The Perception and Production of Glide-Vowel Sequences in Japanese and Korean Learners of English

by

© Hannah Praisley

A thesis submitted to the

School of Graduate Studies

in partial fulfillment of the requirements

for the degree of

Master of Arts

Department of Linguistics

Memorial University of Newfoundland

December 2016

St. John's Newfoundland

Abstract

Learners of a second language (L2) frequently exhibit phonological behaviours that are not fully understood. These behaviours can have a negative impact on intelligibility, sometimes resulting in communicative difficulties. This research looks at one such behaviour, shedding some light on its root causes.

Japanese and Korean learners of English as a second language often omit, or substitute the English glides [w] and [j] in particular vocalic contexts. It has been reported by Tsujimura (2013) and Kang (2014) that Japanese and Korean learners of English either omit the glide [w], or substitute it for [ɯ] or [ʔ], when it precedes a back vowel. Furthermore, Korean learners of English often omit the glide [j] when it is followed by a high front vowel.

In this thesis, I discuss the issue of glide acquisition in Japanese and Korean learners of English in light of two hypotheses. The first hypothesis is based on the assumption that perception is a prerequisite for production, while the second focuses on transfer effects from the L1, which influence L2 productions independently of perceptual abilities. I argue that the root cause of the patterns reported above lies primarily in perception based on empirical data collected in a series of experimental tasks, namely a reading, picture naming and ABX task. However, I acknowledge that transfer effects, which are phonotactic in nature, as well as cross-linguistic markedness, are also relevant to the story.

Acknowledgements

Writing this thesis has been an incredible journey, and it would not have been possible without the help of many people. The love, support and friendship I received during this master's degree really encouraged me to succeed.

Firstly, I want to thank my husband Les for always believing in me, listening to my endless dialogues about this research and still agreeing to marry me. Whenever I felt lost or overwhelmed you were there and never doubted me, for this I am eternally grateful. I dedicate this work to you.

Secondly, I want to thank everyone who agreed to participate in my study. Without your help this research wouldn't have been possible. The enthusiasm that you all showed was really inspiring and I hope that you all enjoyed the experience. ありがとう and 감사합니다. I also want to thank Rebecca for transcribing my sound files and showing so much interest in my research.

Thirdly, thank you to my fellow graduate students. Your kind words helped to keep me focused, and always gave me an excuse to go out for a cup of tea! I am happy that I was able to study alongside you and I wish you all the best with your own theses.

I want to thank all of my teachers at Memorial University, your instruction has given me the knowledge to accomplish this work. Finally, I want to thank my supervisor Dr. Yvan Rose, without your help I could not have completed this thesis. Thank you for your guidance and support.

Table of Contents

Chapter 1. Introduction

1. Introduction

Learners of a second language (L2) frequently exhibit phonological behaviours that are not fully understood by linguists and language teachers. These behaviours, such as sound omission or substitution (Kang 2014; Tsujimura 2013), can have a negative impact on intelligibility, sometimes resulting in communicative difficulties, and a reluctance to speak in the L2 (Juhana 2012). This research looks at one such behaviour, shedding some light on its root causes.

Japanese and Korean learners of English as a second language often omit or substitute the English glides [w] and [j] in particular vocalic contexts. Tsujimura (2013:33-34) and Kang (2012:487-488, 2014:108) report that Japanese and Korean learners of English either omit the glide [w], or substitute it for [ɯ] or [ʔ], when it precedes a back vowel. Furthermore, Korean learners of English often omit the glide [j] when it is followed by [i] (Kang 2014:108).

In this thesis, I discuss the issue of glide acquisition in Japanese and Korean learners in light of two hypotheses. The first hypothesis is based on the assumption that perception is a prerequisite for production, while the second focuses on transfer effects from the L1, which influence L2 productions independently of perceptual abilities. I argue that the root cause of the patterns reported above lies primarily in perception. This argument is based on empirical data collected in a series of experimental tasks. These tasks consist of a reading elicitation task, a picture naming task, and an ABX task. The former enabled me to collect production data which show that omission of the glide in homorganic glide-vowel sequences does occur in Japanese and Korean learners of English. The ABX task allowed me to study the learners ability to perceive the glides in the homorganic glide-vowel sequences, and enabled me to determine, when compared with the production data, that the root cause is indeed perceptual. However, I do acknowledge that transfer effects, which are phonotactic in nature, as well as cross-linguistic markedness, are also relevant to the story.

2. Foundation

In the following sections I outline the theoretical and empirical basis for this study. I then discuss related educational motivations for this study, in terms of how a greater understanding of this behaviour can inform teaching practices, how this phenomenon manifests itself, and how this research provides a contribution to the field of second language acquisition.

2.1. Theoretical Basis

Numerous frameworks designed to capture the complicated nature of second language phonological acquisition are available in the current literature. This thesis looks at four of these frameworks in order to explain the empirical evidence. The Contrastive Analysis Hypothesis (Lado 1957), the Markedness Differential Hypothesis (Eckman 1977), the Speech Learning Model (Flege 1981; Flege, Munro & MacKay 1995) and the Perceptual Assimilation Model for L2 Contexts (Best & Tyler 2007). In addition I discuss the concept of perceptual modification in terms of phonotactic rareness (Massero & Cohen 1983; Dupoux et al. 1999). All of the aforementioned frameworks and concepts provide avenues to explain the cause of this phenomenon.

There are four potential scenarios which can be entertained to explain the particular contribution of perception and production to the L2 acquisition of segments including glide-vowel sequences. Two of these scenarios suggest that the relationship between perception and production go hand-in-hand, in that perceptual abilities will match productive abilities. An alternative scenario is that there is a mismatch between perception and production in that perceptual abilities are good, but the resultant productions are poor. Finally, a fourth scenario would demonstrate that perception does not precede production, in that perceptual abilities are poor, but production abilities are strong (Braidi 1999). This thesis provides evidence for the first three scenarios, but not for the fourth one.

2.2. Empirical Basis

In order to gain insight into the issue of glide omission and substitution I conducted a study involving three experimental tasks. The first two tasks addressed production by using elicitation tasks in the form of a picture naming and a reading task. The third task addressed perception, using an ABX discrimination method. 11 participants who were native speakers of either Japanese or Korean and had high levels of English proficiency took part in the experiments. The data collected provide empirical evidence of the phenomenon of glide omission in homorganic glide-vowel sequences. These data also demonstrate instances of vowel substitution to create non-homorganic sequences with forms like /wʊm/ becoming [wʌm] and /jɪɹ/ becoming [jəɹ], which are considered to be less marked sequences cross-linguistically. However, I observed no cases of the glide substitution patterns (to [ɯ] or [ʔ]) described by Tsujimura (2013) and Kang (2014). Group, language-specific and individual results all demonstrate that lower perceptual abilities result in lower productive abilities, highlighting the relationship between perception and production. Furthermore, instances of vowel substitution present evidence of transfer effects to conform with L1 phonotactics and a preference for less cross-linguistically marked forms.

Personal observations and existing literature (Kang 2014; Tsujimura 2013; Lee 2004) document that the omission and substitution of glides in different glide-vowel contexts yields productions like those shown in Table 1.

Table 1. Target and Predicted Forms

Target Form - Orthography	Predicted Actual Form
/wʊd/- would	[ʊd] or [ɯd]
/woʊnt/ - won't	[oʊnt]
/wʊlf/ - wolf	[ʊlf] or [ɯlf]
/jiːld/ - yield	[iːld]
/jɛɫ/ - yell	[ɛɫ]
/wund/ -wound	[ʔund]

As mentioned above, the results from the current study do not contain occurrences of the substitution patterns shown in Table 1; the results only contain evidence of glide omission. However, the results also reveal cases of vowel substitution to create non-homorganic glide-vowel sequences, which have also been observed in another study focusing on Korean learners of English.

2.3. Contributions to the Field

Segmental acquisition has been well studied in the field of second language acquisition (Aoyama et al. 2004; Aoyama 2003; Sheldon & Strange 1982; Wang 1997; Chen, Robb & Harvey 2001). However, we know that learners realizations of sounds vary depending on various factors, such as the position of the segment within the syllable, phonetic co-articulation and so on. This research focuses on the acquisition of sequences, as opposed to individual segments. Although research has been undertaken on the acquisition of glides by Korean learners, particularly focusing on acoustic properties and production, little attention has been given to glide acquisition for Japanese learners of English. Moreover, to my knowledge, there are no existing studies that look at whether the cause of glide omission in Japanese and Korean learners lies in perception versus production. The current study aims to address this gap.

From an educational perspective, a greater understanding of language specific phonological acquisition difficulties will aid educators and learners to achieve higher levels of attainment. None of the 11 participants in this study identified having difficulties with the target homorganic glide-vowel sequences, yet all of them omitted glides in their productions and had lower perception scores for these sequences compared with their perception scores for non-homorganic glide-vowel sequences. This lack of awareness of the issue, which is likely a result of these sequences never being the focus of a pronunciation class, has resulted in inaccurate productions and reduced intelligibility.

Finally, through this study, I developed a corpus of English words (and sounds) produced by Japanese and Korean learners of English, which will provide a springboard for future research. The stimuli used in this pilot were carefully selected to encompass all English sounds in word-initial, medial and word-final positions. The availability of phonological corpora for research in second language phonology is limited at present; this corpus will therefore provide a valuable resource for the field.

Chapter 2. Theoretical Background

1. Introduction

This chapter addresses the theoretical background relevant to perceptual and articulatory abilities in second language acquisition. In section 1.1, I discuss the theoretical considerations of glide-vowel sequences from a typological perspective focusing on markedness. I then move to a description of Japanese, Korean and English in section 2.1, focusing on glide and vowel phonology and their phonetic manifestations. In section 3, I outline several theoretical frameworks relevant to studies on perception and production, namely the Contrastive Analysis Hypothesis, the Markedness Differential Hypothesis, the Speech Learning Model, the Perceptual Assimilation Model, and the concept of phonotactic rareness. I conclude this chapter by discussing several existing studies looking at the acquisition of glides by Korean learners of English as well as the perception and production of liquids, vowels and consonants (Jang & Cho 2005; Kang 2012; Kang 2014; Lee 2004; Aoyama et al. 2004; Bada 2001; Guion et al. 2000; Do-Seoup Jeong 2006; Sheldon & Strange 1982; Baker & Trofimovich 2006; Minnick Fox & Maeda 1999; Fox et al. 2009; Sperbeck 2010) in section 4, to provide context for the current study.

1.1. Theoretical Considerations

From a typological perspective, homorganic glide-vowel sequences such as /wu/ and /ji/ are not favoured cross-linguistically. This is especially interesting due to the fact that the results of this study demonstrate cases whereby homorganic glide-vowel sequences are altered to non-homorganic sequences with forms like /wum/ becoming [wʌm]. Rose (1999), Ohala & Kawasaki (1984) and Lee (1994) provide typological evidence for the marked nature of these sequences across a variety of languages. Ohala & Kawasaki (1984:122) state that the sequences are rare cross-linguistically due to the minimal acoustic difference between the two sounds of the sequence. In terms of formant frequencies, /w/ is similar to /u/ and /j/ is similar to /i/, resulting in a small perceptual difference between the glide and vowel sounds. Following on from Ohala and Kawasaki's observation, the following hierarchy would be expected in terms of markedness for glide-vowel sequences: $GV_{[Non\text{-}Homorganic]} >> GV_{[Homorganic]}$ (Greenberg 1966; Kang 2014). This hierarchy demonstrates that heterorganic, or non-homorganic glide-vowel sequences are considered to be less marked forms. Although many languages prohibit homorganic glide-vowel sequences, this is not universal. English permits homorganic glide-vowel sequences such as /ji/ and /wu/, which present a challenge for some learners. Interestingly, although these homorganic glide-vowel sequences are permitted in English there are very few words with this combination. In contrast, Japanese and Korean phonotactics prohibit homorganic glide-vowel sequences, while permitting several non-homorganic glide-vowel sequences such as /ja/ and /wa/.

The Markedness Differential Hypothesis (MDH) (Eckman 1977) states that difficulties that a learner will have in their L2 can be anticipated using the differences that exist between the L1 and L2 grammars:

> The areas of difficulty that a language learner will have can be predicted on the basis of a systematic comparison of the grammars of the native language, the target language and the markedness relations stated in the Universal Grammar, such that:
>
> (a) Those areas of the target language which differ from the native language and are more marked than the native language will be difficult.
>
> (b) The relative degree of difficulty of the areas of the target language, which are more marked than those of the native language will correspond to the relative degree of markedness.

(c) Those areas of the target language which are different from the native language, but are not more marked than the those in the native language will not be difficult.

(Eckman 1977:321)

The MDH therefore predicts that the homorganic glide-vowel sequences under investigation in this study will be difficult to acquire, due to their cross-linguistic markedness and absence from the L1. The non-homorganic glide-vowel sequences, although still absent from the L1 in some instances (see section 2.1.2 below for details of the permitted glide-vowel sequences in Japanese and Korean), will be much easier to acquire due to them being unmarked. This relative ease of acquisition is due to a lesser degree of markedness cross-linguistically, which stems from a difference in the acoustic properties between the segments of the sequence.

1.2. Hypotheses

Existing literature on second language acquisition suggests two potential sources of the omission or substitution of /w/ and /j/ in L2 English. The first possibility, formulated in works by Flege (1981) and Flege et al. (1995) on the Speech Learning Model, as well as by Best & Tyler (2007) on the Perceptual Assimilation Model, relates to perceptual effects. These theoretical models lead to the hypothesis that a learner's perception of the glides in relevant contexts is inaccurate, which directly impacts the learner's ability to produce these sounds. The second possible source consists of transfer effects from the L1's phonotactics (Bada 2001), and puts the burden of explanation on the speech articulation of /w/ and /j/, which are not attested within the speakers' native languages in the phonotactic contexts reported in Table 1 (page 4, above).

In this thesis, I entertain both of the above hypotheses by performing a series of perception and production experiments with native speakers of Japanese and Korean who are L2 learners and speakers of English. The data acquired through these experiments provide evidence for the former hypothesis, that the root cause of glide omission lies primarily in perceptual difficulties, but that transfer effects may in fact underlie these difficulties.

1.3. Educational Motivation

In terms of second language education, a greater understanding of the difficulties that students face with glide production will benefit teachers in the field of ESL (English as a Second Language). It is thus my hope that this research will help to provide a foundation to address this issue from an educational perspective by highlighting this issue and shedding light on the cause, in order to help to reduce resultant communication difficulties. It has been reported (Tsai 2015; Aida 1994) that there is a correlation between anxiety and language performance. This correlation suggests that learners with lower levels of anxiety related to their L2 learning achieve higher levels of success in their acquisition. Communication breakdowns are a contributing factor to language learning anxiety, it has been reported that L2 speakers who perceive themselves to have a strong L2 accent, or low proficiency in pronunciation feel apprehensive before verbal interactions (Derwing & Munro 2015). This apprehension is due to the individual being unsure of how successful the up coming verbal exchange will be. Therefore, by reducing communicative failures and the apprehension felt for future communicative exchanges, anxiety levels can be lowered and students can work towards higher levels of attainment.

Traditionally, teaching pronunciation was often neglected (Macdonald 2002). This lack of priority was predominantly due to a lack of understanding of how to teach pronunciation effectively. In light of the Critical Period Hypothesis (Lenneberg 1967) native-like attainment was not deemed achievable for adult learners. Modern approaches, predominantly within the past 20 years, acknowledge that native-like pronunciation is not the most central goal of second language teaching, but that intelligibility is crucial. Effective communication is one of the primary goals of second language learning, and if utterances are unintelligible because of poor pronunciation, communicative ability is compromised. From personal experience, I have witnessed pronunciation classes consisting of drilling individual sounds, with no context and little emphasis on sound sequences or suprasegmental features. It is arguable that teaching individual sound segments has limited efficacy, due to the fact that speech consists of a series of these individual segments co-articulated to form utterances enriched with prosody (Morley 1991; Gilakjani 2012). Derwing et al. (1998) and Derwing & Rossiter (2003) show that students who have received suprasegmental instruction show a significant improvement in their comprehensibility, whereas, those who received segmental instruction did not show any significant improvement. The merits of pronunciation teaching have also been shown by Saito (2012), who states that pronunciation teaching can lead to improved pronunciation proficiency. Therefore, the focus of this research is interesting from a teaching and acquisition perspective, as I am focusing on the acquisition of sequences as opposed to individual segments. By establishing the root cause of the omission patterns in Japanese and Korean learners' productions, I can determine whether discrimination, pure production, or a combination of both activities can be recommended to overcome the issue at hand .

2. Comparison of English, Japanese and Korean

In the following sections I outline several theoretical frameworks which address perception and production. In addition, I give a brief summary of some of the characteristics of both Japanese and Korean that are relevant to the subject of glide omission and substitution.

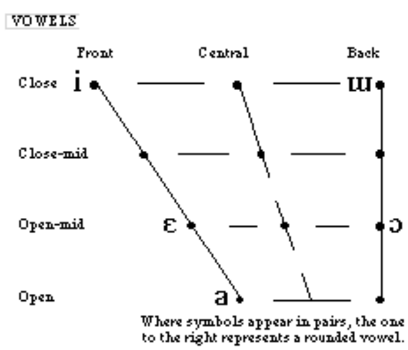
2.1. Language Background

Japanese and Korean are both Altaic languages, with a SOV basic word order (Tranter 2012). English on the other hand is a Germanic language with a SVO word order. Despite these structural differences the three languages have a number of similarities especially when it comes to vowels and glides, which are at the centre of the current study. In the following sections I outline the vowel and glide phonemic inventories, and discuss key similarities and differences between the three languages.

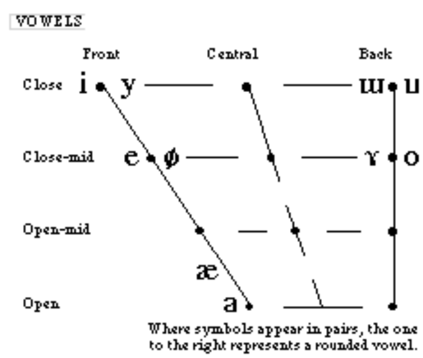
2.1.1. Vowels

In order to understand the phonological development of Japanese and Korean learners of English, some knowledge of each language is required. In Figure 1 I begin with the vowel inventories of Japanese, Korean and English. There is some controversy with the Korean inventory, which I discuss below.

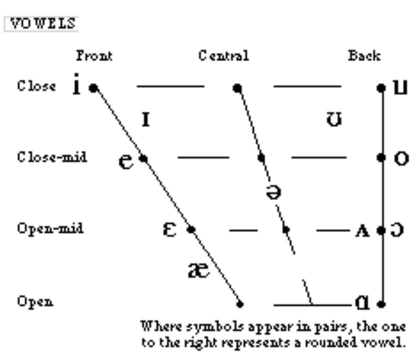
Figure 1. Japanese Korean and English Vowel Inventory
(Source: http://accent.gmu.edu/browse_native.php?function=detail&languageid=33)



Japanese



Korean



English

Figure 1 highlights differences between the vowel inventories of Japanese, Korean and English. Neither Japanese nor Korean have /ɪ/ or /ʊ/ in their inventories, which could explain why, when combined with /j/ or /w/ respectively, these vowels cause difficulties for the learners of English. The presence of /i/ in all three inventories, combined with the difficulties seen when the sound is combined with the glide /j/, provides evidence of the markedness of this sequences (Ohala & Kawasaki 1984; Rose 1999; Lee 1994).

Japanese has five monophthongs, as shown in Figure 1. Vance (1987) reports that there are short and long manifestations of the monophthongs: [a/aː, i/iː, ɯ/ɯː, ɛ/ɛː and ɔ/ɔː], which are mono or bi-moraic in nature. Ingram & Park (1997) state that these long and short vowels are phonologically contrastive.

There is a degree of controversy surrounding the number of monophthongs in Korean, with accounts stating that the language displays anywhere from seven to ten vowels (Kim 1968; Lee 1971; Lee 1973; Lee 1994; Lee 1998; Sohn 1987; Kang 2014; Ahn & Iverson 2007; Hong 1988). Figure 1 shows ten monophthongs, but recent accounts (Kang 2014; Ahn & Iverson 2007) state that there are in fact seven. This is a result of a reduction in contrastive vowel length and the merging of /e/ and /ɛ/ in the Seoul dialect (Ingram & Park 1997). The Seoul dialect is considered to be the standard dialect of South Korea, and is thus the dialect used in educational institutions (Song 2008). The hometowns of the Korean participants in this study are very diverse and not limited just to the Seoul region, as can be seen in Figure 4 (page 32, below). However, all of the Korean participants in this study had studied at a Korean education institution, and therefore would have been exposed to the Seoul dialect. The seven monophthongs of the Seoul dialect (Standard Korean) are shown in Figure 2.

Figure 2. Standard Korean Vowel Inventory (based on Kang 2014)

	Front	**Central**	**Back**
High	i	ɨ	u
Mid	e	ə	o
Low		a	

2.1.2. Glides and Sequences

Japanese, Korean and English all have the glides /j/ and /w/ in their inventories. Flege (1981) remarks that although languages can have sounds identical in terms of the IPA, the phonetic manifestation of these sounds can differ. The Japanese glides have a minimal difference with the English glide in that they are less rounded in their articulation (Labrune 2012). Korean glides are reported to be slightly shorter in duration and also have a lesser degree of lip rounding than their English counterparts (Kang 2012).

English permits both homorganic and non-homorganic glide-vowel sequences without restriction, resulting in words like *yield*, *would, youth* and *wheel*. Japanese and Korean have further restrictions in the permitted combinations of glides and vowels. The permitted sequences are shown in Figure 3 below.

Figure 3. Japanese and Korean Glide-Vowel Sequences

Japanese	wa				ja	jɯ	jɔ		
Korean	wa	wi	we	wə	ja	ju	jo	je	jə

Figure 3 shows that the vowels /i/ and /ɛ/ do not combine with any of the glides in Japanese. Historically there was evidence of a wider distribution of glide-vowel sequences (Pinter 2005; Martin 1976). With the exception of /wo/ and /wu/, it was suggested that the reduction in the number of permitted glide-vowel sequences in Japanese was due to the minimal phonetic difference between the segments in the sequence. The sequence /wu/ was never documented in Japanese, so assumptions have been made that it never existed. Pinter (2005) proposes that the presence of /wo/ may have been due to syllable structure and the requirement of an onset in Middle Japanese, and its disappearance occurred when this requirement ceased to apply to the phonology of the language. Martin (1976) also documented the presence of /je/ in Old Japanese; like the other sequences, this one no longer exists in Modern Japanese. Further elaboration of this is beyond the scope of this study, but the absence of /wu/ and /wʊ/ sequences from Modern Japanese is likely what makes them so difficult to acquire for Japanese learners of English.

The distribution of glide-vowel sequences in Korean is much more diverse than what is seen in Japanese, but still more restricted than English. Korean only prohibits the homorganic sequences such as /ji/ and /wu/ as can be seen in Figure 3. This broader range of glide-vowel sequences suggests that there will be fewer difficulties in the acquisition of non-homorganic glide-vowel sequences for Korean learners than for Japanese learners. In order to investigate the validity of this hypothesis, both homorganic and non-homorganic glide-vowel sequences are included in the stimuli of this thesis.

3. Theoretical Frameworks

Several theoretical frameworks address phonological development in second language acquisition. The Contrastive Analysis Hypothesis (CAH), the Markedness Differential Hypothesis (MDH (discussed above), the Speech Learning Model (SLM), the Perceptual Assimilation Model for L2 contexts (PAM-L2) and the concept of phonotactic rareness are commonly discussed within the acquisition literature (Aoyama et al. 2004; Kang 2014; Escudero & Boersma 2002; Guion et al.

2000 Murphey et al. 2016; Eckman 1977). I present each of these models in turn in the following sections.

3.1. The Contrastive Analysis Hypothesis and the Markedness Differential Hypothesis

Early theories of second language acquisition had their foundations in behaviourism and structuralism. Lado (1957) developed the Contrastive Analysis Hypothesis (CAH), with teaching methods and practices in mind, with the primary aim to improve approaches to language instruction. The CAH is based on the intuitive notion that the parts of a second language that are more difficult to acquire are those which are different from the L1. Therefore, those aspects of an L2 with an L1 counterpart are easier to acquire and are acquired first through a process of "positive transfer" (Saville-Troike 2005:35). Forms without an L1 equivalent must be learned; if there are overlaps in meaning or use, then acquisition will be straightforward. In contrast, the learning of forms not present in the L1 is predicted to be much more complicated and time consuming. With the same viewpoint Lado demonstrated that forms with only partial overlap with the L1 can also cause interference and result in delays in acquisition. Lado exemplified this interference in Spanish learners of English with lexical forms having Spanish cognates, but which have a different meaning such as *embarrassed* vs. *embarazada* (pregnant). These partially overlapping forms are often referred to as lexical false friends in English.

However, this model can be criticized as it makes inaccurate predictions due to it being largely based on the acquisition of individual units (words, speech sounds), and as a result it is unlikely to make accurate predictions about the acquisition of the sound sequences in question. In the context of the current research, it is the sequences of glides and vowels that cause difficulties for learners, rather than isolated sounds. The glide sounds are part of both the Japanese and the Korean phonological systems, but the homorganic glide-vowel sound distributions are not. As reported in

Figure 3 (above), in Japanese, the glide /w/ can only be followed by /a/; in Korean, /w/ is attested in several contexts, but never before back vowels, while /j/ is never followed by front vowels. Therefore for /w+u, ʊ, o/, and /j+i, ɪ, e/, there are no L1 counterparts, which means that in the context of the CAH, acquisition is predicted to be difficult. The apparent limitation of this model led to the development of the Markedness Differential Hypothesis discussed above (Eckman 1977). The MDH explains that forms that are considered marked and differ from the L1 will be difficult to acquire, but that the degree of difficulty in acquisition will correspond with the relative degree of markedness of the form. In contrast to the CAH the MDH states that unmarked forms in an L2, which differ from the L1, will not be difficult to acquire. In terms of the homorganic glide-vowel sequences under investigation, the MDH makes more accurate predictions for their acquisition. The homorganic glide-vowel sequences are considered marked and are absent from the L1, therefore will be difficult to acquire. The limitations of the CAH and further insight on the MDH will be discussed further in Chapter 5 (page 87, below).

3.2. The Speech Learning Model

The Speech Learning Model (SLM) was proposed by Flege (1981;1995). It hypothesizes that L1 experience can cause difficulties for L2 acquisition. Flege criticizes the assumptions behind the CAH, and limits its application to very early stages of acquisition. The CAH states that if a form in an L1 and L2 are alike, these forms should be easy to acquire (Lado 1957:2). Flege counters this by stating that a phone in an L2 which is similar, but not identical, to a phone in an L1, should be more difficult to acquire. This difficulty in acquisition arises from the fact that the learner can misleadingly use an L1 category for an L2 sound, rather than developing a more accurate category. This use of the L1 category for L2 segments is part of the learner's interlanguage, an ever-evolving linguistic system which merges rules and features from both the L1 and the L2.

Flege (1981) discusses a concept known as “Phonological Translation”. This is where “a listener interprets sounds in a foreign language in terms of the categories of the native language”(p. 451). He proposes that the more experience in the L1 that a learner has, the more difficult it is to acquire an L2. Mirroring what is proposed in the Critical Period Hypothesis (Lenneberg 1967; Johnson & Newport 1989, Flege states that the linguistic experience that learners have already gained in their L1 shapes their ability to process and produce speech sounds in an L2.

Moreover, Flege's SLM hypothesizes that sounds without an L1 counterpart are easier for learners to acquire because they create their own phonological category for the sound rather than merging it with an L1 category. Given this hypothesis, Japanese and Korean learners should be able to acquire the homorganic glide-vowel sequences easily, because these sequences have no L1 counterpart. However, as mentioned already, in both Japanese and Korean /w/ and /j/ both exist as segments, as do many of the vowels under observation. The presence of identical phone categories in English, Japanese and Korean may thus cause the confusion predicted by the SLM. In addition, the homorganic glide-vowel sequences are lacking altogether in both Japanese and Korean. This is a potential issue for both the SLM and the CAH. The SLM considers sounds without an L1 counterpart to be new sounds. These new sequences should then be easy to acquire with sufficient practice and exposure, but the results of this study show that even with practice and exposure all of the participants have difficulty with the homorganic glide-vowel sequences.

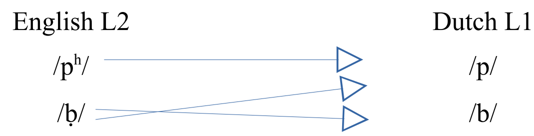
3.3. The Perceptual Assimilation Model – L2

Best & Tyler (2007) expanded on Best’s (1991) Perceptual Assimilation Model (PAM) to account for L2-related phenomena. The PAM states that native language sounds influence how we perceive non-native phonetic contrasts. Best hypothesizes that “articulatory gestures in the speech signal informs the perceiver” (Best 1991:14). This occurs in a way that explicitly links speech perception and articulation; gestural similarities between native and non-native speech sounds may aid the perception of these non-native sounds. This hypothesis means that when an individual listens to the speech of another person, s/he directly perceives the utterance through the movements of related articulatory apparatus. This articulation-based perception process facilitates the decoding of speech and in turn the understanding of the utterance. The PAM claims that speakers develop these finely attuned perceptual abilities during infancy. After we become language-specific listeners, we find it increasingly difficult to perceive the gestural combinations of sounds not present in our language community.

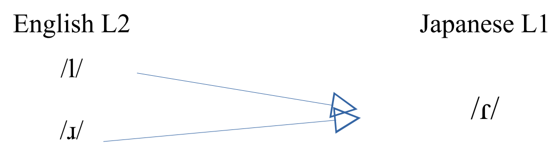
The PAM-L2 shares many features with Flege's SLM. This model looks at both phonetic and phonological levels of speech perception, accounting for dialectal, allophonic and lexical factors. This model states that not all non-native segments or contrasts are equally difficult for L2 learners, and that levels of difficulty can vary depending on the properties of the learners' native language. The principal assumption behind the PAM-L2 is that articulatory gestures in speech production provide a foundation for "language specific phonology" (Best & Tyler 2007:10), as opposed to Flege's proposal that this language specific phonology is based upon acoustic, or perceptual properties of speech. Both the SLM and the PAM-L2 hold the assumption that L2 learners develop an interlanguage, as already discussed in section 3.2 page 17 (above). Best and Tyler state that speech sound segments will be assimilated to be either good or bad examples of the categories that exist in their L1. The PAM-L2 claims that L2 learners categorize the L2 phones into their L1 phonemic system by determining how similar the L1 and L2 phones are. This categorization enables them to perceive the sounds as two examples of a native phoneme: a new uncategorized phoneme which falls between two native phonemes, or a non-assimilable speech sound, which has no identifiable similarity to any native speech sounds (Best, McRoberts & Goodell 2001:776-777).

These methods of categorization can yield the following assimilation scenarios (taken from Best, McRoberts & Goodell 2001:777):

(1) **Two-category Assimilation**: Two non-native phones can be phonetically similar to two native phones, and can assimilate separately to them. Example from Escudero & Boersma (2002:209):



(2) **Single-Category Assimilation**: Both non-native phones can (poorly) assimilate to a single native phoneme. Example:



(3) **Category Goodness:** Although two non-native phonemes can assimilate to a single native category, one will have a better fit than the other. Example: voiceless and ejective velar plosives in Zulu [kh]–[k'], which are perceived by English speakers as being a good and poor exemplars of English /k/ or [k^h] (Reid et al. 2014).

(4) **Uncategorized-Categorized:** One non-native phone is uncategorized and the other is categorized. Example: T1 and T4 in Mandarin when perceived by Cantonese speakers (So 2005).

(5) **Uncategorized:** Both non-native phones are considered to be uncategorized speech segments. Example: Japanese listeners perception of certain Australian English vowels (e.g. /ɜː/, /oː/ and /æ/) (Bundgaard-Nielsen, Best & Tyler 2011).

(6) **Non-Assimilable:** Both non-native phones are unidentifiable to a native speech sound. Example: Zulu clicks are not perceivable as any native speech sound by English speakers, but rather as non-speech sounds like popping a cork out of a bottle (Reid et al. 2014).

The poor performance documented in my experimental results (in Chapter 4) could be due to the fact that this phenomenon falls under S*ingle Category Assimilation* proposed in this model. Best & Strange (1992) suggest that Japanese learners who have difficulties in discriminating between

English /l/ and /ɹ/ are perceiving both sounds as a single phoneme. Following this suggestion and in line with work by Kang (2014) one hypothesis is that Japanese and Korean learners of English, when acquiring the homorganic glide-vowel sequences, are depending on their contrastive perception of these sounds. Under this hypothesis, both phones of the glide-vowel sequences (e.g. /j+i/; /w+u/) may be perceived as tokens of the same native phoneme. This perception could stem from the fact that these sounds could come from a single underlying phoneme in the L1 (Levi 2004; Kang 2014), whereas the glides and vowels in the homorganic sequences are separate phonemes in English.

3.4. Phonemic Segments vs. Sequences

As mentioned above, the CAH and the SLM place their emphasis on acquisition at a segmental level, which is problematic for this study because the focus is on glide-vowel sequences. The PAM focuses on phonemic segments as well, with the explicit suggestion that it can be applied to "syllables and units of meaning in speech" (Best 1991:2). Best & Tyler (2007) also state that the PAM-L2 can be applied to sound sequences, but in the majority of the literature on this model it is only applied to individual speech segments. It is therefore difficult to observe its application to sound sequences. In the next section, I discuss the concept of Phonotactic Rareness, which like the PAM takes sequences into consideration and could provide some insight for this study. I discuss the application of these models to the sound sequences under investigation in the current study in Chapter 5.

3.5. Phonotactic Rareness

The concept of phonotactic rareness suggests that learners will have a bias against perceiving impossible sequences in their L1 (Murphy, Monahan & Grant 2016). This idea of a perceptual bias is based on work by Massero & Cohen (1983) and Dupoux et al. (1999), who argue for the integration of a phonotactic dimension into speech perception models. Massero & Cohen conducted experimental studies to assess how visual and auditory information influences speech perception, to test the effectiveness of the Fuzzy Logical Model of Perceptual Recognition[1] (Massaro 1979). Massero & Cohen discovered that there was a correlation between reaction times in perception and the perceived ambiguity of the input, suggesting that participants were biased towards permissible sequences in their L1s. Dupoux et al. (1999) studied Japanese and French listeners and discovered that phonotactic properties of their L1s guided their perception. The study looked at contrastive vowel length, present in Japanese and absent from French, and consonant clusters present in French and absent from Japanese. The results show that the Japanese participants perceive vowels within consonant clusters when they were absent from the input, and the French learners do not perceive vowel length contrasts absent from their L1s. These findings suggest that the listeners are perceiving the input with a bias towards the phonotactic rules set out in their L1.

4. Existing Studies

Numerous studies look into perception and production in the acquisition of English for Japanese and Korean speakers (Aoyama et al. 2004; Guion et al. 2000; Kang 2014; Cho & Jeoung 2013; Baker & Trofimovich 2006; Minnick Fox & Maeda 1999; Fox et al. 2009; Sperbeck 2010). A majority of these studies are looking at vowel and consonant segments (Aoyama et al. 2004; Guion et al. 2000; Cho & Jeoung 2013; Baker & Trofimovich 2006; Minnick Fox & Maeda 1999; Fox et al. 2009; Sperbeck 2010), as opposed to the glide-vowel sequences in the present study. As there is, to my knowledge, no literature specifically looking into the cause of glide omission in homorganic glide-vowel sequences in either Japanese or Korean speakers of English, or the relationship between

1 The FLMP states that we analyze auditory and visual features of the input and process this information in line with the features of inputs stored in our memory. We then classify this information to create a *best fit* with the features known to us, but this *best fit* may have somewhat fuzzy boundaries.

perception and production of these sequences, I use the following studies to conceptualize my hypotheses, discuss methodological considerations and, more generally, contextualize my own study. I begin by looking at literature on glide acquisition in Korean learners of English, and move on to studies of perception and production in Japanese and Korean learners of English.

4.1. Acquisition of Glides by Korean Learners of English

Several studies have identified difficulties in the acquisition of glides by Korean learners of English (Jang & Cho 2005; Kang 2014; Lee 2004). Lee (2004) identifies that the glide /w/ is omitted in non-initial contexts when followed by non-back vowels in words such as [kwɪz]. Lee proposes that this pattern of omission is a result of transfer; learners are applying the deletion rules (deletion of /w/ when it is a syllable nucleus in non-initial position e.g. in donkey - dangnakwi) to their L2 English productions. The Korean learners of English are perceiving the glide as either a syllable onset or a nucleus. When the glide is perceived as an onset, the [w] is produced; when it is perceived as a nucleus, it is frequently omitted if followed by a non-back vowel. This study does not focus on the homorganic glide-vowel sequences that are the focus of the current thesis, but it does provide evidence that transfer can be an influencing factor in the perception and production abilities of Korean learners of English.

Jang & Cho (2005) and Kang (2012) discuss difficulties in homorganic glide-vowel sequences, which are at the forefront of this thesis. These studies show that the Korean glides are shorter in duration and less constricted in their articulation than their English counterparts, and it is these qualities that could promote the omission patterns seen in English learners due to the glides being similar but not identical to the English glides. Following Ohala & Kawasaki (1984), Rose (1999) and Lee (1994), typologically it is well established that many languages prohibit homorganic glide-vowel sequences. Kang (2012) claims that the weak constriction and short duration in Korean glides result in insufficient contrast in the two sounds in the sequence, and therefore leads to the prohibition of this sequence in Korean. The longer duration and strong constriction creates a significant contrast in English and these properties allow the homorganicity of these sequences to be disregarded, permitting their existence. The differences in the phonotactics of Korean and English with regard to these sequences is discussed as problematic for the acquisition of English by Korean learners, suggesting that L1 characteristics will be used in L2 productions.

Kang (2014) extends his work to study whether Korean EFL students accurately produce homorganic glide-vowel sequences by measuring how the glides differ phonetically between English and Korean. Kang obtained production data and uncovered omission patterns, in line with those observed in the current study. Kang also found substitution of a glottal stop for the glide, a pattern which is unattested in the current research. Kang identifies that 68% of the participants produce inaccurate homorganic glide-vowel sequences, and utilizes four primary repair strategies: glottal stop substitution, glide deletion, vowel shifting, and glide shifting. Of these scenarios, glottal stop substitution and glide deletion are the two most common strategies, but these scenarios give way to target-like productions with higher levels of English proficiency. Kang predicted that the Korean learners of English would transfer their L1 glide characteristics into their L2 productions, and that it was this mismatch in the glides that caused the acquisition difficulties. However, this is not the case; the participants produced glides with accurate L2 characteristics. Based on this finding and the ability to accurately produce the glides in non-homorganic contexts, the author suggests that learners' abilities to perceive the target sequences should be investigated.

The aforementioned studies provide evidence for transfer effects influencing Korean EFL learners' productions of the homorganic glide-vowel sequences, but do not investigate potential perceptual difficulties. The current study fills this gap, and provides empirical evidence that perception is a driving factor in glide omission in both Japanese and Korean learners of English.

4.2. Perception and Production of Liquids

The English liquids /l/ and /ɹ/ are known to be difficult for Japanese and Korean speakers due to the absence of these phonemes from their L1 inventories (Aoyama et al. 2004; Bada 2001; Guion et al. 2000; Do-Seoup Jeong 2006). Aoyama et al. (2004) and Guion et al. (2000) discuss these acquisition difficulties in terms of the Speech Learning Model. They state that the similar, but not identical, nature of Japanese /ɾ/ to English /ɹ/ results in difficulties in perception for the sound. This perceptual difficulty results in poor production and resultant communication breakdowns. Due to the fact that the Japanese and Korean glides are similar, but not identical, to their English counterparts, this could provide evidence that the source of the omission and substitution lies in perception. However, Sheldon & Strange (1982) found the opposite, in that production abilities surpassed perceptual abilities for the English /l/ and /ɹ/. This study does present an argument for the relationship between perception and production being in part due to the way learners are taught, in that the learn the distinction between /l/ and /ɹ/ based on orthography. As we will see in Chapter 4, the results in the current study are much more in line with the remainder of the literature, as opposed to Sheldon & Strange, suggesting that perceptual abilities do drive outcomes in production.

4.3. Perception and Production of Vowels

Cho & Jeoung (2013) investigate the relationship between perception and production for American English vowels in Korean EFL learners. The authors report an average accuracy of just 60% in both perceptual and productive abilities, but also numerous instances where perceptual accuracy appears to be lower than productive abilities. Cho & Jeoung conclude that there is no distinct correlation between perception and production. Baker & Trofimovich (2006), on the other hand, undertook an experimental study focusing on Korean learners of English. These authors conclude that perceptual abilities do drive accurate production, but that self-perception is a critical factor. This self-perception relates to the individual's ability to perceive his/her own speech productions, itself driving a feedback mechanism to modify their productions.

Minnick-Fox & Maeda (1999) look at the perception and production of North American English vowels by Japanese speakers of English. They show that there is a strong correlation between perceptual performance and the intelligibility of productions. In cases where their participants display strong performance in perception, they in turn produce the most intelligible productions. This study again adds weight to the hypothesis that perceptual abilities precede productive abilities in our acquisition journey.

4.4. Perception and Production of Consonants

Fox et al. (2009) look at the acquisition of English sibilant fricatives by Korean learners of English. These scholars observe that there is a relationship between perception and production. In cases where perceptual abilities are high, the resultant productive abilities are also high. These findings offer additional empirical evidence that there is a significant relationship between both perception and production in second language acquisition, and that these skills go hand-in-hand, much like the findings of this thesis.

Sperbeck (2010) undertook an experimental study looking at the perception and production of consonant clusters in Japanese speakers of English. Sperbeck uses a production task and an ABX discrimination task. Sperbeck's results suggest that there is a link between perception and production, in that higher perceptual abilities yield higher productive abilities. This study also reveals that some of the strategies used in production are not related to the L1's phonotactics, suggesting that transfer is not the only factor involved in the emergence of second language proficiency.

4.5. Summary of Existing Studies

Together, the above studies reveal a degree of controversy over whether or not there is a relationship between perception and production. Sheldon & Strange (1982) as well as Cho & Jeoung (2013) suggest that production abilities can surpass perceptual abilities. However, in a majority of the relevant literature, a correlation is found to suggest that perception precedes production in the acquisition process. Furthermore, Kang (2014) acknowledges that perception could be an influencing factor in the acquisition of English glides by Korean learners, and that further research into this is necessary. This is what the current study sets out to do. In the next chapter, I present the methodology used for this research.

Chapter 3. Methodology

1. Introduction

I studied the issue of glide omission or substitution discussed in the previous two chapters by running two experiments, which are subdivided into three parts. The first focuses on a learner's ability to perceive individual sounds and sound sequences (e.g. /w/, /j/, /u/, /i/, /wʊ/, or /ji/). The second focuses on a learner's ability to produce these sounds and sound combinations in different phonological contexts. In order to test each of the hypotheses discussed above, I recruited native speakers of Japanese and Korean who are L2 speakers of English to participate in three experimental tasks, focusing on their ability to perceive and produce the glide-vowel sequences in comparison to other consonant-vowel sequences.

In this chapter, I begin by outlining ethical considerations and describing the research facility for this study, before explaining the recruitment procedure and a profile of each participant. Finally, I discuss the motivation for my selected methods and describe the experimental procedure in detail. As we will see, because of challenges in participant recruitment, which relates to the number of Japanese and Korean speakers available, the experiments described below must be considered as pilots only. However, in spite of this shortcoming, the results obtained show very clear trends, the essence of which form the basis for the ensuing discussion.

2. Ethics Approval

In order to comply with the standards set out by Memorial University's Interdisciplinary Committee on Ethics in Human Research (ICEHR), I successfully completed the process for ethical approval. I submitted an application outlining my research idea, how I planned to conduct the research, how I would minimize any risks and finally what contribution it would make to the field of second language acquisition. In addition I submitted recruitment documents and associated literature, informed consent forms and an introductory letter for participants, all of which were accepted by the committee (ICEHR Number: 20160790AR).

3. Research Facility

This research took place within the Speech Sciences and Language Acquisition Laboratory (SSLAL) at Memorial University. The SSLAL offers all of the research space and equipment required, including audio recording and testing technology in a sound attenuated room. The SSLAL is a private space, which is divided into two rooms, providing a space to complete the required paperwork before commencing the experimental tasks. The experiments were conducted a one-to-one basis and I was present at all times, following the ethical requirements mentioned above.

Due to accessibility issues, one participant was unable to undertake the experiments in the SSLAL. In order to mimic the conditions as closely as possible the participant performed the tasks in a quiet, private space, and identical procedures were followed. The results from this individual are consistent with those of the other 10 participants. I am thus confident that there were no adverse effects from the alternative testing location.

4. Recruitment

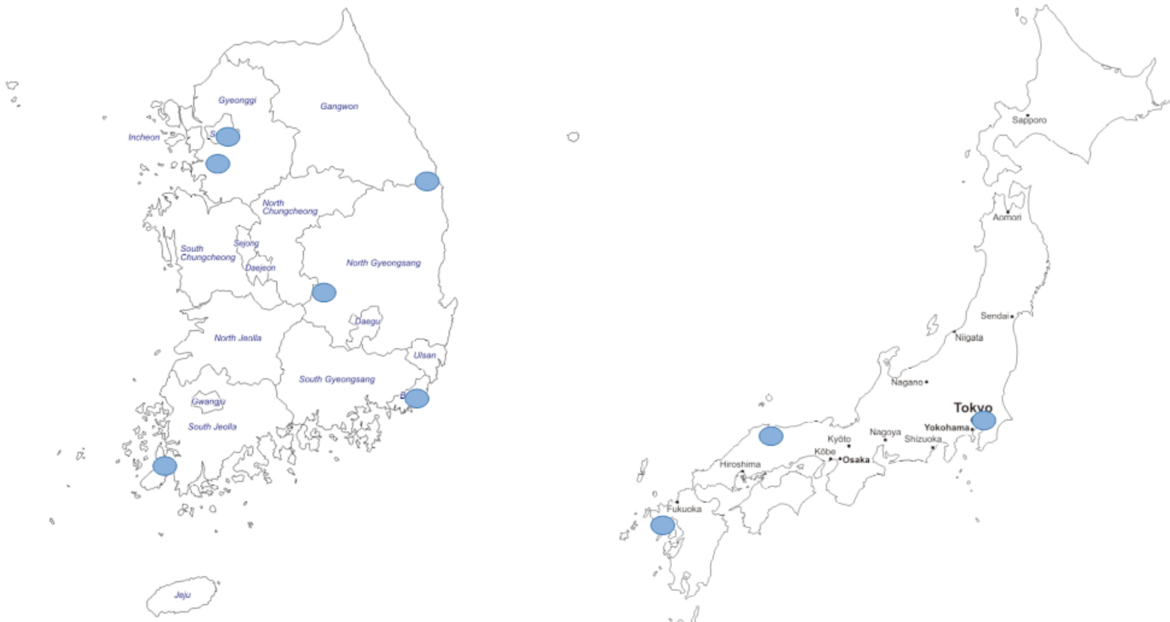
I recruited three native speakers of Japanese and eight native speakers of Korean who were second-language English speakers and resided in St. John's Newfoundland at the time of the experiment. Five males and six females make up the participants who took part in this study. Gender was not a variable in this experiment, therefore the slightly uneven ratio is not a concern. The number of participants is low due to the small population of the target nationalities within the St. John's area, however all efforts were made to recruit the maximum number of participants possible. This low number is far from optimal, but has provided sufficient data for the pilot experiments.

In order to recruit participants, a poster containing relevant information about the study was distributed via the Memorial University International Student Advising Office. The recipients of the poster were members of the international student population, the Japanese and Korean community and related student organizations. Interested individuals then contacted me directly, in adherence with confidentiality and anonymity policies. Informed consent forms were then emailed to participants, so as to give them sufficient time to read and understand the document prior to the experiment. On the day of the experiment, I greeted each participant in the SSLAL and obtained their informed consent. I then familiarized each individual with the procedure and equipment involved in the experiment.

4.1. Participant Profile

The maps in Figure 4 show the approximate hometowns of the participants (indicated by blue circles). As we can see the participants come from various regions of both Japan and Korea. This variation shows that the results shown in chapter 4 are less likely to be a result of a specific language or dialect, and more likely to be phonological in nature.

Figure 4. Hometowns of the Participants (map outline source:dmaps.com)



All of the participants in this pilot have an approximate intermediate to advanced level of proficiency in English, and they have taken either an immigration or university-based entrance exam such as CAEL (Canadian Academic English Language Assessment), TOEIC (Test of English for International Communication), IELTS (International English Language Testing System), or CELPIP (Canadian English Language Proficiency Index Program). Some of the participants have received more formal instruction than others, but they all have a good standard of English proficiency.

Each participant completed a background questionnaire before commencing the experimental tasks. This questionnaire provided details on their language use, language background, age, length of residence in an English-speaking country, occupation and their self-perceived English level. Table 2 (below) provides an overview of each participant's profile.

Table 2. Participant Profiles

Participant	1	2	3	4	5	6	7	8	9	10	11
Nationality J= Japanese K=Korean	J	K	K	K	K	J	J	K	K	K	K
Age	37	26	29	30	25	48	66	35	24	32	35
Years in an English-speaking country	8	2	9	1	10	6	31	6	7	4	6
Hours of English use per week	105	52	47	53	49	14	7	67	13.5	71	26
Gender (M/F)	F	M	F	F	M	F	F	M	M	F	M
Age started learning English	13	13	12	10	15	13	13	13	13	12	12
Self-perceived level of English proficiency	*L=Low I= Intermediate A= Advanced NN= Near Native*										
Listening	A	A	A	I	A	L	L	I	A	A	A
Reading	NN	A	A	A	A	L	L	A	A	NN	NN
Writing	A	I	I	I	A	L	I	A	A	NN	I
Speaking	A	A	A	I	A	L	L	I	A	NN	A

Table 2 shows that, with the exception of one individual, all of the participants use English very frequently in their day-to-day lives. A majority of the participants either work or live in an English-speaking environment, and it is therefore their primary language of communication. The mean number of hours of English use per week for the participants is 46 hours, or approximately seven hours per day.

The average age of the participants is 35 years, with a range of 24-66. The mean number of years in Canada is seven, but the average number of years in an English-speaking country is slightly higher, at eight years. The average number of years in an English-speaking country is higher due to the fact that several of the participants lived in the USA before coming to Canada.

The perceived level of proficiency in English shows some variation, some of it resulting from the participants' cultural background. A majority of the Japanese participants stated that their level of English was low; however, this is not the case. Their level of proficiency is actually approximately intermediate to upper intermediate based on their communicative ability and previous education. From personal experience, I am aware that it is common to underestimate your language ability in Japan. A majority of participants rated their skills as advanced, and some as near native. Although this is their self-perceived ability, I deem a majority of the participants to have acquired English to an advanced level.

In both Japan and Korea children start to learn English in school at around 12-13 years of age, which explains the consistent age at which most of the participants started their ESL journey. However, it should be noted that some participants are likely to have had more formal English instruction than others, and some participants have had breaks in their English learning. The focus of my experiment was not on language ability, therefore I have not put any emphasis on the years of formal instruction or age that their learning commenced. These potential factors are provided for background information; as we will see below, none of these factors appear to influence the outcome of the perception and production experiments.

In informal conversations with the participants, we discussed elements that each individual found difficult when they speak English. Everyone identified sound segments that are commonly problematic when associated with Japanese and Korean learners ([l, ɹ, v, b, f, p]). Interestingly, however, none of the participants identified glides as being problematic, or even made any mention of glides. More generally, I am fairly certain that no participant was aware of my specific research question when they undertook the experiments, which ensured as close a natural performance as possible in the experimental tasks described below. I begin with a description of the stimuli in the next section.

5. Stimuli

The stimuli used in all of the tests were carefully selected to include both homorganic and non-homorganic glide-vowel sequences, frequently occurring English words and a representative sample of all English speech sounds. The stimuli are all monosyllabic and bi-syllabic words with initial stress where possible, to ensure consistency throughout the experimental tasks. A full list of the stimuli used in this research are shown in the Appendix (page 111, below)

In order to establish whether there are any difficulties in the production and perception of glide-vowel sequences in the recruited population, the stimuli contained words with the English glides /w/ and /j/ followed by all possible vowel sounds. The homorganic glide-vowel target stimuli are listed in Table 3 (below). The glide and mid-vowel forms yielded virtually ceiling performance, in line with the control and filler items. I thus categorized them along with the non-target forms in the results sections.

Table 3. Homorganic Glide-Vowel Target Stimuli

/j/ + high front vowel	/j/ + mid front vowel	/w/ + high back vowel	/w/ + mid back vowel
Yield	Yeah	Woohoo	Woven
Yielding	Yay	Whoopi	Won't
Yeast	Yale	Womb	Woke
Yiddish	Yellow	Wounded	Wonder
Yin	Yearn	Woman	Worthwhile
Year		Wool	
Yearly		Would	
Yearbook		Wood	
Yippee		Wolf	

There is a ratio of one homorganic glide-vowel target form to seven foil items, with 14 glide-vowel target forms for each glide and 99 foils. This ratio is to prevent participants from identifying the particular phonotactic contexts under investigation, namely the homorganic glide-vowel sequences. Furthermore, this ratio allowed me to collect sufficient tokens to create a representative corpus of English sounds and words produced by Japanese and Korean learners of English.

The perception and picture naming tasks involved a smaller set of stimuli, but maintain a 1:7 ratio of homorganic glide-vowel target-to-foil stimuli. The smaller number of homorganic glide-vowel target stimuli in the perception task is due to the need of three repetitions of each word (to include the different combinations of stimuli for the ABX trial discussed below), to maintain the ratio with the full list would have made the task too long. The smaller number of target stimuli in the picture naming task is due to the unpicturable nature of some of the stimuli, for example words such as *would* and *yield*.

5.1. Collection of a Corpus for Future Research

The two production tasks also served to create a corpus for future research. I hope that this will be a beneficial contribution, as publicly accessible phonological corpora on second language acquisition are currently relatively limited. The tokens selected for the production experiment are carefully chosen to include all English sounds in word-initial, medial and final contexts, where possible. The corpus will be published through the PhonBank database (http://phonbank.talkbank.org), contributing to potential future research in the field of second language acquisition. Participants were made fully aware of the intention to publish their sound productions through informed consent, and of course had the right to decline this option, but a majority (9 out of 11) were happy to contribute.

6. Motivation for Methodology

In this section, I outline the theoretical and practical motivation for my methodology. Firstly, I discuss my reasoning for the production study and, secondly, for the perception study.

6.1. Production Study

The production experiment, as the name suggests, requires participants to produce speech. There are both elicited and naturalistic methods for this type of research. Naturalistic spontaneous speech has many benefits, however, it also involves a number of logistical issues. For spontaneous speech to be truly naturalistic in its collection there cannot be any elicitation, making it difficult to obtain sufficient productions of the glide-vowel target forms required for this study. In English, the number of words containing the target homorganic glide-vowel sequences (discussed in Chapter 2) are few. For this reason, I did not use spontaneous speech in my study.

Colantoni, Steele & Escudero (2015: 100-101) discuss four variations of elicited speech tasks: narrative, interactional, reading, and elicited imitation. The following descriptions follow these authors' typology. Interactional tasks involve the participants and an experimenter taking part in an interview style scenario. Elicited imitation tasks also require experimenter interaction, with the participants repeating utterances presented to them. For my study, imitation and interaction tasks are not the most appropriate as they could skew the production results. If the difficulties in production do not have roots in perception, the immediate input of a target form from the experimenter could result in (more) accurate productions, overestimating the learners' speech abilities. If there is a suitable degree of complexity, or an appropriate auditory distraction these elicited imitation tasks can be effective. However, there is a need to ensure that a participants echoic memory has been cleared in order to ensure productions are true and not influenced by the previous input. Narrative tasks are independent of the experimenter, but utilize objects, themes or scenarios to prompt productions. These narrative tasks, although effective at eliciting language, can yield very different results due to the participants' interpretation of the stimuli (for example identifying an item by using a brand name versus a common noun). These narrative tasks can be useful to remove orthographic influences, by using images to elicit the required forms. However, by itself, this task would not yield sufficient occurrences of target forms, so an additional method is necessary.

The final method discussed by Colontoni et al. is a reading-based elicitation task. This task involves participants reading pre-prepared texts, which can be designed to ensure a maximal number of productions of the target forms. There are several methods to obtain utterances in reading tasks. Lee (2004) presented targets in a paragraph of text, whereas Diaz-Campos (2004) presented target forms in sentences to take the focus away from the individual stimuli. Brannen (2012) and Eckman (2007) elicited single words, which were presented in a randomized list alongside fillers. For my experiment, I followed Diaz-Campos' method; for both production tasks, I used a carrier sentence containing the homorganic glide-vowel target form.

Reading based tasks are not without limitations. One consideration is that orthography could be an influence. In English, there is a degree of opacity in the orthography of glide sounds, for example in words like *while* [waɪl], *language* [læŋgwɪdʒ] and *acquisition* [ækwɪzɪʃn]. These words either contain extraneous (mute) letters or do not present any written glide, which could affect their pronunciation due to their being no visual representation of a glide. An alternative analysis is that orthography could provide a mechanism to help students identify the presence (or absence) of a segment. In order to alleviate such potential effects, I also used a picture naming task for many of the word forms in my data set. I provide more detail about this procedure in section 7.1 below.

6.2. Perception Study

In order to test the hypothesis that a difficulty to perceive the glides in homorganic glide-vowel sequences may negatively influence a learner's ability to produce /w/ and /j/ in varying contexts, I have reviewed the following experimental methods: the AX, AXB, and well-formedness judgment tests. I discuss these experimental methods below.

Baker et al. (2009) used a well-formedness judgement test to assess how well Korean children and adults could perceive different English vowel sounds. The stimuli were presented aurally under experimental conditions to Korean speakers residing in the USA who had received minimal exposure to English. Baker et al. used a seven point scale *goodness of fit* scale to determine whether English vowels were perceived to be the same as Korean vowels for the Korean learners of English. There is a degree of complexity associated with this task in terms of explaining how the task is administered. When designing my experiments I was aware of the potential variation in English proficiency in the recruited population and decided against this method. Furthermore, this methodology is not optimal at capturing the data needed to address the central question at hand.

Two common methodologies in speech perception research are the *same/different* AX and ABX discrimination tests. There are both advantages and disadvantages to the AX and ABX protocols. The AX discrimination task involves trials of pairs of stimuli (A–A, A–X, and X–X combinations) with participants identifying whether the stimuli are the *same* or *different*. Participants' reaction times and number of correct responses are analyzed in order to determine their ability to discriminate between different pairs of stimuli. Although simple to explain to participants, this protocol often leads to a bias towards the A (*same*) response (McGuire 2010). The ABX discrimination task follows the same principle as that of the AX task; however, there is an additional stimulus to remove the bias towards the A stimulus. In the ABX task, triads of stimuli are used allowing the participants to identify *same/different* stimuli, in line with the AX method described above. A further advantage of this design is that the exact nature of the similarity or difference under investigation does not need to be identified by the participant. The variant AXB task has the same advantages as the ABX task, but Schouten & Van Hessen (1999) showed that participants have a tendency to ignore the third stimulus, which would be problematic for my study. For the purposes of the current pilot, I have utilized the ABX methodology, a detailed description of which is presented in section 7.2 below.

7. Experimental Protocol

In this section, I outline the protocol used for the three experimental tasks used in this study. I begin with the two production tasks, the reading and the picture naming tasks. I then move on to the perception task.

7.1. Production Studies

In order to establish if the root cause of glide deletion or substitution lies in a learner's (in)ability to produce the homorganic glide-vowel target sequences, I conducted a production experiment to obtain data for a comparison with the perception study. To obtain utterances with the homorganic glide-vowel sequences, I used the elicited production methods discussed in section 6.1 (page 37 above). To address potential influences associated with orthography already discussed above, as well as the unpictureable nature of some stimuli, the experiment consisted of two parts: a reading-based elicitation task in line with work by Baker et al. (2009) and Brannen (2012), and a picture naming task, following works by Eckman (2007); Lee (2004); Diaz-Campos (2004) and Brannen (2012). In order to ensure that the relevant homorganic glide-vowel contexts could be studied, I selected multiple tokens with minimal pairs involving the homorganic glide-vowel targets (e.g. [wʊd] - [kʊd]) and additional fillers consisting of common English words. These tokens were fully randomized to prevent the participants from knowing which forms were the focus of the study. In the following section, I outline the detailed procedures for the production experiments.

7.1.1. Procedure

Both of the production tasks consisted of two parts, a practice phase and a test phase. The protocol for both phases was identical for both tasks. Participants first completed the picture naming task, and then proceeded to the reading task. This ordering ensured that the orthographic input in the reading task did not affect the productions in the naming task.

7.1.2. Practice Phase

The inclusion of a practice phase allowed me to ensure that all of the participants fully understood, and felt comfortable with, the task. The practice phase for both production tasks involved six stimuli that were not related to my research question, and were discarded before any analysis took place. For each task, in my role as the experimenter, I began by verbally explaining and modelling the task. The participants then began the practice phase. This phase involved the participants producing sentences aloud, into a recording device. I used unobtrusive audio recording equipment (Roland R-05 Wave/MP3 Recorder with a built-in microphone), with the aim of minimizing any potential stress or anxiety.

For each task, a carrier sentence and either a picture or a written stimulus was presented on slides using Microsoft PowerPoint. The participant then produced the carrier sentence with the inserted stimulus. After each production, the participants pressed the space bar to move to the next slide. When the practice stimuli were completed the participants were given the opportunity to take a short break and ask questions. All of the participants decided to proceed straight to the test phase.

7.1.3. Test Phase

For the test phase, the protocol was identical for both production tasks. The full set of stimuli (homorganic glide-vowel targets and fillers presented in a random fashion, without the possibility that two experimental stimuli occur immediately after one another) were given to participants. In order to elicit the productions from this task, as in the practice phase, I asked participants to produce the stimulus as part of a carrier sentence. For the naming task, participants had to complete the sentence by inserting the word associated with the picture they saw, and then utter that sentence aloud. In the reading task the participants were given a full sentence containing the stimuli word, and they then said it aloud into the recording device. The carrier sentence varied slightly for the two tasks, for example:

Task	**Carrier Sentence**
Picture Naming	"I see *wood,* now"
Reading	"I say *wood,* now"

In the above examples the sentences lend a natural isolation of the target word, as the word preceding it ends with a vowel, and the following comma elicits a minimal pause. This facilitated the isolation and later transcription of the produced forms.

With the picture naming task there were cases where the participants did not know the required stimulus. In these cases, I attempted to elicit the form by asking questions or getting the participant to complete a sentence examples are shown in Table 4. In cases where they were still unable to produce the required stimulus, it was given to them and they repeated the word. In these instances the participants' productions were removed from the results. For the reading task, participants were not prompted; they were instructed before the task commenced to say the word even if they were unsure of how to produce it.

Table 4. Examples of Prompting in the Picture Naming task

Required Form	Given Form	Prompt to Elicit Required Form
Time	Clock	What does a clock tell us?
Water	Droplet	What is the droplet made of?

Several of the target stimuli were difficult for the participants to produce. For the reading task all of the stimuli were nonetheless elicited and recorded. For the picture naming task some stimuli, such as *yearbook* and *wool* were problematic and some productions (9 utterances of *wool*) had to be removed from the results.

7.2. Perception Study

In line with work by Flege (2003), Tremblay & Kamiyama (2009), and Guion et al. (2000), I used an ABX discrimination task to collect the perception data in order to establish if the root cause of omission in homorganic glide-vowel sequences is a result of perception. As I have mentioned above in an ABX discrimination design, three stimuli are presented in a series and the listener indicates which stimulus, *A* or the *B,* is the same or most similar to the *X* stimulus. The ABX task took place after the production experiments, in order to ensure that the aural input did not affect a participants production of the homorganic glide-vowel targets during these prior tasks.

7.2.1. Procedure

The ABX experiment was created using OpenSesame, version 3.0.7 (http://osdoc.cogsci.nl/), an open-source software program used to create psychology and neuroscience experiments (Mathôt, Schreij & Theeuwes 2012). This software enabled me to not only measure the number of correct responses, but also the reaction times for responses.

The stimuli for the ABX task were produced by a Canadian native speaker of English, representative of the L2 learners' general linguistic context of immersion at the time of the experiment. The Canadian speaker was a 37 year old male who grew up in British Columbia. While it could be argued that the St. John's dialect should have be chosen, given that the second language background of the participants included more than St. John's English, it was deemed beneficial that a more generally neutral dialect was used. For all intents and purposes, there is also no clear differences between St. John's and Canadian English concerning the phonological contexts at stake. The Canadian speaker produced all of the stimuli into a recording device during a single recording session, in order to minimize potential variability across his productions. The produced stimuli were then segmented into individual word tokens and uploaded to OpenSesame.

In this task, both homorganic glide-vowel target and filler tokens were presented aurally to the participants from a recording, which ensured that they all received identical stimuli. The stimuli consisted of a target and two further stimuli (A and B). These stimuli allowed me to see if the listener was able to discriminate if one form is different to the target (X form). An example of a triad is as follows:

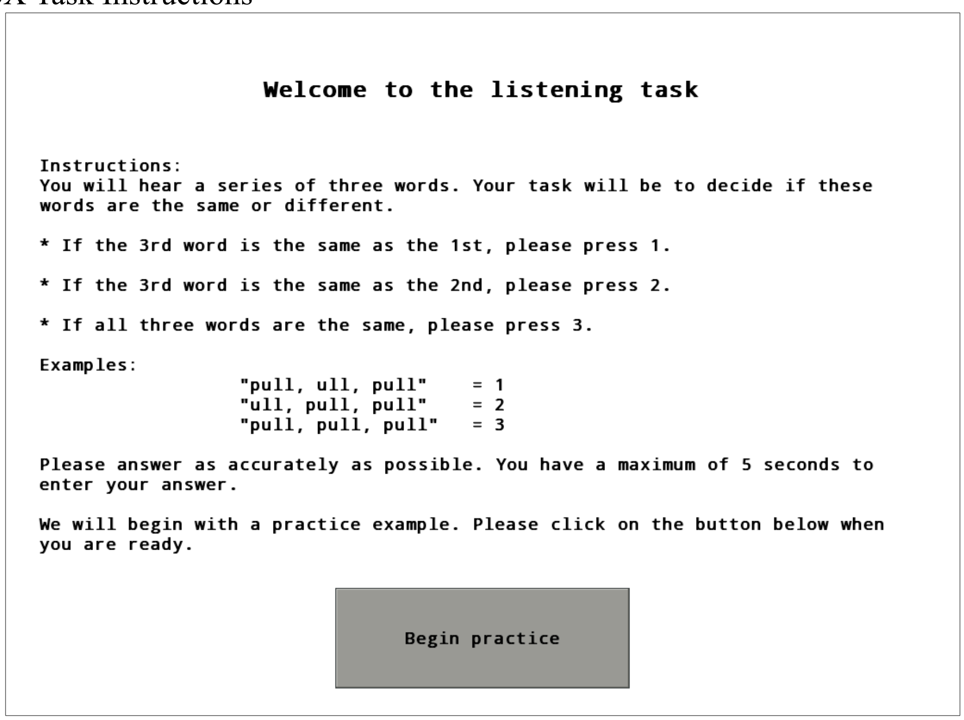
Stimulus A: /wʊd/ Stimulus B: /ʊd/ Stimulus X (TARGET): /wʊd/

The interval between the stimuli was 350ms. Existing literature (Edwards & Zampini 2008; Bohn & Munro 2007; Stackhouse et al. 2007) acknowledges that ABX tasks can be demanding on participants in terms of memory load. In order to address this concern, I divided the experimental stimuli into two blocks, with an optional short (2 minute) break in between. Each experimental block commenced and ended with four independent stimuli, which were removed from the results, to minimize lead-in, or lead-out task effects.

7.2.2. Practice Phase

The practice phase began with a verbal explanation of the task ahead, and as with the production tasks, in my role as the experimenter I modelled the task. After my brief demonstration the participants were given a set of headphones (Sennheiser HD 230 Pro 64) and instructed to set the volume to a comfortable level. Instructions concerning how to complete the task were displayed on the computer screen (Figure 5 below), and each participant was asked to read these instructions carefully and ask any questions before starting the practice phase.

Figure 5. ABX Task Instructions

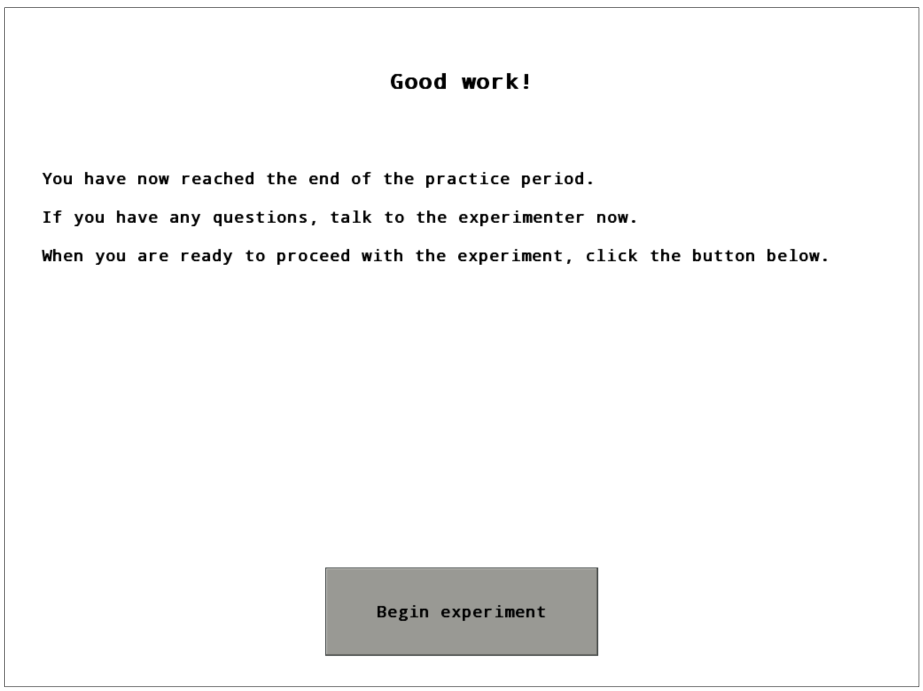


When ready, the participant clicked the button displayed at the bottom of the screen. The stimuli triads were then presented. After each triad, the participants were instructed to respond to the stimuli by pressing a number key, either 1, 2, or 3. Participants pressed 1 if the first and third stimulus were the same, 2 if the second and third stimulus were the same, and 3 if all three stimuli were the same. In line with the production experiment, the practice phase stimuli were independent to my research question and were discarded before any analyses took place.

7.2.3. Test Phase

After the practice phase was completed, the participants were presented with a screen (Figure 6) instructing them to ask any questions and to proceed to the experiment when they were ready.

Figure 6. ABX Test Phase Screen



The test phase followed the same protocol as the practice phase, and the software was programmed to give participants a five-second time window to respond. If the participants did not respond within this time frame the next triad began, and their response was marked as incorrect, and a reaction time in excess of 5000ms was recorded. The triads of stimuli used in this experiment were selectively randomized to ensure there are not any adjacent homorganic glide-vowel target forms, or more than two consecutive trials that required the same numerical response. To ensure consistency with the production experiments I kept a 1:7 ratio between test and filler items.

In the next chapter I present descriptive results collected in the three experimental tasks outlined above. I describe these results independently, and then provide a comparison of the perception and production results in order to shed light on the root cause of glide omission in homorganic glide-vowel sequences.

Chapter 4. Results

1. Introduction

In this chapter, I describe the results from all of the experiments in this pilot study. After a discussion with members of Memorial University Mathematics Department, I have established that the small population size does not satisfy the requirements of most statistical tests. Therefore, I will only be using descriptive statistics to present the results of this pilot study. Due to the small population size and the similarities in their L1 phonology, I combined the results from of all of the participants; where relevant, I break down the data into language groups.

Firstly, I outline the data categories that I use to describe my results in section 2. I then continue by describing the results of the production studies in section 3. I begin by looking at accuracy, focusing only on glide omission. Thirdly, I outline the results of the perception study in terms of accuracy and reaction times. In section 4, I discuss the perception and production trends for the Japanese and Korean participants separately. I finish this chapter by discussing potential lexical effects and individual patterns in sections 5 and 7.

2. Data Categories

The table below (Table 5) shows the categories that were used to describe the results of this pilot. The results show that there is only a minimal difference between the filler and control groups and the glide+mid-vowel categories e.g /j + ɛ/. As a result, these glide+mid-vowel categories will be grouped with the non-target results.

Table 5. Data Categories

<table>
<tr><th>Category</th><th colspan="2">Description</th></tr>
<tr><td>Consonant + Vowel</td><td>Filler</td><td rowspan="4">Non-Target</td></tr>
<tr><td>Non-Target Glide + Vowel</td><td>Control</td></tr>
<tr><td>[j]+mid front vowels</td><td rowspan="2">Glide mid-vowel</td></tr>
<tr><td>[w]+mid back vowels</td></tr>
<tr><td>[j]+high front vowels</td><td colspan="2" rowspan="2">Target (Homorganic glide-vowel)</td></tr>
<tr><td>[w]+high back vowels</td></tr>
</table>

3. Production Results

In the next sub-sections, I describe the results from my two production tasks, firstly from the picture naming and then from the reading task. In some instances in the picture naming task, as I have mentioned, participants were unable to produce the required form; in these instances the utterance has been omitted from the results.

3.1. Picture Naming Task Results

The results I describe in this section focus on accuracy (whether the glide is deleted, or not, in homorganic glide-vowel sequences). As we will see, these data reveal many cases of glide omission in homorganic glide-vowel sequences. Other production patterns, such as vowel substitution, are discussed in a subsequent section.

3.1.1. Accuracy across Phonotactic Contexts

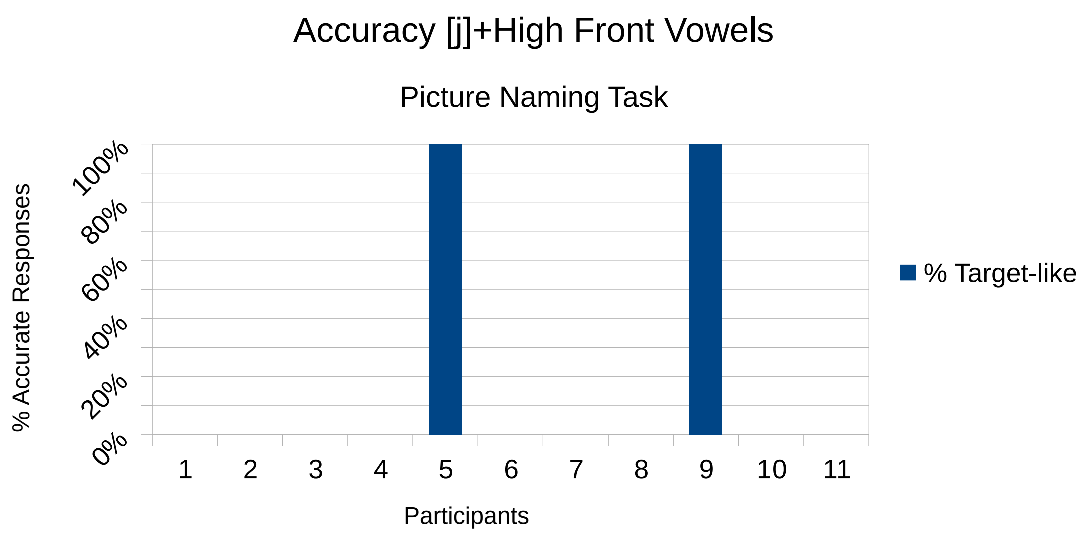
Productions that include the target glide, even if the following vowel is not what would be expected from a native English speaker (vowel substitution patterns are discussed in section 6), are considered target-like. For the picture naming task the participants performed at ceiling level with 100% accuracy in the consonant and [j]+mid front and [w]+mid back vowel categories as well as the filler category. The control sequences also display very high rates of accuracy, which given the small number of tokens per context can also be considered as ceiling performance, as shown in Table 6.

Table 6. Accuracy in Non-Target Glide-Vowel Sequences

Sequence Type	**Participant**										
	1	2	3	4	5	6	7	8	9	10	11
Filler (n=28)	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Control (n=11)	100%	100%	100%	91% (10/11)	100%	100%	100%	91% (10/11)	100%	100%	100%
[j]+mid front vowel Sequence (n=1)	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
[w]+mid back vowel Sequence (n=1)	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Table 6 shows the correct responses as a percentage, as well as the number of correct responses from the total number of stimuli in each category in cases where omission occurs. The near perfect accuracy shown in the table is considerably higher than the accuracy in homorganic glide-vowel sequences. Figure 7 shows a lower rate of accuracy for both the [j] and [w] homorganic glide-vowel target sequences, with many omissions in the productions. In the [j] target sequences there is a 0% rate of accuracy across a majority of the participants. By comparison, in the [w] target sequences the group mean accuracy is higher at around 61%, but still strikingly different from what is seen in the non-target contexts. It must be noted there are very few tokens of both homorganic glide-vowel target forms in this task (5 stimuli; two for [j] and three for [w]).

Figure 7. Accuracy in Homorganic Glide-Vowel Target Sequences



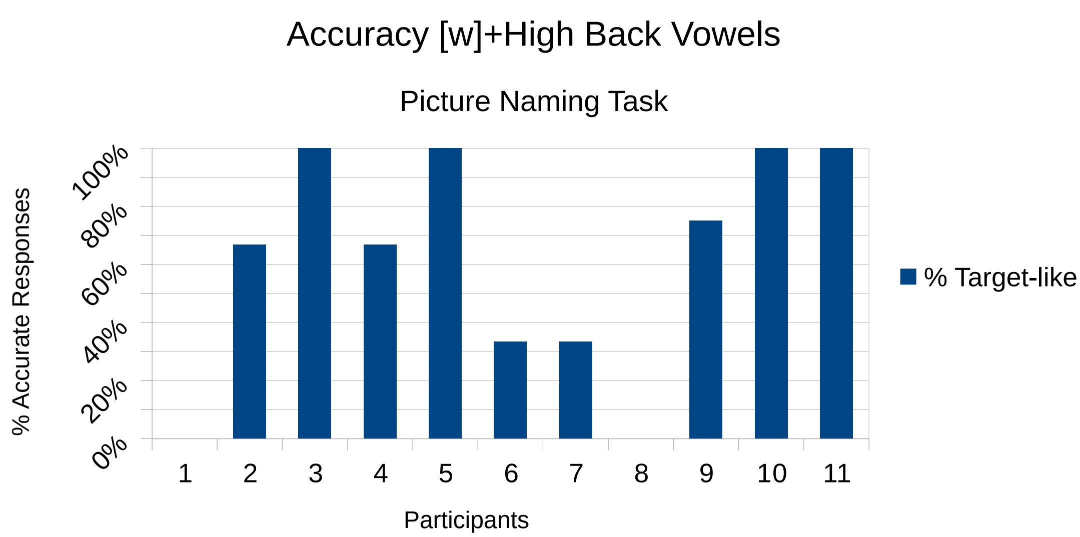
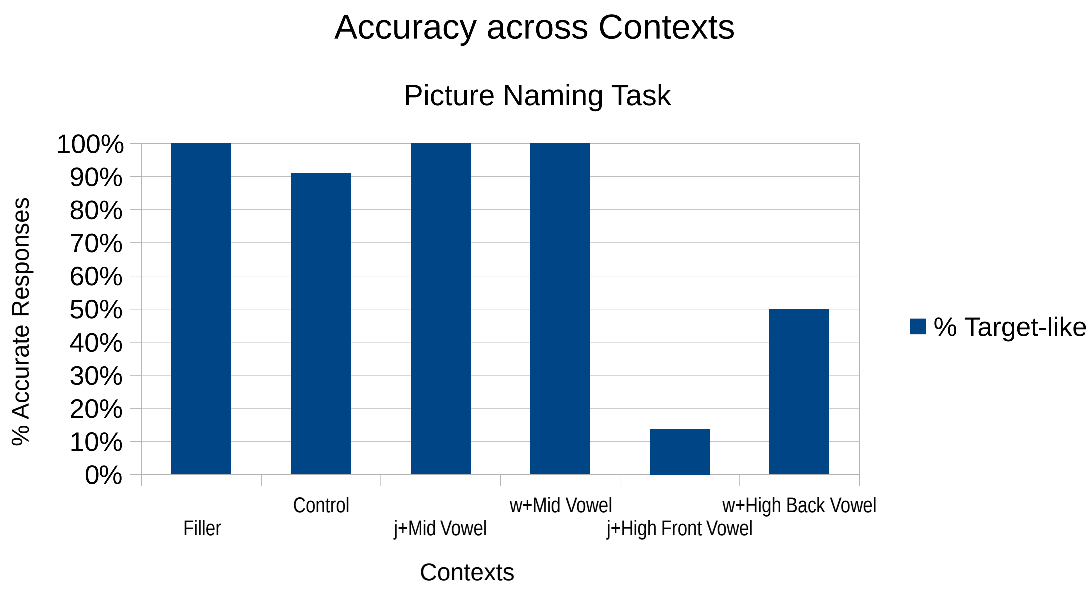


Figure 8 shows the average accuracy for all participants in each of the contexts. The chart shows a striking difference, with the homorganic glide-vowel target contexts showing extensive glide omission. These data are also consistent with the results from the other two tasks described in subsequent sections.

Figure 8. Accuracy across Contexts



In order to establish if the trend shown in Figure 8 is caused by a single word or group of words, rather than the glide-vowel sequences, a word-level analysis is required. I turn to a detailed description of the production patterns of the two homorganic glide-vowel contexts in the next section.

3.1.2. Production Patterns

As we saw above, with the exception of the [j]+high front vowel and [w]+high back vowel sequences, performance is generally at ceiling. With this accuracy in mind, I focus my description on the target homorganic glide-vowel contexts. I begin in Figure 9 with a description of accuracy at the word level for the [j]+high front vowel context.

Figure 9. Accuracy in [j]+High Front Vowel Sequences

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
Year (j+ɪ)	jɪ ↔ i	1	1	1	1	0	1	1	1	0	1	1	9
	jɪ ↔ jə	0	0	0	0	1	0	0	0	1	0	0	2
	Target-Like	0	0	0	0	1	0	0	0	1	0	0	2
	Omission	1	1	1	1	0	1	1	1	0	1	1	9
	% Target-Like	0%	0%	0%	0%	100%	0%	0%	0%	100%	0%	0%	18.18%
	% Omission	100%	100%	100%	100%	0%	100%	100%	100%	0%	100%	100%	81.82%

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
Yearbook (j+ɪ)	jɪ ↔ i	1	0	1	1	0	1	1	1	0	1	1	8
	jɪ ↔ jə	0	0	0	0	1	0	0	0	0	0	0	1
	jɪ ↔ ɹ	0	1	0	0	0	0	0	0	0	0	0	1
	Target-Like	0	0	0	0	1	0	0	0	0	0	0	1
	Omission	1	1	1	1	0	1	1	1	0	1	1	9
	% Target-Like	0%	0%	0%	0%	100%	0%	0%	0%	N/A	0%	0%	10.00%
	% Omission	100%	100%	100%	100%	0%	100%	100%	100%	N/A	100%	100%	90.00%

As we can see in these data (Figure 9)[2] the accuracy for this phonetic context is very low in contrast to the near 100% accuracy in non-target contexts (shown in Figure 8). Given the fact that there are only two stimuli for this context, a comparison with the reading elicitation and perception data is necessary, which I discuss in sections 3.6 and 8.2. Figure 9 also reveals that the participants are fairly consistent in their behaviours. For example, Participant 5 uses the same strategy of producing the glide with a schwa in both productions. The other participants simply omit the glide and produce the remaining sounds of the word. These data also show similar rates of omission for both of the *year* and *yearbook* forms (cf. results of the reading task in Figure 14, page 59 below).

The results for the [w]+high back vowel context are shown in Figure 10. Again there are only a few stimuli for this sequence, but the results are striking when compared to those from the other non-target contexts presented in this pilot. A striking difference can also be seen between the forms *woman* and *wood,* with *woman* yielding 73% accuracy compared with *wood* only yielding 45%.

2 Participant 9 was unable to produce the word *yearbook* without being given the form by the experimenter; as a result it has been removed from the results.

Figure 10. Accuracy in [w]+High Back Vowel sequences[3]

	Participant	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**Total**
Woman (w+ʊ)	wʊ ↔ wu	0	0	0	0	1	0	0	0	0	0	0	1
	wʊ ↔ wɔ	0	0	0	0	0	1	0	0	0	0	0	1
	wʊ ↔ wʊ	0	0	1	1	0	0	0	0	0	1	1	4
	wʊ ↔ wʌ	0	1	0	0	0	0	0	0	1	0	0	2
	wʊ ↔ ʊ	1	0	0	0	0	0	1	1	0	0	0	3
	Target-Like	0	1	1	1	1	1	0	0	1	1	1	8
	Omission	1	0	0	0	0	0	1	1	0	0	0	3
	% Target-Like	0%	100%	100%	100%	100%	100%	0%	0%	100%	100%	100%	72.73%
	% Omission	100%	0%	0%	0%	0%	0%	100%	100%	0%	0%	0%	27.27%

	Participant	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**Total**
Wood (w+ʊ)	wʊ ↔ wu	0	0	0	1	1	0	0	0	0	0	0	2
	wʊ ↔ wʊ	0	0	1	0	0	0	0	0	0	0	1	2
	wʊ ↔ wʌ	0	0	0	0	0	0	0	0	0	1	0	1
	wʊ ↔ ʊ	1	1	0	0	0	1	1	1	1	0	0	6
	Target-Like	0	0	1	1	1	0	0	0	0	1	1	5
	Omission	1	1	0	0	0	1	1	1	1	0	0	6
	% Target-Like	0%	0%	100%	100%	100%	0%	0%	0%	0%	100%	100%	45.45%
	% Omission	100%	100%	0%	0%	0%	100%	100%	100%	100%	0%	0%	54.55%

	Participant	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**Total**
Wolf (w+ʊ)	wʊ ↔ wʊ	0	1	1	0	1	0	1	0	1	1	1	7
	wʊ ↔ ʊ	1	0	0	1	0	1	0	1	0	0	0	4
	Target-Like	0	1	1	0	1	0	1	0	1	1	1	7
	Omission	1	0	0	1	0	1	0	1	0	0	0	4
	% Target-Like	0%	100%	100%	0%	100%	0%	100%	0%	100%	100%	100%	63.64%
	% Omission	100%	0%	0%	100%	0%	100%	0%	100%	0%	0%	0%	36.36%

All participants except 3, 5, 10, and 11 omit the glide in at least one of the forms shown in Figure 10. Recall that all participants except 5 and 9 omit the [j] in the forms shown in Figure 9; Participant 5 is thus the only one who produced all forms accurately. Everyone else omitted at least one glide in at least one context, with the majority omitting glides in both homorganic glide-vowel target contexts. Finally, note that most of the forms considered to be accurate in terms of glide production also displayed conspicuous patterns of vowel substitution. I return to this topic in Chapter 5.

3 In Figure 10, I have omitted the word *wool* from the data. This omission is because only two of the 11 participants successfully produced the word, without needing to repeat the production of the experimenter.

3.2. Reading Task Results

The results described in this section are consistent with those from the picture naming task. The stimuli which contain [j]+high front vowels, as well as [w]+high back vowels show strikingly lower rates of accuracy than all other word forms.

3.2.1. Accuracy across Phonotactic Contexts

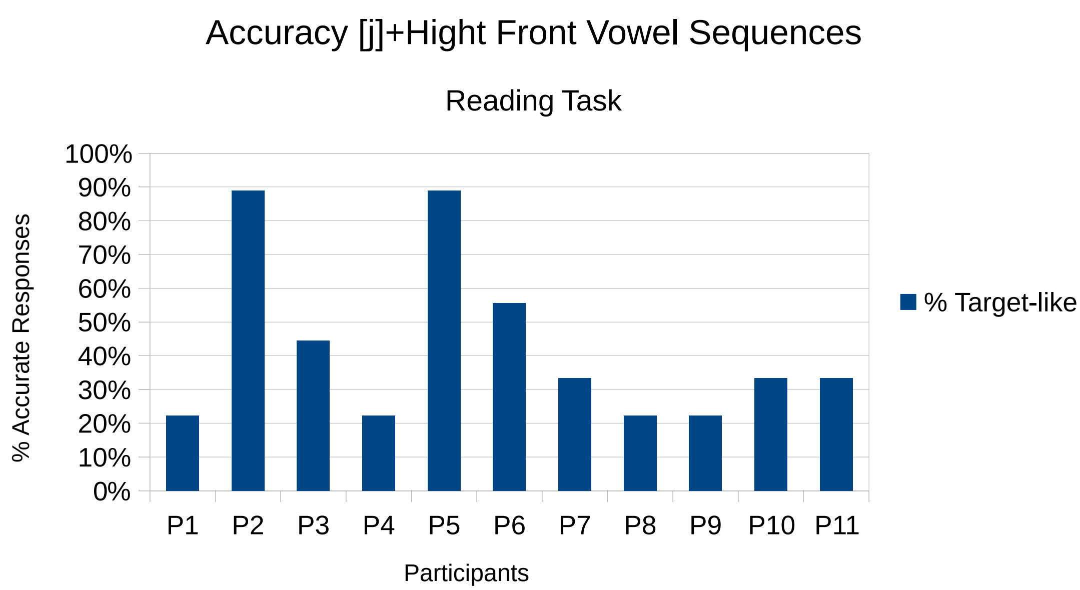
The data in Figure 11 display excellent accuracy across the four non-target contexts. Concerning the [w]+mid-glide sequence, there is marginally lower accuracy. However, the lowest rate of accuracy observed remains high, at 80% (4/5), also in line with the results from the picture naming task, which are consistently high in this context.

Figure 11. Accuracy for Non-Target Contexts

Sequence Type	**Participant**										
	1	2	3	4	5	6	7	8	9	10	11
Filler (n=61)	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Control (n=24)	100%	96% (23/24)	100%	100%	100%	100%	100%	92% (22/24)	100%	100%	100%
[j]+mid front vowel Sequence (n=5)	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
[w]+mid back vowel Sequence (n=5)	80% (4/5)	100%	100%	80% (4/5)	100%	80% (4/5)	80% (4/5)	80% (4/5)	100%	80% (4/5)	100%

Figure 12 (below) shows that the homorganic glide-vowel target productions from the reading task are also consistent with the results of the picture naming task, with a general rate of accuracy lower than for the non-target contexts. Accuracy for the [w]+high back vowel context is slightly higher than that for the [j]+high front vowel context with seven of the 11 participants producing more accurate responses, again mirroring the results from the picture naming task. Figure 12 also shows some degree of consistency among individuals, who tend to pattern in similar ways for both of the target [w] and [j] contexts (with participants 1, 4, 6, 7, 8, and 9 having lower accuracy in both).

Figure 12. Accuracy in Homorganic Glide-Vowel Target Contexts



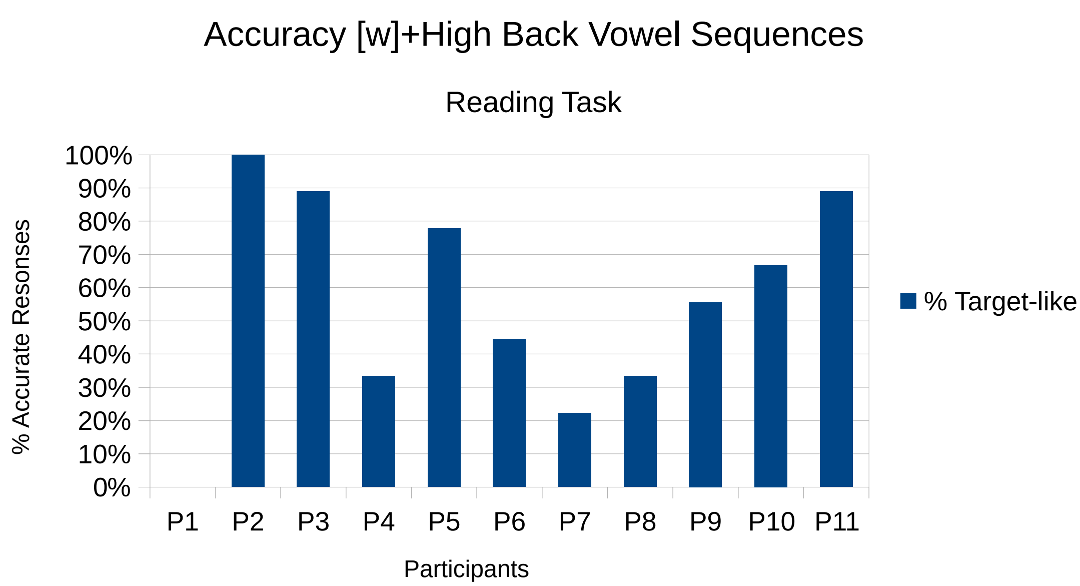
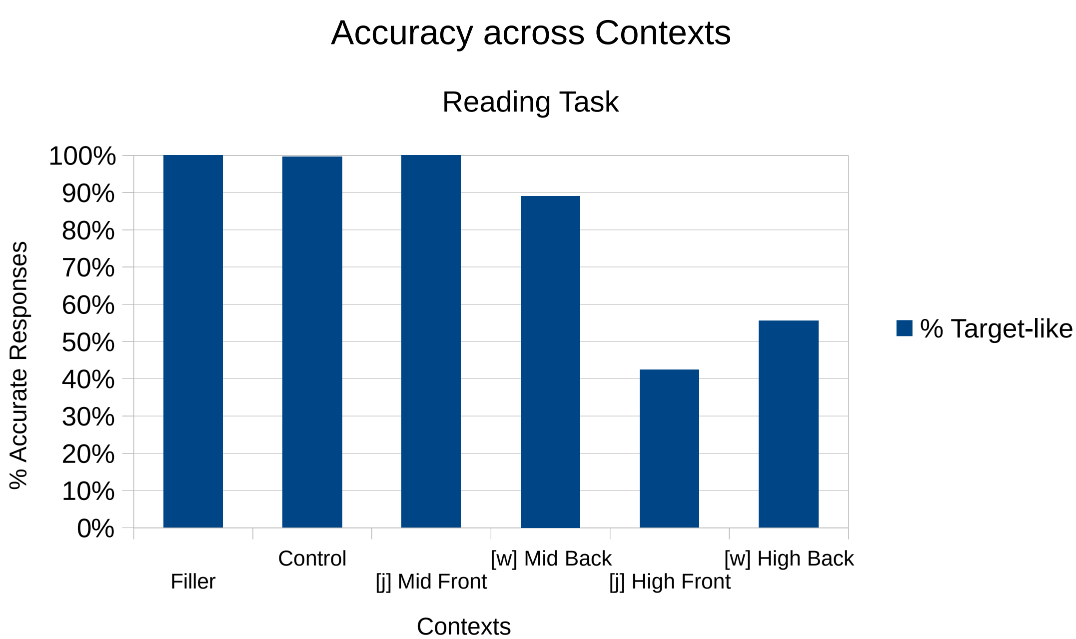


Figure 13 gives an average of all of the participant data across the different contexts. In combination with Figure 8, the results from both the picture naming and the reading tasks reveal a range of average accuracy rates of 50-55% for the target [w] context. The [j] context shows more variation, with the accuracy for the picture naming task being at just over 10%, whereas in the reading task it is just over 40%. The lower result for the picture naming task may be related to the small number of tokens (n=5 homorganic glide-vowel targets). In comparison, the reading task has considerably more tokens with this structure (n=18). These data could thus be considered more representative.

Figure 13. Accuracy across Contexts



3.2.2. Production Patterns

Figure 14 provides a breakdown of the participants' production patterns for the two homorganic glide-vowel target contexts.

Figure 14. Analysis for [j]+High Front Vowels

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
Yiddish (j+ɪ)	jɪ ↔ i	0	0	0	0	0	0	1	0	1	1	0	3
	jɪ ↔ jɪ	1	1	1	0	1	1	0	0	0	0	0	5
	jɪ ↔ wi	0	0	0	0	0	0	0	1	0	0	0	1
	jɪ ↔ ɪ	0	0	0	1	0	0	0	0	0	0	1	2
	Target-Like	1	1	1	0	1	1	0	1	0	0	0	6
	Omission	0	0	0	1	0	0	1	0	1	1	1	5
	% Target-Like	100%	100%	100%	0%	100%	100%	0%	100%	0%	0%	0%	54.55%
	% Omission	0%	0%	0%	100%	0%	0%	100%	0%	100%	100%	100%	45.45%

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
Yin (j+ɪ)	jɪ ↔ i	0	0	1	0	0	0	1	1	1	1	0	5
	jɪ ↔ ju	0	0	0	0	0	1	0	0	0	0	0	1
	jɪ ↔ jɪ	0	1	0	0	1	0	0	0	0	0	0	2
	jɪ ↔ ɪ	1	0	0	1	0	0	0	0	0	0	1	3
	Target-Like	0	1	0	0	1	1	0	0	0	0	0	3
	Omission	1	0	1	1	0	0	1	1	1	1	1	8
	% Target-Like	0%	100%	0%	0%	100%	100%	0%	0%	0%	0%	0%	27.27%
	% Omission	100%	0%	100%	100%	0%	0%	100%	100%	100%	100%	100%	72.73%

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
Year (j+ɪ)	jɪ ↔ i	1	0	0	1	0	1	1	1	0	0	1	6
	jɪ ↔ ji	0	1	0	0	0	0	0	0	0	0	0	1
	jɪ ↔ jə	0	0	0	0	1	0	0	0	0	1	0	2
	jɪ ↔ jɪ	0	0	1	0	0	0	0	0	1	0	0	2
	Target-Like	0	1	1	0	1	0	0	0	1	1	0	5
	Omission	1	0	0	1	0	1	1	1	0	0	1	6
	% Target-Like	0%	100%	100%	0%	100%	0%	0%	0%	100%	100%	0%	45.45%
	% Omission	100%	0%	0%	100%	0%	100%	100%	100%	0%	0%	100%	54.55%

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
Yearly (j+ɪ)	jɪ ↔ i	0	0	1	0	0	0	0	1	1	1	1	5
	jɪ ↔ jɑ	0	0	0	0	0	0	1	0	0	0	0	1
	jɪ ↔ jɪ	0	1	0	1	1	0	0	0	0	0	0	3
	jɪ ↔ ɪ	1	0	0	0	0	1	0	0	0	0	0	2
	Target-Like	0	1	0	1	1	0	1	0	0	0	0	4
	Omission	1	0	1	0	0	1	0	1	1	1	1	7
	% Target-Like	0%	100%	0%	100%	100%	0%	100%	0%	0%	0%	0%	36.36%
	% Omission	100%	0%	100%	0%	0%	100%	0%	100%	100%	100%	100%	63.64%

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
	jɪ ↔ i	1	0	1	0	0	1	0	1	0	0	0	4
	jɪ ↔ jə	0	0	0	0	1	0	0	0	0	0	0	1
	jɪ ↔ jɪ	0	1	0	1	0	0	1	0	1	0	1	5
Yearbook (j+ɪ)	jɪ ↔ jʌ	0	0	0	0	0	0	0	0	0	1	0	1
	Target-Like	0	1	0	1	1	0	1	0	1	1	1	7
	Omission	1	0	1	0	0	1	0	1	0	0	0	4
	% Target-Like	0%	100%	0%	100%	100%	0%	100%	0%	100%	100%	100%	63.64%
	% Omission	100%	0%	100%	0%	0%	100%	0%	100%	0%	0%	0%	36.36%

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
	jɪ ↔ jɪ	1	1	1	0	1	1	1	0	0	0	1	7
	jɪ ↔ wɪ	0	0	0	0	0	0	0	1	0	0	0	1
	jɪ ↔ ɪ	0	0	0	1	0	0	0	0	1	1	0	3
Yippee (j+ɪ)	Target-Like	1	1	1	0	1	1	1	1	0	0	1	8
	Omission	0	0	0	1	0	0	0	0	1	1	0	3
	% TargetlLike	100%	100%	100%	0%	100%	100%	100%	100%	0%	0%	100%	72.73%
	% Omission	0%	0%	0%	100%	0%	0%	0%	0%	100%	100%	0%	27.27%

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
	ji ↔ i	1	0	0	0	1	0	1	1	1	1	1	7
	ji ↔ jæ	0	0	0	0	0	1	0	0	0	0	0	1
	ji ↔ jɪ	0	1	0	0	0	0	0	0	0	0	0	1
Yield (j+i)	ji ↔ ɪ	0	0	1	1	0	0	0	0	0	0	0	2
	Target-Like	0	1	0	0	0	1	0	0	0	0	0	2
	Omission	1	0	1	1	1	0	1	1	1	1	1	9
	% Target-Like	0%	100%	0%	0%	0%	100%	0%	0%	0%	0%	0%	18.18%
	% Omission	100%	0%	100%	100%	100%	0%	100%	100%	100%	100%	100%	81.82%

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
	ji ↔ i	1	0	0	1	0	0	1	1	0	0	0	4
	ji ↔ ji	0	0	0	0	1	0	0	0	0	0	0	1
	ji ↔ jɑ	0	0	0	0	0	1	0	0	0	0	0	1
	ji ↔ jə	0	0	0	0	0	0	0	0	0	0	1	1
	ji ↔ jɛ	0	0	0	0	0	0	0	0	0	1	0	1
Yielding (j+i)	ji ↔ jɪ	0	1	1	0	0	0	0	0	0	0	0	2
	ji ↔ ɪ	0	0	0	0	0	0	0	0	1	0	0	1
	Target-Like	0	1	1	0	1	1	0	0	0	1	1	6
	Omission	1	0	0	1	0	0	1	1	1	0	0	5
	% Target-Like	0%	100%	100%	0%	100%	100%	0%	0%	0%	100%	100%	54.55%
	% Omission	100%	0%	0%	100%	0%	0%	100%	100%	100%	0%	0%	45.45%

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
Yeast (j+i)	ji ↔ i	1	1	1	1	0	1	1	1	1	1	1	10
	ji ↔ jɪ	0	0	0	0	1	0	0	0	0	0	0	1
	Target-Like	0	0	0	0	1	0	0	0	0	0	0	1
	Omission	1	1	1	1	0	1	1	1	1	1	1	10
	% Target-Like	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	9.09%
	% Omission	100%	100%	100%	100%	0%	100%	100%	100%	100%	100%	100%	90.91%

The results in Figure 14 show variation in the accuracy of some of the homorganic glide-vowel target forms, with *yearbook* and *yippee* yielding higher rates of accuracy. The accuracy for *yearbook* in the picture naming task was only 10%, compared with the 63% we observe for the reading task. The higher accuracy for *yearbook* could be due to the fact that most participants needed prompting to produce this word in the picture naming task. The data for *year* in both tasks also show variation, but it is not as extreme.

In addition, Figure 14 shows that there is a high degree of variation between some forms with identical stems. The difference in the accuracy of *yield* and *yielding* is striking, with omission occurring in 81.82% of cases for *yield* and 45.45% of cases for *yielding*. The word level breakdown also shows considerable variation in the forms of *year, yearly,* and *yearbook*, with *year* showing glide omission at 54.55%, *yearly* at 63.64%, and *yearbook* at 36.36%. Although not visible across the board, there does appear to be higher rates of accuracy in the longer words. These apparent lexical effects will be discussed in more detail in section 5 on page 78 (below).

Figure 15. Accuracy [w]+High Back Vowels

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
Woohoo (w+u)	wu ↔ u	1	0	1	1	0	1	1	1	0	1	0	7
	wu ↔ wu	0	0	0	0	1	0	0	0	0	0	0	1
	wu ↔ wə	0	1	0	0	0	0	0	0	0	0	1	2
	wu ↔ wʊ	0	0	0	0	0	0	0	0	1	0	0	1
	Target-Like	0	1	0	0	1	0	0	0	1	0	1	4
	Omission	1	0	1	1	0	1	1	1	0	1	0	7
	% Target-Like	0%	100%	0%	0%	100%	0%	0%	0%	100%	0%	100%	36.36%
	% Omission	100%	0%	100%	100%	0%	100%	100%	100%	0%	100%	0%	63.64%

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
Whoopi (w+u)	wu ↔ hu	0	0	0	0	0	0	1	0	0	1	0	2
	wu ↔ u	1	0	0	0	0	1	0	1	0	0	1	4
	wu ↔ wu	0	0	0	1	0	0	0	0	0	0	0	1
	wu ↔ wʊ	0	1	0	0	1	0	0	0	0	0	0	2
	wu ↔ wʌ	0	0	1	0	0	0	0	0	1	0	0	2
	Target-Like	0	1	1	1	1	0	0	0	1	0	0	5
	Omission	1	0	0	0	0	1	1	1	0	1	1	6
	% Target-Like	0%	100%	100%	100%	100%	0%	0%	0%	100%	0%	0%	45.45%
	% Omission	100%	0%	0%	0%	0%	100%	100%	100%	0%	100%	100%	54.55%

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
Womb (w+u)	wu ↔ u	1	0	0	1	0	0	1	0	0	0	0	3
	wu ↔ wu	0	0	0	0	0	0	0	0	0	1	0	1
	wu ↔ wɑ	0	0	0	0	0	1	0	0	0	0	0	1
	wu ↔ wʊ	0	0	0	0	0	0	0	1	0	0	0	1
	wu ↔ wʌ	0	1	1	0	0	0	0	0	1	0	1	4
	wu ↔ ʊ	0	0	0	0	1	0	0	0	0	0	0	1
	Target-Like	0	1	1	0	0	1	0	1	1	1	1	7
	Omission	1	0	0	1	1	0	1	0	0	0	0	4
	% Target-Like	0%	100%	100%	0%	0%	100%	0%	100%	100%	100%	100%	63.64%
	% Omission	100%	0%	0%	100%	100%	0%	100%	0%	0%	0%	0%	36.36%

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
Woman (w+ʊ)	wʊ ↔ u	1	0	0	0	0	0	0	0	0	0	0	1
	wʊ ↔ wo	0	0	0	0	0	1	0	0	0	0	0	1
	wʊ ↔ wu	0	0	0	1	0	0	0	0	0	0	0	1
	wʊ ↔ wɑ	0	0	0	0	0	0	0	0	1	0	0	1
	wʊ ↔ wʊ	0	1	0	0	1	0	0	0	0	1	1	4
	wʊ ↔ wʌ	0	0	1	0	0	0	0	0	0	0	0	1
	wʊ ↔ ʊ	0	0	0	0	0	0	1	1	0	0	0	2
	Target-Like	0	1	1	1	1	1	0	0	1	1	1	8
	Omission	1	0	0	0	0	0	1	1	0	0	0	3
	% Target-Like	0%	100%	100%	100%	100%	100%	0%	0%	100%	100%	100%	72.73%
	% Omission	100%	0%	0%	0%	0%	0%	100%	100%	0%	0%	0%	27.27%

	Participant	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**Total**
Wool (w+ʊ)	wʊ ↔ u	1	0	0	0	0	0	0	0	0	0	0	1
	wʊ ↔ wu	0	0	0	1	0	0	0	0	0	0	0	1
	wʊ ↔ wʊ	0	1	1	0	1	0	0	0	0	1	1	5
	wʊ ↔ ʊ	0	0	0	0	0	1	1	1	1	0	0	4
	Target-Like	0	1	1	1	1	0	0	0	0	1	1	6
	Omission	1	0	0	0	0	1	1	1	1	0	0	5
	% Target-Like	0%	100%	100%	100%	100%	0%	0%	0%	0%	100%	100%	54.55%
	% Omission	100%	0%	0%	0%	0%	100%	100%	100%	100%	0%	0%	45.45%

	Participant	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**Total**
Would (w+ʊ)	wʊ ↔ u	0	0	0	1	1	0	0	0	0	0	0	2
	wʊ ↔ wu	0	1	0	0	0	0	0	0	0	0	0	1
	wʊ ↔ wʊ	0	0	1	0	0	0	0	0	0	0	1	2
	wʊ ↔ ʊ	1	0	0	0	0	1	1	1	1	1	0	6
	Target-Like	0	1	1	0	0	0	0	0	0	0	1	3
	Omission	1	0	0	1	1	1	1	1	1	1	0	8
	% Target-Like	0%	100%	100%	0%	0%	0%	0%	0%	0%	0%	100%	27.27%
	% Omission	100%	0%	0%	100%	100%	100%	100%	100%	100%	100%	0%	72.73%

	Participant	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**Total**
Wood (w+ʊ)	wʊ ↔ u	0	0	0	1	0	0	0	0	0	0	0	1
	wʊ ↔ wu	0	1	0	0	1	0	0	1	0	1	0	4
	wʊ ↔ wʊ	0	0	1	0	0	0	0	0	0	0	1	2
	wʊ ↔ ʊ	1	0	0	0	0	1	1	0	1	0	0	4
	Target-Like	0	1	1	0	1	0	0	1	0	1	1	6
	Omission	1	0	0	1	0	1	1	0	1	0	0	5
	% Target-Like	0%	100%	100%	0%	100%	0%	0%	100%	0%	100%	100%	54.55%
	% Omission	100%	0%	0%	100%	0%	100%	100%	0%	100%	0%	0%	45.45%

	Participant	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**Total**
Wolf (w+ʊ)	wʊ ↔ wɔ	0	0	0	0	0	0	1	0	0	0	0	1
	wʊ ↔ wʊ	0	1	1	0	1	1	0	1	0	1	1	7
	wʊ ↔ wʌ	0	0	0	0	0	0	0	0	1	0	0	1
	wʊ ↔ ʊ	1	0	0	1	0	0	0	0	0	0	0	2
	Target-Like	0	1	1	0	1	1	1	1	1	1	1	9
	Omission	1	0	0	1	0	0	0	0	0	0	0	2
	% Target-Like	0%	100%	100%	0%	100%	100%	100%	100%	100%	100%	100%	81.82%
	% Omission	100%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	18.18%

In the [w] target context, we find similar accuracy rates between the picture naming and reading tasks. The fact that the accuracy rates are similar in both production tasks, and that the reading task contains orthographic forms of the stimuli and the picture task does not, suggests that orthography may not be considered as a significant factor in this study. However, orthography should be considered in any subsequent research to provide confirmation that it bears no influence. There is a striking difference in the accuracy of the noun *wood* and the modal *would,* with the modal showing a higher rate of glide omission at 72.73% compared with omission in the noun form at 45.45%. This variation will be discussed in section 5, which is dedicated to lexical effects.

Together, Figure 14 and Figure 15 show that all participants, with the exception of Participant 2, omit both of the glides [j] and [w] in at least one production. Participant 2 only omits [j], in a single word form (*yeast*). These figures also revealed a tendency for the participants to omit the glide in forms with a tense vowel, namely [j+i] and [w+u] sequences, with those containing a lax vowel undergoing glide omission less systematically. This omission of the glide in homorganic glide-tense vowel sequences could provide evidence in favour of a perceptual cause to this phenomenon, as it is in these segments that the two phones are the most similar, and are thus prone to perceptual assimilation. However, an articulatory explanation cannot be discounted.

In the following section I will describe the results of the perception task, with the hope of establishing whether the cause of the omission is purely articulatory, or have origins in perception, as suggested by the omission of glides in homorganic glide-vowel sequences.

3.3. Perception Results

As mentioned in Chapter 3, I analyzed both the accuracy and the reaction times recorded through the ABX perception experiment. In some cases, the participants were unable to select a response within the time frame given for the ABX task (5000ms). In these cases, I have kept the result for accuracy because for the purposes of this study it is considered an error as the participant was unable to give a response, this is consistent for both target and non-target stimuli. However, I have removed these from the reaction time calculations, as they compute as 5001ms response times, which would make the averaged reaction times for actual responses appear considerably slower. As we will see below, perceptual difficulties were observed across the relevant contexts.

3.3.1. Accuracy across Phonotactic Contexts

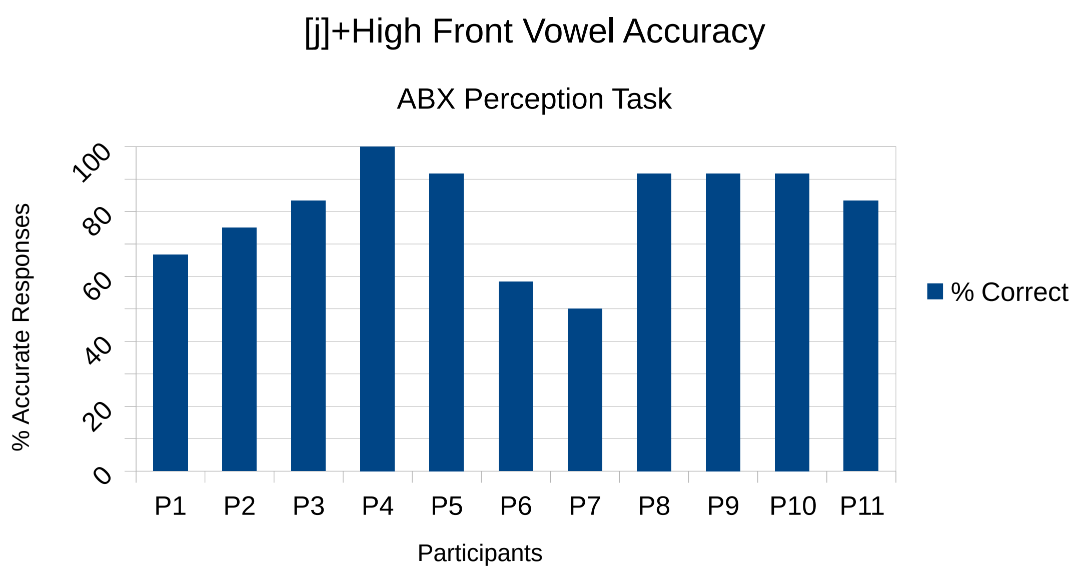
Table 7 shows accuracy of the four non-target contexts. Performance is close to ceiling, however with Participant 8 showing scores that are slightly lower than that of the other participants. For the most part, we can relate these scores to a task effect for this participant, who often failed to offer a response within the given time frame of 5000ms. Finally, the glide+mid vowel contexts show slightly lower accuracy than the filler and control groups, but when the scores are averaged, the difference is not important, so these contexts are grouped with the other non-target contexts.

Table 7. Accuracy in the Non-Target Contexts

Sequence Type	**Participant**										
	1	2	3	4	5	6	7	8	9	10	11
Filler (n=69)	97% (67/69)	93% (64/69)	100%	99% (68/69)	100%	88% (61/69)	94% (65/69)	74% (51/59)	100%	99% (68/69)	100%
Control (n=39)	85% (33/39)	95% (37/39)	100%	100%	100%	92% (36/39)	92% (36/39)	74% (29/39)	97% (38/39)	100%	100%
[j]+mid front vowel Sequence (n=9)	100%	100%	100%	100%	100%	100%	89% (8/9)	56% (5/9)	100%	100%	100%
[w]+mid back vowel Sequence (n=6)	100%	67% (4/6)	100%	100%	83% (5/6)	100%	100%	100%	100%	100%	100%

Figure 16 gives charts of the target homorganic glide-vowel sequence contexts (12 triads for /j/ + high front vowels and 9 triads for /w/ + high back vowels). As we can see the perceptual accuracy of the homorganic glide-vowel target forms is considerably lower than the scores in Table 7.

Figure 16. Accuracy in Target Homorganic Glide-Vowel Contexts



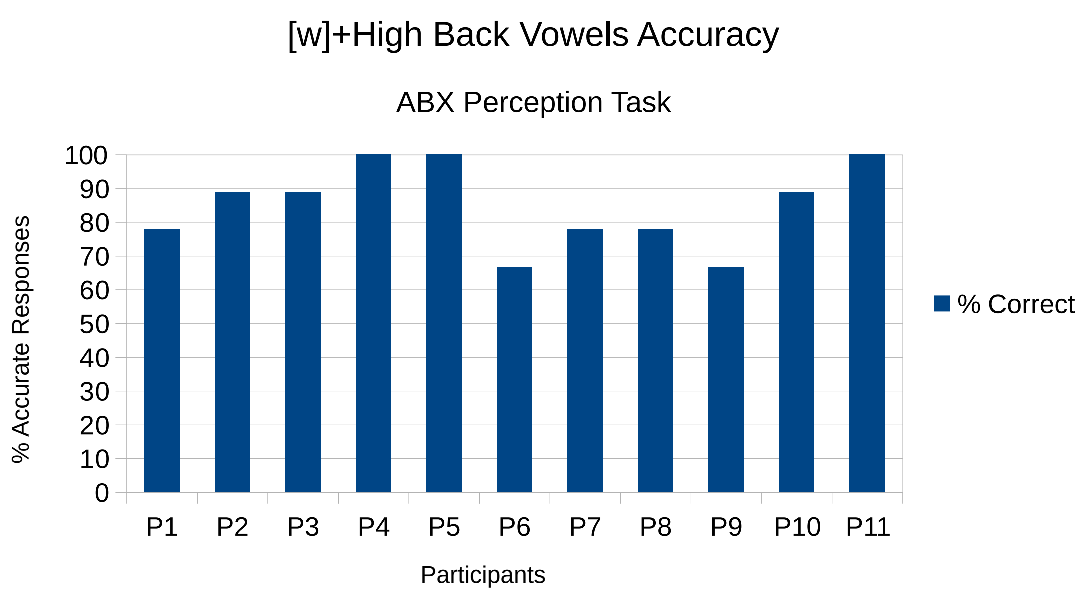
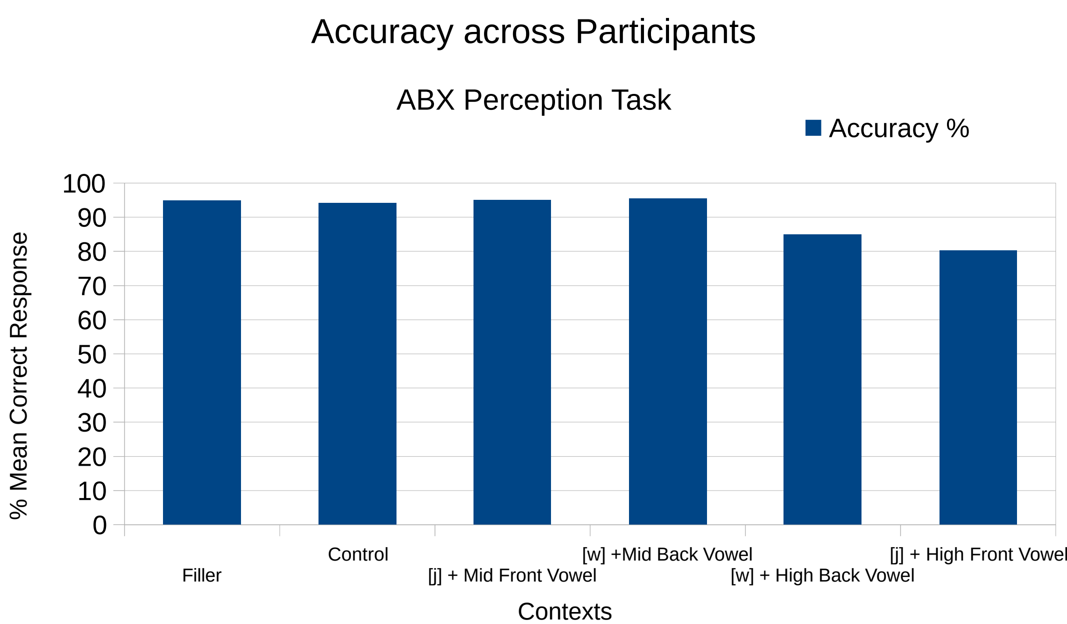


Table 7 and Figure 16 present evidence of individual variation. With the exception of Participant 4, no participant achieves ceiling performance across all of the homorganic glide-vowel target contexts. In contrast, several participants are hitting 100% accuracy in the non-target contexts.

In order to provide a clear comparison of the contexts, I present the average performance of all 11 participants in Figure 17 (below). As we can see in Figure 17, the target forms [ji], [jɪ], [wu] and [wʊ] have a lower percentage of accuracy than all other forms, with an average of around 80-85% across all the participants. The filler, control, as well as the glide+mid vowel contexts behave in similar ways, with scores revolving around the 95% mark. In contrast to this, the homorganic glide-vowel target sequences average a 10-15% higher error rate, which is striking when compared to the almost ceiling performance of the other contexts.

Figure 17. Average Performance for each Category



3.4. Perception: Word Level Breakdown

In the following section I provide a word-level breakdown for accuracy of the homorganic glide-vowel target stimuli used in the ABX perception task.

3.4.1. Individual Glide-Vowel Sequences

In order to ascertain whether the higher error rate in the homorganic glide-vowel target forms is caused by isolated stimuli, this section looks at the stimuli with homorganic glide-vowel sequences individually.

Table 8 (below) shows the number of correct responses by each participant, for each word (presented three times in the ABX task). These correct responses are important as they enable us to observe whether the third trial (response/ triad type 3) was easier for participants and, more importantly, to verify that the participant could perform the task. The numbers in parentheses in the table represent the correct response for the ABX task, the responses that correspond to these numbers are demonstrated below:

[jild] – [ild] – [jild] = 1 [ild] – [jild] – [jild] = 2 [jild] – [jild] – [jild] = 3

Table 8 illustrates the homorganic glide-vowel target stimuli for the ABX task. The accuracy on all other stimuli was close to ceiling, therefore a detailed breakdown of all stimuli is not necessary.

Table 8. Homorganic Glide-Vowel Sequences: Accuracy

Sequence	Stimulus (triad type)	Participant												Mean
		1	2	3	4	5	6	7	8	9	10	11	Total Correct	%
j+i	yield (3)	1	1	1	1	1	1	1	0	1	1	1	10	91%
	yield (1)	0	1	1	1	0	0	0	1	1	1	1	7	64%
	yield (2)	1	0	1	1	1	0	0	1	1	0	0	6	55%
j+ɪ	yippee (3)	1	1	1	1	1	1	1	1	1	1	1	11	100%
	yippee (1)	1	0	0	1	1	1	1	1	1	1	1	9	82%
	yippee (2)	1	1	1	1	1	1	1	1	1	1	1	11	100%
j+i	yeast (3)	1	1	1	1	1	1	1	1	1	1	1	11	100%
	yeast (1)	0	0	0	1	1	0	0	1	1	1	1	6	55%
	yeast (2)	0	1	1	1	1	0	0	1	0	1	1	7	64%
j+ɪ	year (3)	1	1	1	1	1	1	1	1	1	1	1	11	100%
	year (1)	0	1	1	1	1	1	0	1	1	1	1	9	82%
	year (2)	1	1	1	1	1	0	0	1	1	1	0	8	73%
w+ʊ	wolf (3)	1	1	1	1	1	1	1	1	1	1	1	11	100%
	wolf (1)	1	1	1	1	1	1	1	1	1	1	1	11	100%
	wolf (2)	1	1	1	1	1	0	1	1	1	1	1	10	91%
w+ʊ	would (3)	1	1	1	1	1	1	1	1	1	1	1	11	100%
	would (1)	0	1	1	1	1	1	1	1	0	1	1	9	82%
	would (2)	1	1	1	1	1	1	1	1	0	1	1	10	91%
w+u	womb (3)	1	1	1	1	1	1	1	1	1	1	1	11	100%
	womb (1)	1	0	0	0	1	0	0	0	1	0	1	4	36%
	womb (2)	0	1	1	1	1	0	0	0	0	1	1	6	55%

Table 8 shows that all of the *same* trials (response/triad type 3) have a considerably higher accuracy rate than the *different* trials (responses/triad types 1 and 2) across all participants. The accuracy of the *different* trials is similar, with only around a 10% difference between them. With the exception of *year* participants showed higher accuracy with triad type 2 (e.g. ield, yield, yield), this triad type shows the two *same* stimuli as being adjacent. The proximity of these same stimuli suggests that this task used both phonological knowledge and phonological memory, giving weight to the hypothesis that perception is the root cause of the omission patterns. These data also clearly show that the [ji] and [jɪ] stimuli have the lowest rate of accuracy, followed by the [wu] and [wʊ] stimuli, this is an opposite trend to what was observed in the production results in section 3 (above).

3.4.2. Reaction Times

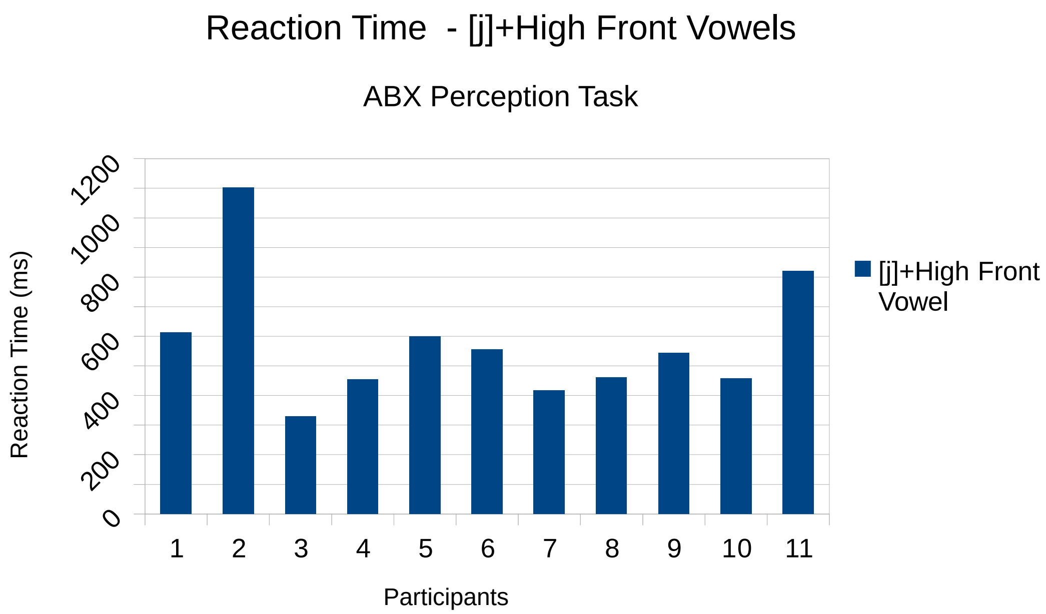
Recall from above that I have omitted the timed-out responses from my report. The software lists these timed out trials as 5001ms and this would skew the averages, making reaction times appear slower for the homorganic glide-vowel target forms. Table 9 shows the average reaction time across all participants for all of the non-target categories. The reaction times for the non-target contexts range from around 250-750ms, with a majority of people on average reacting in about 400ms.

Table 9. Reaction Time in Non-Target Contexts

Sequence Type	**Participant**										
	1	2	3	4	5	6	7	8	9	10	11
Filler	505	661	328	385	333	375	359	251	459	573	378
Control	435	649	283	365	509	428	388	402	414	498	479
[j]+mid front vowel Sequence	306	661	301	409	750	282	398	520	273	589	327
[w]+mid back vowel Sequence	399	408	312	548	294	389	454	386	356	554	484

The charts in Figure 18 (below) reveal noticeably slower reaction times in the homorganic glide-vowel target contexts for several participants, compared to those for the non-target contexts in Table 9. These charts also show a high degree of individual variation with reaction times ranging from approximately 300 to 1100 ms for both of the homorganic glide-vowel target contexts.

Figure 18. Reaction Time in Homorganic Glide-Vowel Target Contexts



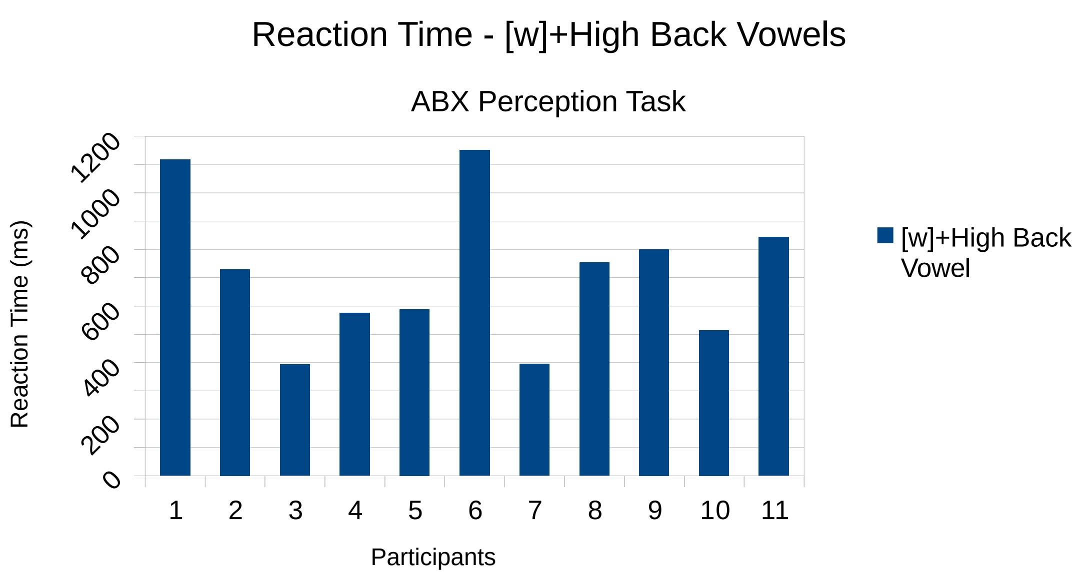
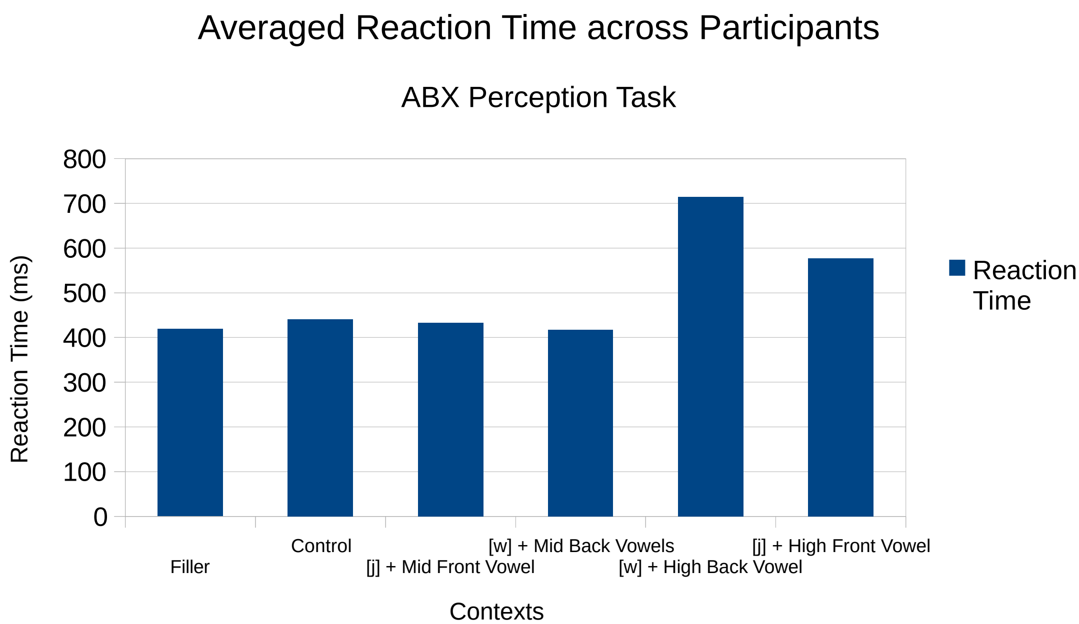


Figure 18 shows that the homorganic glide-vowel sequences yield slower reaction times, with many participants reacting in more than 500ms for [j] and 700ms for [w], this mirrors the patterns observed thus far. Note also that for Participant 8 two reaction times have been removed due to this participant not responding within the required time frame.

These observations are confirmed by the inspection of Figure 19, which shows the average reaction times across participants for each of the relevant contexts. The averaged reaction times show close to equal performance across all of the non-homorganic contexts, but slower reaction times for the target homorganic glide-vowel sequences, at around 600-700ms.

Figure 19. Averaged Reaction Times



3.4.3. Homorganic Glide-Vowel Target Reaction Times by Lexical Item

Table 10 shows a breakdown of the reaction times across the homorganic glide-vowel target words. As already mentioned, timed-out responses have been removed to avoid misrepresenting reaction times on actual responses.

Table 10. Homorganic Glide-Vowel Perception: Reaction Times

Sequence	**Stimulus (triad type)**	**Participant**											**Mean**
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	
j+i	yield (3)	427	460	224	435	55	170	68	--	394	1280	1217	473
	yield (1)	709	1072	171	316	2441	205	93	155	443	244	385	567
	yield (2)	606	2271	384	552	332	725	597	260	295	294	1455	706
j+ɪ	yippee (3)	351	239	406	520	334	23	169	149	144	609	371	301
	yippee (1)	131	2094	329	504	367	226	500	444	729	698	212	567
	yippee (2)	468	401	653	515	489	407	534	303	1018	508	1696	636
j+i	yeast (3)	408	310	290	360	40	1036	182	162	251	403	412	350
	yeast (1)	379	567	329	354	952	463	313	1589	1278	473	1896	781
	yeast (2)	399	1481	234	1033	970	2128	788	1516	499	257	444	886
j+ɪ	year (3)	326	1348	363	136	728	753	275	49	423	180	102	426
	year (1)	2369	1789	208	315	45	380	538	200	1016	112	985	723
	year (2)	782	1195	359	423	432	148	959	246	23	438	662	515
w+ʊ	wolf (3)	499	272	406	442	78	13	150	456	355	795	79	322
	wolf (1)	275	1049	653	421	1493	944	653	582	800	650	263	708
	wolf (2)	265	92	254	447	374	2310	304	637	398	754	3573	855
w+ʊ	would (3)	484	256	360	549	159	227	420	70	317	746	303	354
	would (1)	2440	709	413	402	335	1016	610	349	1201	522	1031	821
	would (2)	3208	1014	449	532	231	1856	600	1376	1872	317	377	1076
w+u	womb (3)	361	1330	92	551	834	532	246	270	401	181	439	476
	womb (1)	1907	765	216	595	1523	2051	190	--	623	337	1176	938
	womb (2)	606	1079	686	1235	261	1403	376	2347	1231	326	345	899

Table 10 shows that there is a slower mean reaction time for a majority of the homorganic glide-vowel target forms than we observe in the non-target sequences. In the ABX *same* trials, there are slightly faster reaction times; however, these are not as pronounced as the accuracy scores in Table 8 (page 69, above) could have suggested.

3.5. Comparison of Accuracy and Reaction Time: Perception Data

The data from the perception experiment show similar trends between the accuracy and reaction time scores for the homorganic glide-vowel sequences, in that these sequences have lower accuracy also have a slower mean reaction times than their non-target counterparts. The 11 participants were consistent in their accuracy and reaction time for the non-target sequences. This consistency clearly highlights the difficulty that the participants have with both the [j+i/ɪ] and [w+u/ʊ] sequences.

3.6. Summary of Experimental Results

In order to highlight the consistent difficulty that participants had with the homorganic glide-vowel target sequences, the table below shows all of the homorganic glide-vowel target form results for the three experiments.

Table 11. Averaged Results for all Experiments[4]

		All Participants							
		Perception (averaged)			Production (averaged)				
					Reading Task		Naming Task		
Word	**Context**	**Correct**	**Error**	**Reaction Time**	**Glide Produced**	**Glide Deleted**	**Glide Produced**	**Glide Deleted**	**Frequency**
Yield	j+i	70%	30%	574	18%	82%			206
Yielding	j+i				55%	45%			22
Yeast	j+i	73%	27%	587	9%	91%			44
Yiddish	j+ɪ				55%	45%			37
Yin	j+ɪ				27%	73%			168
Year	j+ɪ	85%	15%	573	45%	55%	18%	82%	14174
Yearly	j+ɪ				36%	64%			45
Yearbook	j+ɪ				64%	36%	10%	90%	241
Yippee	j+ɪ	94%	6%	615	73%	27%			72
Woohoo	w+u				36%	64%			14
Whoopi	w+u				45%	55%			82
Womb	w+u	64%	36%	790	64%	36%			177
Wounded	w+u				64%	36%			992
Woman	w+ʊ				73%	27%	73%	27%	22166
Wool	w+ʊ				55%	45%	100%	0%	161
Would	w+ʊ	91%	9%	737	27%	73%			90162
Wood	w+ʊ				55%	45%	45%	55%	1377
Wolf	w+ʊ	97%	3%	611	82%	18%	64%	36%	1034

4 Only two participants (5 and 9) gave unprompted productions for *wool* in the picture naming task.

Concerning the results of the production experiment, Table 11 shows that in three (*yield*, *yeast* and *year*) of the four cases with a higher rate of errors (15% or higher), there are also high rates of omission in the production tasks. These contexts showing lower rates of accuracy are also matched with longer reaction times recorded during the perception experiment.

One issue that needs to be mentioned here is the lack of stimuli that overlap in all three tasks. This was an oversight on my part and something that should be amended in future research. The degree of individual variability seen with certain stimuli shows that individual lexical items can affect an individuals phonological behaviour. The lack of overlap limits the amount of comparison that can take place between the perception and production experiments in order to fully understand these lexical effects. I will elaborate on this oversight further in section 8.2 (Chapter 5).

Table 11 also provides the frequency of occurrences of each homorganic glide-vowel target word, as reported in the SUBTLEXus database (http://subtlexus.lexique.org/moteur2/index.php) (Brysbaert & New 2009). These frequency scores vary widely between words ranging from 14 to 90162. I discuss the implications of these frequency scores will be discussed in Chapter 5 (below).

These results also present various lexical effects in forms such as *year, yearly,* and *yearbook.* These forms show considerable differences in accuracy, despite the fact that the stem of each word is identical in its orthography and pronunciation. These lexical effects extend to forms which vary in their orthography and use, such as *would* and *wood*, even though in terms of pronunciation they are indistinguishable.

The forms that present lexical effects also vary in their frequency ranking. Some of these forms such as *would,* are very frequent yet show low accuracy, whereas some infrequent forms show higher accuracy. The results presented in this study suggest that there is no direct correlation between frequently occurring forms and the perceptual and productive ability in the Japanese and Korean participants.

Up to this point, my data description collapsed all data for both Japanese and Korean speakers. I discuss potential differences between the two groups below.

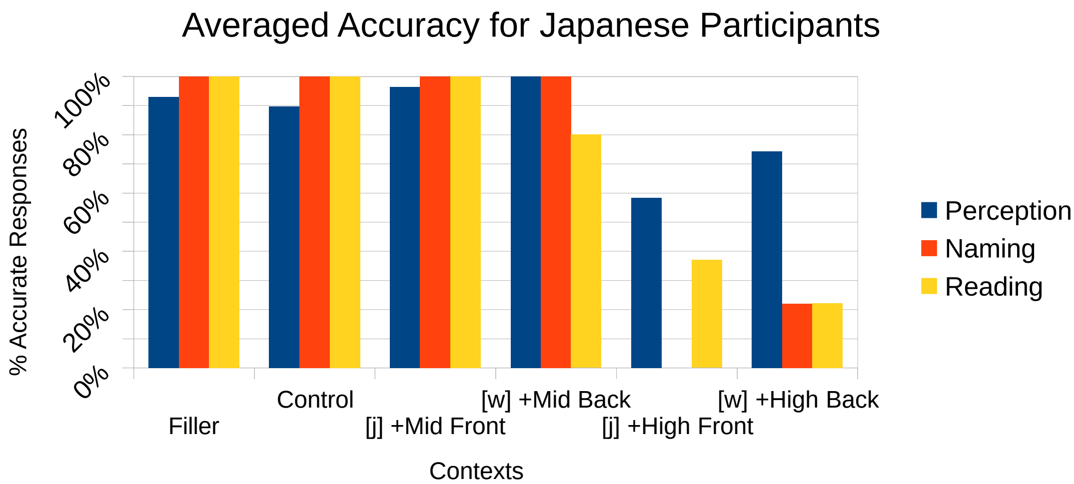
4. Language Specific Results

In this section, I discuss the results from the three tests for the Japanese participants and the Korean participants. In the interest of brevity, and due to the small sizes of both groups of participants, I am only providing the mean scores of all participants for the three tests. Recall that I could only recruit three participants in the Japanese group and eight participants in the Korean group.

4.1. Summary of Japanese Data

Figure 20 provides a summary of the accuracy results for the Japanese group for the three experiments. The chart shows a clear difference between the accuracy rate in the homorganic glide-vowel targets versus non-target contexts with all three tests, with the percentage of accurate results being much lower for the homorganic glide-vowel target contexts than for the non-target contexts.

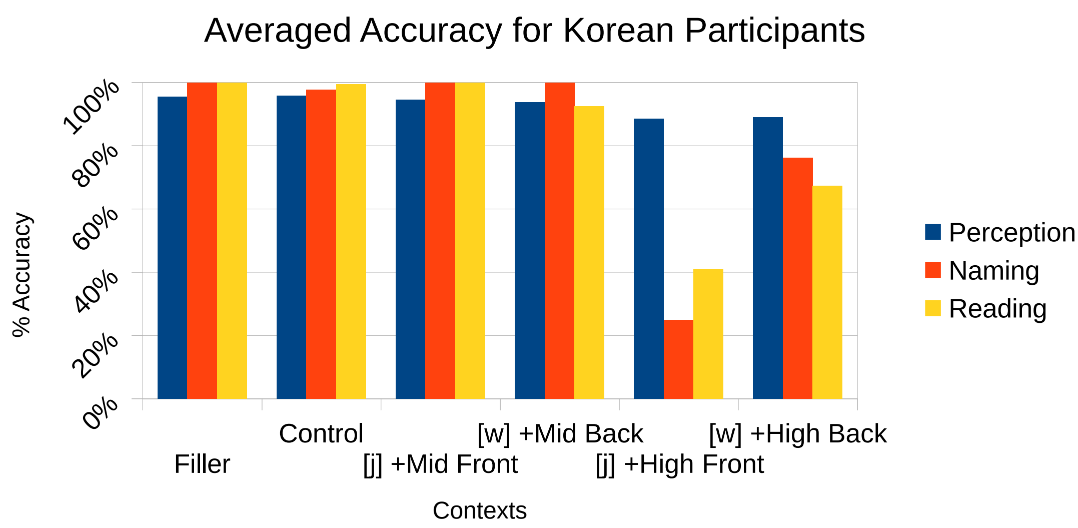
Figure 20. Summary of Japanese Data



4.2. Summary of Korean Data

Figure 21 displays similar scores for the Korean participants. The chart shows close to ceiling performance for the non-target contexts and a lower rate of accuracy for the homorganic glide-vowel target contexts. In spite of generally higher perception scores in the two homorganic glide-vowel target contexts than we observed in the data for the Japanese group, production scores are still low, especially for the [j] target sequences.

Figure 21. Summary of Korean Data



In previous sections I have mentioned some lexical effects observed in these data. I elaborate on these effects in the next section.

5. Lexical Effects

Figure 22 illustrates the production of the modal and noun forms for *would/wood* ([wʊd]) from the reading task. While these production patterns are similar across both words, we observe a difference in accuracy between these forms, with the modal only yielding 27% of the target-like responses, and the noun form showing 55% of target-like productions. Participants 5, 8 and 10 produced the noun form in a target-like fashion, but they omit the glide in the modal form.

Figure 22. Reading Task Productions of *Would* and *Wood*

	Participant	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**Total**
Would (w+ʊ)	wʊ ↔ u	0	0	0	1	1	0	0	0	0	0	0	2
	wʊ ↔ wu	0	1	0	0	0	0	0	0	0	0	0	1
	wʊ ↔ wʊ	0	0	1	0	0	0	0	0	0	0	1	2
	wʊ ↔ ʊ	1	0	0	0	0	1	1	1	1	1	0	6
	Target-Like	0	1	1	0	0	0	0	0	0	0	1	3
	Omission	1	0	0	1	1	1	1	1	1	1	0	8
	%Target-Like	0%	100%	100%	0%	0%	0%	0%	0%	0%	0%	100%	27.27%
	% Omission	100%	0%	0%	100%	100%	100%	100%	100%	100%	100%	0%	72.73%

	Participant	1	2	3	4	5	6	7	8	9	10	11	Total
Wood (w+ʊ)	wʊ ↔ u	0	0	0	1	0	0	0	0	0	0	0	1
	wʊ ↔ wu	0	1	0	0	1	0	0	1	0	1	0	4
	wʊ ↔ wʊ	0	0	1	0	0	0	0	0	0	0	1	2
	wʊ ↔ ʊ	1	0	0	0	0	1	1	0	1	0	0	4
	Target-Like	0	1	1	0	1	0	0	1	0	1	1	6
	Omission	1	0	0	1	0	1	1	0	1	0	0	5
	%Target-Like	0%	100%	100%	0%	100%	0%	0%	100%	0%	100%	100%	54.55%
	% Omission	100%	0%	0%	100%	0%	100%	100%	0%	100%	0%	0%	45.45%

There are also several forms which have identical stems, but vary in their accuracy. *Year, yearly,* and *yearbook* show different rates of accuracy both within and across the experimental tasks. *Yield* and *yielding* also show similar behaviours. The table below shows differences in accuracy across these forms.

Table 12. Lexical Effects and Accuracy

Stimuli	Picture Naming Accuracy	Reading Accuracy	Perception Accuracy
Year	18%	45%	85%
Yearly		36%	
Yearbook	10%	64%	
Yield		18%	70%
Yielding		55%	

There was also a large degree of individual variation in the accuracy and reaction times for the word *womb.* After investigating whether there is a potential gender bias in individual performances, I found that an equal number of males and females showed 100% accuracy in their productions of this word. Furthermore, due to parenthood being a sociolinguistic factor, I looked at whether accuracy related to an individual being a parent. Of those individuals with ceiling performance, there was a 50:50 ratio of parents and non-parents. From this I can conclude that these individual results are not a result of sociolinguistic factors. In order to explain why this word yielded such variation in accuracy, further investigation is needed.

6. Vowel Substitution

Recall that homorganic glide-vowel sequences are known to be marked cross-linguistically (Ohala & Kawasaki 1984; Rose 1999; Lee 1994). Evidence of this can also be found in the vowel substitution patterns, which make the segments in the target homorganic glide-vowel sequences more acoustically different from one another. Figure 9, Figure 10, Figure 14, and Figure 15 all show evidence of vowel substitutions in the target-like productions. The table below summarizes the vowel substitutions observed in these data for all 11 participants. In the naming task results for the [w] target context, there are 35 tokens; this is due to there only being two utterances of the form *wool*, due to a majority of participants being unfamiliar with this word.

Table 13. Vowel Substitution in Homorganic Glide-Vowel Target Forms

Reading Task			Naming Task		
Vowel Substitution	Number of occurrences		Vowel Substitution	Number of occurrences	
ji ↔ jæ	1		jɪ ↔ jə	3	
ji ↔ jɑ	1				
ji ↔ jə	1				
ji ↔ jɛ	1				
jɪ ↔ ju	1				
jɪ ↔ jʌ	1				
ji ↔ jɪ	4				
jɪ ↔ ju	1				
jɪ ↔ ji	1				
jɪ ↔ jə	3				
jɪ ↔ jɑ	1				
Total	**16**	**Total Utterances = 99**	**Total**	**3**	**Total Utterances = 22**
wu ↔ wə	2		wʊ ↔ wu	3	
wu ↔ wʊ	5		wʊ ↔ wɔ	1	
wu ↔ wʌ	10		wʊ ↔ wʌ	3	
wu ↔ wɑ	1				
wʊ ↔ wo	1				
wʊ ↔ wu	7				
wʊ ↔ wɑ	1				
wʊ ↔ wʌ	2				
Total	**29**	**Total Utterances = 99**	**Total**	**7**	**Total Utterances = 35**

Finally, as mentioned in various places above, individual participants tended to show a high level of consistency. I provide more detail about this in the next section.

7. Individual Patterns

Results for individual speakers confirm that the patterns observed in both group and language specific data are consistent with each participant's individual scores. In the Appendix (page 107 below), charts are provided for each participant's performances for each task. These charts clearly indicate that when perceptual accuracy is high, the resultant production accuracy is also high, and that when perception is lower, production accuracy is also lower.

However, some asymmetries are visible when the overall data are considered for each participant. There are five cases (one in the Japanese group, four in the Korean group) that suggest good perceptability for a given homorganic glide-vowel target form yet a low score in production. These asymmetries clearly suggest that good perception is a condition, but not a guarantee, for accurate production.

In the next chapter I discuss all of the data described above in light of the theoretical background, my hypotheses and research question. I then present my conclusion based on the empirical evidence presented in this thesis, from a language-specific and group perspective.

Chapter 5. Discussion and Conclusion

1. Introduction

The previous chapters have addressed the phenomenon of glide omission in Japanese and Korean learners of English. Based on the existing literature, I entertained two hypotheses to explain this phonological process. The first hypothesis is that difficulties in perception cause difficulties in production. The second hypothesis is that transfer effects from the L1 result in poor production, due to the lack of homorganic glide-vowel sequences in the speakers' L1s. In the following sections, I discuss the relationship between the results of the three experimental tasks presented in Chapter 4 and existing theoretical frameworks, in order to discuss the implications of my findings, conclude this work, and offer a foundation for future studies.

I argue that despite some literature stating there is no relationship between perception and production (Sheldon & Strange 1982; Cho & Jeoung 2013), the root cause of the glide omission pattern at the centre of this thesis, primarily lies in perceptual difficulties, which in turn result in inaccurate productions. I also acknowledge that transfer effects, themselves driven by phonotactic effects, cannot be discounted and do offer an explanation for the variation in rates of omission between the two glides as well as for the pattern of vowel substitution in the target-like forms. It has been widely acknowledged that homorganic glide-vowel sequences are marked cross-linguistically. The participants in this study display two strategies to make these sequences less marked forms and more faithful to the phonotactics of their native language: they either omit the glide altogether, or substitute the vowel in ways that make it easily distinguishable from the preceding glide.

2. Perception Drives Production

Four potential scenarios are proposed in the literature for the relationship between perception and production, as discussed in Chapter 1, section 2.1(page 2, above). Firstly, if perceptual abilities are poor, then production abilities will also be poor. Secondly, that perceptual abilities are good, but individuals have poor productive abilities. Thirdly, in line with the first scenario, if perceptual abilities are good, then productive abilities will also be good, because the two skills go hand in hand. The fourth scenario is unattested in this study, whereby perceptual abilities are poor, but production is good.

The empirical evidence collected in this pilot shows substantial evidence highlighting a correlation between accuracy in perception and in production. This correlation is visible through individual, language-specific and whole-group data, as demonstrated in Chapter 4. For the most part, the data from the current study evidence the first and third scenarios, namely that there is a relationship between perception and production and that these two skills go hand in hand, or that perception is a prerequisite of production.

There is also evidence for the second scenario, in particular from participants 4, 5 and 8, each of whom shows excellent accuracy in perception, when compared with the control and filler forms, but poorer production. This apparent lag between perception and production can be captured in the recently proposed A-Map Model (McAllister Byun, Inkelas & Rose 2016), which provides a framework to explain potential mismatches between perception and production. The A-Map is a model of articulatory reliability which involves interactions between a speaker's phonological grammar and the mapping of articulatory representations that correspond to perceptual units (e.g. sounds, sound sequences, syllables, words). This model predicts variability in cases where the perceptual target is ill-defined in the speaker's representations, but more systematic variation as the different aspects of the perceptual unit (e.g. manner or place of articulation) are acquired by the learner. Recall that homorganic glide-vowel [ji] or [wu] sequences involve rapid and subtle transitions between two closely-related sounds. As this sequence is grammatically possible in English but not in Japanese or Korean, in the present case, the grammatical component of this model imposes a bias against the English glide-vowel sequences. This bias may weigh on every component of the system, from perception to articulation. The data from the present pilot study are consistent with this general prediction.

The results reported in Chapter 4 also reveal a certain degree of individual variation, with some participants showing near perfect accuracy across all of the tests. Even in these participants we find evidence relating to perceptual abilities driving production, albeit to a lesser degree than in other participants. However, given that transfer effects may affect both perception and production, it also blurs our outlook on both components of the system. I elaborate on this complicated relationship in the following sections.

2.1. Evidence for Perception as a Root Cause

Results from this study indicate that in all contexts in which perception accuracy is low, there are higher rates of glide omission in production. In the non-target forms, namely the filler, control, and glide+mid vowel sequences, the perception scores average at around 95% accurate. I consider this to be ceiling performance, allowing for 5% human error. Figure 16 shows the accuracy in perceiving the homorganic glide-vowel target stimuli for most participants is lower, with an average of 85% accuracy for [w] targets, and 80% for [j] targets. This suggests that the homorganic glide-vowel target sequences are harder to perceive than all other contexts examined in this study, for both the Japanese and Korean participants. This provides strong evidence for the above hypothesis.

The individual scores further suggest that the Japanese participants have in general more difficulty that the Korean participants in their perception of homorganic glide-vowel sequences. This perceptual difficulty for the Japanese individuals is likely due to the fact that they have fewer glide-vowel contexts in their L1, and fewer of the target vowels in their L1 inventory, than the Koreans. This difficulty in perception is likely to be a result of transfer, due to the of the lack of these sequences in the L1 and the universal markedness of the target homorganic glide-vowel sequences. The Japanese participants, in most cases, also had a lower self-perceived English proficiency, and given that transfer is more likely to be influential with lower target language proficiency, there is additional evidence for transfer effects. I elaborate on these factors in a subsequent section.

2.2. Reaction Time

The reaction time scores collected in the ABX perception task also offer evidence for the root cause of glide omission to be perceptual. With the exception of Participant 10, everyone shows a faster reaction time for the non-target contexts, suggesting that these contexts are easier to perceive. In eight of the 11 participants [w] target sequences result in a slower reaction time suggesting that these sequences are harder to perceive. However, eight participants of the 11 showed higher accuracy in perception and ten showed higher accuracy in production for these same [w] target sequences. In order to establish why there is a mismatch between reaction time and accuracy, further research will need to be conducted.

3. Implications for Existing Theories

In chapter 2, I outlined several theoretical frameworks designed to explain the empirical findings of this study. In the following sub-sections I revisit each of these theories in light of my findings, and discuss their applicability to the phenomenon of glide omission in Japanese and Korean speakers of English.

3.1. The Contrastive Analysis Hypothesis and the Markedness Differential Hypothesis

The Contrastive Analysis Hypothesis (CAH; Lado 1957) predicts that segments in an L2 which are different from the L1 are harder to acquire. As I have discussed above, neither Japanese nor Korean have [ɪ] or [ʊ] in their inventory, or permit [jɪ] or [wʊ] glide-vowel sequences. [ji] and [wu] sequences are also unattested in both L1s, but the vowels do exist ([i] in both Japanese and Korean, and [u] in Korean). However, there are fewer cases of omission in homorganic glide-vowel target forms containing [jɪ] or [wʊ] than in the forms containing [ji] or [wu] sequences, despite these sequences also being prohibited in the L1s. This suggests that the sequences with vowels which are not present in the L1 are in fact easier to acquire, yielding higher rates of accuracy and lower rates of omission. This has problematic implications for the CAH. However, the CAH makes predictions based on the acquisition of segments as opposed to sequences, so it cannot be fully applied to this context. The MDH (Eckman 1977) offers a better explanation for the patterns observed in this study. The model predicts that forms which are absent from the L1 and are considered to be marked forms are difficult to acquire. As I have discussed above the homorganic glide-vowel sequences are considered marked cross-linguistically due to the minimal acoustic differences in the two sounds of the sequence (Ohala & Kawasaki 1984). This markedness combined with the absence of these sequences from the L1 provides an account for the difficulties in acquisition observed in this study.

3.2. The Speech Learning Model

The Speech Learning Model (Flege 1981) states, in contrast to the CAH, that segments that have no L1 counterpart are easier to acquire. This prediction may offer an explanation as to why the [jɪ] and [wʊ] sequences have higher rates of accuracy, due to the absence of the vowels and the sequences from the L1. However, Japanese and Korean both have [i] in their inventories, and Korean also has [u] in its inventory, but prohibit [ji] and [wu] sequences. This means that there is a degree of contradiction between what is predicted for segments and sequences. As the sequences [ji] and [wu] have no L1 counterpart, they should be easy to acquire, but clearly are not. Therefore, although this framework offers an explanation for some of the omission contexts, it cannot accurately capture all of the patterns seen in these data.

3.3. The Perceptual Assimilation Model

The Perceptual Assimilation Model (Best 1991; Best & Tyler 2007), as I have discussed, predicts that articulatory gestures aid perception. This model also states that a phone in an L2 will be assimilated into good or bad examples of phones in the L1. This framework offers an explanation as to why glide omission occurs, as a predicted outcome of the Single Category Assimilation hypothesis. This concept of single category assimilation suggests that [ji] and [i] may in fact be perceived as two examples of the native category [i]. Therefore, the speaker assimilates the sequence into an L1 category resulting in a production like [ild] instead of [jild]. However, this idea of Single Category Assimilation in this context is based on the assumption that due to the gestural similarities of [j] and [i], they are perceived as a single segment when adjacent. In this sense, it remains unclear whether the patterns we observe in the data may represent canonical causes of perceptual assimilation between two segments. This model also fails to account for the vowel substitution patterns concomitant with glide production.

3.4. Phonotactic Rareness

As discussed in section 3.5 of Chapter 2 (page 21 above), Massero & Cohen (1983) and Dupoux et al. (1999) discuss the concept of phonotactic rareness, which suggests that learners will have a bias against perceiving impossible sequences in their L1 (Murphy, Monahan & Grant 2016). Therefore, learners require L2 sequences to conform to the phonotactics of their L1 to ensure effective perception. I argue that this in itself cannot offer a satisfactory explanation for the results seen in

Chapter 4. As I have discussed above, Japanese and Korean prohibit homorganic glide-vowel sequences, and Japanese has very few glide-vowel sequences in its phonotactic inventory. This means that for the phenomenon of phonotactic rareness to be a factor, the Japanese participants should also have difficulty perceiving the other glide-vowel sequences not present in their L1, depending on their language proficiency. The empirical evidence in the current study shows close to ceiling performance in all non-homorganic glide-vowel sequences, suggesting that there is no bias against these sequences, which are also not present in the L1.

4. Frequency Effects

There is a long-standing debate as to whether frequency in itself can explain patterns of phonological language acquisition. In order to assess this hypothesis I have considered the frequency of the homorganic glide-vowel target stimuli. From the frequency scores shown in Table 11 (page 74, above), no correlations could be used to explain the patterns seen in the data. A good illustration of the inadequacy of frequency as an explanation is in the words y*ield* and *yielding*. *Yield* shows higher rates of omission than *yielding*, even though *yield* is much more frequent within the SUBTLEXus database (n=206) than *yielding* (n=22). Furthermore, corpus data also demonstrates that the word *yield* is present in almost a third of all engineering and science based textbooks used in 3rd year university courses (Nesi et al. 2007). This means that a majority of the participants in this study could have been exposed to *yield* (based on their occupations), far more frequently than *yielding*, yet their accuracy for this item is considerably lower.

Frequency can also be ruled out as a cause of the lexical effects already discussed for the words *wood* and *would*. As I stated in section 5 of Chapter 4, there are much higher rates of omission in the modal form, which is much more frequent than the noun form. This again provides evidence that frequency in itself cannot explain the variation in accuracy among the homorganic glide-vowel target stimuli studied in this pilot.

5. Markedness

As I have discussed in previous sections, homorganic glide-vowel sequences are considered to be marked cross-linguistically (Ohala & Kawasaki 1984; Rose 1999; Lee 1994), with many languages including Japanese and Korean prohibiting them. In line with the predictions of the Markedness Differential Hypothesis (Eckman 1977) (see Chapter 2 section 1.1), these target homorganic glide-vowel sequences are considered to be very marked structures, and differ to the sequences in the L1, thus posing difficulties in acquisition. Following work by Greenberg (1966) and Kang (2014), for Japanese and Korean the following hierarchy could be expected: CV>>V (resulting in CV structures being less marked) and for these CV structures a further breakdown to $GV_{[Non\text{-}Homorganic]}$>>$GV_{[Homorganic]}$ could be made. This hierarchy would therefore predict that substitution for a non-homorganic glide-vowel sequence is more likely than omission. This is interesting as the data clearly demonstrate that omission is a more common strategy than vowel substitution. Moreover, Wang (1995) states that learners will produce unmarked structures when their L1 and L2 compete, meaning that in this case vowel substitution is again more likely than omission.

Although not the most common strategy used by participants, both the Japanese and the Korean speakers of English did display behaviour of vowel substitution in the target homorganic glide-vowel sequences, are shown in Table 13 (page 81, above). I argue that this is a result of transfer effects combined with the cross-linguistic markedness of these homorganic glide-vowel sequences. As both Japanese and Korean prohibit the target homorganic glide-vowel sequences investigated in this pilot, substituting the vowel is an expected strategy to conform with the L1's phonotactics. Vowel epenthesis is also an expected strategy, but this is unattested in this research. These data show consistency with an existing study (Kang 2014), in that this strategy of vowel substitution to create less marked forms is employed as well as glide omission.

6. Individual Patterns

The individual data presented in the Appendix show a degree of asymmetry in the homorganic glide-vowel target results. This asymmetry shows a higher perception accuracy score for one target, and a higher production score for the other target. This asymmetry does not affect the conclusion that perception drives production, but it does add another dimension to the results, and provide evidence of the complicated relationship between perception and production. Several of the Korean participants (4, 8 and 10) show higher accuracy in perception for [j] targets, but higher production scores for [w] targets than [j] targets. Korean, as already discussed, has the same number of [w] and [j] vowel sequences in the L1, therefore this asymmetry cannot be attributed to transfer effects. Participant 1, who is Japanese, shows higher accuracy in her perception of [w] targets, but higher production accuracy in the [j] targets. In this instance, it could be argued that because Japanese permits three [j]+vowel sequences and only one [w]+vowel sequence this mismatch could be a result of transfer. However, the fact that Participant 5, who is Korean, shows this same asymmetry with higher perception of [w] and higher production of [j] means that a firm conclusion cannot be drawn from these data. Therefore, I suggest that this asymmetry should be investigated further in future research.

The data show some evidence that production accuracy is higher in the longer homorganic glide-vowel sequence target forms than in the shorter forms. This production trend is not visible across the board, but can be seen in forms such as *yield* and *yielding*. During the reading task, some participants took a noticeable pause before uttering some of the longer stimuli. This could indicate that these words are not commonly used, or were unknown to the participants. The additional time taken to process these words before producing them could account for the higher accuracy seen in these results. Again here, further research is in order.

7. Educational Insight

As has been discussed by Morely (1991) "intelligible communication is an essential component of communicative competence" (p. 488). Therefore, awareness of the difficulties students have in pronunciation is paramount in order to assist them to maximize their communication skills.

There is extensive research into the segmental and suprasegmental features that build our phonological system of spoken languages. Our conscious and unconscious knowledge of these features is developed within our native language systems. The way we then apply this knowledge to our L2 can lead to inaccuracies in production. Therefore, when learning a second language, we must modify our segmental and suprasegmental knowledge to meet the needs of the new language. This need for modification means that we must be able to discriminate what differences there are. The need to balance discrimination and production tasks in L2 pronunciation is now widely acknowledged, with a majority of English language learning textbooks taking this approach (Ricketts 2014). However, these texts only provide limited notes to the teacher. This lack of information, coupled with the fact that few ESL teaching courses provide any modules on how to teach pronunciation (Foote 2011 in Ricketts 2014), means that some teachers are not fully equipped to effectively instruct their students. Research also demonstrates that a majority of teachers prefer to teach segments, either because they are deemed as easier to teach, or because there is a lack of support and instruction on how to teach anything else (Burns 2006; Macdonald 2002; Burgess & Spencer 2000). This trend towards teaching only isolated segments means that some teachers are disadvantaging their students. As educators, we must all strive to deliver the most effective lessons we can, even if it pushes us into uncharted waters, or requires teachers to undergo professional development.

In section 1.3 (page 10, above), I discuss that there is limited efficacy in the teaching of pronunciation only at the segmental level without any suprasegmental awareness. Providing detailed descriptions of how we articulate sounds often presents a mammoth challenge for learners, who now must grapple with complex terminology and concepts in a second language with very limited context to assist them. When presented with individual sounds, one may confuse or neglect suprasegmental effects. For example, how can we teach the schwa sound of English as an isolated unit, without it then becoming stressed? Although some teachers do this with excellent levels of success, these issues can provide a challenge for both learner and teacher. While I am not discounting the value of learning to accurately articulate sounds, I am emphasizing that pronunciation teaching should not be limited to drilling individual phones in isolation. Rather attention should be paid to sequences, suprasegmental features, perception, articulation as well as isolated segments.

Furthermore, this research reinforces the idea that we need to combine discrimination and production activities in the pronunciation classroom. The current study shows that if our perception of the L2's phonological system is inaccurate, our resultant productions will also be inaccurate. From the findings in this research, targeted materials to tackle the problem of glide acquisition in Japanese and Korean learners of English can be created.

8. Suggestions for Future Research

In the course of this research, I have made a number of errors, and these oversights were purely my own. I outline these issues in the following sections to ensure transparency, and with the hope that the same errors will not be made in any follow-up research.

8.1. Duration of ABX Sound Files

Firstly, when creating the stimuli for the ABX discrimination task I did not record the sound samples with a specific time interval between words. When the stimuli were recorded, the speaker produced the words at a pre-established rhythm, however this was not a timed interval. Further, when the stimuli were prepared, I did not ensure that they were all samples of the same duration. This lack of attention to duration resulted in some sound files being shorter than others, particularly the shorter stimuli (two non-target stimuli making a total of 6 triads). When these files were uploaded to OpenSesame, version 3.0.7 (http://osdoc.cogsci.nl/) to create the ABX task an interval of 350Ms was established, but due to the slight inconsistency in the initial sound files some trials were completed more quickly than others. Although there are no visible signs (the results from these stimuli are consistent with the others in their category), this inconsistency in sound files could have had a negative effect on the results presented in this pilot.

8.2. Stimuli Distribution

For this pilot, I examined 28 glide-vowel target stimuli, of which 18 (those with homorganic glide-vowel sequences) became my central focus, as the remainder behaved at ceiling level, similar to the non-target forms. Table 11 (page 74, above) lists of all of the homorganic glide-vowel target context results across the three experiments, where we can observe a lack of overlap across the stimuli forms. The reading task used a full set of stimuli, and the picture naming and ABX tasks used a subset of those stimuli, due to methodological issues preventing the use of the full set, such as the unpicturable nature of some stimuli and the need of three repetitions of the stimuli in the ABX task. When selecting the subsets of stimuli for the picture and ABX tasks, I failed to consider the need for overlap; I should have used identical subsets for both tasks. My failure to do this has resulted in an inability fully compare my results across the three tasks.

8.3. Number of Participants

The number of participants recruited for this study was sadly very low. This low number relates to the small number of Japanese and Korean individuals living in St. John's. In order to provide enough data to verify the findings of this study a larger sample should be recruited. This sample could also be followed longitudinally in order to capture any changes in behaviour over time in order to establish if exposure to the target language in an immersion context increases accuracy.

9. Conclusion

Throughout this thesis, I have presented evidence for a strong relationship between perception and production in the acquisition of glide-vowel sequences. In this section, I offer a brief summary of this study as well as some concluding remarks.

The phenomenon of glide omission in homorganic glide-vowel sequences for Japanese and Korean learners of English is a complicated one. Through this study, I have shown that although the root cause of this phenomenon lies within perception, there is an influence from transfer effects, which are phonotactic in nature. I also showed that existing theories in the field of L2 acquisition discussed in this thesis cannot fully capture the empirical evidence presented in this study, as they focus primarily on individual phones as opposed to phone combinations.

In Chapter 4, I presented empirical evidence that the relationship between perception and production, for the most part, go hand-in-hand. In cases where perceptual accuracy is high, productions are also high, and vice-versa. Despite this relationship, there is a degree of asymmetry in the results, in cases where perceptual abilities yield higher accuracy for one target, whereas the productive accuracy is higher for the other. This asymmetry cannot be explained by this research, and therefore warrants further investigation.

In Chapter 4, I also discussed evidence of vowel substitution to create non-homorganic glide-vowel sequences, which are considered to be less marked and more favoured cross-linguistically. To my knowledge, this finding has been attested for Korean speakers of English in another study (Kang 2014), but has never been observed in experimental results for Japanese speakers of English.

Due to the small population size and the similarity in the results, for the most part this pilot study focuses on individual and group results, rather than language specific results. These data show that the Japanese participants have more difficulty with both perception and production of the homorganic glide-vowel sequences than the Koreans. However, this is likely a result of the fewer number of contexts of glide-vowel sequences permitted in Japanese, giving evidence for the phonotactic transfer effects. For the Korean participants, there is a clear difference in accuracy between the target homorganic glide-vowel contexts. The Koreans have much lower productive abilities in the [j] target sequences than the [w] target sequences, despite similar perceptual abilities. This cannot be explained by phonotactic transfer effects or markedness, therefore calling for further research.

In conclusion, this research provides evidence that difficulties in perception are the root cause of the glide omission patterns observed in Japanese and Korean learners of English. Phonotactic transfer effects also play a part, leading to vowel substitution and variation in accuracy between the two homorganic glide-vowel contexts. Finally, in terms of education, it is apparent from this pilot that discrimination activities, followed by contextual productive tasks, are required to assist learners with the difficulties they face with these sequences.

References

Ahn, Sang-Cheol & Gregory Iverson. 2007. Structural Imbalances in the Emergence of the Korean Vowel System. In Joseph Salmons & Shannon Dubenion-Smith (eds.), *Historical Linguistics*, vol. 284, 275–293. (Current Issues in Linguistic Theory). Amsterdam: John Benjamins Publishing Company.

Aida, Yukie. 1994. Examination of Horwitz, Horwitz, and Cope's Construct of Foreign Language Anxiety: The Case of Students of Japanese. *The Modern Language Journal* 78(2). 155–168.

Aoyama, Katsura. 2003. Perception of Syllable-Initial and Syllable-Final Nasals in English by Korean and Japanese Speakers. *Second Language Research* 19(3). 251–265. doi:10.1191/0267658303sr222oa.

Aoyama, Katsura, James E. Flege, Susan G. Guion, Reiko Akahane-Yamada & Tsuneo Yamada. 2004. Perceived Phonetic Dissimilarity and L2 Speech Learning: The Case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics* 32(2). 233–250. doi:10.1016/S0095-4470(03)00036-6.

Bada, Erdogan. 2001. Native Language Influence on the Production of English Sounds by Japanese Learners. *The Reading Matrix* 1(2). 1–15.

Baker, Wendy & Pavel Trofimovich. 2006. Perceptual Paths to Accurate Production of L2 Vowels: The Role of Individual Differences. *IRAL – International Review of Applied Linguistics in Language Teaching* 44(3). 231–250. doi:10.1515/IRAL.2006.010.

Baker, Wendy, Pavel Trofimovich, James E. Flege, Molly Mack & Randall Halter. 2009. Child-Adult Differences in Second-Language Phonological Learning: The Role of Cross-Language Similarity. *Language and Speech* 51(4). 317–342. doi:10.1177/0023830908099068.

Best, Catherine T. 1991. Emergence of Language-Specific Constraints in Perception of Non-Native Speech: A Window on Early Phonological Development. In Bénédicte de Boysson-Bardies, Scania de Schonen, Peter Jusczyk, Peter McNeilage & John Morton (eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, 289–304. (NATO ASI Series 69). Amsterdam: Springer Netherlands.

Best, Catherine T., Gerald W. McRoberts & Elizabeth Goodell. 2001. Discrimination of Non-Native Consonant Contrasts Varying in Perceptual Assimilation to the Listener's Native Phonological System. *The Journal of the Acoustical Society of America* 109(2). 775–794.

Best, Catherine T. & Winifred Strange. 1992. Effects of Phonological and Phonetic Factors on Cross-Language Perception of Approximants. *Journal of Phonetics* 20(3). 305–330.

Best, Catherine T. & Michael D. Tyler. 2007. Nonnative and Second-Language Speech Perception: Commonalities and Complementarities. In Ocke-Schwen Bohn & Murray J. Munro (eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, 13–34. Amsterdam: John Benjamins Publishing Company.

Bohn, Ocke-Schwen & Murray J. Munro. 2007. *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*. Amsterdam: John Benjamins Publishing.

Braidi, Susan M. 1999. *The Acquisition of Second Language Syntax*. New York: Oxford University Press.

Brannen, Kathleen Joan. 2012. The Perception and Production of Interdental Fricatives in Second Language Acquisition. McGill Ph.D. Dissertation.

Brysbaert, Marc & Boris New. 2009. A New and Improved Word Frequency Norm for American English: Subtlexus. *SUBTLEXus*. http://subtlexus.lexique.org/moteur2/index.php.

Bundgaard-Nielsen, Rikke L., Catherine T. Best & Michael D. Tyler. 2011. Vocabulary Size Is Associated with Second-Language Vowel Perception Performance in Adult Learners. *Studies in Second Language Acquisition* 33(3). 433–461. doi:10.1017/S0272263111000040.

Burgess, John & Sheila Spencer. 2000. Phonology and Pronunciation in Integrated Language Teaching and Teacher Education. *System* 28(2). 191–215. doi:10.1016/S0346-251X(00)00007-5.

Burns, Anne. 2006. Integrating Research and Professional Development on Pronunciation Teaching in a National Adult ESL Program. *TESL Reporter* 39(2). 34–41.

Cho, Mi-Hui & Soonyong Jeoung. 2013. Perception and Production of English Vowels by Korean Learners: A Case Study. *Studies in Phonetics, Phonology and Morphology* 19(1). 155–177.

Colantoni, Laura, Jeffrey Steele & Paola Escudero. 2015. *Second Language Speech*. Cambridge: Cambridge University Press.

Derwing, Tracey M. & Murray J. Munro. 2015. *Pronunciation Fundamentals: Evidence-Based Perspectives for L2 Teaching and Research*. Amsterdam & Philadelphia: John Benjamins Publishing.

Derwing, Tracey M., Murray J. Munro & Grace Wiebe. 1998. Evidence in Favour of a Broad Framework for Pronunciation Instruction. *Language Learning* 48. 393–410.

Derwing, Tracey M. & Rossiter, Marian. 2003. The Effects of Pronunciation Instruction on the Accuracy, Fluency and Complexity of L2 Accented Speech. *Applied Language Learning* 13. 1–18.

Díaz-Campos, Manuel. 2004. Context of Learning in the Acquisition of Spanish Second Language Phonology. *Studies in Second Language Acquisition* 26(2). 249–273.

Do-Seoup Jeong. 2006. Teaching Pronunciation of /r/-initial Words in Perception and Production to Korean Adult Learners. University of Kansas M.A. Thesis.

Dupoux, Emmanuel, Kazuhiko Kakehi & Yuki Hirose. 1999. Epenthetic Vowels in Japanese: A Perceptual Illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25(6). 1568–1578.

Eckman, Fred R. 1977. Markedness and the Contrastive Analysis Hypothesis. *Language Learning* 27(2). 315–330.

Eckman, Fred R. 2007. Hypotheses and Methods in Second Language Acquisition: Testing the Noun Phrase Accessibility Hierarchy on Relative Clauses. *Studies in Second Language Acquisition* 29(2). 321–327. doi:10.1017/S0272263107070179.

Edwards, Jette G. Hansen & Mary L. Zampini. 2008. *Phonology and Second Language Acquisition*. . Vol. 36. (Studies in Bilingualism). Amsterdam: John Benjamins Publishing.

Escudero, Paola & Paul Boersma. 2002. The Subset Problem in L2 Perceptual Development: Multiple-Category Assimilation by Dutch Learners of Spanish. In Barbora Skarabela, Sarah Fish & Anna Do (eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development*, 208–219. Somerville, MA: Cascadilla Press.

Flege, James E. 1981. The Phonological Basis of Foreign Accent: A Hypothesis. *TESOL Quarterly* 15(4). 443–455. doi:10.2307/3586485.

Flege, James E., Murray J. Munro & Ian R. A. MacKay. 1995. Factors Affecting Strength of Perceived Foreign Accent in a Second Language. *The Journal of the Acoustical Society of America* 97(5). 3125–3134. doi:10.1121/1.413041.

Flege, James E. 1995. Second Language Speech Learning: Theory, Findings, and Problems. In Strange Winifred (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 233–277. Timonium, MD: York Press.

Flege, James E. 2003. Methods for Assessing the Perception of Vowels in a Second Language. In Fava, Elisabetta & Mioni, Alberto (eds.), *Issues in Clinical Linguistics*, 19–44. Padova, Italy: Unipress.

Fox, Robert Allen, Ewa Jacewicz, Fred R. Eckman, Gregory K. Iverson & Soyoung Lee. 2009. Perception versus Production in Korean L2 Acquisition of English Sibilant Fricatives. *Current Issues in Unity and Diversity of Languages*. 2661–2680.

Gilakjani, Abbas P. 2012. A Study of Factors Affecting EFL Learners' English Pronunciation Learning and the Strategies for Instruction. *International Journal of Humanities and Social Science* 2(3). 119–128.

Greenberg, Joseph. 1966. *Language Universals*. The Hague: Moulton.

Guion, Susan G., James E. Flege, Reiko Akahane-Yamada & Jesica C. Pruitt. 2000. An Investigation of Current Models of Second Language Speech Perception: The Case of Japanese Adults' Perception of English Consonants. *The Journal of the Acoustical Society of America* 107(5). 2711–2724. doi:10.1121/1.428657.

Hong, Yunsook. 1988. A Sociolinguistic Study of Seoul Korean. University of Pennsylvania Ph.D. Dissertation.

Ingram, John CL & See-Gyoon Park. 1997. Cross-Language Vowel Perception and Production by Japanese and Korean Learners of English. *Journal of Phonetics* 25(3). 343–370.

Jang, Tae-Yeoub & Yun Jeong Cho. 2005. A Comparative Experimental Study on English On-Glides Produced by Korean Learners and Native Speakers: Focusing on [yi], [wu] and [w]. *The Journal of Linguistics Science* 34(1). 309–328.

Johnson, Jacqueline S & Elissa L Newport. 1989. Critical Period Effects in Second Language Learning: The Influence of Maturational State on the Acquisition of English as a Second Language. *Cognitive Psychology* 21(1). 60–99. doi:10.1016/0010-0285(89)90003-0.

Juhana. 2012. Psychological Factors that Hinder Students from Speaking in English Class: A Case Study in a Senior High School in South Tangerang, Banten, Indonesia. *Journal of Education and Practice* 3(12). 100–110.

Kang, Sang Kyun. 2014. The Acquisition of English Glides by Native Speakers of Korean. University of Iowa Ph.D. Dissertation.

Kang, Sang-Kyun. 2012. Differences in the Duration of Constriction between Korean and English Glides. *Language Research* 48. 487–507.

Kim, Chin-Wu. 1968. The Vowel System of Korean. *Language* 44(3). 516–527.

Labrune, Laurence. 2012. *The Phonology of Japanese*. Oxford: Oxford University Press.

Lado, Robert. 1957. *Linguistics Across Cultures; Applied Linguistics for Language Teachers.* Ann Arbor, MI: University of Michigan Press.

Lee, Byung-gun. 1973. Underlying Segment in Korean Phonology. Indiana University Ph.D. Dissertation.

Lee, Duck-Young. 1998. *Korean Phonology: A Principle-Based Approach*. Munich: Lincom Europa.

Lee, Hyun-Bok. 1971. The Vowel System of Seoul Korean. *Language Research* 7. 19–24.

Lee, Jiyeon. 2004. L1 Transfer in L2 Variation: The Case of Disappearing /w/ in the English of Korean Learners. *University of Wisconsin Linguistics Students Organization Working Papers in Linguistics* 4(1). 53–64.

Lee, Yongsung. 1994. Onset Analysis of Korean On-Glides. In Kim-Renaud & Young-Key (eds.), *Theoretical Issues in Korean Linguistics*, 134–156. Pusan: CSLI.

Lenneberg, Eric H. 1967. *Biological Foundations of Language*. New York: Wiley.

Levi, Susannah. 2004. The Representation of Underlying Glides: A Cross-Linguistic Study. University of Washington Ph.D. Dissertation.

Macdonald, Shem. 2002. Pronunciation: Views and Practices of Reluctant Teachers. *Prospect* 17(3). 3–18.

Martin, Samuel. 1976. *A Reference Grammar of Japanese*. New Haven, Connecticut: Yale University Press.

Massaro, Dominic W. 1979. Letter Information and Orthographic Context in Word Perception. *Journal of Experimental Psychology: Human Perception and Performance* 5(4). 595–609. doi:10.1037/0096-1523.5.4.595.

Massaro, Dominic W. & Michael M. Cohen. 1983. Phonological Context in Speech Perception. *Perception & Psychophysics* 34(4). 338–348. doi:10.3758/BF03203046.

Mathôt, Sebastiaan, Daniel Schreij & Jan Theeuwes. 2012. OpenSesame: An Open-Source, Graphical Experiment Builder for the Social Sciences. *Behavior Research Methods* 44(2). 314–324. doi:10.3758/s13428-011-0168-7.

McAllister Byun, Tara, Sharon Inkelas & Yvan Rose. 2016. The A-Map Model: Articulatory Reliability in Child-Specific Phonology. *Language* 92(1). 141–178.

McGuire, Grant. 2010. A Brief Primer on Experimental Designs for Speech Perception Research. *Laboratory Report* 77(1). 2–19.

Minnick Fox, Michelle & Kazuaki Maeda. 1999. Perception and Production of American English Tense and Lax Vowels by Japanese Speakers. *University of Pennsylvania Working Papers in Linguistics* 6(1). 299–315.

Morley, Joan. 1991. The Pronunciation Component in Teaching English to Speakers of other Languages. *TESOL Quarterly* 25(3). 481–520.

Murphy, Patrick, Philip Monahan & Margaret Grant. 2016. Affrication Patterns and Perceptual Tendencies in Canadian and European French. http://cla-acl.ca/wp-content/uploads/resumes-2016-abstracts/MurphyMonahanGrant.pdf.

Nesi, Hilary, Sheena Gardner, Paul Thompson, & Paul Wickens. 2007. British Academic Written English Corpus. Text. http://ota.ox.ac.uk/desc/2539.

Ohala, John & Haruko Kawasaki. 1984. Prosodic Phonology and Phonetics. *Phonology Yearbook* 1(1). 113–127.

Pinter, Gabor. 2005. Where Have the Japanese Velar Glides Gone? *Linguistics @ Oxford Postgraduate Conference*, 1–8. Oxford: Oxford University Press.

Reid, Amanda, Denis Burnham, Benjawan Kasisopa, Ronan Reilly, Virginie Attina, Nan Xu Rattanasone & Catherine T. Best. 2014. Perceptual Assimilation of Lexical Tone: The Roles of Language Experience and Visual Information. *Attention, Perception, & Psychophysics* 77(2). 571–591. doi:10.3758/s13414-014-0791-3.

Ricketts, Deborah. 2014. Analysis of Tasks and Activities in ESL Pronunciation Books. University of Alberta M.A. Thesis.

Rose, Yvan. 1999. A Structural Account of Root Node Deletion in Loanword Phonology. *Canadian Journal of Linguistics* 44(4). 359–404.

Saito, Kazuya. 2012. Effects of Instruction on L2 Pronunciation Development: A Synthesis of 15 Quasi-Experimental Intervention Studies. *TESOL Quarterly* 46(4). 842–854. doi:10.1002/tesq.67.

Saville-Troike, Muriel. 2005. *Introducing Second Language Acquisition*. Cambridge: Cambridge University Press.

Schouten, Marten & Arjan Van Hessen. 1999. Categorical Perception as a Function of Stimulus Quality. *Phonetica* 56. 56–72.

Sheldon, Amy & Winifred Strange. 1982. The Acquisition of /r/ and /l/ by Japanese Learners of English. *Applied Psycholinguistics* 3(3). 243–261.

So, Connie K. 2005. The Influence of L1 Prosodic Background on the Learning of Mandarin Tones: Patterns of Tonal Confusion by Cantonese and Japanese Naïve Listeners. *Proceedings of the 2005 Annual Conference of the Canadian Linguistic Association*, 1–18. http://westernlinguistics.ca/Publications/CLA-ACL/So.pdf.

Sohn, Hyang-Sook. 1987. Underspecification in Korean Phonology. University of Illinois Ph.D. Dissertation.

Song, Jae-Jung. 2008. *The Korean Language: Structure, Use and Context*. London: Routledge.

Sperbeck, Mieko. 2010. The Production and Perception of English Consonant Sequences by Japanese-Speaking Learners of English. *Proceedings of Meetings on Acoustics*, vol. 9, 1–12. Baltimore: Acoustical Society of America. http://scitation.aip.org/content/asa/journal/poma/9/1/10.1121/1.3651080.

Stackhouse, Joy, Maggie Vance, Michelle Pascoe & Bill Wells. 2007. *Compendium of Auditory and Speech Tasks: Children's Speech and Literacy Difficulties 4 with CD-ROM*. New York: John Wiley & Sons.

Tranter, Nicolas. 2012. *The Languages of Japan and Korea*. New York: Routledge.

Tremblay, Marie-Claude & Takeki Kamiyama. 2009. Perception of Second-Language (L2) Production by First-Language (L1) Speakers of Different Dialectal Backgrounds: The Case of Japanese-Speaking Learners' /u/ Perceived by French and Quebec Native Speakers. *The Journal of the Acoustical Society of America* 125(4). 2776–2777. doi:10.1121/1.4784766.

Tsai, Hui-Ju. 2015. Exploring College Students' Motivational Beliefs in Ability-Grouped English Classes in Taiwan. University of Durham Ph.D. Dissertation.

Tsujimura, Natsuko. 2013. *An Introduction to Japanese Linguistics*. New York: John Wiley & Sons.

Vance, Timothy J. 1987. *An Introduction to Japanese Phonology.* Albany, New York: State University of New York Press.

Wang, Chilin. 1995. The Acquisition of English Word-Final Obstruents by Chinese Speakers. New York State University Ph.D. Dissertation.

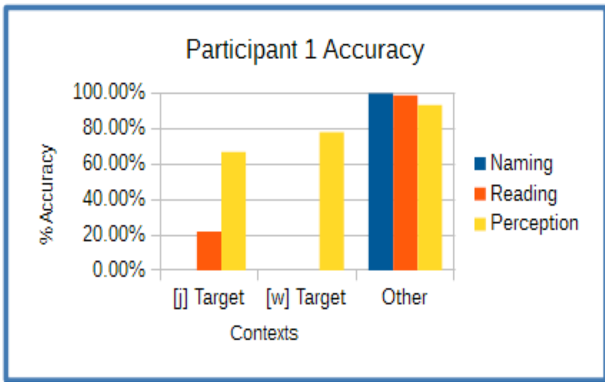
Wang, Xinchun. 1997. The Acquisition of English Vowels by Mandarin ESL Learners: A Study of Production and Perception. Simon Fraser University M.A. Thesis.

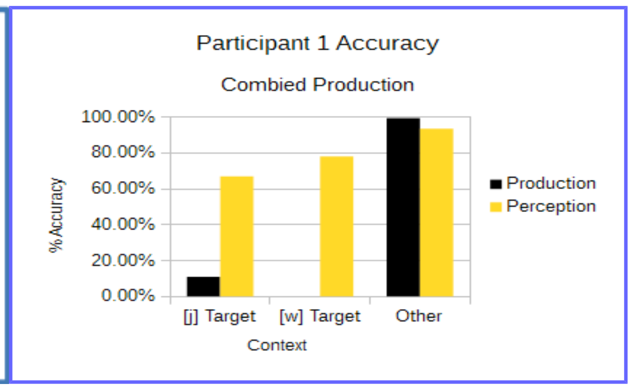
Yang Chen, Robb, Michael & Gil, Harvey. 2001. Vowel Production by Mandarin Speakers of English. *Clinical Linguistics & Phonetics* 15(6). 427–440. doi:10.1080/02699200110044804.

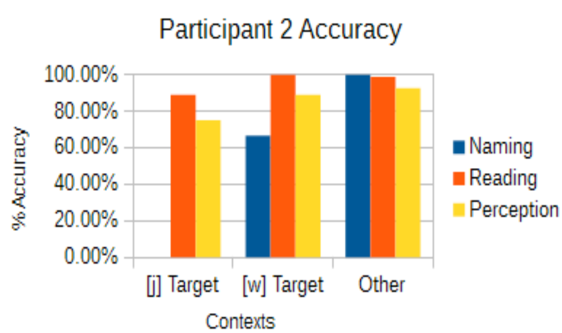
Appendix

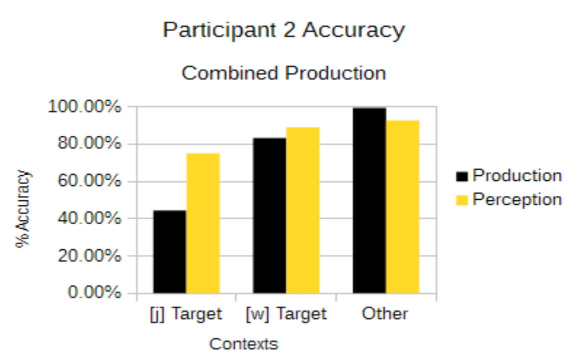
1. Individual Results for All Participants [5]

The following graphs show the individual results for all three experimental tasks. The graphs with combined production scores are designed to show a clear comparison in perception and production.

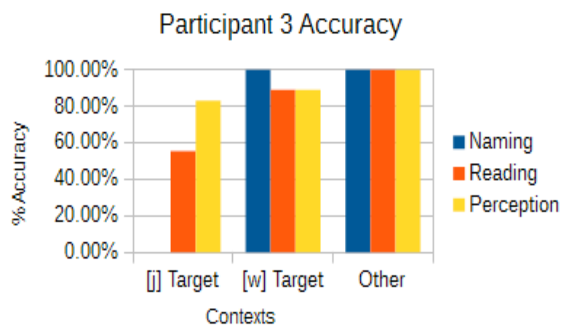




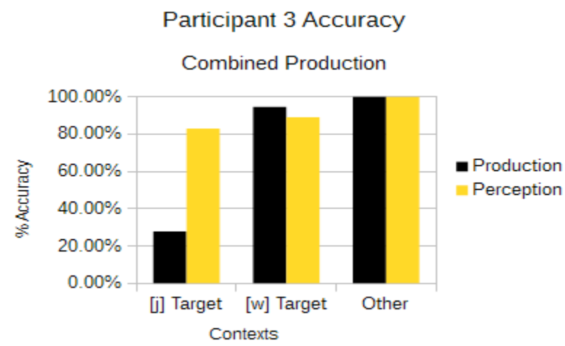




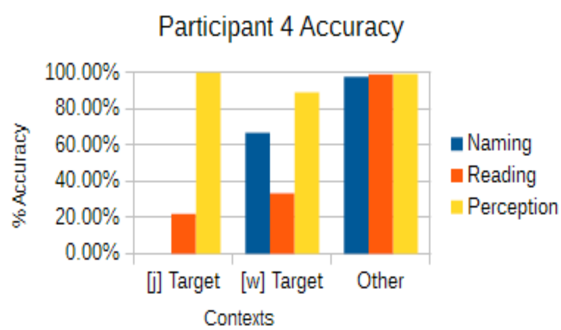
5 Charts with a border indicate the Japanese participants, all other charts are for Korean participants.



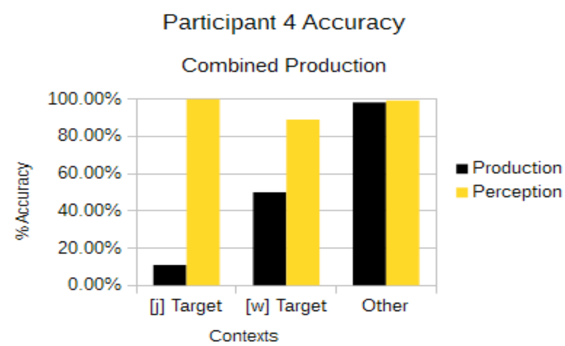
Participant 3 Accuracy
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Naming
Reading
Perception



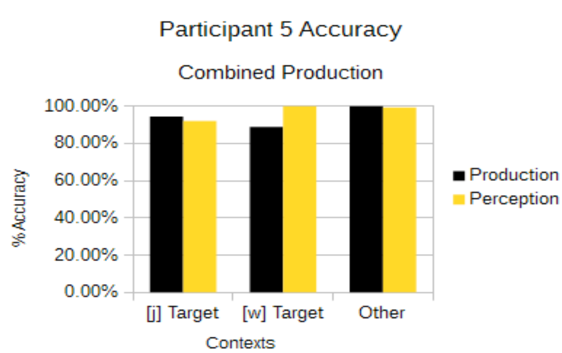
Participant 3 Accuracy
Combined Production
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Production
Perception



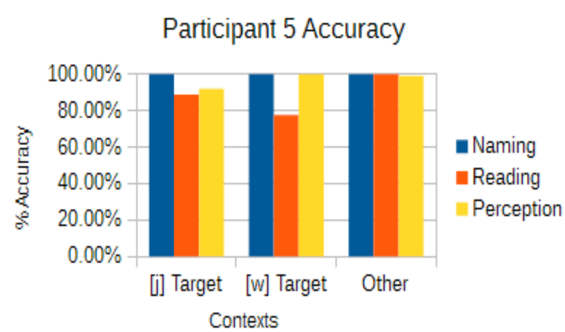
Participant 4 Accuracy
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Naming
Reading
Perception



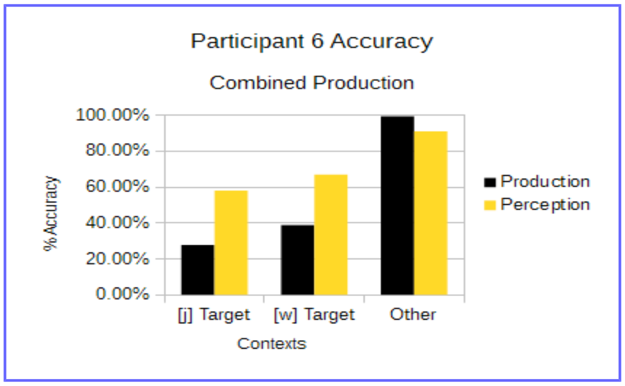
Participant 4 Accuracy
Combined Production
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Production
Perception

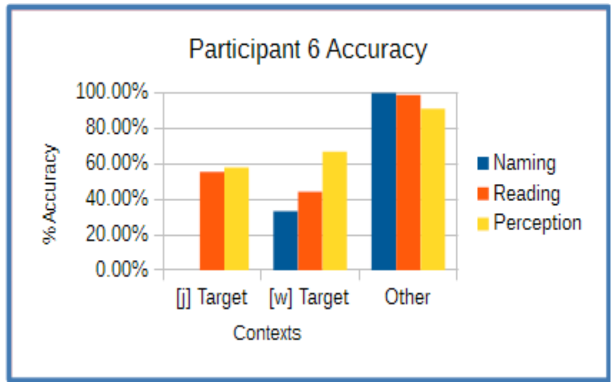


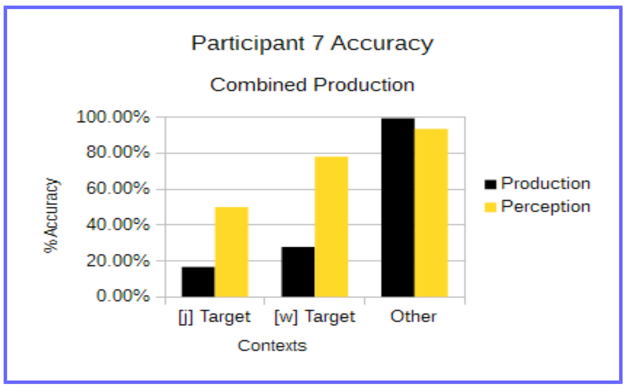
Participant 5 Accuracy
Combined Production
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Production
Perception

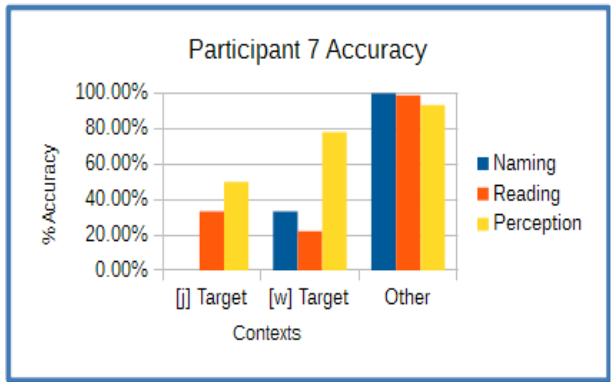


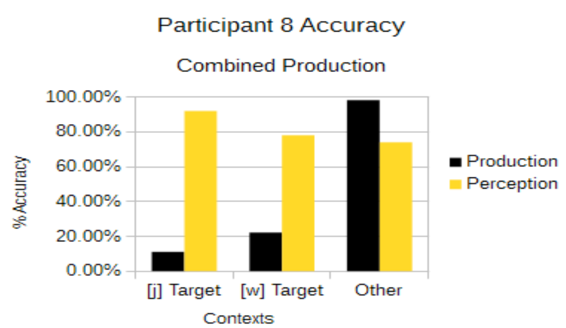
Participant 5 Accuracy
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Naming
Reading
Perception

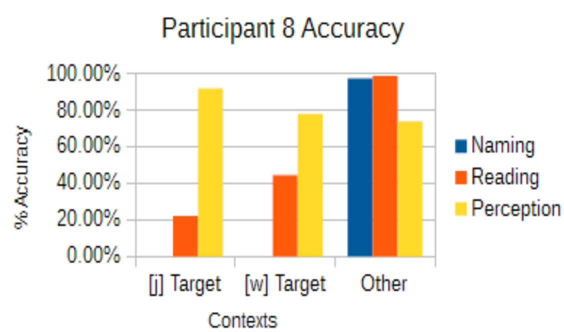

Participant 6 Accuracy
Combined Production
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Production
Perception

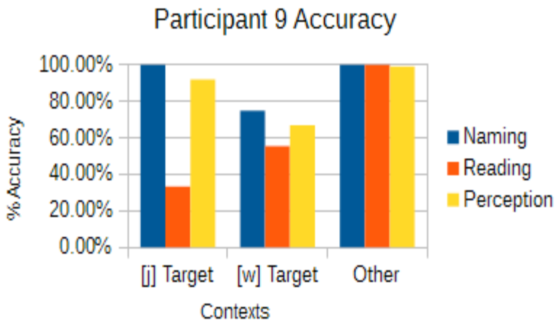

Participant 6 Accuracy
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Naming
Reading
Perception

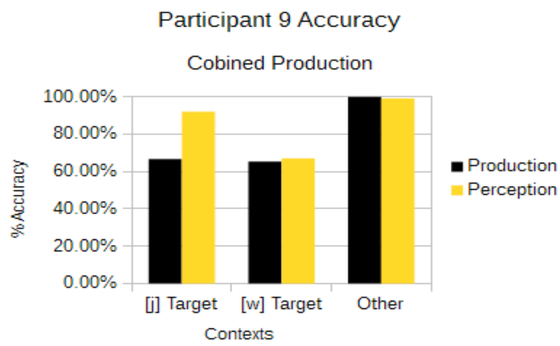

Participant 7 Accuracy
Combined Production
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Production
Perception

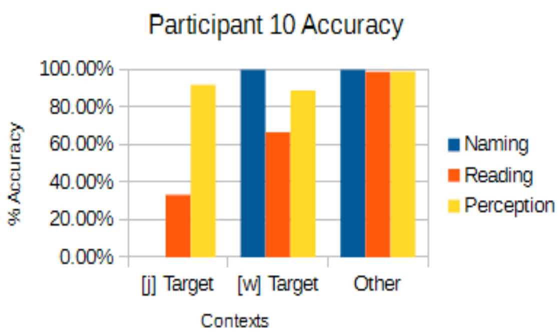

Participant 7 Accuracy
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Naming
Reading
Perception

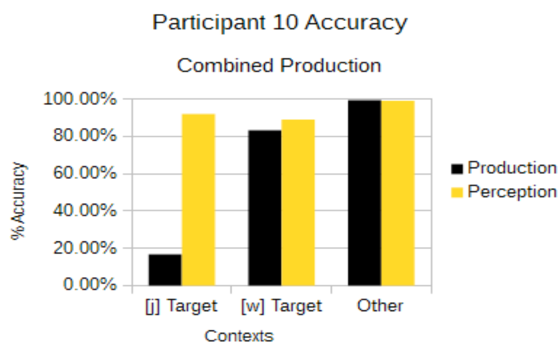

Participant 8 Accuracy
Combined Production
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Production
Perception

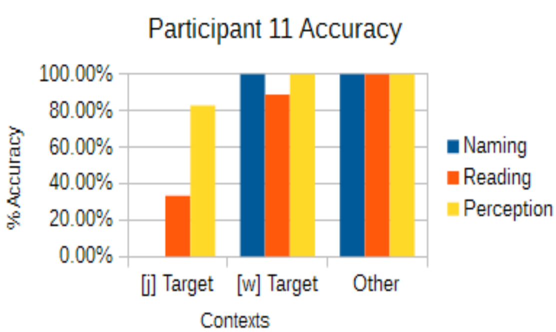

Participant 8 Accuracy
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Naming
Reading
Perception

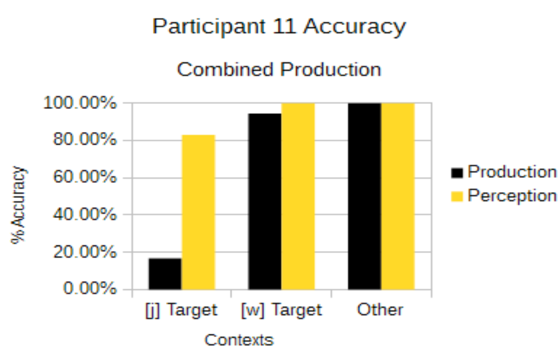

Participant 9 Accuracy
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Naming
Reading
Perception


Participant 9 Accuracy
Cobined Production
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Production
Perception


Participant 10 Accuracy
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Naming
Reading
Perception


Participant 10 Accuracy
Combined Production
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Production
Perception


Participant 11 Accuracy
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Naming
Reading
Perception


Participant 11 Accuracy
Combined Production
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
% Accuracy
[j] Target
[w] Target
Other
Contexts
Production
Perception

2. Full Stimuli List

Stimuli List				
always	hear	shiver	whoopi	year
badger	honey	should	wife	yearbook
bathe	jeans	slim	window	yearly
beach	kiss	smile	witch	yearn
beauty	Korea	snow	woke	yeast
beige	language	study	wolf	yellow
broken	log	swap	woman	yiddish
card	look	table	womb	yield
child	make	tablet	won't	yielding
chili	manage	they	wonder	yin
could	north	thing	wood	yippee
cute	orange	think	woohoo	yolk
daughter	pale	three	wool	young
dishes	paper	time	worm	your
don't	pen	treat	worthwhile	yours
dove	phone	tree	would	yoyo
English	pitcher	unique	wounded	zoo
field	pleasure	veil	woven	
fifty	question	wash	wow	
flower	quiz	water	Yale	
forehead	rich	weather	yard	
germ	roses	whale	yawn	
go	rusty	wheat	yay	
google	school	wheel	yeah	