (ranging from 1 to 12 years).

The participants came from all over Japan and had TOEIC scores with a broad range from 280 to 640 (with 6 out of 36 participants omitting their scores). This means that the participants' ability to understand written and spoken English ranged from *basic user* to *fairly proficient user*. The average TOEIC score of those in the sample who provided their results was 415.

The institution at which this study was executed is very international; there are students from Indonesia, India, China, Taiwan, Vietnam, Russia, Thailand, as well as local students. However, this study was intentionally limited to Japanese speakers. The decision to restrict the study to only Japanese students meant that they all came from the same language background. It is important to control this variable because every nationality included in the study would have introduced different forms of first language interference. Limiting the participants meant that the study could better focus on one particular set of issues.

To summarise, the characteristics shared by the participants were their nationality, the amount of time they had lived in New Zealand before the survey, all were at least false beginners and all were learning English as a second language. Their English ability and the amount of daily exposure to English varied as did their background in terms of study in private institutions.

3.3 Paired sessions

The way that the sessions were conducted would determine the anxiety levels of the participants. Therefore, I attempted to reduce the levels of anxiety and increase the potential for relaxed and natural speech samples, which would be likely to contain more sandhi. This was done by pairing the students and

having them talk with each other, using familiar environments and familiar topics and maintaining a friendly demeanour during the sessions.

Pairing the participants had disadvantages as it meant that the interviewer was not able to guide the conversation towards the target language. Pairing meant that some participants had to be paired with participants that they did not know in cases where a class had an odd number of participants. Pairing some participants with a stranger surely had implications in terms of the level of anxiety experienced by the participants and therefore the speed and possibly the accuracy with which they spoke. However, the reason that the students were paired was simple. Pairing the participants and having them speak to one another instead of an unknown authority figure minimised the anxiety of the participants. Especially the lower-level students may potentially have experienced increased levels of anxiety as they may feel inadequate using a language that they are not yet proficient in using. Of course, the participants who were paired with a student from another class may not have enjoyed this reduction in anxiety, but on average, the anxiety levels of the interviews were lower.

The atmosphere created at the time of the sessions played an essential part in getting the learners relaxed enough to use sandhi. When speaking with or in front of people they identify as a teacher, learners tend to speak slower or more carefully than they would with their friends through fear of correction and therefore loss of face. Pairing the learners with a peer meant that the participants could feel freer to use colloquial English; however, they were still being observed by a teacher and recorded.

Past research has debated whether or not sandhi is exclusively a feature of fast speech. Researchers (Marks, 1999; Pennington & Richards, 1986; Richards, 1983; Weinstein, 2001) define connected speech as something that happens in fast, informal, relaxed or casual speech. Although others such as Kaisse (1985) disagree stating that sandhi is always present in varying degrees. To play it safe, it was decided to remove potential affective barriers to connected speech by creating a relaxed environment. In this environment, the participants discuss familiar topics so that they are not pressured to come up with unfamiliar language with peers, not educators.

3.4 The pilot study

During the design phase of this project, I decided that the best way to decide what elements should be included in the sessions was to do a pilot study. So I conducted a session with two students of the institution who were arbitrarily chosen to avoid bias. The sessions were very enlightening in terms of how much data could be collected in a practical sense.

The first decision that was made based on the pilot test was the types of sandhi to be included. The pilot test included 3 different forms of sandhi: intrusive sounds, linking sounds and gemination. Gemination (AKA twinning) is the production of one sound in an utterance where the same sound occurs in both the coda position in the preceding syllable and the onset of the following syllable. An example of this is the utterance "stop pushing". The /p/ sound terminates the preceding word and initiates the following word. Because the /p/ sound is a *plosive*, it is made by building pressure behind the lips and then releasing to let the air explode from the mouth in one burst. However, it is

ungainly to produce a plosive and then reset the mouth in order to produce another plosive so the coda becomes an unreleased /p/ sound while the onset becomes the aspiration (or the "explosion").

When gemination was included in the pilot study, it was quickly noticed that gemination was not an issue that the Japanese ESL learners really struggled with and when they were speaking fast enough to have the two /p/ sounds in succession, they usually pronounced the two /p/ sounds as one with a slight pause in the middle. Although there may have been slight differences in the timing of the gemination, it was decided that the measurement of such timings was not in the scope of this project. The content of the project was therefore limited to the intrusives and the linkers.

The second set of decisions that were made on the basis of the pilot study was regarding the amount of data to be collected. The original intention of the author was to collect as much data as possible with regards to how the participants produced the token. Originally, the responses of the participants were to be designated as one of the following:

- **A standard link/connection** (the learner added a consonant sound which was the sound that a native English speaker would add)

- **Not linked/connected due to a glottal stop** (natural)

- **Not linked/connected due to a glottal stop** (elongated)

- **Not linked/connected due to an unnatural pause**

    (learner paused to consider sentence structure)

- **Not linked/connected due to a natural pause**

    (learner paused to consider content)

- **A non-standard link/connection** (the learner added a consonant sound which was not the sound that a native English speaker would add)

In her study, Alameen (2009) discovered that the native speaker participants sometimes produced a glottal stop that was less audible than the non-native speaker participants when they did not link; however, the occasions in which they did not link were far less frequent than the non-native speakers. Therefore, the duration of the glottal stop and also the analysis of the pauses were of interest to the author of the present study. Eventually, it was decided that a simpler more targeted approach was best.

It was decided that the tokens would be designated as follows.

- **a standard link/connection** (the learner added a consonant sound which was the sound that a native English speaker would add)

- **not linked** – the participant paused (naturally or not) and therefore did not insert the sound that a native English speaker would add

- **a non-standard link/connection** (the learner added a consonant sound which was not the sound that a native English speaker would add)

The pilot session provided some other information that helped me to decide which questions to include in the session. It was found that one question in Part Three was found to have two possible answers and one question in Part Two was found to be simply too difficult for the participants to pronounce because of difficult words used in the utterance. Pronunciation issues in the latter mitigated any chance of linking occurring.

3.5 The sessions

There were three parts to the session. The first part was called "Part One: Free Conversation" where the participants paired up and conversed about their families for 1-2 minutes. The second part was "Part Two: Reading" The participants read a list of sentences, once at natural speed and once at an accelerated speed. Two readings for each of the two students took 3-4 minutes in total. The third part was "Part Three: Listening". Part Three: Listening took approximately 2 minutes. All sessions were held in unused classrooms.

Part One: Free Conversation was executed in pairs, as the participants were encouraged to engage each other in conversation about their families. This topic was chosen, as was the format, to minimise anxiety and get them speaking as they normally would to peers outside the classroom as it was thought that speaking with an authoritative stranger may slow down speech and/or encourage students to speak more formally; both situations would lead to less tokens recorded. The topic was also chosen because of the expected high frequency of linking /r/s that may be found because of the number of kinship terms ending in /r/, such as 'father', 'mother', 'brother', 'sister'.

It was found during the pilot session that getting the participants to draw a stick figure of their family before entering into a conversation about them helped the participants to speak more fluently. In a way, the picture was like the outline that is done before the essay is written, it provided the participants with direction in their conversation and reduced the time required for consideration of content.

The objective of Part Two: Reading was to capture samples of participants as they read 12 sentences aloud in a reading task. This was to compare reading

with conversation to see how they differed (orthographic visualization of the utterance may reduce tokens as students may feel less inclined to link when it contradicts the written text). Also, in this part, participants were asked to read the sentences twice; they read once at natural speed and then they were asked to read the text again but faster. This repetition was to see if faster production led to more tokens of sandhi.

Part Three: Listening aimed to see if participants could identify intrusive sounds made by a native New Zealand English speaker. In Part Three: Listening, they were given very little instruction. They were simply told that the sentences they were about to hear all had an extra sound and that they were to listen to the sentences and identify the extra sound in each one. The participants' attention was drawn to the fact that they were multiple-choice questions (all having the three options; /w/, /j/ and /r/).

Part One: Free Conversation and Part Two: Reading were recorded. The answer sheet from Part Three: Listening was collected from the participants. The recordings were then analysed by playback in slow motion and moderation was done with a research supervisor at Massey University. The moderation process involved taking a selection of recordings for each part and both the researcher and the moderator independently analysing them and comparing the results. Any discrepancies were discussed and this introduced a few points for discussion and issues as discussed below.

A *potential token* was considered to be a *standard token* when the speaker made the consonant sound that a native speaker would usually make to resolve the hiatus in question. If the participant resolved the hiatus with an unexpected consonant, it was deemed *non-standard* and discussed. The strength of the sound

46

produced was not relevant only its production. Also, the pronunciation of the consonant sound had to fairly resemble how a native English speaker would say it (so the intrusive /r/ must sound like an /r/ not an /l/).

As mentioned earlier, the use of sandhi is not consistent even in the speech of native speakers. This led to the need for a session with two native English speakers (who were unaware of the purpose of the study) in order to establish a base line. When discussing how far the participants vary from native English speech it is important to compare them to this baseline rather than comparing them to an unrealistic perfection.

3.5 Session design decisions

There were a few issues that were discussed during the design phase of the sessions. These issues were pauses, glottal stops, nonstandard sandhi and contractions.

3.5.1 Pauses

When speakers paused, they paused for a number of possible reasons. Many of these reasons are natural even to a native speaker. All of the cases that were mentioned earlier (Hieke, 1984) for a glottal stop can also be extended in duration to pauses.

1.   In deliberate speech for reasons of extra clarity

2.   Where special stress assignment overrides absorption phenomena

3.   Where plus juncture is phonemic and thus obligatory

4.   After silence (with no prior syllable to draw on)

On completing the sessions, however, it was my feeling that most of the pauses that occurred in the sessions were not for the reasons above but for a reason

more specific to learners of another language. Most of the pauses appeared to be because the participant had paused to consider grammar, structure and/or vocabulary. When native speakers are talking, speakers think about what they want to say, but they spare little cognitive energy on the grammatical structures that are being used because they are usually so familiar with them. Learners on the other hand, being less familiar with the language need time to plan structures, choose appropriate words and remind themselves of pronunciation before speaking. The lower the student's current level of study, the more time is required to produce an utterance. This may represent a significant barrier to the production of sandhi.

3.5.2 Glottal stops

The second issue was whether glottal stops should be recorded differently from pauses. In order to quantify the glottal stops produced by participants, it would be necessary to differentiate not only between a pause and a glottal stop but also to define *creaky voice* as well. Creaky voice (also known as vocal fry or laryngealisation) is a slow vibration of only one end of your vocal cords (Crystal, 2003). This is different to a glottal stop which is the sound made when the speaker closes the glottis (an aperture between your vocal cords) (Crystal, 2003). However, distinction between these two phenomena is difficult. Mompean and Gomez (2011) analysed potential cases of /r/ linking in the speech of BBC broadcasters. They demonstrated that there were far more cases of creaky voice than true glottal stops. I decided that the distinction between these three devices was a distraction from the main aims of this project.

3.5.3 Non-standard sounds

The third issue was the production of a non-standard sound. This became one of the main focuses of the Discussion section. The production of a non-standard sound only occurs when the standard sound would have been an intrusive /r/. As mentioned earlier, the glides (/j/ and /w/) can be described as *low-level articulatory transitional phenomena* (Heselwood, 2006), which means that the sound is made simply by the air passing through the articulators (lips and tongue) while they are transitioning from one position to the next. The intrusive /r/ is not a *low-level articulatory transitional phenomenon* and therefore the participants need to be aware of which sound needs to be inserted in order to correctly resolve the hiatus. In all of the situations in which this error could be made by the participants, there was the orthographic distracter of a semivowel in the spelling of the word. So when the learner sees the words "saw a", they might pronounce it as /sɔː(w)ə/ instead of /sɔː(r)ə/. If the learner is unsure of which consonant sound they should insert to resolve this hiatus, it is understandable that they insert a /w/ since the written forms of the word "saw" contains a 'W'. The written form of the utterance is a distracter. How the participants coped with this distracter is discussed further below.

3.5.4 Contractions

Another issue that was discussed in the moderation process was whether a contraction should be deemed to be a correct use of sandhi. The utterance "he is" could be pronounced /hiː(j)ɪz/ or can be contracted and pronounced /hiːz/. Although contraction is a type of sandhi and /hiːz/ is a correct usage of contraction, it is not the /j/ sandhi that is being targeted in this study. So by

determining that /hiːz/ is not the target language, this study is not implying that it is "incorrect".

3.5.5 Summary

After deliberation, I decided that it was best to categorise each potential token as either a token or not. The only other category I allowed was for non-standard tokens. Allowing a category for natural/unnatural pausing requires that I determine the reason for the pause or assign a time period which I would deem a "natural" pause. For practical reasons, I decided to exclude the issues of pausing, glottal stops and contractions but to include the issue of nonstandard tokens.

Chapter Four:

Results

In this section, I will first review the research questions, then I will present the results from the NEST sessions and finally, I will present the results from the participant sessions part by part.

4.1 Restated research questions

At this point, it is necessary to revisit the research questions so that they are fresh in mind as the report examines the results of the sessions.

1. Do ESL learners use intrusives and linkers when engaged in conversation or when reading aloud? Is the extent that they use them comparable with native English speakers?

2. Can Japanese ESL learners identify intrusive sounds when spoken by a native English speaker?

3. Is there a relationship between the learners' usage of sandhi and the learners' current level of study? In other words, do learners in more advanced levels of study link more?

4. Is there a relationship between the learners' usage of sandhi and the learners' speed of production? In other words, if the participant is asked to speak faster, does he or she link more?

It was hypothesised that the token rate of the glides will increase as the speaking pace of the participant is accelerated, but this will not be true of the intrusive /r/.

It was also hypothesised the participants' current level of study will weakly correlate with the token rate of sandhi because the current level of study is a weak indicator of how long the participant has been studying and how much exposure to English the participant has had.

4.2 Setting the benchmark (NEST sessions)

Before delving into the results from the learner samples, it is important to set the benchmark to establish a realistic standard against which the learners can be compared. Because native speakers do not use sandhi with a strict consistency, this report needs to determine to what extent they do use sandhi. In order to set this benchmark, this report will now turn to the results of the two native speakers.

Since all native speakers in this research were teachers, the acronym NEST (Native English Speaking Teachers) will be used to refer to them. It is important to remember that even though teachers were chosen for the sessions, they were not aware of what the purpose of the research was and therefore were unable to call upon their teaching backgrounds to assist them in the sessions. They were simply asked for a speech sample that would be comparable to learners in terms of simplicity but spoken at a natural rate.

NEST One was a Kiwi male in his late twenties who described his English as "educated Kiwi" English. NEST Two was a well-travelled American female, in her early thirties. She explained that she did not associate with any particular accent because of her well-travelled background both within the U.S. and overseas. However, it is important to note that NEST Two had a rhotic accent while NEST One did not. This allows us to compare the performance of the learners with both a rhotic and non-rhotic accent.

The NESTs were selected because they had not been involved in any discussions about the project as one NEST was new to the establishment and the other worked in a different department.

One issue to consider in this comparison is that for the NESTs, a conversation at this level of simplicity with a stranger may have seemed more contrived than it would for a student who has six months' experience in an English class, getting accustomed to the role play style activities of a communicative English classroom. This may cause discomfort or raise the level of anxiety or otherwise affect the performance of the NESTs although I had resolved to ensure, as much as possible, that the NESTs and the participants had a similar experience when they were participating in the sessions.

4.2.1 The NEST sessions – Part One: Free Conversation

Part One: Free Conversation of the NEST sessions proved that as Mompea and Mompea (2009) had earlier discovered, native English speakers do not use sandhi consistently. Even with a small amount of data, it is clear that native English speakers allow some word connection opportunities to pass by without

capitalizing on them. With only two participants, the number of potential tokens was as shown in Figure 1 (overleaf).

	NEST 1	NEST 2	Total
Linking /j/	0	2	2
Linking /w/	0	0	0
Linking /r/	7	11	18
Intrusive /j/	1	2	3
Intrusive /w/	2	2	4
Intrusive /r/	0	0	0
	10	17	27

Figure 1: NEST sessions Part One: Free Conversation – Occurrence of potential tokens

With such low numbers, it is impossible to draw any conclusions from the data. So it was decided that the solution would be to increase the number of NEST participants to six. It is important to remember that the purpose of the NEST sessions is to act more like a case study to compare the learners with rather than a quantitative study in its own right.

The four additional NEST participants came from New Zealand, America and Canada. Only NESTs Two and Three had rhotic accents, both being American. It is important to consider this when looking at the quantity of potential linking /r/ tokens. As discussed in Section 2.2 of this study, a rhotic accent means that tokens that look like linking /r/ may just be the production of the rhotic /r/ sound. With six participants the data increased to the numbers shown in Figure 3.

	NEST 1	NEST 2	NEST 3	NEST 4	NEST 5	NEST 6	Total
Linking /j/	0	2	1	2	2	0	7
Linking /w/	0	0	2	0	0	0	2
Linking /r/	7	11	1	3	4	3	29
Intrusive /j/	1	2	0	2	0	0	5
Intrusive /w/	2	2	2	2	1	1	10
Intrusive /r/	0	0	0	0	0	0	0
	10	17	6	9	7	4	

Figure 2: NEST sessions Part One: Free Conversation – Occurrence of potential tokens with six NESTs

Again the absence of intrusive /r/s is the most striking feature of the table. It should be noted that linking /w/ has very few tokens also. Linking /r/ is again the most prevalent of the six with intrusive /w/ coming a close second.

The token rates of Part One of the NEST sessions are displayed in Figure 4.

	Potential tokens	Tokens	Token rate
Linking /j/	7	5	71.4%
Linking /w/	2	2	100%
Linking /r/	29	17	58.6%
Intrusive /j/	5	4	80.0%
Intrusive /w/	10	7	70.0%
Intrusive /r/	0	0	N/A
	53	**35**	**66.0%**

Figure 3 NEST sessions Part One: Free Conversation – Token rate

Linking /r/ is the best example to look at as it has the largest sample of potential tokens. With 29 potential tokens analysed, 17 were linked. This means that the NEST participants linked using an /r/ sound on 58.6% of the 29 possible

occasions (or 50% of the 24 possible occasions if you factor out the NESTs with a rhotic accent).

It is also interesting to note that the speech of the native English-speaking teachers was not without pronunciation errors. NEST One at one point in the session neglected to use the weak form of *the* before a word that began with a vowel. This was not a regular function of the NESTs accent but possibly a departure from the planned utterance. The result of this was that the expected intrusive /j/ sound was not produced.

4.2.2 The NEST sessions – Part Two: Reading

By its nature, Part Two: Reading was set in the number of data points it would provide us. There were 12 utterances that the NEST participants read aloud and therefore there were 12 data points. The table below summarises the performance of the two NESTs in both a natural reading speed and a faster reading speed.

	Number of correctly liaised tokens	Percentage of correctly liaised sentences
Nest 1 (natural speed)	11	91.7%
Nest 1 (faster)	12	100.0%
Nest 2 (natural speed)	8	66.7%
Nest 2 (faster)	9	75.0%

Figure 4 NEST sessions Part Two: Reading – Natural vs. accelerated reading speed

It is clear that both participants used more word connection when they were asked to read faster and that the token rate is much higher than in Part One. This is probably because just like the learners, the native speakers pause to consider content as they speak although they did not consider the construction of their sentences. The native speakers tended to consider what they wanted to say, but the learners considered both what to say well as how to say it. However, when given a pre-determined utterance, all participants spoke a little faster and with more confidence.

It may also be noteworthy that question seven in Part Two which was only connected by one participant in the study, also had a low token rate with the native English-speaking teachers. In only one out of four occasions this sentence was correctly joined with an intrusive /r/. The removal of this item was considered before the sessions, but finally, I decided to include it. This was to include some less familiar language in order to see if unfamiliarity affected the rate of connection.

4.2.3 The NEST sessions – Part Three: Listening

Part Three: Listening was quite telling of the NESTs ability to identify the intrusive sounds and linking sounds in a given utterance. NEST One correctly identified 60% of the extra sounds, while NEST Two correctly identified 30%. By the end of the exercise, NEST One admitted to knowing what the test was about, but NEST Two did not.

4.2.4 Reflection on the NEST sessions

Overall, the NEST sessions failed to glean enough data to set a baseline for the learner participants. Part Two and three gave a strong indication that the

native speakers are loosely following any guidelines that teachers would teach to learners and it is important to realise that even native speakers use sandhi inconsistently.

4.3 Part One: Free Conversation

We now turn our attention to the sessions involving the Japanese learner participants. In Part One: Free Conversation, the learner participants were given a topic to discuss freely with a peer (the sessions were all held with a partner who was usually known to the participant). The topic given was *family*, a topic chosen for its high likelihood of producing tokens. The tone of the session was kept friendly and casual in order to increase the chance of tokens.

The sample dialogues that the participants produced were recorded and transcribed. The potential tokens were identified and then judged to be either tokens, non-standard tokens or not tokens. In Part One, two aspects of the data were examined. The first is how many potential tokens appeared in the sample (*occurrence of tokens*) and the second was the quantity of tokens versus non-tokens (*token rate*).

4.3.1 Occurrences of sandhi in free speech

The rate at which each sandhi feature occurs in the samples gives an insight into the rate at which the feature may occur in the participant's day-to-day life. To illustrate this, consider the intrusive /w/ which resolves the hiatus in such word combinations as "do it" or "glue ends". These are two examples of potential tokens that require an intrusive /w/. Now, assuming that the sample of speech collected fairly represents the rate of usage in normal conversation, if a given sample has a high rate of occurrence of potential tokens for intrusive /w/,

it might be extrapolated that the participant encounters such potential tokens often in normal conversation. If they are faced with the hiatus often, the frequency of facing that hiatus may increase their chance of resolving it. If, on the other hand, the participant never faces that particular kind of hiatus, they may be less likely to have the skills to resolve it.

As stated above, the topic was specifically chosen because there was a high chance of linking /r/ occurring. This is because many of the familiar family words such as *father, mother, brother,* and *sister*, will be likely to be used as subjects followed by the verb 'to be' (*is* and *are*). Because of the topic choice, the linking /r/ is overrepresented in the samples of speech collected, and therefore, in the case of linking /r/, the samples do not represent the population of linkers in normal conversation. As for the rest of the linkers and intrusives, there is no reason to believe that the quantity of each feature does not proportionately represent the quantity in normal conversation.

For the purposes of brevity, each sandhi feature was given a two-letter code. Intrusives were denoted 'N' and linkers were denoted 'L'. The second letter in the code was the sound. For example, NJ is intrusive /j/ and LW is a linking /w/.

In their two-minute sample dialogues the participants, as a group, produced 265 potential tokens. This consisted of 184 potential linker tokens and 81 potential intrusive tokens. See the table overleaf.

NJ	57
NW	22
NR	1
LJ	43
LW	23
LR	118

Figure 5 Participants Part One: Free Conversation - Number of potential tokens

As was to be expected, the most common of the sandhi features was the linking /r/. There were 118 potential tokens which could have been resolved with a linking /r/. This is likely to be because the topic was chosen for its high probability of linking /r/ sounds in such utterances as "there are (…in my family)"/ "father is"/ "brother is" etc.

But what *was* surprising was the number of potential tokens that would require an intrusive /r/ sound to resolve the hiatus. In approximately 44 minutes of analysed speech, only one such combination occurred. It was "grandma and". It is clear that of all of the potential tokens faced by the participants in producing their samples, the one faced least often was the intrusive /r/. The implication of so rarely facing this kind of hiatus is unfamiliarity. Learners may be so unfamiliar with this kind of hiatus that they are likely to be unsure of how they should resolve it.

This finding fits with findings later in this report, which suggest that of all six of the sandhi features the intrusive /r/ is produced least often. The fact that it is the type of sandhi least often produced in speech may mean it is the least often

heard by the learners although this is a dangerous assumption to make and further research is necessary to prove that this is the case. Assumptions on the production of native English speakers cannot be extrapolated from a sample of learners.

However, if it is true that the intrusive /r/ is seldom produced in native English speech, and intuition suggests that it is, perhaps this is one area where teachers can help by giving more exposure to this type of sandhi to ensure that students are equipped to resolve it when they encounter it.

4.3.2 Initial observations

There are many situations or sound environments where it would be natural for any speaker to pause or otherwise not use sandhi to resolve hiatus. Some of the environments noticed by the researcher are listed below.

- Pausing to consider content

- Pausing between clauses (especially when using the contraction "*and*")

- Pausing to separate items on a list

- Using elision instead (how about => how 'bout)

- Using a contraction instead (he is => he's)

- Pausing for dramatic effect

It was assumed that the most common reason for participants not linking was none of the above. The researcher had expected the lower levels to pause more to consider the grammar and vocabulary that they needed to construct their utterances. However, this did not seem to be happening. Especially amongst the students in the lower-ability classes, discourse is constructed in their heads one word at a time. It is a very piece-by-piece process. Learners require time to

consider which vocabulary to use, how to form the word and how it should be pronounced. This "on the fly" consideration creates pauses, but it did not seem to inhibit word connection as much as expected.

The participants appeared to pause, but they did not pause in places that disrupted word connection except in cases that could be explained by one of the other reasons above. I would speculate that words which can be connected have a kind of bond which then makes it an inappropriate place to pause. Pausing in any place which unnecessarily breaks a word linking opportunity could be analogous to pausing in the middle of a word.

Also, pauses that came between words such as "father is" often did not prevent the production of the /r/ sound. Sometimes the participant would hold the vowel sound while thinking, sometimes it would appear that the participant had paused, but upon resuming the speech, the participant started with an /r/ sound.

If pausing to consider sentence structure led to no word connection, the lower-level participants would resolve fewer cases of hiatus and as later results tell us, there is no significant correlation between study level and rate of successful linking. This paradox is discussed further later.

4.3.3 Experimental observations

With only one observation for intrusive /r/, there are obviously no valid conclusions to be drawn. For the other sandhi features, it can be seen that, in general, the linking sounds enjoyed a higher token rate than the intrusive counterparts, as shown in the table below.

Sound		Potential tokens	Correct tokens	Token rate
Linkers	LJ	43	12	27.91%
	LW	23	19	82.61%
	LR	117	18	15.38%
Intrusives	NJ	57	30	52.63%
	NW	22	4	18.18%
	NR	1	0	0.00%

Figure 6 Participants Part One: Free Conversation – Token rate

The exception is /j/ sandhi. 52.63% of potential tokens that could be resolved using intrusive /j/ were deemed tokens while only 27.91% of potential tokens of linking /j/ were correctly linked.

When compared to the NEST sessions it is noted that while the NESTs linked using the linking /r/ on 58.6% of the possible occasions, the learners linked only 15.38% of possible occasions. While the NESTs linked using an intrusive /w/ in 70% of the tokens, the learners used an intrusive /w/ on only 18.8% of occasions. This appears to demonstrate a significant difference between the performance of the learners and the six-person control group.

4.3.4 A comparison between lower levels and higher levels

The nature of the data dictates that only non-parametric testing is valid, but a Spearman's Rho test can be conducted. A parametric test cannot be used with these data because the data cannot be assumed to be of a normal distribution and the assumption of *independence of observation* is violated by the fact that the sessions were held in pairs. However, two levels of groups can be created by

combining all of the lower levels together (groups 3, 4, and 6) and all of the higher levels together (groups 7, 8, 9, and 10). In order to create two levels with equal numbers of participants, group one has been excluded from these calculations. A Spearman's Rho test can now determine if there is any significant difference between the number of tokens and potential tokens between the lower-level groups and the higher-level groups.

First, the number of tokens linked in a standard fashion in the lower-level groups was compared to the number of tokens linked in a standard fashion in the higher-level groups. The Spearman's Rho test showed that there was no significant difference between the two levels ($r=-0.15544$, $p=0.55135$). This suggests that the participants from the lower study levels use /j/, /w/ and /r/ sandhi as much as those in the higher study levels. It suggests that if there is a learning process for such sandhi, it is learnt very early on.

Next, the number of potential tokens was also compared between the lower-level groups and the higher-level groups. The two groups were compared to see if the number of occasions which allowed the usage of sandhi was different between the two groups i.e. do the lower-level groups have less opportunity to create sandhi? The Spearman's Rho test again determined that there is no significant difference between the number of potential tokens in the lower-level groups and the number of potential tokens in the higher-level groups ($r=-0.16929$, $p=0.51598$). As expected, this suggests that the number of potential tokens created by the participants was equal regardless of the level at which they are currently studying.

4.3.4 Discussion on Part One

This activity did not use any written material so it is fair to say that the participants were not affected by any visual representation of the language. Some may argue though that particularly the speech of a visual learner could still be influenced by having initially learned a word combination through text. It is possible that some visual learners imagine the, previously-learnt, written form of the language as they speak.

For the /r/ sandhi, the results generated from the NEST sessions were insignificant, but this report can compare the usage of the /r/ sandhi to a similar study done with native English speakers (non-rhotic). Mompea and Mompea (2009) found that the average token rate of linking /r/ for native speakers was between 56-59% depending on the demographic of the speaker. The participants in the present study had a token rate of 33%. The average token rate for the intrusive /r/ was 24-35% in their study whereas the in the current study, no intrusive /r/s were recorded. The token rates for both linking and intrusive /r/ are much lower than reported in the study by Mompea and Mompea (2009).

It can be concluded even at this early stage that learners are indeed using linking techniques. The learners may be hampered by the need to pause more in order to construct sentences, but this effect is less prevalent than expected. It is too early to say whether the rates of success are comparable with their native speaking counterparts, but perhaps Part Two will provide enough data to make this conclusion.

4.4 Part Two: Reading

The focus of Part Two was on the difference between the sandhi usage when participants are allowed to read a sentence at their leisure and when they are asked to speed up the production of the sentence. Each participant read the list of sentences twice; they were asked to read faster the second time. The written form of the language was visible to them this time and because the students are not being asked to produce original language, pausing for consideration time should have less of an effect.

It is important to remember here that the NESTs linked sentences between 66.7% and 91.7% of the time (on the first reading). On the second reading, a NEST increased to 100% of correct sandhi use.

4.4.1 /j/ sandhi

The linking /j/ sound was explored using the utterance "She likes to play a dangerous game". 49 participants correctly inserted the /j/ sound between the word "play" (which usually ends with the vowel sound /eɪ/ unless followed by another vowel sound) and the vowel sound /ə/ from the word "a". Only 25 participants did not insert the /j/ sound. Most of the participants who did not insert the /j/ sound failed to eliminate the hiatus and used a glottal stop to separate the vowel sounds instead.

The intrusive /j/ sound was explored using the utterance, "He is a very happy man." 46 participants correctly inserted the /j/ between the words "he" and "is". Only 4 failed to insert the /j/ sound while 26 avoided the need for insertion by using "he's a", which is phonetically /hiːzə/. This means that allowing for the avoidance, there was a 92% token rate. If avoidance is deemed

to be a negative result, there is a 60.5% token rate. The author considers the use of a contraction to be a success in terms of speaking naturally and therefore would promote consideration of the 92% token rate.

4.4.2 /w/ sandhi

The intrusive /w/ sound was correctly produced on 36 occasions (not produced on 39 occasions) in one utterance and on 66 occasions (not produced on 9 occasions) in the second utterance. The linking /w/ was not assessed in Part Two.

In the first utterance, "go (/w/) and see what is happening", the issue here is that the word "go" is often mispronounced as the mid-back rounded vowel (/o̞/) due to first language interference. The /w/ sound is only produced after high-back vowels in English. This means that it would become unnatural to pronounce the intrusive /w/. So in the case of this particular utterance, it seems that the mispronunciation of the surrounding vowels decreases the probability of the participant producing the intrusive linker.

The second utterance is quite different. This utterance uses simple vocabulary, which the participants are highly familiar with. The phrase "do (/w/) it now" was produced quickly by participants and was correctly linked most often. 66 participants correctly inserted the /w/ and only 9 failed to do so. This is a token rate of 88.0%.

So if the hypothesis that being forced to speak faster leads to more linking and connection is true, then it can be extended to allow for the possibility that short simple phrases that learners are quite familiar with and confident in

pronouncing can also lead to more correct linking. This follows on the assumption that linking is a natural phenomenon that results from fast speech.

4.4.3 /r/ sandhi

4.4.3.1 Linking /r/
Before a discussion about the linking /r/ sound is embarked upon, it is important to be reminded that the students would likely have had varying degrees of contact with both rhotic and non-rhotic teachers. The more influenced the students are by their rhotic teachers, the more they are likely to produce /r/ sounds in the terminal position where vowel reduction in non-rhotic speech would reduce the sound to a schwa (/ə/). This means that the influence of rhotic speakers in their background is potentially a very influential factor in whether participants use linking /r/ or not.

The phrases used to explore the linking /r/ sound were, "The teacher is a woman", "The car is over there" and "We looked far and wide". The token rate on these utterances is quite low. The token rates for the three utterances were 12/20/12 respectively. On the other hand, 63/56/61 participants did not insert the /r/ sound (with 1/0/3 utterances excluded from analysis because an unexpected speech event occurred which made the link impossible). This gave token rates of 16.0%/26.3%/16.4%. Most of the participants who did not insert the /r/ sound in these utterances used a glottal stop and therefore did not remove the hiatus that interrupted their flow of speech.

4.4.3.2 Intrusive /r/
As expected the /r/ sound was the one that stood out as different. The /r/ sound was explored using the utterances "I like draw(/r/)ing people" and "I saw (/r/) a man in the window". The first is an internal linker (the /r/ connects two

vowel sounds within the same word) and the second is external (across the word boundary).

In both utterances, there were very few cases of successful /r/ intrusion. In most cases, the participant actually replaced the /r/ sound with a /w/ sound. This happened on 56 occasions for utterance 1 (73%) and 36 occasions for utterance 2 (47%). The reason for this seems obvious. The participant was reading the sentences; when a sound was needed, the participant used the sound of the letter that was visually represented in the text that they were reading. Both "drawing" and "saw" contain the letter 'W'. In the absence of contradicting instruction, it seems like a logical step that the participant would assume that a /w/ sound would be the linker here.

This is where the intrusive /r/ differs from the intrusive /w/ and /j/. The glide consonants are a by-product of the articulation of the vowels that surround it (as discussed in Section 1.5 in the Introduction). The articulation of an /r/ sound is additional to the vowels around it.

The reason that the participants used a /w/ instead of an intrusive /r/ in the utterance "saw a" could be that the participants did not know which consonant was required in order to resolve the hiatus (as discussed in the Discussion section). However, it is also important to consider that the main articulator in making the /w/ sound is the lips, but the main articulator in making the /r/ is the tongue. Mispronunciation sometimes boils down to different cultural preferences. It is possible that the Japanese participants are simply more accustomed to using their lips as articulators rather than their tongue.

In pronouncing Japanese syllables, the tongue does not protrude the lips as it does in English /θ/ and /ð/. Therefore, it is possible that the Japanese are a little more averse to using their tongues and prefer to use their lips if possible.

For native English speakers, the production of the /r/ is faster and requires less effort as it involves just a small movement of the tongue. Using a /w/ to resolve the "saw a" hiatus involves a more pronounced movement of the lips. The lips are rounded at the front in order to make a /w/ sound.

In conclusion, the purpose of sandhi is to make the utterance faster and easier to pronounce. The production of an intrusive /r/ might be the easiest way for a native English speaker to resolve a hiatus, but it might not be the easiest way for people who come from a different culture or a language which more commonly uses different sounds.

4.4.4 Two effects at play

It is important to realise when we look at the results for this section that the difference between the first and the second reading for each participant is not solely explained by the speed of production alone. The other effect that influences the results is the familiarity effect. As the participants approach the sentence for the first time, there is more apprehension as the sentence may contain words that are unfamiliar or difficult to pronounce. The second time they approach the sentence they have the experience of having read it once before so they are more relaxed and more likely to produce it fluently.

Because of the nature of the experiment, it is impossible to measure the proportion of the change difference in scores which is due to the familiarity effect versus the change in the difference in scores due to the speed of

production. Even changing the sentences to remove the familiarity effect will not solve the problem because then it would be comparing unlike observations.

Because it is not possible to apportion the familiarity effect from the difference resulting from the production speed, it is essential that the familiarity effect is borne in mind when discussing these results.

4.4.5 Summary

The table below summarises the results from Part Two.

	Reading 1	Reading 2
LJ	60.5%	68.4%
LW	N/A	N/A
LR	12.3%	26.3%
NJ	63.2%	57.9%
NW	55.3%	78.9%
NR	6.1%	11.4%

Figure 7 Participants Part Two: Reading – Token rate

According to this table, almost all features increased their token rate on the second reading due to the participants speaking faster. This provides evidence to contradict the hypothesis that only the glides will increase in token rate as the learner speaks faster. Surprisingly, the intrusive /r/ increased in token rate in the faster reading, but the intrusive /j/ sound decreased.

In order to determine whether there was a significant statistical difference between the medians of the two sets of data, a Wilcoxon Signed Rank test was used. The Wilcoxon Signed Rank test was used to determine whether the median of the results from the first readings and the second readings were significantly different. The test concluded that there was a significant difference between the first and second readings of the participants ($z=-3.9697$, $p \leq 0.01$). This provides

71

evidence that when participants were asked to read faster, they produced more corrected linked phrases.

It appears that the orthographic representation of the utterance seems to influence how Japanese ESL learners use linkers and intrusive sounds. This is suggested by the learners using a /w/ sound in place of an /r/ sound in combinations which contain the letter 'W' in the text but actually require an /r/ sound to resolve the hiatus naturally.

It was theorised that ESL learners easily incorporated the glide consonants into speech because the glides are a by-product of the articulation of the surrounding vowel sounds. However, the intrusive /r/ is not a by-product of the articulation of the surrounding vowels, nor is it orthographically represented. It is therefore necessary for the students to notice that the native speakers around them use an /r/ linker in such cases. The identification of this intrusive /r/ sound could possibly be expedited through teachers raising awareness of these phonological features. Whether direct teaching of the glides and the linking /r/ sound expedites correct pronunciation is doubtful to this writer, but it is perhaps a question for further research.

4.5 Part Three: Listening

Part three was a recognition test whereby participants listened to 9 utterances and they were instructed to identify the extra sound. It was multiple choice and the participants we asked to circle /j/, /w/ or /r/. The NESTs achieved scores of 60% and 30%. In this particular test, I wonder if being a native speaker of English is much of an advantage. It is clear that the ability to speak English does not necessarily indicate an awareness of grammar or

phonological processes. People who learned to speak English as a child may have had no cause to study grammar and therefore may be unable to recognise tenses. It seems likely that the same is true with phonology. The ability to speak English does not necessary lead to an awareness of linking sounds.

The scores of the learners in this test ranged from 20% to 90%. The mean score was 55.8%. The real lessons learnt, however, come from examining the data question by question. The utterances that the participants listened to are listed below with the epenthesized sound in brackets.

1. Do (/w/) I care?

2. He (/j/) is off now.

3. At the ki(/j/)osk.

4. Draw (/r/) in your book.

5. I can't see (/j/) in the dark.

6. I saw (/r/) it.

7. Six grasshoppers and two (/w/) ants.

8. I (/j/) entered quickly.

9. ~~I (/j/) ate raw (/r/) eggs.~~

10. I didn't go (/w/) in.

The data collected from question nine was rejected from the set because there are two possible answers. Unfortunately, this was not noticed until after the sessions were completed.

Question 1	53%
Question 2	34%
Question 3	58%
Question 4	55%
Question 5	32%
Question 6	61%
Question 7	58%
Question 8	50%
~~Question 9~~	~~66%~~
Question 10	92%

Figure 8 Participants Part Three: Listening – Question analysis

Question 2 and question 5 stand out in Figure 8 as being the lowest results and they are both /j/ sandhi questions. In fact, if we look at the token rate of the three sounds, the /j/ sandhi was the most difficult for the participants: they correctly identified /j/ sandhi only 43.4% of the time. They correctly identified /w/ sandhi the most at 67.5% of the time and /r/ sandhi 57.9%.

It is important to note here that the participants achieved a particularly high average score in Question Ten. This may be the result of a practice effect. As the participants listened to more sentences, they improved at the task.

In the following section, these results will be discussed. This will include a comparison between the participants' free speech and reading results, the relationship between the sound and the token rate, a comparison between the participants' and the NESTs' results and a discussion about some of the issues that arose in this study.

Chapter Five:

Discussion

5.1 A comparison between free speech and read production

The critical difference between Part One: Free conversation and Part Two: Reading was that in conversation, the participant was required to contribute the ideas and language. When the participants were reading, the ideas and language were given, the participant only needed to reproduce them by reading them off the page. This section will examine the differences between the language that the participants constructed and the language they simply read.

When the participants were asked to speak in Part One: Free Conversation, they were required to construct the language from what they had previously learnt. This means that the linkers and intrusive sounds that occur in simple sentence structures and structures that are commonly taught in the lower-ability levels are potentially highly over-represented in the analysis of Part One.

The utterances in Part Two were designed by the researcher in order to elicit specific sounds. The researcher chose sentences where the linker or intrusive sound was in the middle of the sentence to increase the chances of tokens. However, when constructing the sentences themselves, the participants

chose to use sentences where the linker appears before the start of a new clause, between items in a list and other places where it would be natural for even a native English speaker to pause.

The table below shows a comparison of Parts One and Two in order to determine whether there is a significant difference between linking in reading activities and linking in free speech activities.

	Reading 1	Reading 2	Free speech
LJ	60.5%	68.4%	28.6%
LW	N/A	N/A	82.6%
LR	12.3%	26.3%	15.5%
NJ	63.2%	57.9%	54.7%
NW	55.3%	78.9%	19.0%
NR	6.1%	11.4%	0.0%

Figure 9 Comparison of Part One: Free Conversation and Part Two: Reading – Token rate

We see that all of the sounds except intrusive /j/ (in red) increase in the second, faster reading. This gives support to the theory that the rate of speech is a factor in the token rate. In other words, participants use sandhi more when they speak faster.

In the case of intrusive /r/, although participants correctly produced it 6.1% of the time in the first reading and 11.4% of the time in the second, faster reading, in the free speech exercise, it barely occurred at all. Perhaps this lack of

practice leads to a low rate of success or perhaps the low rate of success leads to avoidance of combinations that require it.

The token rates in the free speech samples tend to be lower than in the reading samples and that could be because the participants were given the utterance. Being given the utterance means that the participants don't need to consider the content, grammar, and vocabulary in order to construct a sentence. The participants are likely to be more confident reading a sentence that they know to be correct because it was written by a native speaker and that confidence probably means that they read it faster.

5.2 The relationship between the sound and the rate of identification

Despite the potential of the practice effect being a highly influential factor, it seems that there may be a relationship between the sound and the rate of identification. To examine the rate of identification, we will return to the results from Part Three: Listening.

Intrusive sound	Number of occurrences	Token rate
/w/	3	67.5%
/j/	4	43.4%
/r/	2	57.9%

Figure 10 Participants Part One: Free Conversation - Identifying sounds by phoneme

Figure 10 (above) shows the number of times each sound occurred in the test and the token rate for each sound. The table provides further argument that the /j/ sound is more difficult for Japanese ESL learners to identify and that the

/w/ sound is easier. If this is compared to a similar table made from the data in Part Two: Reading, some differences become clear.

Intrusive sound	Number of occurrences	Token rate
/w/	2	67.1%
/j/	1	60.5%
/r/	3	8.8%

Figure 11 Participants Part Two: Reading - Production of sounds by phoneme

It is clear that the /r/ sound was more difficult for the participants to produce than identify. At least when they are instructed to listen for it, participants can hear the /r/ sound in a sentence and the theory suggests that noticing should lead to production. Two possible reasons exist that could explain why awareness does not lead to production. Firstly, they had never been instructed to listen for the sound before, and it is possible that since the session, participants have become aware of this sound and have begun the process of incorporating the /r/ sound into their speech. Unfortunately, a follow-up session to test this is not possible, as participants have since returned to Japan. Secondly, some intervening factor means that in this case, awareness does not result in adoption of the intrusive sound. This intervening factor may be difficulty in pronouncing the /r/ sound (entirely possible in this case given the notorious difficulties Japanese ESL learners have with learning to produce /l/ and /r/ sounds). Alternatively, it may be issues with assimilating the /r/ sound into the

sound environment or even social factors. The production of /r/ sandhi may be looked upon by others as "lazy" or "incorrect" speech.

It is clear that the ability to produce a sound and the ability to identify the sound are not as linked as one might assume. This is especially true with regards to /r/ sandhi. The issues that Japanese ESL learners have with producing sandhi were discussed earlier in Chapter One.

5.3 Comparing participants' token rates to NESTs'

For all three parts, the token rates for the participants were significantly different from those of the NESTs. In Part One: Free Conversation, there were more tokens for all three sounds during the NEST sessions than the participant sessions. The average difference between the two sessions for each sound was 36.6% in favour of the NESTs.

For the entirety of Part Two: Reading, the participants realised 41% of the tokens whereas the NESTs realised 83.4% of the tokens. This demonstrates that in speaking, both free conversation and when reading text, the NESTs used linkers and intrusive sounds more than the learners.

Part Three: Listening was the exception. In this part, the participants outperformed the NESTs. The participants achieved an average score of 55.8% whereas the NESTs achieved an average score of 45%.

These results illustrate that although the participants are using linking and intrusive sounds, they are not using them with the frequency with which the native English speaking teachers use them. It is debatable, however, whether this is because the participants were learners or whether the participants used sandhi less because of the teacher/student relationship. It is possible that the

teacher/student relationship, the taboo on "lazy" speech or any of a range of other possible sociocultural factors prevented the learners from using more sandhi in the sessions.

5.3 The relationship between the students' level of study and the rate of successful linking

The participants hail from a tertiary institution in New Zealand where the students (all arriving at the same time) have spent about 6 months living and experiencing New Zealand culture. They all study at the same institution and do a similar workload. However, their experience before coming to New Zealand is varied and, combined with the different personalities of the participants, this means that their level of study is varied between what this report has termed *false beginner* to *upper intermediate*. The participants are assigned to classes which are labelled Group One to Group Six (Group One being the lower-level students and Group Six being the higher-level students). The students are assigned to classes based on their TOEIC® score and their performance in institutional language proficiency tests.

One of the research questions was *does the students' ability to identify or produce intrusive sounds correlate with the level at which they are currently studying English?*

In Part One, it is clear simply from the inspection of the data on participant token rate in Figure 12, below, that there is no clear correlation between study groups and token rate of sandhi in free conversation. This means that a participant from the lowest group (Group One) was just as likely to resolve a hiatus as a participant from the highest group (Group Ten).

Group One	36.8%
Group Two	No data
Group Three	13.9%
Group Four	23.8%
Group Five	No data
Group Six	45.5%
Group Seven	48%
Group Eight	30.7%
Group Nine	14.3%
Group Ten	25%

Figure 12 Participants Part One: Free Conversation – Average results by group (Students'
study levels)

In Part Two, it is again clear that the mean token rate of the lowest level
was not significantly different from the mean token rate of the highest level of
study. The provides some evidence against the assumption that word connection
is developed alongside other English abilities as the student receives more
exposure to English. Instead, it appears that whenever participants are able to
string two potentially token-forming words together with some degree of
fluency, the intrinsic relationship between the two words may encourage the
learner to link them. The average token rates for participants in Part Two:
Reading are shown in Figure 13 overleaf.

Group One	6.0
Group Two	No data
Group Three	5.7
Group Four	5.3
Group Five	No data
Group Six	5.7
Group Seven	5.3
Group Eight	6.1
Group Nine	5.3
Group Ten	5.5

Figure 13 Participants Part Two: Reading – Average results out of ten by group
(Students' study levels)

One factor that may help to explain this is the notion of function vs. content language. It appears that there is a strong link between sandhi and the purpose of the word. Simply stated, we can categorise each word as either a content word (which gives information such as nouns, verbs and adjectives) or a function word (which serves a grammatical purpose such as prepositions, pronouns and determiners).

Alameen (2007) demonstrated that sandhi occurs more frequently with function words than with content words. In Part 2: Reading, the participants were asked to read twelve sentences. If we remove the two that tested for gemination and one outlier (see below), the students read three sentences that involved only content words and six that included both content and function words. The outlier was the sentence that few participants produced fluently and no one managed a standard connection. About fifty percent of the sentences that contained both content and function words were linked or connected with a standard sound. Only 12.3% of the content-only word combinations were connected.

It is possible that the participants are more likely to pause before a content word as they require time to recall the vocabulary. However, the function words are more commonly used so "there are" is likely to be more familiar to the participant and therefore more likely to be linked than "to engineers" which may require a moment to recall the noun. In fact, it could also be true of native speakers that functional words are linked more because of the stress-timed nature of the English language. Being a stress timed language, English tends to speed up the production of groups of non-stressed words in order to maintain the rhythm.

The results of Part Three support the findings above that there is no significant correlation between level of study and linking token rate. Part Three: Listening results show that there is no significant correlation between the groups in terms of ability to hear the epenthesized consonants.

Group One	70.0%
Group Two	
Group Three	56.7%
Group Four	52.5%
Group Five	
Group Six	50.0%
Group Seven	43.3%
Group Eight	55.6%
Group Nine	73.3%
Group Ten	45.0%

Figure 14 Participants Part Three: Listening – Average token rate by group (Students' study levels)

The results shown in Figures 12, 13 and 14 provide compelling evidence that the usage and perception of sandhi is not correlated with the current study

level of the participant. This means that there is no evidence to show that as a learner develops his or her ability to produce English, the ability to link develops alongside it. Equally in the lower levels and the higher levels, the participants could already link frequently but not as often as their native English counterparts. Therefore, if any "learning" occurs in order for learners to be able to correctly link /j/ sandhi and /w/ sandhi, it must occur before the learner becomes a beginner-level student.

5.4 Error vs. Mistake?

Penny Ur (2002) makes the distinction between an *error* and a *mistake* in the production of ESL learners. She claims that an *error* represents a gap in the learner's knowledge. An *error* is made because the learner does not have sufficient knowledge to correctly make the utterance. So a learner who says, "I have go to the park" is making an *error* if they made this error because they have not yet studied the present perfect tense. A *mistake*, however, is a slip caused by lack of concentration, speaking too fast or a mid-thought change in sentence structure.

This framework of distinguishing between a variance due to lack of knowledge and a variance due to lack of correct application of the knowledge is a useful one. It is important to realise that in the current study it would be inappropriate to refer to any usage or lack of usage of sandhi as a *mistake* or an *error* except, arguably, in the case of when a non-standard consonant is used. Bear in mind that there is variance in the way native English speakers use /r/ sandhi in terms of rhoticity.

The NEST sessions provided evidence that even native English speakers make mistakes when they are speaking. NEST One used the weak form of the word *the* before a word that started with a vowel. Usually the weak form /ðə/ is used before a consonant and the strong form /ðiː/ before a vowel. Later questioning confirmed that NEST One, being an English teacher, knew the rule and was capable of applying it so it can be confirmed that the usage of the wrong form was a *mistake* as opposed to an *error*. This is evidence that even native speakers make mistakes in their speech due to lapses in concentration etc.

It therefore becomes a consideration whether unlinked tokens were not linked because of lack of sufficient knowledge to make the sandhi or if the token was not linked because of a slip or a natural pause in speech comparable to a native speaker.

It appears to be the case that intrusive /r/ was often not connected because the participants did not know which consonant needed to be inserted in order to resolve the hiatus. When the participant made the utterance "Grandma is", they failed to insert the intrusive /r/ and this is likely because they had no idea that it was an /r/ that was necessary to make this utterance sound natural. Similarly, in Part Two, when the participants were asked to say "drawing" and "saw a", they did not know that they should insert an intrusive /r/ into these and 70.9% and 45.6% (respectively) inserted a /w/ sound instead. I would suggest that this was an error not a mistake in many of these cases as the participants lacked the knowledge to produce the standard sandhi.

In terms of the glides (/j/ and /w/), it could be argued that because they are a bi-product of the transition of the articulators, there is no possibility of the participants lacking the knowledge to produce them. It should be a natural

acquisition. In other words, because the /w/ sound is naturally made when the speaker moves their mouth from the /əʊ/ to the /aʊ/ in "go out", they will not need to be instructed to make this sound.

In conclusion, it is likely that the intrusive /r/ alone could be seen as sometimes being omitted in error (when a non-standard consonant is used) where missed opportunities to use the glides would more likely be simple non-compliance as is consistent with native speakers.

5.5 Wrong or different?

The occurrence of the /w/ in the utterance "saw a" represents a departure from the way in which the native English speakers pronounced the utterance. This poses the question of whether the learner's version of the utterance is incorrect and the native English speakers' version is correct. This section discusses the possibility of making an allowance for non-standard speech.

Stating that any departure from the way in which a native speaker would pronounce something is taking a prescriptive view on the situation. A prescriptive perspective on language is to look upon grammar as a set of rules that prescribe how a language should be spoken. This contrasts with a descriptive perspective under which grammar is a set of rules, which explains how language is actually used. The question is, does the usage determine the grammar or does the grammar determine the usage?

Just because the participants' version differs from the "natural" version, does this mean that it is incorrect and therefore needs correction in order to be acceptable? Or should we take the descriptive perspective and state that when the participants produce the utterance "saw a" with a /w/, they are simply

creating their own version of the language. *Learner language* is a term that many ESL teachers use to define the language that learners use that differs from that of native speakers. Throughout this study, I have used the term *non-standard* to describe learner variations from native pronunciation.

There is variation between different dialects of English. American English differs from British English, and even within countries different dialects exist. The English spoken in Auckland, New Zealand differs somewhat from the English heard in the Canterbury Plains (New Zealand). Therefore, without a single standard with which to compare speech, it could be argued that the whole concept of *correctness* is inappropriate in this context.

More and more we are tending to recognise different accents of English, including those from countries which primarily speak other languages. After all, Harmer (2007) states that to the majority of competent English speakers, English is a second language. English has been enriched with dialects such as Indian English and Pilipino English.

As our recognition of different forms of English expands, perhaps the /w/ sound in the utterance "saw a" will become recognised as a feature of the way Japanese people speak their version of the English language rather than an error that needs to be corrected. The utterance, after all, is still comprehensible when pronounced either way.

5.6 Easy to pronounce, easy to miss?

There is a paradox evident with regards to the usage and identification of /j/ sandhi. On the one hand, it was the most difficult for the participants to identify when listening to utterances produced by native English speakers.

87

However, it was also the most easily produced by the participants when reading set sentences. The intrusive /j/ was also the second most often produced by participants in free speech with the linking /j/ third.

It is easy to look at the simplicity with which the /j/ sound is produced and wonder if it is perhaps overlooked in the identification test because it is such a subtle sound. The production of a /j/ sound requires little movement on the part of articulators. The rise and fall of the tongue is all that is required to articulate a /j/ sound. Also, the articulating movement of the tongue is concealed within the mouth. In face to face conversation, the /w/ sound would be easier to identify because of the wide movement of the articulators, the rounded lips. However, Part Three: Listening was conducted using a recording of the sentences so these visual clues were absent during the sessions.

Another consideration is how different the intrusive sound is to the sound created by the surrounding vowels. The /j/ sound does not contrast with the surrounding vowels as much as an intrusive /r/ does. Also, because the /r/ sound is not expected in the utterance "saw a", it might therefore catch the listener's attention more than /j/ sound.

It seems that because the glides are *low-level articulatory transitional phenomena* (Newton and Wells, 2002), they are likely to be the path of least resistance between the two vowel sounds. By the path of least resistance, I mean that a consonant sound is produced with the least amount of movement of the articulators. It appears that the semivowels are sounds made by transitioning the articulators, which has been assigned the title of *consonant* even though they are not entirely similar to the other consonants.

The semivowels are not only similar in sound to the surrounding vowels but also are bi-products of the articulatory transition whereas the /r/ sound does not appear to occur because of the articulatory process of the surrounding vowels. This brings into question the /r/ sound and why it is used to resolve hiatus at all. In non-rhotic languages there is an underlying /r/ in word final position which is only articulated when the following word begins with a vowel (Hay & Sudbury, 2005). Therefore, if the theory that non-rhotic languages are a divergence from rhotic languages is true, this theory explains the linking /r/ but not the intrusive /r/. The mystery of of the intrusive /r/ is compounded by the fact that in the utterance "saw a", the /w/ sound could just as easily resolve the hiatus as an intrusive /r/. These factors mean that the intrusive /r/ is adventitious in the utterances, which make them easier for the participants to identify.

It is possible that the less a consonant epenthesis contrasts with the vocalic environment in which it occurs, and the less movement is required of the articulators compared to the preceding and following vowel sounds, the more it blends in with its environment and the more difficult it is for learners to build awareness of it.

In conclusion, there is a disparity between the likelihood of a participant perceiving linking and intrusive sounds and the likelihood of the participant producing them. The probability of a participant perceiving linking and intrusive sounds is likely to be linked to the contrast between the consonant sound and the surrounding vowel sounds and the amount of movement required to articulate that sound in contrast to the vowel sounds before and after it. The probability of a participant producing the linking and intrusive sounds is likely to

be dependent on a range of factors including their speaking speed and degree of formality.

5.7 Concurrence with previous research

In this section, I will compare previous research with the current study. Hieke (1984) and Alameen (2009) are the two most significant studies in this area because their research is based on ESL learners.

Hieke's (1984) quantitative study found that 80% of the potential tokens were linked or connected by native English speakers yet just over half of the potential tokens were linked or connected by learners. These figures are comparable with the discoveries of this study. As mentioned in Section 5.3, 41% of potential tokens in Part Two were realised by the participants whereas 83.4% of the potential tokens were realised by the NEST participants.

Alameen (2009) demonstrated a statistically significant difference between the use of consonant-attraction and vowel-to-vowel linking by native American English speakers and non-native English speakers. In contrast, however, she also found no significant difference between spontaneous and reading speech styles. The present study noted both a difference between the rates of success in free speech and reading speech styles for the glides, but it did not note a difference for the anomalous /r/ sandhi.

Alameen also found no significant difference between the two groups that she investigated. One was a beginners' class and the other was an intermediate class. No significant difference suggests that the rate at which the participants used linking was not directly proportional to the amount of time they had spent studying English. This concurs with the results of the present study in that there

was no correlation between the current level of study and the rate of correct sandhi usage. Alameen explained that since the textbooks used in the course from which her participants were drawn had no focus on linking, it was logical that the intermediate group would fare no better than the beginner group. However, in the present study, the participants had all been in attendance at the institution an equal length of time; specifically, six months. Their current level of study is based more on the length of study before coming to New Zealand and their personal capacity for learning English.

5.8 Is the same token rate actually less sandhi?

One thing to consider when looking at the results of the sessions is if there is no significant difference between the groups, then is it prudent to assume that there is no actual difference.

Given that leaners in the lower levels do pause more to consider and plan their sentences as they produce them, it is fair to expect that the higher levels would connect their words more. Therefore, any statistical analysis that concludes that there is no difference between the word connection of the two levels despite the lower-level learners pausing more is actually suggesting that the higher levels connect less. The possible reasons for higher-level leaners connecting less therefore need to be considered.

Firstly, it is possible that the Japanese language uses word connection and the connection that is apparent in the lower-level learners is actually a kind of first language interference which causes both accidentally standard connections and non-standard connections. The learners may therefore decrease the amount

of word connection they use to remove the non-standard cases as they develop their English.

Secondly, it is possible that the learners develop their skill at linking their speech in appropriate contexts, but they perceive a session at which a teacher is present to be a situation in which linking is inappropriate. As discussed in the literature review, even in English-speaking communities, the intrusive /r/ is stigmatised to varying levels. If this is the case, learners may tend to revert to using language they consider to be more formal any time they feel they are being observed by authority. A clear example of this is *gonna*. *Gonna* is spoken English and considered to be slang. It is a form of sandhi which ESL teachers tell their students is only to be used in an informal situation; for example, only with friends. So when learners are being recorded as they speak to a teacher or with a teacher present, it is likely that they would use the more formal version of *going to*. It may be the same with epenthesized consonants.

The lack of evidence for the hypothesis that there is a significant difference between the groups of different study levels raises questions about why there is no difference. It is logical to expect that the higher the fluency of the participant, the less they pause and therefore the more they connect their words into a stream of speech. It is possible that there is a counter-acting effect which negates any advantage to the higher-level participants. Alternatively, it may be that lower-level participants do not pause when there is a link to be made.

It is possible that there is an intrinsic connection between two words which have a high chance of being linked that even lower-level participants are unlikely to break with a pause. This raises a question for further research: do the pauses of a learner correlate with the links that are made by native English speakers?

92

Are learners less likely to pause in a word combination that is likely to be linked by a native speaker, or a learner for that matter?

5.9 Summary

It appears from the data above that the level at which the participants were studying at the time of participating has no correlation to their performance in either sandhi production or sandhi identification tests.

There does seem to be a relationship between the sound and the participants' ability to identify it in the listening test. It seems that the /j/ sound is more difficult to identify and the /w/ sound is easier to identify. It is also important to note that the /r/ sound was not produced in many cases in Part Two, but it was identified often in Part Three.

Chapter Six:

Conclusion

This investigation made the following discoveries. Japanese ESL learners use and understand sandhi but not to the same extent as native English speakers. These discoveries are perhaps not very surprising. However, it was also discovered that some sounds are easier for the learners to identify in speech and these sounds are not the sounds that they produce the most. While speaking speed is a factor in the rate of sandhi usage, the study level of the learner is not. This section will examine each of these points in turn.

In response to research question number one "Do Japanese ESL students use sandhi when speaking in conversation or reading?" This report offers the following conclusions. It is clear that the participants did indeed use sandhi techniques to varying degrees. It is clear that learners use the /w/ and /j/ sounds much more commonly than the /r/ sounds with the intrusive /r/ being used the least. This is a logical finding since, as discussed in the introduction, the /w/ and /j/ sounds are by-products of the transition of the articulators between the two surrounding vowel sounds so it is quite natural for the participants to start to produce these sounds quite early on in their English studies. Also, the

sounds that create a hiatus requiring a glide to resolve it are more common than the sounds that create a hiatus that require an /r/.

However, this study also concurs with Hieke (1984) that there is a substantial difference between the quantity of tokens produced by learners and by native speakers. One factor in explaining this difference is the knowledge gap where learners are not aware that an intrusive /r/ is necessary; however, since very few cases of intrusive /r/ potential tokens were recorded, it is likely that this factor contributes little to the difference.

In response to the second research question, "Can Japanese ESL students identify intrusive sounds when spoken by a native English speaker?" Again, the answer is 'yes' in varying degrees. It is encouraging to see that the average score out of nine in the listening test was 5.6. This makes the average percentage of correctly identified sounds 62%. On average, the participants could listen to a complete sentence and identify the correct sandhi 62% of the time. This is significantly more than the 33% they could have attained by guessing. This demonstrates some skill on the part of the learners.

In the same listening test, the two NESTs got 30% and 60%. The NEST obtaining a score of 30% admitted that she had not fathomed the objective of the listening test, the other NEST had. This suggests that being fluent in English was not an advantage for the NESTs. Instead, the advantage lies in understanding the focus of the activity.

In response to the third research question, "Does the frequency of sandhi use correspond with the participants' current level of study?" the answer is a resounding 'no'. In all 3 sections of the study, there was no correlation found between the results and the participants' current level of study.

This means that sandhi is not a skill acquired as the learner progresses though the levels, rather it is a phenomenon that occurs with inconsistency from the time the learner starts confidently producing multiple word utterances. Pausing does not interrupt sandhi as much as expected, possibly because the intrinsic connection between linkable sounds makes any linkable word combination an unlikely place for a speaker to pause.

In response to the final research question, "Does the frequency of sandhi use correspond with the speed of production?" I conclude that participants increase their token rate when they increase the speed at which they produce an utterance. This increase was demonstrated in all of the six types of sandhi except for one, which provides further evidence that the rate of speech is a factor in whether a learner will use sandhi on a particular occasion but their length or level of study is not.

The first hypothesis was, "... that the probability of the /j/ and /w/ sounds occurring will naturally increase as the production speed of the learners is artificially increased suggesting that the learners will increase the probability of sandhi usage as they develop their fluency."

This hypothesis was partially correct. The /j/ and /w/ sounds did indeed increase as the participants increased the speed of their production. However, the /r/ was noticeably absent in the hypothesis as it was assumed that the /r/ would be different to the /r/ sound is not a by-product of the transition of the articulators. No evidence was found to suggest that the token rates of either linking or intrusive /r/ does not increase as the participants' speech was sped up. However, the token rate of the /r/ sandhi was lower initially as the participants were often unaware that an /r/ was required to resolve the hiatus

in "saw a". In conclusion, if a learner is aware that an /r/ is required, they are more likely to produce the /r/ at faster rates of speech. However, if the learner is unaware that /r/ is required, they will be unable to connect the utterance at any speaking speed.

The participants produced more of each type of sandhi at faster speaking speeds but since no link between sandhi usage and level or sandhi usage and length of English study was discovered, it is no longer logical to assume that sandhi increases as a learner develops their fluency.

The second hypothesis was that the rate of sandhi usage will correlate with the student's current study level, but this correlation is expected to exist more because the participant's study level is a weak indicator of the amount of English to which they have been exposed. The correlation is suspected to be quite weak in order to allow for individual experiences, differing profiles amongst the students (in terms of their various abilities) and varying levels of exposure to English environments. As discussed above, there was no correlation found between the students' current study level and their usage of sandhi.

One discovery was made during the course of this project. It was discovered that the participants were using a /w/ sound as a linker in place of an intrusive /r/ in words like drawing. This is because the participants were being influenced by the orthographic representation of the word. When they see a 'W' written in the word, they say a /w/ when an intrusive /r/ is actually required.

The results of this study also support the theory that students naturally "pick up" sandhi without direct instruction. This is supported by the fact that the participants used sandhi regardless of their current level of study; they used sandhi more with short sentences that they were highly familiar with and they

increased the rate of successful sandhi use when asked to speak faster. However, I would also like to point out the exception here and recommend that ESL instructors increase awareness of /r/ sandhi. There does seem to be a knowledge gap with intrusive /r/s that does not exist with the glides because as Newton and Wells (2002) stated they are *low-level articulatory transitional phenomena.* When a learner does not know which sound is required to resolve a hiatus, a knowledge gap is present which can be eliminated by an instructor. Otherwise, it seems that there is no need to spend time teaching linking and intrusive glides and linking /r/s to Japanese learners as they have already begun using them even at the beginner level of study.

6.1 Limitations of the current study

Upon reflection, there are a few things that I would do differently with regards to the session design. The omission of any texts that include a linking /w/ in Part Two was an error. Also, one question in Part Three was removed from the results because the utterance, "I ate raw eggs" actually contains two different forms of sandhi: an intrusive /j/ can be used to resolve the hiatus between "I ate" and an intrusive /r/ can be used to link "raw eggs".

In Part Three, the participants were instructed to guess if they were not sure of any answer. This was to ensure that the participants' uncertainty did not reduce the number of correct answers. However, it also meant that one third of the guessed answers were potentially correct and this artificially inflated the number of correct answers.

6.2 Recommendations for further research

The purpose of this research was to help to build understanding of sandhi and how it is used by non-native speakers. There is still a lot of research to be done before we can answer some of the important questions in this field.

- Does direct instruction on sandhi help learners to sound more natural?
- Is sandhi pedagogy a time-effective activity for the classroom?
- Or is sandhi something that is best *acquired* rather than *learnt*?

But these are not the only questions that could be researched. The glottal stop is one of the avoidance strategies when learners do not know the correct consonant to insert. Native speakers use glottal stops as well, but according to Alameen (2009) native speaker glottal stops are less frequent and less audible than learners' glottal stops. There is more that can be uncovered here with research into the duration of glottal stops.

We assume that a difference between native speech and learner speech is an error and needs to be corrected, but surely correction is only necessary when the difference inhibits the understanding of the audience. Another question for the future researchers might be *does a learner's failure to produce sandhi inhibit the understanding of native speakers listening to that learner?* Or other learners listening to that learner for that matter.

Also differences in other forms of sandhi such as gemination and assimilation may lead to a better understanding of the differences between learner English and native English.

Finally, one further research question might be to investigate whether the student's internal motivation for studying English is related to their ability to

produce connected speech. Some learners perceive the goal of studying English to be achieving a native-sounding accent, but for other learners the goal is simply to communicate their ideas and to be understood. It may be the case that there is correlation between the learners' internal motivations for studying and their ability to connect their speech. This means that students who covet the *native* sound may spend more effort on listening and mimicking the sounds as they hear them.

6.3 Implications for the ESL classroom

For most readers of this report, the takeaway points will be how all this impacts the education of pronunciation. Whether or not sandhi needs to be directly instructed to Japanese ESL learners is key. This reader would assert that the answer to this question is a general, but not comprehensive, 'no'. The linkers do not need direct instruction because there is no confusion over which consonant is required in order to resolve the hiatus. In terms of linkers, the necessary consonant is the silent letter at the end of the preceding word. In terms of intrusives, the intrusive /j/ and /w/ also do not need direct instruction because they are *low-level articulatory transitional phenomena* (Newton and Wells, 2002). Because of this, the intrusive /j/ and /w/ need no instruction as the participants will naturally start making these sounds as they produce the sounds around them correctly and build the confidence and fluency to speak at *natural* speeds.

As expected, the intrusive /r/ was the anomaly. The participants often made an error by resolving the hiatus with a non-standard consonant. They often used the consonant that appeared in the written form of the utterance rather

than the consonant which (to a native speaker, at least) required the least effort to produce.

This implies that the participants would benefit from some kind of intervention on the part of their teacher. So what forms might such interventions take? A thorough educator might suggest that in order to fully understand how to use the intrusive /r/, the students must first develop their understanding of the English system of vowels (to understand which vowels are low and middle vowels, for example).

Another educator might suggest that too much background knowledge might just confuse and demotivate students and therefore take the approach that simply building the students' awareness of the intrusive /r/ will allow the student to begin to hear the intrusive /r/ in context and therefore the student may start to imitate the word combinations that include an intrusive /r/. Once the student has imitated some phrases with intrusive /r/, they will start to produce others with the intrusive /r/.

A third educator might argue that the students have much bigger issues to worry about in their language development. The intrusive /r/ would make a small impact on the students' comprehensibility and the amount of information that needs to be taught to build the context for the student would take a considerable amount of time so it would likely be prioritised right out of the curriculum in most cases. In the end, most teachers tend to focus on the biggest blocks that prevent their students from being understood and intrusive /r/ did not make the participants in this study much less comprehensible.

As discussed earlier, one of the pronunciation issues that is common with Japanese ESL learners is the pronunciation of the /r/. Speakers of Japanese often

consider the two liquid consonant sounds (/r/ and /l/) to be allophones (variations of the same sound). Learners may therefore be confused by the existence of two distinct letters and two distinct sounds. This may be a contributing factor to why there were few intrusive /r/s produced in the sessions.

It can no longer be assumed that the intrusive /r/ did not often occur in the sessions solely because of the knowledge gap (the participant not knowing which consonant to use to resolve the hiatus). It may be that the production of the English /r/ is simply more difficult than the production of a /w/ and the participants, being human, took the path of least resistance. It remains unanswered whether the participants would have been capable of producing the /r/ sound required to resolve a hiatus with a likeness to that of a native speaker.

6.4 Summary

This investigation into the sandhi use of Japanese ESL speakers has deepened our understanding of the factors that arise in the development of this skill. It has shown us that the development of sandhi skills requires knowledge only for intrusive /r/. It has provided evidence that sandhi is not acquired in a linear fashion from the time the learner starts learning English to the time he/she gains a mastery of it. Sandhi, for the most part starts when the learner starts to confidently produce multiple word utterances and depends on situation, rate of speech and the function of the words produced.

Bibliography

Alameen, G. (2007). *The use of linking by native and non-native speakers of American English.* Ann Arbor, MI: UMI Microform.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgements of nonnative pronunciation and deviance in segmentals, prosody and syllable structure. *Language Learning, 42*, 529-555.

Antila, A., & Cho, Y. (1998). Variation and change in optimality theory. *Lingua, 104*, 31-56.

Bloomfield, L. (1933). *Language history: from language* (Rev. ed.). Ann Arbor, MI: Holt, Rinehart and Winston.

Broadbent, J. (1991). Linking and intrusive r in English. *UCL working papers in linguistics, 3*, 281-302.

Brown, A. (1988). Linking, intrusive and rhotic /r/ in pronunciation models. *Journal of the International Phonetic Association, 18*, 144-51.

Brown, A. (1988). Linking, intrusive, and rhotic /r/ in pronunciation models. *Journal of the international phonetic association, 18*, 144-51.

Brown, G. (1977). *Listening to spoken English.* London, UK: Longman.

Brown, H. D. (1994). *Principles of language learning and teaching.* Englewood Cliffs, NJ: Prentice Hall Inc.

Brown, J. D. (2012). *New ways in teaching connected speech*. Virginia, VA: TESOL International Association.

Brown, K., & Miller, J. (2013). *The Cambridge dictionary of linguistics.* Cambridge, UK: Cambridge University Press.

Bussmann, H. (2006). *Routledge dictionary of language and linguistics* (Vol. 1). (G. P. Traumath, K. Kazazzi, Eds., G. P. Traumath, & K. Kazazzi, Trans.) New York, NY: Routledge.

Content, A., Kearns, R., & Frauenfelder, U. (2001). Boundaries versus onsets in syllabic segmentation. *Journal of memory and language, 45*, 177-199.

Collins, B. and Mees, I. M. (2013). *Practical phonetics and phonology*. New York, NY: Routledge.

Crystal, D. (1992). *An encyclopedic dictionary of language & languages.* Oxford, UK: Blackwell Publishers.

Crystal, D. (2003). *A dictionary of linguistics & phonetics.* Oxford, UK: Blackwell Publishing Ltd.

Cutler, A., & Norris, J. (1988). The role of strong syllables in segmentation for lexical access. *Journal of experimental psychology, 14*, 113-121.

Derwing, T. M., & Munro, M. J. (1998). Evidence in favour of a broad framework for pronunciation instruction. *Language learning, 48*(3), 393-410.

Dumay, N., Frauenfelder, U., & Content, A. (2002). The role of syllable in lexical segmentation in French: Word- spotting data. *Brain and language, 81*, 144-161.

Fayol, M., Dugua, C., & Chevrot, J. (2009). Liaison acquisition, word segmentation and construction in French: A usage-based account. *Journal of child language, 36*(3), 557-96.

Fox, A. (1978). To 'r' is human? Intrusive remarks an a recent controversy. *Journal of the international phonetic association, 8*, 72-74.

Gick, B. (1999). A gesture-based account of intrusive consonants in English. *Phonology, 16*, 29-54.

Gieve, S., & Miller, I. K. (2006). *Understanding the language classroom.* Hampshire, UK: Palgrave Macmillan.

Harmer, J. (2007). *The practice of English language teaching.* Essex, UK: Pearson Education Limited.

Hay, J., & Margaret, M. (2012). /r/-sandhi in early century New Zealand English. *Linguistics, 50*(4), 745-763.

Hay, J., & Sudbury, A. (2005). How rhoticity became /r/-sandhi. *Language, 81*(4), 799-823.

Henrichsen, L. E. (1984). Sandhi-variation: A filter of input for learners of ESL. *Language Learning, 34*(3), 103-129.

Heselwood, B. (2009). R vocalisation, linking R and intrusive R: Accounting for final schwa in RP English. *The philological society, 107*(1), 66-97.

Hieke, A. E. (1984). Linking as a marker of fluent speech. *Language and speech*, 343-354.

Hill, C., & Beebe, L. M. (1980). Contraction and blending: The use of orthographic clues in teaching pronunciation. *TESOL Quarterly, 14*(3), 299-323.

Johansson, S. (1973). Linking and intrusive /r/ in English: A case for more concrete phonology. *Studia Linguistica*, 53-68.

Kaisse, E. M. (1985). *Connected speech: The interaction of syntax and phonology.* California, CA: Academic Press.

Kennedy, S., & Blanchett, J. (2014). Language awareness and perception of connected speech in a second language. *Language Awareness, 23*(1-2), 92-106.

Krashen, S. D. and Terrell, T. (1983). *The natural approach-Language acquisition in the classroom*. Oxford, UK: Alemany Press.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages.* Oxford, UK: Blackwell.

Leather, J. (1983). Second-language pronunciation learning and teaching. *Language teaching* 16(3), 198-219.

Lewis, J. W. (1975). Linking /r/ in the general British pronunciation of English. *The journal of the international phonetic association, 5*, 37-42.

Mahootian, S. (1997). Persian. *Anthropological linguistics, 41*(4), 574-577.

McMahon, A., Foulkes, P., & Tollfree, L. (1994). Gestural representation and lexical phonology. *Phonology*, 227-316.

McQueen, J. (1998). Segmentation of continuous speech using phonotactics. *Journal of memory and language, 39*, 21-46.

Mompea, P., & Mompea, J. (2009). /r/-liaison in English: An empirical study. *Cognitive linguistics, 40*(4), 733-776.

Mompean, J. A., & Gomez, F. A. (2011). Hiatus resolution strategies in non-rhotic English: The case of /r/-liaison. *Proceedings from the 17th International Congress on Phonetic Sciences (ICPhS).* (pp. 2-5). Hong Kong: University of Murcia, Spain and University College, London.

Newton, C., & Wells, B. (2000). Between-word junctures in early multi-word speech. *Journal of child language, 29*, 275-299.

Norris, J., McQueen, J., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive psychology, 34*, 191-243.

Ohata, K. (2004). Phonological differences between Japanese and English: Several potentially problematic areas of pronunciation for Japanese ESL/EFL learners. *(Unpublished doctoral thesis)*. Indiana University of Pennsylvania, Indiana, PA.

Pennington, M. C. (1996). *Phonology in English language teaching*. Essex, UK: Addison Wesley Longman Limited.

Pennington, M. C., & Richards, J. C. (1986). Pronunciation revisited. *TESOL Quarterly*, 1545-7249.

Pring, J. T. (1976). More thoughts on the /r/ link business. *Journal of the international phonetic association, 6*, 92-95.

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

Richards, J. C., Platt, J. T., & PLatt, H. K. (1992). *Longman dictionary of teaching and applied linguistics* (Vol. 1). Essex, UK: Longman.

Roca, I., & Wyn, J. (1999). *A course in phonology.* Oxford, UK: Blackwell Publishers.

Rosa, M. (2002). Don'cha know? A survey of ESL teachers' perspectives on reduced forms instruction. *Second language studies, 21*, 49-78.

Rudes, B. (1976). Lexical representation and variable rules in natural generative phonology. *Glossa*, 111-115.

Selkirk, E. (1995). Sentence prosody: Intonation, stress, and phrasing. In J. A. Goldsmith, *The handbook of phonological theory.* (pp. 550-569). Cambridge, MA: Blackwell.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics 30*(3) 510-532. doi:10.1093/applin/amp047

Stern, H. H. (1992). *Issues and options in language teaching*. Oxford, UK: Oxford University Press.

Swan, M., & Smith, B. (2001). *Learner English.* Cambridge, UK: Cambridge University Press.

The Economist. (2011, December 31). *Japan's cramming schools: Testing times*. Retrieved from http://www.economist.com/node/21542222

Trask, R. L. (1995). *A dictionary of phonetics and phonology* (Vol. 1). Oxon, UK: Routledge.

Trubetzkoy, N. S. (1969). *Principles of phonology: Translated by A. M. Baltaxe.* Los Angeles, CA: University of California Press.

Uffman, C. (2007). Intrusive [r] and optimal epenthetic consonants. *Language sciences*, 451-476.

Venneman, T. (1972). Rule inversion. *Lingua*, 209-242.

Weinstein, N. (1982). *Whaddaya say? Guided practice in relaxed spoken English.* Culver City, CA: ELS Publications.

Weinstein, N. (2001). *Whaddaya say? Guided practice in relaxed speech.* New York, NY: Addison Wesley Longman.

Wells, J. C. (1982). *Accents of English.* Cambridge, UK: Cambridge University Press.

Zahedi, H., Sahragard, R., & Zahra, N. (2007). The effects of phonological features on Iranian EFL learners listening comprehension. *Pan-Pacific association of applied linguistics, 11*(2), 115-130.

Appendices

Appendix A – Hand-out used in sessions

1. Draw a picture of your family. Use stick figures like Glenn did. Use your picture to introduce your family to your partner.

2. Read the following sentences:
1) I told him to stop pushing.

2) Go and see what is happening.

3) I like drawing people.

4) He is a very happy man.

5) I saw a man in the window.

6) I can't trust him.

7) I found a crab claw and a fish bone on the beach.

8) Do it now.

9) The teacher is a woman.

10) The car is over there.

11) We looked far and wide.

12) She likes to play a dangerous game.

3. Listen to the following phrases. What is the hidden sound? Tick one box for each question.

	/w/	/j/	/r/
a)			
b)			
c)			
d)			
e)			
f)			
g)			
h)			
i)			
j)			

Appendix B - Background survey

Name: _____   Date of birth: _____

Nationality: _____   Gender: Male   or   Female

1. How many hours per week do you study English in class? _____ hours

2. How many hours per week do you study English outside of class? _____ hours

3. How long have you lived in New Zealand?

   _____ years _____ months

4. How long have you been studying English in total? _____ years

5. Have you ever studied English at a private school or cram school? YES   NO

   a. How long? _____ years

   b. How many hours per week? _____ hours

6. Have you lived in any other English speaking countries?   YES   NO

   a. How long? _____ years

7. Which prefecture of Japan are you from? _____

8. What languages are regularly spoken in your household?
   _____

9. Why do you study English?

10. What is your highest test mark?

    _____ in TOEIC   or   _____ in IELTS