COVERT CONTRAST IN THE SPEECH OF AN ADOLESCENT WITH APRAXIA OF SPEECH: A CASE STUDY

KIMBERLEY CHURCHILL

Covert contrast in the speech of an adolescent with Apraxia of Speech: a case study

by

© Kimberley Churchill

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

August, 2009

St. John's, Newfoundland

Abstract

This is a case study of Marshall, a 15-year-old with Apraxia of Speech (AoS). AoS is an articulatory disorder affecting ability to program speech musculature. This thesis examines Marshall's production of the approximants /r, l, w, j/. In singleton onsets Marshall's /r/ sounds like /w/, 'red' [wɛd]. For reasons outlined in the thesis, we expect Marshall to produce a contrast between /w/ and /r/ in singleton onsets, however, Marshall appears to produce both as [w]. This apparent lack of contrast prompted this investigation into the possibility of a "covert" contrast between /r/ and /w/ (a difference that is measurable but not perceivable to listeners). Marshall's approximants in singleton onsets were acoustically analysed. Statistical analysis revealed contrasts between all four approximants. Furthermore, acoustic correlates of /r/ and /w/ were statistically different indicating covert contrast. Verifying the existence of covert contrast in an individual with AoS may have implications on the diagnosis and treatment of such speech disorders.

Acknowledgements

I would like to first express a sincere thank you to my supervisor, Dr. Carrie Dyck. It is your unfailing patience and endless encouragement that has allowed me to persist in finishing my thesis. I would also like to thank Gavin Willins for graciously allowing me access to his speech recordings and his research.

I would like to acknowledge Dr. Yvan Rose for his invaluable contribution to this research. I am forever grateful for his original idea that I should look for covert contrast in Marshall's speech, thus, giving me the foundation of my thesis. I would also like to extend my gratitude to Dr. Catherine Penney of the Department of Psychology for her advice and guidance with my statistical analysis.

I would like to thank the School of Graduate Studies for awarding me the financial support of a School of Graduate Studies Fellowship (2004, 2005). I would also like to express my thanks to the Department of Linguistics for the further support of a Graduate Assistantship position (2004, 2005). I have gained immeasurable experience and knowledge working with the supportive professors in the department.

I would like to extend a special thanks to my participant and his parents. Without their kindness and cooperation this thesis would not have been possible.

A sincere thank you to my family and friends, especially Mom, Dad, Wendy, Dave, and Krista, for their endless support and encouragement.

Last but certainly not least I would like to thank Adam for his unconditional love, support, and patience. It is your unwavering belief in me that is behind all of my successes and, as with everything in my life, this thesis is dedicated to you.

Table of Contents

List of Tables

List of Figures

List of Appendices

1 Introduction

This paper is a case study of the speech of Marshall (pseudonym), a 15-year-old male with Apraxia of Speech (AoS), looking specifically at Marshall's production of the approximants /r, /l/, /w/, and /j/.

Marshall's overall approximant system is non-adult-like. He produces /w/ in an adult-like manner in both singleton and complex onset positions, i.e. 'week' [wik] and 'quick' [kwɪk]. /w/ does not occur in a coda position in English, hence, his overall /w/ productions are consistently adult-like. Marshall produces a target /j/ in singleton onsets, 'yes' [jɛs], however, he omits /j/ from complex onsets, i.e. 'cute' is produced as [kut] rather than the target [kjut]. Again, /j/ does not occur in coda position in English.

Table 1 shown below describes Marshall's production of the liquid semivowels /l/ and /r/.

Table 1: Marshall's Liquid Production Pattern

	Onset		**Coda**	
	/l/	**/r/**	**/l/**	**/r/**
Singleton	[l]	[w]	[ʊ]	[:] [ə]
Complex	[Cl] [Cw]	[Cw]	[ʊC]	[:C]

In singleton onsets, Marshall produces an adult-like /l/, i.e. 'let' [lɛt], but he varies in his pronunciation of /l/ in the second consonant of complex onsets, sometimes producing [l], i.e. 'flag' [flæg], other times producing [w] (approximately 36% of the time), i.e. 'plate' [pwet]. In coda position, however, Marshall produces an /l/ sound, although it does vary somewhat from an adult /l/ in that it tends to be a vocoid (vowel-like) variant, i.e. 'pill' [pɪᵒ], 'milk' [mɪᵒk]. In singleton and complex onset positions Marshall's /r/ sounds like an adult /w/ to listeners, i.e. 'red' [wɛd] and 'frog' [fwɑg]. In singleton and complex coda position, Marshall's /r/ is produced like a speaker with an r-less dialect with a vocoid (vowel-like) variant, i.e. 'car' [kɑː], 'hear' [hɪᵊ], 'dark' [dæːk]. (Willins, 2005)

When Marshall's overall approximant system is observed, it is clear that his production of approximant contrasts is non-adult-like. He has a systematic contrast between /r/ and /l/ in coda position, i.e. [kɑː] 'car' versus [kɑᵒ] 'call', yet no evidence of a contrast between /r/ and /w/ in singleton or complex onset position, i.e. [wɛd] 'wed' versus [wɛd] 'red'. For reasons discussed in §5, this is an unexpected pattern.

However, in this paper I show that Marshall actually has a contrast between /r/ and /w/ in singleton onset position, but that it is a "covert" contrast, meaning that the difference between /r/ and /w/ is measurable and statistically significant but it is not perceivable to the naked ear. I discuss how these findings shed light on the nature of AoS.

In §2-§4, I present the background of relevant concepts, including a definition of AoS, covert contrasts, and the acoustic correlates of the approximants /l/, /r/, /w/, and /j/. Next, in §5, I present the hypothesis of the paper in detail, that Marshall has a covert contrast between /r/ and /w/. Methodology and procedures are presented in §6. In this section I discuss the motive for the use of a case study design and describe the data, as well as how I measured the acoustic correlates of the approximants. In §7, I present the results and reveal that there is a main effect of consonant type in singleton onsets, indicating Marshall has a contrast between /r/, /l/, /w/, and /j/; also, I describe how the acoustic correlates of /r/ and /w/ are shown to be statistically different, indicating the presence of a covert contrast. Finally, in §8 I present a discussion of the implications these findings may have on the diagnosis and treatment of speech disorders, including AoS.

2 Apraxia of Speech

Marshall, the participant of this case study, has been diagnosed with AoS. I define AoS and discuss the theoretical background of the disorder in this section.

2.1 Description

AoS is an articulatory disorder resulting from the impairment of one's ability to program the positioning of speech musculature. Apraxic speakers experience difficulty in the sequencing of muscle movements for the production of phonemes (Darley1969, in Fabbro 1999:269). However, the same muscles affected by AoS, function normally in nonlinguistic tasks, such as eating, breathing, etc. This would imply that AoS is not just a simple motor disorder, but a more specific type of language impairment. There are two distinct types of AoS: acquired apraxia of speech and developmental apraxia of speech. Acquired AoS is the result of a brain injury, while developmental AoS is a developmental disorder present since childhood. The participant in this study, Marshall, has developmental AoS.

AoS is often confused with dysarthria, which is caused by impairment of muscle strength, tone, range of motion and/or coordination as the result of damage to the central nervous system. Due to this damage, dysarthria can affect phonation, resonance, articulation or prosody (Darley et al., 1975). Although AoS can also be caused by damage to the central nervous system (in the case of acquired AoS), it primarily affects articulation, not resonance or phonation (Darley et al. 1975). Also, as opposed to individuals suffering from aphasia, individuals with AoS show no receptive difficulties understanding spoken language.

The most salient characteristics of an individual with AoS are as follows: (1) difficulty in initiating spontaneous speech; (2) struggling to position the articulators correctly; in doing so, speakers are seen to visibly and audibly grope as they attempt to form a word; (3) completed articulations are often off target; frequently the speaker will be aware of this and make an effort to correct the error; and (4) errors in AoS occur in similar environments, although the errors are often quite variable from trial to trial (Darley et al. 1975:250).

2.2 Theoretical Background: Phonological vs. Motor Disorder

There are opposing theories as to the underlying nature of AoS. AoS has been largely understudied; in fact, research spans back only about 20 years (Shriberg et al., 1997). Originally, AoS was seen primarily as a motor programming deficit (Darley et al., 1975). It was defined neurologically as "An articulatory disorder resulting from impairment of the capacity to program the positioning of speech musculature and the sequencing of muscle movements for the volitional production of phonemes" (Darley 1969, in Fabbro 1999:269). In other words, AoS was assumed to be a surface disorder primarily affecting the muscle coordination that had little or nothing to do with underlying mental processes such as the phonological system.

Subsequent theories have emerged suggesting the deficit affecting individuals with AoS may involve deeper phonological processes (Aichert and Zeigler, 2004; Jacks et al., 2006). Evidence supporting these theories indicates that AoS is sensitive to syllabic factors, such as syllable structure, and coarticulation. Aichert and Zeigler (2004) found that the rate of segmental errors of apraxic speakers was influenced by the syllable frequency; that is, frequent syllables were better preserved than less frequent syllables. They also found that error rates on consonant clusters were dependent on the position of the syllable boundary; consonants separated by a syllable boundary (i.e. CVC.CVC) are reduced less frequently than consonant clusters not separated by a syllable boundary (i.e. CCVC or CVCC). Similarly clusters in the onset position (CCVC) appear more vulnerable to reduction than clusters in the coda position (CVCC). Jacks et al. (2006) found that children with AoS showed impaired ability to construct accurate word shapes. They suggest that the frequent consonant omissions found in their data represent a consistent pattern of syllabic error. Jacks et al. (2006) suggest that AoS may be the result of a deficit in syllabic construction rather than sound-specific errors.

However, there is still mounting evidence supporting the theory that AoS is a speech motor programming deficit rather than a phonological disorder (Bahr, 2005; Maas et al., 2008; Peter and Stoel-Gammon, 2008). Bahr (2005) compared children with AoS to children with phonological disorders and found that the AoS children had more trouble coordinating complex articulatory gestures (i.e. a production involving the velum and the lips rather than the velum alone) than the phonologically disordered children. Maas et al. (2008) studied the affect of AoS on motor programming in the framework of a two-stage model of speech production: a preprogramming stage and a sequencing process (assigning an order to the relevant motor programs). Maas et al. found impairment in the preprogramming stage for AoS speakers compared to normal speakers but no such impairment for the sequencing stage. Maas et al. suggest that individuals with AoS are able to program sound sequences but that they need more preprogramming time to do so. These findings support the theory that the underlying deficit in AoS lies in the motor programming of speech.

There is merit to both the motoric disorder approach and the phonological approach; there has not been enough evidence collected to rule out one theory over the other. In this paper, I show that Marshall has an adult-like phonemic inventory that is obscured by his production of covert contrast. This evidence seems likely to support the theory that AoS is a surface (motor) disorder, however, further research is certainly required.

3 Covert Contrast

At this point, it is necessary to discuss the concept of covert contrast in more detail. The following sections present a description of covert contrast, as well as evidence supporting its existence.

3.1 Description

A 'covert' contrast (Macken and Barton, 1979) occurs when two sounds are perceived as homophones, even though there is a measurable acoustic or articulatory difference between them. Covert contrasts often arise in the normal development of speech in children, as well as in the speech of children and adults with phonological disorders. Child speech transcriptions tend to show neutralization of phonological contrasts disproportionate to that of adult speech. Sometimes a child is able to produce a contrast before listeners can perceive it because the acquisition, development and co-ordination of acoustic cues occur gradually over time, rather than instantaneously. In order for a distinction to be perceivable, the child must learn to coordinate the production and timing of numerous phonetic cues, which interact to convey the contrast (Scobbie et al., 1998:147-148).

3.2 Evidence of Covert Contrasts

Numerous studies have provided evidence for the existence of covert contrasts. These studies investigate several contrasts including voiced versus voiceless stop consonants, velar versus alveolar stop consonants, /θ/ versus /s/, and the omission of /s/ in initial /s/ clusters, to name a few. The following sections discuss studies of three such contrasts: the voiced versus voiceless stop contrast; the alveolar versus velar consonantal contrast; and coronal contrasts.

3.2.1 Voiced vs. Voiceless Stop Contrasts

Macken and Barton's (1979) longitudinal study of children's acquisition of English voicing contrast in word-initial stops was among the first studies of covert contrasts. They measured voice onset time (VOT), or the lag between the pronunciation of a consonant and the subsequent voicing of the following segment. They measured VOT for word-initial voiced and voiceless stops in English children with normal speech development. Macken and Barton (1979) identified a stage during which normal children produced acoustically measurable differences in the VOT values between target voiceless and voiced stops which were perceived by listeners as voiced homophones. Even though the VOT values for these segments were measurably different, all measures were within the average VOT values for voiced stops in normal adult speech; hence the contrast was not perceived by the listener and all segments were perceived as voiced.

3.2.2 Alveolar vs. Velar Contrasts

There have been a number of studies investigating the presence of covert contrast for the alveolar versus velar stop contrast. Covert contrasts have been found in children with phonological disorders who displayed substitution errors such as velar target /k/ being perceived as alveolar [t]. These studies showed that the children produced measurable articulatory or acoustic differences between alveolar and velar stop consonants, even though listeners did not perceive the contrast.

Forrest et al. (1990) found measurable differences between the target word-initial alveolar /t/ and word-initial velar /k/ for a child with a phonological disorder. It was assumed that the child did not have an alveolar/velar contrast because all of the productions were perceived as [t]. However, The investigators found that they could distinguish between /t/ and /k/ 87% of the time on the basis of information measured from the acoustic signal.

Gibbon (1990) studied two sisters. One sister clearly produced an alveolar/velar contrast (i.e. /g/→[g] and /d/→[d]) and the other appeared to neutralize the contrast (i.e. /g/→[g] and /d/→[g]). Using electropalatographic (EPG) data, the investigators made two main observations. First, both children had different tongue contact patterns for target alveolar /d/ than for target velar /g/ and second, neither had the same tongue contact patterns for /d/ as an adult speaker would exhibit, even though one sister appeared to produce an adult-like target /d/. Both children produced a double articulation in the alveolar and velar regions during their production of /d/. Gibbon (1990) stated that the sister who seemed to produce the contrast had a more adult-like sequence of articulation events than the sister who did not seem to produce a contrast; yet, neither sister produced these articulations accurately. The study concluded that listeners were actually influenced by the sequence of the articulatory events in the double articulation rather than by the accuracy of articulation.

Edwards et al. (1999) observed three children with phonological disorders (all of whom neutralize the alveolar/velar contrast at least sometimes) and three normally developing children (all of whom produce the target alveolar/velar contrast). Word initial /t/ and /k/ segments were acoustically analysed and the results were compared with impressionistic transcriptions of the productions. In the normally developing children, the results were similar to adults in that there was a clear distinction between the two sounds in the acoustic analysis. However, in the phonologically disordered children, some of the productions that were transcribed as being different sounds were found to be acoustically similar upon acoustic analysis. Also, some of the productions that were transcribed as the same sound, in fact, were acoustically different. This study shows that perceived distinctions between alveolar/velar contrasts are not categorical. It also questions the validity of relying on impressionistic transcriptions alone.

Gibbon et al. (1993) had a similar finding in their study. They used simultaneous EPG and acoustic data from a child with a phonological disorder. The data was presented to 20 listeners, all trained in the International Phonetic Alphabet (IPA), who were asked to judge the correctness of alveolar target sounds. There was a great amount of disagreement amongst listeners as to which sounds were correct target alveolars and which were incorrect velar substitutes. All of the target alveolar sounds had a great amount of tongue contact in the velar region but less in the alveolar region. Hence, the targets were all abnormally produced, and yet listeners judged many of them as correct. Gibbon et al. (1993) counseled caution when using perceptual judgements of disordered speech.

3.2.3 Coronal Contrasts

Scobbie et al. (1998) examined the target productions of word-initial /t/, /d/, and /st/ in a phonologically disordered child, DB. All of these coronal targets were perceived as [t] by listeners. Scobbie et al. (1998) found similar VOT duration values for all of DB's [t] productions, explaining the perceived homophony. This observation pointed to an immature phonological system that did not contain the /t/, /d/, and /st/ contrast. However, Scobbie et al. found that the VOT values were also influenced by the height of the following vowel; the duration of the VOT was found to increase as the vowel height increased. This evidence indicated that DB had acquired some of the characteristics of a mature phonological system but not enough for listeners to perceive a contrast (Scobbie et al., 1998:150).

Scobbie et al. (1998) also found that DB had acquired differences in phonation type after the pronunciation of /t/ when compared with /d/ and /st/. As in the target system, due to the large glottal opening required for the aspiration in the release of /t/, the subsequent voicing of the following vowel is achieved more slowly, giving the first few glottal cycles of the vowel a "breathy or murmured phonation" (Scobbie et al., 1998:151). Hence, Scobbie et al. (1998) found DB's phonation pattern to be similar to an adult's target system, which indicates evidence of a covert contrast in DB's speech. DB had apparently acquired some of the aspects of the mature contrast, but not all. Scobbie et al. (1998) concluded that, "phonetic maturity is reached only when the entire constellation of motorically based cues transmitting the contrast have reached adult values, each in their own time" (Scobbie et al., 1998:153).

3.2.4 Summary of Covert Contrast

A common theme in the studies discussed above is that there is evidence of inconsistency in perceptual judgements made by listeners. First, this evidence brings into question the veracity of listener judgements for non-adult speech. In these cases the inconsistency seems to occur primarily when the speaker is in the process of acquiring a contrast. Second, learning a contrast involves being able to produce several different components, and these components can often develop independently of each other.

The preceding sections have shown that there is evidence for the existence of covert contrasts in children's productions, in both disordered and typical speech, and that these contrasts have been observed in the production of various types of phonemes. Such contrasts are not perceived by listeners; however, they can be detected through acoustic or articulatory analysis. The presence of covert contrast suggests a more mature phonological system than could be expected through impressionistic observation alone.

This paper will provide evidence that Marshall has acquired some, but not all, of the differences between /r/ and /w/, and that his phonological patterning has influenced how his contrasts are perceived (or in this case not perceived) by listeners.

4 Approximants

This section describes the class of approximants, a prerequisite to discussing the hypothesis and methodology of this paper. This section summarizes the defining features of the approximants, outlines their corresponding acoustic correlates, and identifies the measurable acoustic parameters for each correlate.

Approximants can be broken down into two subclasses: glides, consisting of /j/ and /w/, and liquids, comprised of /l/ and /r/. Both types of approximants have a sonorant quality; "both liquids and glides have a well-defined formant structure associated with a degree of vocal tract constriction that is less severe than that for the obstruents (stops, fricatives, and affricates)" (Kent and Read, 1992:138).

4.1 Glides

Glides are produced with gradual articulatory motions that occur when the vocal tract is narrowed but there is no closure (Kent and Read, 1992). As mentioned above, the consonants /w/ and /j/ make up the inventory glides of English. Glides show properties of both consonants and vowels. They are vowel-like in their articulation but they function more like consonants in that they are never the nucleus of a syllable and are necessarily adjacent to vowels (Ladefoged and Maddieson, 1996). They are articulated by moving relatively slowly, when compared to other consonants, from a vocal tract with a degree of narrowing to a vocal tract that is appropriate for the pronunciation of the following vowel (Kent and Read, 1992).

In the production of /w/ there are two narrowings in the vocal tract, one at the lips (lip rounding) and another between the tongue dorsum and the velum. Because of the location of the narrowings, /w/ is considered a labio-velar glide. The production of /j/ involves a narrowing around the palatal region of the vocal tract, with the tongue body in a high front position; thus, /j/ is referred to as a palatal glide. Glides are produced with constant motion of the articulators; therefore, their formants (described later in §4.3) transition very smoothly into adjacent vowel formants (Espy-Wilson, 1992).

/w/ is very similar in articulation to the English high back vowel /u/, while /j/ is very similar to the English high front vowel /i/ (Ladefoged and Maddieson, 1996). Espy-Wilson (1992) point out that the vocal tract positioning for /w/ and /j/ is quite similar to /i/ and /u/, however, the constriction in the vocal tract of glides is more extreme than in the vowels.

4.2 Liquids

The class of liquids is further divided into two types: *lateral*, to which /l/ belongs, and *rhotic*, to which /r/ belongs. The two share certain phonetic and phonological similarities; for example, they are among the most sonorous of all the English oral consonants; they both occur in complex onsets and codas; and both can be syllabic. Liquids are also similar to stops, in that their articulatory movements can be quite rapid. This rapid change occurs faster in /l/ than /r/, giving /l/ a briefer duration than /r/. Although liquids are associated with rapid change, they are also potentially sustainable in their articulation and can be syllabic, word-finally and after a consonant in English. Acoustic information about liquids can be obtained from their steady state production, as well as from their transitions in connected speech (Kent and Read, 1992).

4.2.1 Laterals

In the production of lateral /l/, the tongue tip makes an occlusion at or near the alveolar ridge along the mid-line of the vocal tract. This closure allows air to flow laterally around one or both sides of the tongue.

There are at least two allophones of English /l/, light and dark. Light /l/ is typically found at the beginning of a syllable, while dark /l/ is typically in syllable final or syllabic position. According to Sproat and Fujimura (1993), the articulatory differences between the two allophones is that the tongue dorsum is less retracted for light /l/ than for dark, consequently making light /l/ more acoustically *front*; as well, the middle of the tongue is relatively higher for light /l/ which makes it more acoustically *high* than dark /l/.

4.2.2 Rhotics

Unlike laterals, rhotics are best described acoustically, rather than being defined by manner or place of articulation. American/Canadian English has different articulatory types of /r/, the paragraph below discusses dialectal or idiolectal variations in how r/ is produced. As with the other three approximants, however, there is no full closure of the articulators, only a narrowing.

One articulatory variation of /r/ has a narrowing at the alveolar or post-alveolar region. Another articulation occurs behind the alveolar ridge, for retroflex /r/. The narrowing for both the alveolar and postalveolar /r/ is generally produced by raising the tongue tip or blade. Another, more articulatorily complex, /r/ is the 'bunched-r'. This rhotic is produced with constrictions in the lower pharynx and at centre of the palate but has no raising of the tongue tip or blade. A final variant of /r/ is a syllabic /r/. This /r/ is produced much like the bunched-r. Regardless of whether /r/ is retroflex or bunched, lip rounding may occur when /r/ is prevocalic or intervocalic and before a stressed vowel. (Ladefoged and Maddieson, 1996)

Delattre and Freeman (1968) point out that American /r/ generally has the same type of acoustic signals regardless of the articulatory position used. In other words, bunched /r/, retroflex /r/ or any articulation in between produces the same acoustic signal.

4.3 Acoustic Correlates of the Approximants

Given that /r/ can be produced in a variety of fashions with the same acoustic results, acoustic cues are more reliable than articulatory ones. Consequently, I chose to measure the acoustic correlates of Marshall's approximant system. This section describes the defining features of the approximants as well as their corresponding acoustic correlates. Espy-Wilson (1992) proposed that the defining features distinguishing the approximant consonants from each other are *high*, *back*, *front*, and *retroflex*. Espy-Wilson uses the term *retroflex* to refer to all rhotics, whether the articulation is retroflex, bunched, etc. I will replace the term *retroflex* with the term *rhotic*, which is more articulatorily neutral. Before defining the features that distinguish between the approximants and the corresponding acoustic correlates I will first explain the acoustic terms used throughout this paper.

The human vocal tract is a natural resonator, set into motion by vibrations from the vocal folds. As the vocal tract changes shape for different articulations, it naturally resonates at varying frequencies. In the resulting sound energy, shaped by these resonances of the vocal tract, some frequencies are muted while others are enhanced. These enhanced frequencies, or peaks of resonance, are called *formant frequencies*. The formants are named in ascending order and this paper deals with the first formant (F1), the second formant (F2), and the third formant (F3). Another frequency measurement dealt with in this paper is the *fundamental frequency* (F0). The F0 is the lowest frequency in a human voice and it largely influences the listener's perception of the speaker's pitch. When a speaker changes the intonation in his/her voice, it is his/her F0 that is changing.

F0 is part of the reason why men's voices sound lower than women's voices and what differentiates one individual's voice from another. Both formant frequencies and F0's can be seen and measured in the acoustic signal using a wide-band spectrogram. (Borden et al., 2003)

Table 2 below shows the defining features of the four approximants.

Table 2: Features discriminating among the approximants. Data modified from Espy-Wilson (1992).

	High	Back	Front	Rhotic
/w/	+	+	-	-
/j/	+	-	+	-
/r/	-	-	-	+
/l/ (prevocalic)[1]	-	-	-	-

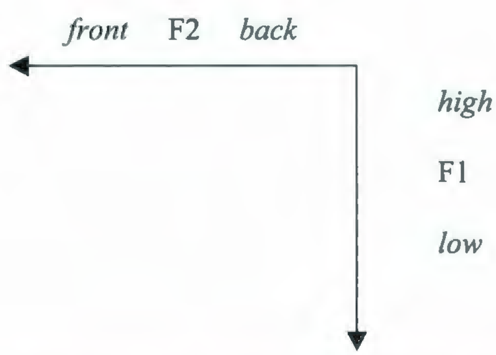

Figure 1: *High* versus *Low* and *Front* versus *Back* in relation to F1 and F2

As sown in Figure 1 above, the feature *height* is negatively correlated with F1 frequency. Hence, *high* sounds have a low F1 frequency, while *non-high* sounds have a

[1] In her study Espy-Wilson (1992) necessarily distinguishes between prevocalic /l/ and postvocalic /l/ as they are both a part of her data set. The data used in the present study contains only word-initial prevocalic /l/, thus, I will exclude postvocalic /l/ from this discussion. Any reference to /l/ can therefore be assumed to be a prevocalic /l/ unless otherwise specified.

high F1 frequency. *Frontness*[2] has a direct correlation with F2 frequency values, therefore, *front* sounds have a high F2 frequency, while *back* sounds have a low F2 frequency.

As for the feature *rhotic*, it seems no matter which variant of English /r/ is utilized, the distinctive property across the board is a lowered F3. In fact, of all English sounds, /r/ has the lowest F3, making it the most distinctive spectral property of /r/. F3 is so low that it is generally quite close to the F2 (Ladefoged and Maddieson, 1996). Therefore, there are two acoustic correlates for *rhoticity*: a low F3 and the close proximity of F3 and F2. I will discuss what it means to have a lowered F3 in more detail in §7.2.

For each acoustic correlate, an acoustic measurement must be calculated. Table 3, below, lists all the defining features, acoustic correlates and acoustic parameters for the approximants.

---

[2] Both *back* and *front* are used so that prevocalic /l/ can be defined as [–*back*] and [-*front*] because the F2 frequency values for prevocalic /l/ lie between those for *front* /j/ and *back* /w/.

Table 3: Defining Features, Acoustic Correlates and Acoustic Measurements for English Approximants

Defining Features	Acoustic Correlate	Acoustic Measurements	/l/	/r/	/w/	/j/
High	low F1 frequency	F1-F0	*high*	n/a	*high*	*high*
Back	low F2 frequency	F2-F1	n/a	n/a	*back*	n/a
Front	high F2 frequency	F2-F1	*front*	n/a	n/a	*front*
Rhoticity[3]	low F3 frequency; close proximity of F2 and F3	F3-F0; F3-F2	n/a	*rhoticity*	n/a	n/a

All of the acoustic correlates discussed above are relative measures (i.e. high F2, low F1, close proximity of F3 and F2, etc.); there is no definite threshold frequency value that determines if the formant is high or low; formant height can only be determined relative to another segment or another formant. Therefore, Espy-Wilson (1992) defined acoustic parameters in order to compare the formant frequency measurements for the approximants. For the feature *high*, the acoustic correlate is a low F1 frequency. To calculate this the acoustic parameter F1-F0 is used. For *back/front* feature, the acoustic correlate is a high/low F2, thus, the acoustic parameter is F2-F1 for both. As for the feature *rhotic*, there are two acoustic parameters, each corresponding to one of the two acoustic correlates of *rhotic*. For a low F3 frequency the parameter F3-F0 is used and to measure a close proximity of F3 and F2 the parameter F3-F2 is used.

[3] F3-F0 measures the value of F3 in absolute terms while F3-F2 measures of F3 in relative terms (relative to the speaker's F2 value). An additional method used to measure rhoticity is discussed further in § 7.2.

I have defined AoS, covert contrast, the approximants and their acoustic correlates and I have presented a way of measuring the existence of covert contrast in Marshall's approximant system. In §5 that follows, I will describe the hypothesis of my paper in more detail.

5 Hypothesis

Marshall's error patterns concerning the class of approximants, suggests that covert contrasts do exist in his speech. Marshall produces an /r/ that sounds like [w] in singleton onsets, i.e. 'red' [wɛd], yet, he also produces an adult-like target /l/ in singleton onset position, i.e. 'let' [lɛt]. In complex onsets, Marshall's /r/ also sounds like [w], i.e. 'crack' [kwæk] but Marshall's /l/ in complex onsets varies from utterance to utterance with no discernable pattern; most often it is a target /l/, i.e. 'flag' [flæg], but sometimes it is a [w], i.e. 'clock' [kwɑk]. However, Marshall produces a systematic difference between codas containing /r/ and /l/. Marshall's /l/ in coda position is a vocoid variation, i.e. [pɪᵘ] 'pill' and [mɪᵘk] 'milk'. Marshall's /r/ in a coda is like that of a speaker from an r-less dialect, i.e. [kɑ:] 'car', [hɪᵊ] 'hear', dark' [dæ:k]. The significance of these observations can only be appreciated if we place them within the context of normal acquisition. Acquisition findings are thus overviewed in the following sections.

5.1 Stages in the Normal Development of Approximants

In the basic acquisition pattern of approximants, glides are generally acquired before liquids (Bernhardt and Stemberger, 1998). In fact, [w] frequently appears in infant babbling. Glides are fairly stable in that patterns affecting them are uncommon and /w/ and /j/ are rarely interchanged. Liquids on the other hand are less stable and may develop over a longer period of time. /l/ is usually acquired before rhotics, but both orders of acquisition are found in English (Bernhardt and Stemberger, 1998). In acquisition, /l/ and /r/ are commonly replaced by glides in onsets, and by vowel-like elements in codas (which appears to be the basic pattern observed in Marshall). In onsets, /l/ becomes [j] and /r/ often becomes [w] and sometimes [j] (/r/ also becomes [l] in onsets but this is rare and occurs when glides are not produced in the data; in this situation /j/ also becomes [l] due to the impossibility of glides) (Bernhardt and Stemberger, 1998). In codas, /ɫ/ is often replaced by [ʊ] and /r/ is replaced by vowel lengthening [:] (as in an r-less dialect). The general conclusion to be drawn from the acquisition data presented above is that glides are acquired before liquids.

In adult speech, there are different target allophones for /l/ and /r/ in coda versus onset position. Table 4 below shows the allophones of adult English liquids.

Table 4: Liquid Allophones of Adult English. (Modified from Bernhardt and Stemberger (1998)).

Onsets	Codas
l	ɫ
ɹ [+rounded]	ɹ [-rounded] after unrounded vowels

Because children acquiring English are exposed to varying allophones for the liquids depending on its position in the syllable, syllable position could influence the order of acquisition in codas and onsets (Bernhardt and Stemberger, 1998). In other words, the acquisition of glides and liquids can be further nuanced by syllable position.

5.2 Covert Contrast in Normal and Disordered Development

One indication that would suggest the presence of covert contrast in AoS was discussed above in §2.2, Bahr (2005) found that children with AoS had more trouble coordinating complex articulatory gestures than children with phonological disorders. Also, in §3.1 it was mentioned that sometimes both normal and disordered children are able to produce a contrast before listeners can perceive it because the acquisition, development and co-ordination of acoustic cues occurs gradually and the child must learn to coordinate the production and timing of numerous phonetic cues in order for the contrast to be perceivable (Scobbie et al., 1998). Due to their difficulty coordinating gestures, children with AoS would be more likely to have covert contrasts which listeners assume to be contrast neutralization.

5.3 Main Hypothesis

The above observations lead to the hypothesis that (a) Marshall should have a contrast between /w/ and /r/ in singleton onsets and that (b) since we cannot hear a contrast, it must be covert. Acoustic analysis should reveal a covert contrast between Marshall's /r/ and /w/ in singleton onsets.

6 Methodology

In this section of the paper I describe how this investigation was conducted and recount what led to the results discussed in §7. As well, I explain the rationale for the specific procedures chosen. This sections starts with a discussion of the design of the study in §6.1, followed by a description of the participant, Marshall, in §6.2. Next, §6.3 describes where the data presented here originated and how it came to be used in the current study. § 6.4 describes the instruments and tools used in the data collection and analysis. §6.5 explains the procedures used in the data collection, extraction, acoustic analysis and subsequent statistical analysis.

6.1 Design

The use of a case study is appropriate in this situation because the participant displays a behavior that is rare; therefore, the results cannot be compared with the general population. The data obtained in a case study is valid in its own right, regardless of how well it represents a population. Also, a study involving apraxic speakers requires a flexible methodology, such as that found in case studies, due to a lack of agreement about the underlying impairment involved in AoS. Another motivating factor in selecting the design of the study was the complexity of AoS. Errors produced by apraxic speakers vary from one individual to another, as well as from one moment to another for one individual. Using a group study would complicate the research because extreme caution would have to be taken to ensure the speakers displayed comparable errors. (Wray et al., 1998:190)

On the other hand, one of the strengths of case studies is also a weakness. A small sample size may stand well on its own but it does make it difficult to compare to the results with a population (Wray et al., 1998:190). However, I follow other researchers, some of which are described in §3.2, in choosing the case-study methodology.

6.2 Participant

The participant in my study is Marshall (pseudonym), a fifteen-year-old male from St. John's, Newfoundland. A speech-language pathologist diagnosed Marshall with AoS at the age of ten and he attended speech therapy sessions for approximately one year after diagnosis. Marshall's speech therapy sessions focused on decreasing his speech rate, improving his articulation of consonant and vowel pairs, improving articulation of function words and improving sound awareness and rhyming.

Marshall was referred by his school to the dyslexia reading clinic at Memorial University in 2001at the age of thirteen due to academic difficulties. At that time, Marshall began literacy tutorial sessions with different instructors, including Gavin Willins, then an undergraduate student in the Linguistics Department, using a rhyme-based approach to reading called the Glass Analysis technique (Glass, 1976), described later in § 6.3.

When Willins began his Master of Arts degree with the Department of Linguistics in 2003 he decided to focus on AoS for his M.A. thesis and examine Marshall's speech. He looked specifically at the remediation of speech impairment through reading instruction. Willins began recording the tutorial sessions with Marshall in February 2004 and continued to do so until the end of May 2004. Marshall's age at the time of the initial recording was 15;05.03 and he was 15;08.13 at the time of the final recording.

When I began my Master of Arts degree, I was interested in studying disordered speech and obtained permission from Marshall's parents and ethical approval from the university to use Willins' recordings of Marshall's tutorial session for my thesis research.

In the fall of 2005, when Willins was no longer available to tutor Marshall, I began literacy tutorial sessions with Marshall once a week using the same rhyme-based approach that was employed by Willins. I have been working with Marshall for the past three years and continue to work with him. I have witnessed an improvement in his reading skills over that time, as I have periodically administered the Slosson Oral Reading Test (SORT), which is a test designed to asses word recognition levels. The results of Marshall's initial SORT test, in October 2005, indicated that he was reading at a level of approximately a grade two student. The most recent SORT test, administered in July 2009, indicated that Marshall is currently reading at approximately a fourth grade level.

6.2.1 Error Patterns

As discussed in § 4.3, some of the salient characteristics of AoS are: largely unintelligible speech; physical groping; a high consonant error rate (i.e. substitutions, omissions, additions, etc); and inconsistency of errors. Marshall's speech is highly unintelligible; it is very difficult for an unfamiliar listener to comprehend what he is saying. Marshall also exhibits persistent physical groping and frequent hesitations in his productions. Marshall's speech was examined closely and found to contain many different error patterns, mostly affecting consonants, which is typical of AoS speakers (Jacks et al., 2006). Marshall's vowel system appeared to be near the adult target system. Many of Marshall's consonant error patterns are not always categorical. In contrast, other errors are categorical, occurring every time Marshall attempts to produce a certain sound.

Table 5 (modified from Willins (2005)) shows a summary of Marshall's phonemic inventory in singleton onset productions.

Table 5: Marshall's singleton onset phonemic inventory. Percentage of production rate given due to high inconsistency in errors. Modified from Willins (2005)

Onset	Production	Target Example	Erroneous Production
p	/p/ → [p] (99.6%) /p/ →[d] (0.4%)	paper	dʌɪpɚ
b	/b/ →[b] (99.5%) /b/ →[d] (0.5%)	beep	dip
f	/f/ →[f] (100%)	-	-
v	/v/ →[f] (76.9%) /v/ →[v] (23.1%)	van	fæn
m	/m/ →[m] (97.9%) /m/ →[b] (2.1%)	monopoly	bɑnɑpli
*θ	/θ/ →[t] (100%)	think	tɪŋk
*ð	/ð/ →[d] (97.3%) /ð/ →[ð] (2.7%)	that	dæt
t	/t/ →[t] (100%)	-	-
d	/d/ →d] (100%)	-	-
s	/s/ →[s] (90.9%) /s/ →[z] (9.1%)	soup	zup
z	/z/ →z] (56.9%) /z/ →[s] (43.1%)	zebra	sibɹə
n	/n/ →n] (100%)	-	-
l	/l/ →[l] (100%)	-	-
ɹ	/ɹ/ →[w] (100%)	right rug	wʌɪt wʌg
tʃ	/tʃ/ →[ʃ] (100%)	cheap	ʃip
dʒ	/dʒ/ →[ʃ] (100%)	jumpins	ʃʌmpɪnz
ʃ	/ʃ/ →[s] (100%)	shoot	sut
k	/k/ →[k] (100%)	-	-
g	/g/ →g] (100%)	-	-
h	/h/ →h] (100%)	-	-
w	/w/ →[w] (100%)	-	-
j	/j/ →[j] (100%)	-	-

*Dialectal effects (discussed below)

In singleton onsets Marshall simplifies affricates into fricatives so that /dʒ/ and /tʃ/ are realized as [ʒ] or [ʃ] (i.e. 'cheap' /tʃip/ sounds like 'sheep' [ʃip]). In turn, postalveolar fricatives are produced as their alveolar counterparts so that /ʃ/ is realized as [s], and /ʒ/ as [z] (i.e. 'shut' /ʃʌt/ sounds like 'sut' [sʌt]). These are both systematic errors, occurring 100% of the time. Marshall also has difficulty with voicing contrasts. These errors are not systematic in his speech; for example, /v/ is almost always realized as [f] (76.9%); /f/ is always produced correctly as [f]; /z/ is realized as [s] almost 50% of the time, yet /s/ is realized as [z] 9.1% of the time. Another systematic error pattern exhibited in Marshall's speech is within the class of approximants. The liquid /r/ is realized as the semivowel [w] in singleton and complex onsets 100% of the time.

One "error" mentioned in Marshall's speech is likely not an error, but a dialectal characteristic. As mentioned above, Marshall is from St. John's, Newfoundland. Marshall is amongst the category of middle class, young male St. John's speakers. One of the dialectal features typical of this group of speakers is to replace the voiced and voiceless dental fricatives /θ/ and /ð/ with the corresponding voiced and voiceless alveolar stops (i.e. /θ/→[t] and /ð/→[d]) (Paddock, 1982). Although it is possible that this may be an error due to AoS, it is far more likely that this substitution is an effect of Marshall's local dialect.

As discussed above, Marshall displays all of the salient characteristics of an individual with AoS. He has difficulty initiating speech; his groping is both visually and audibly apparent; his vowel system is very close to normal, yet his consonant articulations are frequently off target; and these errors occur in similar environments, although the types of errors may vary from trial to trial.

I will next describe how I analysed Marshall's speech.

6.3 Data Provenance

Data was taken from a recorded database of literacy tutorial sessions with Marshall compiled by Gavin Willins (2005). The recordings took place at Memorial University of Newfoundland in a private reading room located in the Science Building. The recorded tutorial sessions contained two types of speech: oral reading and spontaneous speech. The oral reading portions of the recordings are part of the rhyme-based approach Willins utilized in the lessons. This approach is called the Glass Analysis technique (Glass, 1976). It is used for students who cannot segment words into phonemes. The goal of Glass Analysis is to teach the student letter-sound correspondence for units larger than phonemes such as the syllable onset and rhyme (an advantage to focusing on the rhyme is that the spelling-sound correspondence is more regular at the level of syllable and rhyme than at the level of individual phonemes). In this approach, words are taught in groups (or families) where the groups have common rhymes. For example, one group could be: "bat, cat, hat, mat", while another group could be: "will, pill, sill, dill".

The oral reading sections of Willins' recorded tutorials are made up of two distinct reading tasks. The first task is comprised of oral reading passages from reading-level-appropriate books. The second task involves oral reading of randomized word lists.

Many different word lists were used throughout the tutorial sessions. An example of a typical word list would be: "see, dog, stamp, clock, bat, dirt, golf, lamp" etc. The subset of words from the randomized word list that I used for my analysis (CɛC words) can be seen in Appendix A. The spontaneous speech portions of the recordings are also made up of two types of speech: flashcard identification and true spontaneous speech. The flashcard identification speech is not part of the literacy tutorials; it is a task Willins used to elicit desired singleton and complex onsets for the purpose of his study. Flashcards consisted of a picture of an object with no other information and Marshall was asked to name the object. For the purpose of his study, Willins was looking to elicit a number of different singleton and cluster onsets. An example of a flashcard used to elicit the word 'jet' [d͡ʒɛt] is pictured below in Figure 2. Marshall's response to this flashcard was [ʃɛt] as he simplifies affricates into fricatives[4].

[4] It is of interest to note Marshall's avoidance of problematic words as described by Willins (2005). "Marshall will often make use of an avoidance strategy. In order to avoid difficult articulations, he may substitute a word giving him difficulty for a synonymous replacement such as answering "pickup" for "truck" or "plane" for "jet". This was, and continues to be, a factor in the presentation of the flash cards. If Marshall chooses to abandon a difficult articulation, I will ask for another name for that item. If the new answer does not produce the desired target, the card will be removed for the next trial and a new flash card inserted in its place to elicit the same target." (Willins, 2005)

Figure 2: Example of flashcard used in tutorial sessions to elicit word 'jet' [d̠ʒɛt]

Finally, the true spontaneous speech is just that, spontaneous utterances Marshall produced during his sessions, mostly in conversation with Willins.

6.4 Apparatus

Willins recorded Marshall's tutorial sessions using a Sony DAT recorder as a preamp with a direct line to a computer. The sound was then digitized by Willins using Amadeus audio recording software (http://www.hairersoft.com/Amadeus.html) and saved as digital sound files. For transcription purposes, Willins linked each sound file to a written file using the CLAN (Computerized Language Analysis) program (http://childes.psy.cmu.edu/). Once I obtained the files, I used the CLAN program to extract individual words with approximant onsets from the recordings. For example, from the sentence "I wonder what it's like out?" I extracted the words 'wonder' 'what' and 'like'. When the desired tokens were extracted using CLAN, I analysed each word acoustically, using Praat software (http://www.praat.org). Praat is free software that can perform a wide range of analysis and manipulations including spectral, pitch, formant and intensity analysis. The acoustic analysis performed using Praat is discussed further in §6.5. I then entered all of the raw data into the Microsoft Excel program, which was used for statistical analysis purposes by utilizing the data analysis function of Excel. The statistical analysis, including the types of tests performed, are described in detail below in §6.5.

6.5 Data For This Study

The data used in this study came from a total of twelve recorded tutorial sessions, each session lasted approximately one hour. I extracted CVC words with an approximant in the singleton onset position from Marshall's recorded speech. I initially extracted a total of 1564 CVC words with approximant singleton onsets in the recorded sessions; however, the four approximants were not equally represented in all possible vocalic contexts. Only the front vowel /ɛ/ context (CɛC) provided adequate representation of all four approximants and so I focused on CɛC shaped words. There were a total of 230 (CɛC) words with an approximant in the onset position (/w, r, l, j/). Some examples of CɛC words used include: 'ledge', 'leg', 'less', 'red', 'ren' (a nonsense word[5] used by Marshall), 'well', 'wet', 'yes', and 'yet'. Refer to Appendix A for a complete inventory of the CɛC words extracted.

[5] There are three nonsense words contained in the data that have either /r/ or /w/ in the onset position (two with /r/ and one with /w/). I determined which segment Marshall was intending to produce based on the context in which the word occurred. All nonsense words occurred during reading drills when Marshall came upon a word that was difficult for him. He would guess the word, often times making up a nonsense word in the process, but he would include the correct onset in his guesses. Willins always spoke the correct target word after Marshall's guesses, therefore making it possible to determine Marshall's target word while listening to the recordings.

The dependent variables in this study were the different approximant consonants, while the independent variable was the environment occurring before [ɛ]. Possible confounding independent variables in the study are the speech rate and speech type. As discussed above in § 6.3, there are two main speech types in the data recordings: oral reading and spontaneous speech. Willins observed that Marshall's speech was more intelligible during oral reading speech than it was during spontaneous speech. I did not control for the type of speech Marshall used during the tutorial sessions because I did not have enough tokens of the target singleton onsets in each speech type. Given that there is a marked difference between the two speech types, this decision may introduce a confounding independent variable: the speech type could affect the quality of the dependent variable, the approximant consonants. However, since Marshall produced [w] for /ɹ/ regardless of the task, speech type is likely not a relevant factor.

Even though the focus of the investigation is specifically on /ɹ/ and /w/, I included the two other English approximants /l/ and /j/ in the analysis to get a more complete picture of Marshall's approximant category and his overall sound inventory. Recall from earlier discussion in §6.2.1 that Marshall appears to produce /w/, /l/ and /j/ in a relatively adult-like manner. By analysing Marshall's entire approximant inventory I was able to compare his approximant productions to those of normal adult speech and determine if they were in fact similar in the cases of /w/, /l/, and /j/.

6.6 Data Analysis

Once I extracted the desired data, I analysed the acoustic values of the approximant consonants. In the onset of each of the 230 extracted words I measured the F0 as well as the F1, F2, and F3 frequencies (920 measurements in all).

Following Espy-Wilson's (1992) methodology, I measured the formants for each approximant as follows /w/ during the F2 minimum value, /j/ during the F2 maximum value, /l/ during the F2 minimum value and /r/ during the F3 minimum value. Whenever possible, I picked the steadiest point in the formant during each of these minimum or maximum values as the measurement point. I also utilized the formant transitions as visual guides to determine where each approximant ended and the following vowel began, though the transitions themselves were not part of any measurements (formant transitions occur at the edges of the segment and are due to coarticulation with adjacent segments). To measure the formants and durations, the settings used in Praat were kept constant. The view range was 0-4000 Hz, the window length was 0.03 s (30 ms), and the dynamic range was 60 dB.

I used the formant frequency values for each segment to calculate the acoustic measurements that correspond to the distinguishing features for the approximants. As discussed in §4.3 the distinguishing features for the approximants are *height*, *backness* and *rhoticity*. The acoustic measurement corresponding to *height* is F1-F0, for *backness* F2-F1 and for *rhoticity* F3-F0 and F3-F2. In order to calculate F1-F0 for a particular segment, the F0 measurement for that segment was subtracted from the F1 measurement for the same segment. For F2-F1, the F1 measurement was subtracted from the F2 measurement, and so on and so fourth. I calculated these parameters for each of the 230 segments for a total of 920 measurements.

I also measured the duration of the entire /CɛC/ syllable, as well as, the duration of the onset consonant. I obtained durations so that the ratio of consonant length versus syllable length could be calculated, while this measurement could have possibly revealed some significant differences among the approximants, during the analysis of the measurements, I determined that the durational ratio provided no relevant information and hence I decided not to use this measure in my analysis.

Finally, for reasons discussed in §7.2, I also measured the F3 frequency of a sample of the vowels /i/, /e/, /ɑ/, /o/, and /u/ in Marshall's CVC syllables. I extracted additional words from Marshall's corpus in order to carry out these measurements. The full inventory of tokens extracted for this purpose is presented in Appendix B.

I did not include stutters or hesitations in the measurements. For example, if there was a stutter at the beginning of the word 'let' [lə-lɛt], only the [lɛt] portion was measured. However, I did include sustained segments. For example, if the onset for the word 'let' was prolonged as in [lllllllllllllɛt], I included the entire [l] in measurements for both frequency and duration.

Once I measured all of the formants and calculated the acoustic parameters for all 230 segments, I performed a statistical analysis on all raw scores to determine if there was a significant difference between the four approximants, particularly between /r/ and /w/. An alpha level of 0.05% was used for all statistical tests. First, I obtained the descriptive statistics for all the formants in each approximant environment. The descriptive statistics included the mean, standard deviation, standard error, variance and the sample size. All descriptive values are listed in Table 7 found below in §7. Second, I calculated an analysis of variance (ANOVA) with all of Marshall's approximant consonants and discovered a main effect of the consonant type. Third, I performed two-sample *t*-tests assuming unequal variances for each possible paired comparison for /r/ versus /w/ and determined a significant difference between the formant values obtained for the two phonemes. The results of the acoustic parameter *t*-tests are shown in Table 13-Table 16 in §7.

7 Results

I calculated descriptive statistics for all of the formant values (F0, F1, F2, and F3) for each approximant (/l/, /j/, /w/ and /r/). Before presenting them, however, Table 3 detailing the defining properties of the class of approximants from §4.3 is reproduced below in Table 6 for ease of reference.

Table 6: Defining Features, Acoustic Correlates and Acoustic Measurements for English Approximants (Reproduced from Table 3)

Defining Features	Acoustic Correlate	Acoustic Measurements	/l/	/r/	/w/	/j/
High	low F1 frequency	F1-F0	*high*	n/a	*high*	*high*
Back	low F2 frequency	F2-F1	n/a	n/a	*back*	n/a
Front	high F2 frequency	F2-F1	*front*	n/a	n/a	*front*
Rhoticity	low F3 frequency; close proximity of F2 and F3	F3-F0; F3-F2	n/a	*rhoticity*	n/a	n/a

Listed below in Table 7 are the results calculated for the descriptive statistics of each formant measure for all of Marshall's approximant consonants. A discussion follows in §7.1.

Table 7: Descriptive Statistics for Formant Frequencies of Marshall's /l/, /r/, /w/, and /j/

		/l/	/r/	/w/	/j/
F0	Mean	121.218	120.348	133.162	115.313
	Standard Deviation	32.913	31.304	38.619	28.384
	Standard Error	4.322	4.615	4.790	3.634
	Sample Variance	1083.255	979.909	1491.466	805.633
	Sample Size	58	46	65	61
F1	Mean	436.920	433.797	487.201	438.542
	Standard Deviation	107.967	107.331	84.564	112.586
	Standard Error	14.177	15.8250	10.489	14.4152
	Sample Variance	11656.944	11519.848	7151.055	12675.610
	Sample Size	58	46	65	61
F2	Mean	1568.853	1261.845	970.243	2117.681
	Standard Deviation	219.132	522.157	279.663	280.678
	Standard Error	28.773	76.988	34.688	35.937
	Sample Variance	48019.028	272647.708	78211.192	78779.984
	Sample Size	58	46	65	61
F3	Mean	2845.339	2508.767	2645.717	2928.331
	Standard Deviation	289.136	267.288	207.649	353.605
	Standard Error	37.965	39.409	25.756	45.274
	Sample Variance	83599.750	71442.936	43118.174	125036.225
	Sample Size	58	46	65	61

7.1 Formant Frequency Means

The means of the formant frequencies for Marshall's approximant consonants are shown in Figure 3 below.

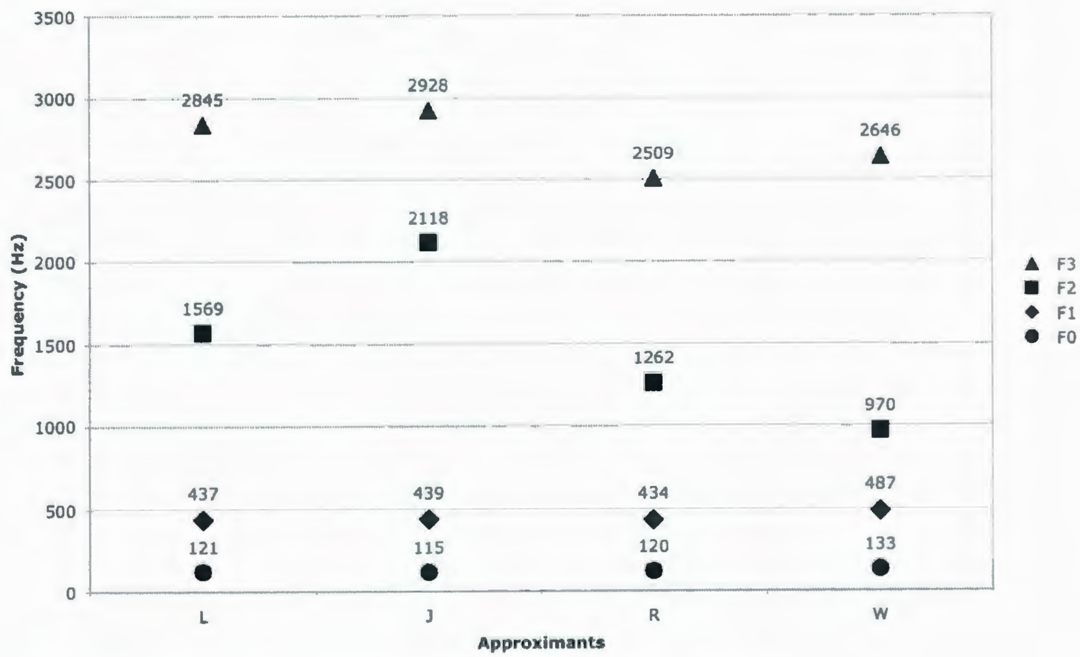

Figure 3: Marshall's approximant formant frequency means

The F0 frequency in Figure 3 stayed relatively stable across all contexts, as one would expect when analysing utterances by a single speaker. Figure 3 suggests that Marshall produces a distinct formant pattern for each approximant, indicating that he does have some form of covert contrast between /r/ and /w/.

To determine if Marshall's formant frequencies for his approximants were significantly different, a two-factor ANOVA with replication was performed to determine the effect of the consonants. There was a statistically significant main effect of consonant type ($F(6, 540)=49.14$, $p<0.0001$) indicating that Marshall does indeed produce four distinct approximant consonants in singleton onsets. More importantly, this statistic confirms that he is producing a covert phonemic contrast for /r/ and /w/ in singleton onsets.

Though Marshall has four distinct approximants, each one is not necessarily produced in the same manner as an adult speaker would produce them. Some, but not all, aspects of Marshall's approximant formant distributions in Figure 3 are as predicted based on the acoustic correlates that correspond to the defining features of the approximants presented above in Table 6. In Figure 3, F2 behaves as expected across all approximants. It has a relatively low frequency value for /w/, indicating /w/ is more *back*, and a relatively higher frequency value for /j/, indicating that /j/ is more *front*. However, F1 was relatively stable across all approximants; one would expect that /w/ and /j/ would have a lower F1 frequency, because they are typically more *high*, relative to /r/ and /l/. As expected from the acoustic correlates, F3 is lower in /r/ than in Marshall's other three approximants, indicating *rhoticity*. However, a close proximity of F2 and F3 was also expected for /r/ when compared to the other approximants, yet, /j/ had the closest F2 and F3 frequencies of all the consonants, a fact discussed further after Table 8/Figure 5. From looking at Figure 3, we can tell that some aspects of Marshall's approximant productions seem to resemble typical adult-like target approximants, while other aspects seem to differ from the adult target.

In order to determine how Marshall's approximants compare to normal adult productions, we must next look at a normal adult approximant system. Figure 4 below shows the means of formant frequencies for approximant consonants spoken by normal adult speakers. This data is presented in Espy-Wilson's (1992) study of features that distinguish the class of approximants from the other consonant classes, as well as the features that distinguish the four different approximants consonants from each other.

Espy-Wilson (1992) presented an average of approximant formant frequencies across all speakers (two male and two female adult speakers). Marshall's frequencies would presumably range somewhere in between that of a typical male adult speaker and a typical female adult speaker, since he was fifteen years old at the time of the recordings. Marshall's voice was not quite into the average adult male frequency range at that point in time, for this reason an average which includes both male and female adult speakers would probably compare quite well to Marshall's frequency range. Hence, I compared Marshall's approximant productions with the mean formant frequencies of all speakers presented in Espy-Wilson's (1992) data.

The four speakers used in the Espy-Wilson study spoke American English. The female speakers were from the northeast and the males were from the Midwest. The difference in dialect between Marshall and the participants of Espy-Wilson (1992) is a potential barrier to directly comparing Marshall's productions to the American English productions (as stated earlier in §6.2.1, Marshall speaks an English dialect from St. John's, Newfoundland). Another possible source of discrepancy between the two sets of data is the environment in which the consonants were found. As previously mentioned, all of Marshall's productions were in the singleton onset position in CVC syllables. As well, the context of the following vowel was controlled for by including only the front vowel /ɛ/. The Espy-Wilson (1992) data set, however, contained prevocalic approximants, only some of which were in word-initial singleton onset position. Others were found in clusters, both word initial and word-medial. As well, the vowel context was not controlled for in the Espy-Wilson (1992) data. Other sources containing data on English formant frequency means include Edwards (2003), Stevens (1998), and Kent and Read (1992).

With these caveats in mind, Figure 4 shows normal adult formants for the approximants.

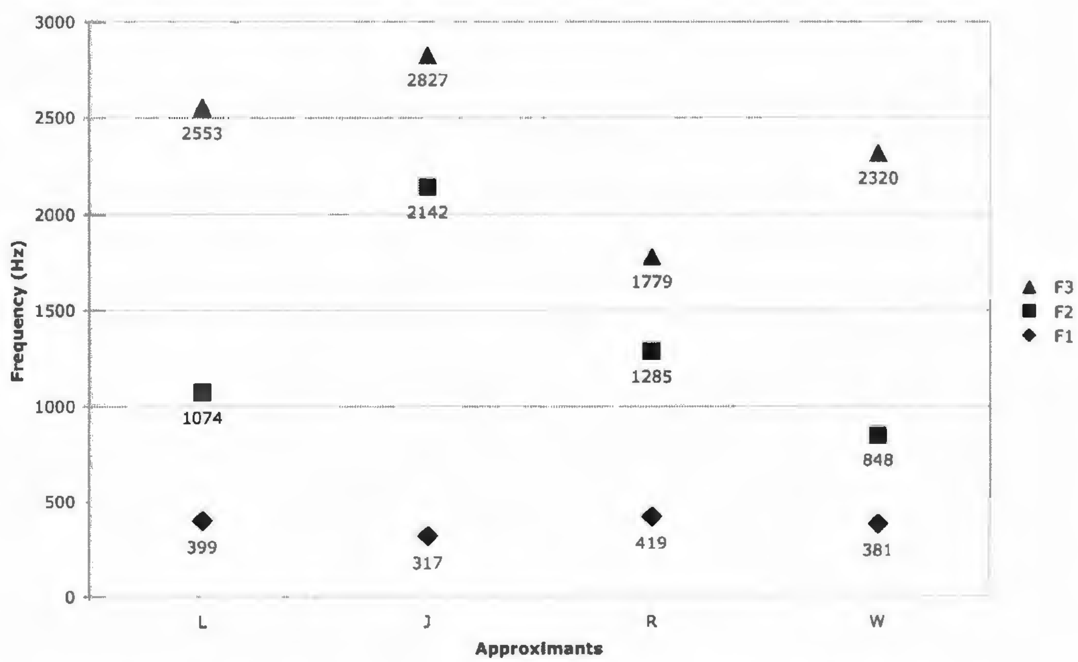

Figure 4: Normal adult formant mean values from Espy-Wilson (1992)

When Marshall's productions in Figure 3 are compared directly with the normal adult speaker productions in Figure 4 it is obvious that although Marshall does produce four distinct approximants, he does not produce all four in an adult-like manner.

Each of Marshall's approximants must be compared individually to that of normal adult speakers in order to get a full picture of his approximant system. Table 8-Table 11 below compare the formant frequency mean values for the normal adult speakers and Marshall. For ease of comparison, Figure 5-Figure 8 show Marshall's values compared directly with those of the normal adult speakers in Espy-Wilson (1992) in line graph format. There are some differences between Marshall's productions and the normal adult productions in all the approximants, especially with respect to /r/. The data for Marshall and the Espy-Wilson (1992) adults for /j/ are presented below in Table 8 and Figure 5.

Table 8: Marshall and normal adult speakers frequency means and difference for /j/

		F1	**F2**	**F3**
/j/	Marshall	439	2118	2928
	Normal Adult	317	2142	2827
	Difference[6]	**122**	**-24**	**101**

[6] The raw scores and standard deviations for the normal adult data obtained from Espy-Wilson (1992) were not available, only the formant frequency means were presented in the paper, thus, I could not conduct a statistical test to compare the normal adult formant means with Marshall's formant means for any of the approximant consonants. The only comparison possible was to calculate the differences between the means, however, caution must be taken when making comparisons such as this as they are not necessarily statistically significant differences.

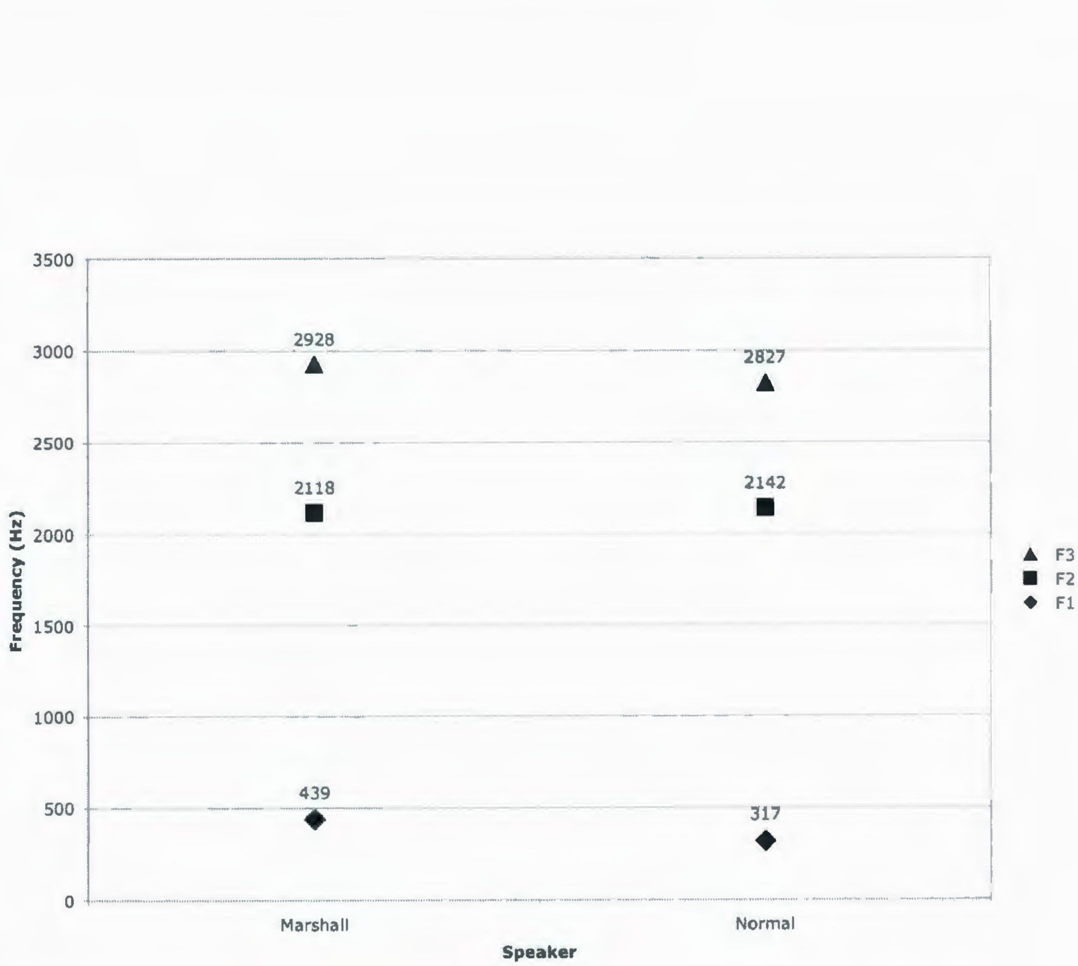

Figure 5: Marshall's formant means versus normal adult formant means for /j/; adult values from Espy-Wilson (1992)

There are no appreciable differences between Marshall's /j/ and that of normal adult speech. The largest difference in the formant means, 122 Hz, is found in F1. As F1 correlates negatively with *height*, this indicates that Marshall's /j/ is slightly less *high* than the normal adult production. Otherwise, Marshall's /j/ is very much like a target adult production. In Figure 3, it was mentioned that the proximity between F2 and F3 for Marshall's approximants was expected to be closest for /r/, when it was actually closest for /j/. In Figure 5 above, it is clear that the proximity of F2 and F3 in Marshall's /j/ is comparable to the normal adult productions. This impression is confirmed in the discussion that follows Table 12 in §7.2.

The data for /w/ is presented below in Table 9 and Figure 6.

Table 9: Marshall and normal adult speakers frequency means and difference for /w/

		F1	F2	F3
/w/	Marshall	487	970	2646
	Normal Adult	381	848	2320
	Difference	**106**	**122**	**326**

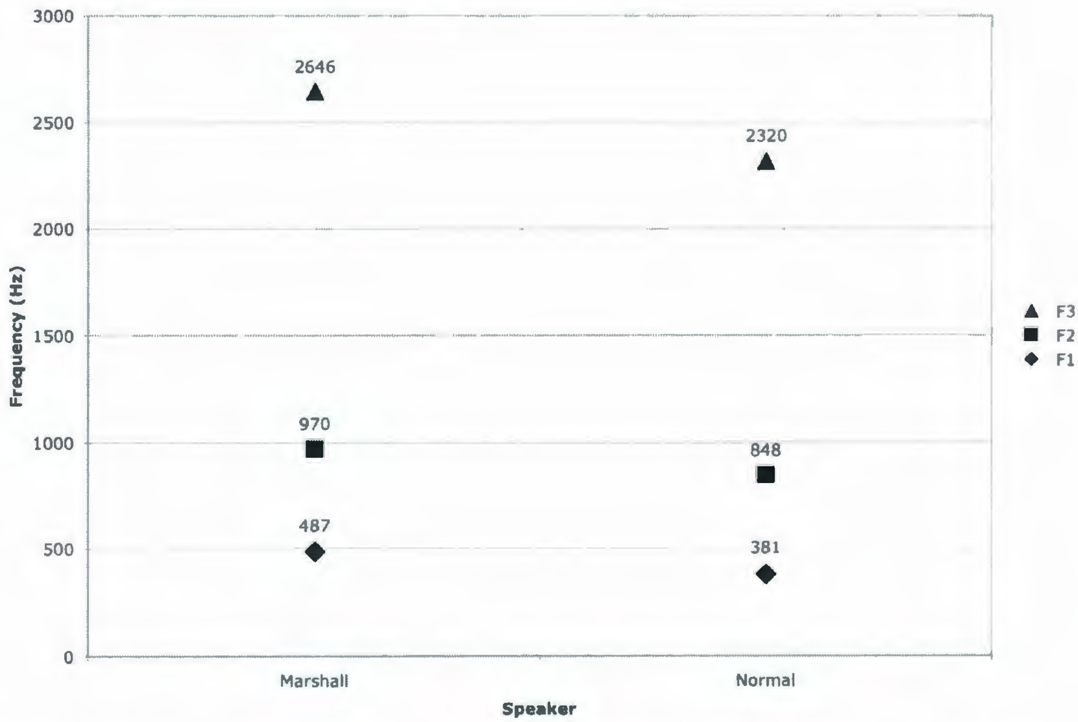

Figure 6: Marshall's formant means versus normal adult formant means for /w/; adult values from Espy-Wilson (1992)

The major difference between Marshall's /w/ and that of the normal adult speakers occurs in F3. Although Marshall's F1 and F2 frequencies are somewhat higher than the normal adults', his F3 frequency is 326 Hz higher than the normal adult F3. There is no doubt that even though it sounds very similar to an adult /w/, Marshall's /w/ is acoustically different. Marshall's F3 values will be discussed in more detail in §7.2

below. Because Marshall's F2 is higher than the adult F2 (122 Hz) it could be that his /w/ is less rounded than the adult /w/. Lip rounding affects English *back* vowels by lowering the F2, which in turn makes the segment sound more *back* (Ladefoged, 2005). Because Marshall's F2 for /w/ is higher, it is less *back* and/or less rounded than the adult normal speakers'.

The /l/ data for Marshall and the Espy-Wilson (1992) adults is presented below in Table 10 and Figure 7.

Table 10: Marshall and normal adult speakers frequency means and difference for /l/

		F1	**F2**	**F3**
/l/	Marshall	437	1569	2845
	Normal Adult	399	1074	2553
	Difference	**38**	**495**	**292**

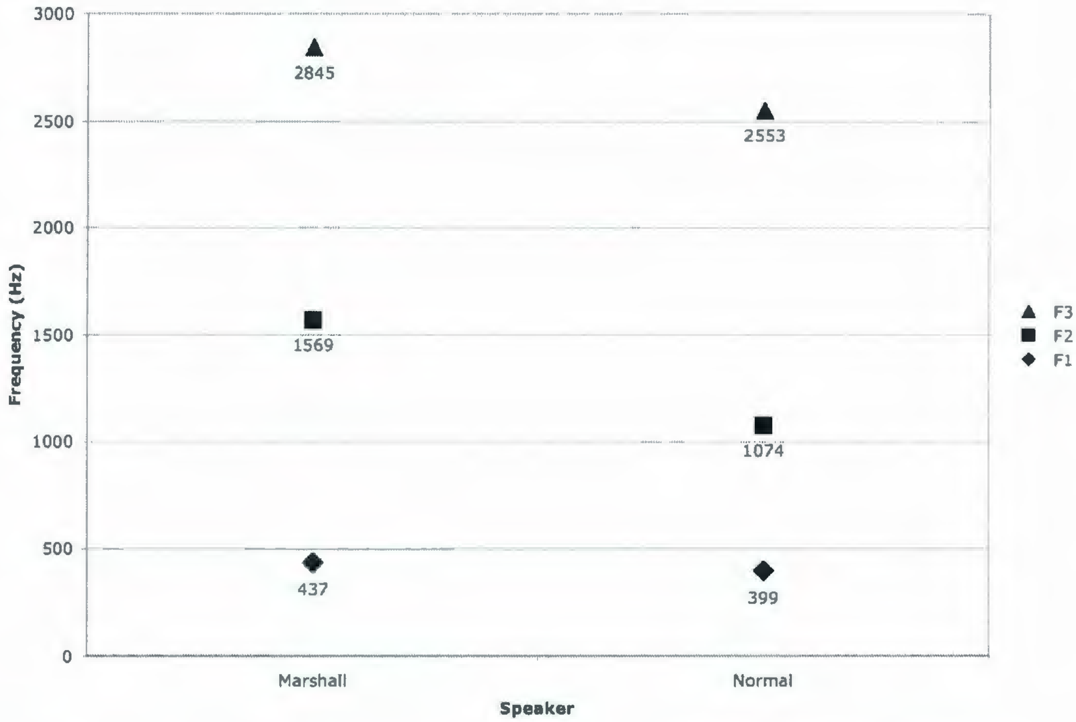

Figure 7: Marshall's formant means versus normal adult formant means for /l/; adult values from Espy-Wilson (1992)

It is evident in Figure 7 that Marshall's /l/ production differs from the normal adult prevocalic /l/ with a higher F2 and F3. As shown in Table 6, F2 correlates directly with backness. A high F2 value indicates a more acoustically *front* sound. Therefore, we can assume that Marshall's /l/ is more acoustically *front* than the target /l/ (but not necessarily more articulatorily *front*). Again, Marshall's overall F3 values will be discussed below in §7.2. As discussed above in §4.2.1, there are two common allophones of English /l/, light /l/ and dark /ɫ/. Light /l/ is typically produced in prevocalic position and is more acoustically *front* than dark /ɫ/ which is typically found in postvocalic position. In addition, dark /ɫ/ is velarized, or produced with a secondary *high back* articulation. Since all of Marshall's /l/ productions were in word initial positions in singleton onsets, it can be assumed that all of his /l/ productions were light /l/ allophones. This, in fact, was confirmed impressionistically in the recording of Marshall's speech. Hence, it makes sense that Marshall's /l/ productions were more acoustically *front*. In contrast, the Espy-Wilson (1992) data used here contains prevocalic /l/'s, which were not exclusively word initial. Many of the Espy-Wilson (1992) prevocalic /l/ tokens were found in complex onsets (either word initial or word medial). Therefore, the acoustic differences seen between Marshall and the normal adult speakers may be due to differences in contexts in the two data sets. The normal adult /l/'s found in clusters may be more acoustically *back* than those found in word-initial singleton onsets. This may account for Marshall's more acoustically *front* /l/ articulations.

The formant means and differences for Marshall and the Espy-Wilson (1992) adults for /r/ are listed below in Table 11 and shown in Figure 8.

Table 11: Marshall and normal adult speakers frequency means and difference for /r/

		F1	**F2**	**F3**
/r/	Marshall	434	1262	2509
	Normal Adult	419	1285	1779
	Difference	**15**	**-23**	**730**

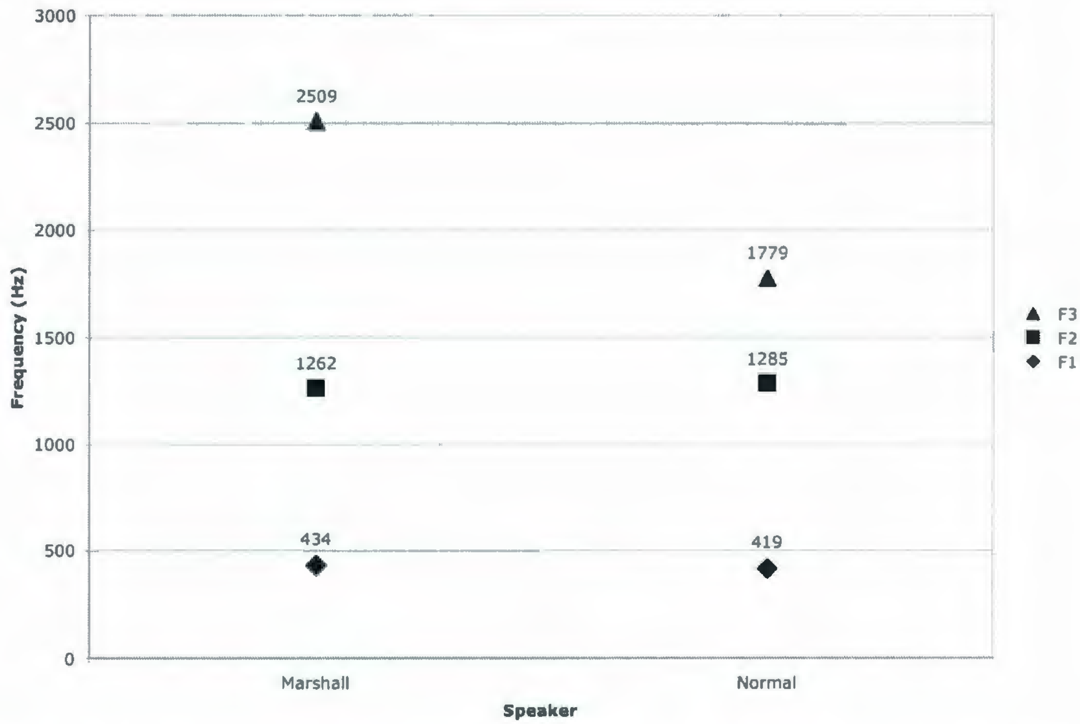

Figure 8: Marshall's formant means versus normal adult formant means for /r/; adult values from Espy-Wilson (1992)

The most striking difference between any of Marshall's approximants and the normal adult speakers is clearly seen in /r/ in Figure 8, more specifically in the F3 of Marshall's /r/. As shown in Table 6 above, the defining feature that separates /r/ from the other approximants is *rhoticity*. The two acoustic correlates for *rhoticity* are a low F3 frequency combined with a close proximity of the F2 and F3 frequencies. These correlates are measured by the two acoustic parameters F3-F0 and F3-F2 respectively. It is evident from Figure 8 that Marshall's F3 frequency is dramatically higher than that of the normal adult speakers; the difference between the two is 730 Hz. In contrast, his F2 is on par with the adult production. Marshall's F2 is a mere 23 Hz lower than the normal adults. Thus, the proximity of F3 and F2 frequencies in Marshall's /r/ production is much farther apart. By looking at Figure 8, it does not appear that Marshall's /r/ displays either of the two correlates for *rhoticity*. It is clear that there is something very different about Marshall's articulation of /r/ compared to the normal adult speakers, especially for F3. In order to examine this observation further, Marshall's overall F3 pattern will be examined more closely in the following section.

7.2 Marshall's F3

It is not just Marshall's F3 for /r/ that differs from the target productions. With the exception of /j/, all of Marshall's approximants have higher F3 frequencies than that of the normal adult counterparts from the Espy-Wilson (1992) data. For /w/, Marshall's F3 is 326 Hz higher, in /l/ it is 292 Hz higher, and in /r/ it is 730 Hz higher. Due to this pattern of high F3 frequencies, Marshall's overall F3 pattern requires closer examination.

Hagiwara (1995) posed the question of what constitutes a "lowered" F3. He called attention to the fact that the lowered F3 of American /r/ is typically discussed in terms of some critical frequency cutoff point (usually around 2000 Hz) where anything below that point is considered to be a lowered F3 and anything above it is not. Hagiwara also pointed out that this is an inappropriate way to look at the situation; given individual speaker variation and that the formant values of vowels are never discussed in such terms. He proposed a method of examining the F3 of /r/ relative to the speaker's own neutral F3 value. To achieve this, the F3 frequency values for the speaker's plain vowels are measured and the overall mean is calculated. It is thought that this average F3 value closely corresponds to an individual's 'neutral' F3 for their particular vocal tract. Then, the speaker's F3 of /r/ is compared to this neutral F3 value. This is done by dividing the F3 of /r/ by the 'neutral' (vowel) F3 and multiplying it by 100, obtaining the percentage of the neutral F3 that the F3 of /r/ represents. Hagiwara states that the F3 in an American /r/ allophone is expected to be approximately 60-70% of the neutral F3. This manner of looking at F3 values relative to the speaker's neutral F3 allows for productions of /r/ allophones that otherwise might not be within the critical frequency cutoff point and hence, not considered to be an /r/. Hagiwara suggests that perception experiments using such a calculation would help to determine a "scale of rhoticity" (Hagiwara 1995: 121). He proposed that this scale would be useful for measuring /r/ productions in both child language acquisition and disordered speech, two areas where comparing F3 productions to adult norms is often misleading (Hagiwara 1995: 121).

In order to compare the F3 value of Marshall's /r/ to his neutral F3, I extracted tokens of /i/, /e/, /ɑ/, /o/, and /u/ from the tutorial recordings in CVC syllables. Ten tokens of each vowel were obtained. The F3 frequency was measured for each token and the mean value for all of the vowel F3 frequencies was calculated to determine Marshall's neutral F3 value, which was 2916.32 Hz. Next, the percentage of the neutral F3 value was calculated for all four of the approximant F3 frequency means (as mentioned above, it is the F3 pattern for the entire approximant system that is of interest, not just the F3 for /r/). The results are shown below in Table 12. The percentages range from 100% (meaning the F3 was not lowered) for /j/ to 86% for /r/.

Table 12: Marshall's Percentages of Neutral F3 Values

Segment	j	l	w	r
Approximant F3	2928 Hz	2845 Hz	2646 Hz	2509 Hz
Neutral F3	2916.32 Hz	2916.32 Hz	2916.32 Hz	2916.32 Hz
% of neutral F3	100.4%	97.55%	90.73%	86.03%

When Marshall's F3 frequencies were compared to normal adult values, they seemed high. However, Table 12 shows that Marshall is exhibiting an adult-like pattern of F3 lowering across his approximants. For /j/ and /l/, Marshall does not seem to alter his F3 frequency from his neutral F3. In fact, Marshall's F3 for /j/ is exactly the same value as his neutral F3 at 100%. For /l/, even though his F3 is almost 300 Hz higher than the normal adult F3, it is actually just under 100% of his own neutral F3 value. The F3 frequencies for /j/ and /l/ are not expected to rise or fall from the speaker's neutral F3. Therefore, Marshall's productions in this case are displaying normal F3 proportions.

In the case of /w/, when comparing Marshall's productions to the normal adult /w/, his F3 appears high because it is 326 Hz above the adults, but it in fact is almost 10% lower than his neutral F3 frequency, §7.3 deals with the formant frequencies of /w/ and /r/ in more detail. The F3 for /r/ in normal adult English is typically lowered more than any other approximant F3. This is also the case for Marshall; his F3 value for /r/ is about 14% lower than his neutral F3 value even though his F3 for /r/ is 730 Hz higher than the adult normal group. Hagiwara (1995) states that although an American /r/ allophone is expected to be approximately 60-70% of the neutral F3, certain exceptions do occur and that some speakers do not lower their F3 that much, while others lower it much more.

Marshall's higher than normal F3 value is quite likely the reason his /r/ is heard as a /w/. Perhaps, it is because his F3 value for /r/ is closer to the normal adult /w/ F3 that his /r/ is heard as a /w/.

7.3 [r] vs. [w]: A covert contrast

It can be seen from looking at the formant means in Table 12 that Marshall does indeed produce two different segments for /w/ and /r/, and also, that Marshall's /r/ segment, while different from his /w/, is also quite different from /r/ in normal adult speech. It is necessary to determine whether these differences are in fact statistically significant. This was achieved through the use of independent sample t-tests.

Independent sample *t*-tests were performed for each pairwise comparison (F1, F2, F1-F0, F3-F0, etc) for /r/ and /w/. All tests used an alpha level of p=0.05. As expected when utterances from the same speaker are analysed, the F0 for Marshall's /r/ and /w/ are not significantly different. However, all of the other *t*-tests performed showed a significant difference between /w/ and /r/. The results for the formant measurement tests were as follows: F1 $t(82)= -2.81$, $p=0.006$ meaning /r/ has a significantly lower F1; F2 $t(63)=3.45$, $p=0.001$ meaning /r/ has a significantly higher F2; F3 $t(81)=-2.91$, $p=0.005$ meaning /r/ has a significantly lower F3. Result tables for the formant *t*-tests can be seen in Appendix C. The results of the acoustic parameter *t*-tests are shown in table format below.

Table 13 shows the results of the *t*-test for F1-F0, the acoustic parameter associated with *height*.

Table 13: R vs. W Independent *t*-test for F1-F0 (*height*)

		/r/	/w/
F1-F0 t-test	Mean	313.5	354.0
	Variance	10702.2	10253.0
	Standard deviation	103.45	101.26
	Sample size	46	65
	Degrees of Freedom	96	
	t Stat	-2.05	
	p value	0.043	

Because *height* has a negative correlation with F1 frequency values, the lower the F1-F0 value, the *higher* the consonant. It is evident in Table 13 that Marshall's /r/ is higher than his /w/ because the F1-F0 value is significantly lower for /r/ than for /w/. This is unusual, considering /r/ is usually produced around the mid-central vowel area (corresponding to the *mid* vowel [ə]) while /w/ is a high consonant, corresponding to the *high back* vowel [u]. I discuss this observation further after Figure 9.

Table 14 shows the results of the t-test for the acoustic parameter F2-F1, which corresponds to *backness*, for Marshall's /r/ and /w/.

Table 14: R vs. W Independent *t*-test for F2-F1 (*backness*)

		/r/	/w/
F2-F1 t-test	Mean	828.0	483.0
	Variance	250287.6	67023.3
	Standard deviation	500.29	258.89
	Sample size	46	65
	Degrees of Freedom	62	
	t Stat	4.23	
	p value	0.00006	

F2 has a direct correlation with *backness*, meaning a high value for the acoustic parameter F2-F0 indicates that the segment is more *front*, whereas a low value for F2-F0 indicates that the segment is more *back*. The results in Table 14 show that Marshall's /r/ is more *front* than his /w/ because /r/ has a significantly higher F2-F0 value. This finding is not surprising in itself because one would typically expect an /r/ to have more *front* articulation than a /w/ (which again corresponds with the *high back* vowel [u]).

When the results of Table 13 and Table 14 are considered together, it is clear that Marshall's /r/ is both more *high* and more *front* than his /w/. This is interesting given that both articulations sound the same to listeners. Because his /r/ is more *high* and more *front*, it is similar to a palatal segment, like /j/. Figure 9 below plots the F1 and F2 values for /j/, /w/, and /r/ in a type of graph designed to be comparable to the 'vowel space' charts. It shows that the *height* and *backness* of Marshall's /r/ has more in common with his /j/ than with /w/ and that his /r/ is therefore more palatal-like.

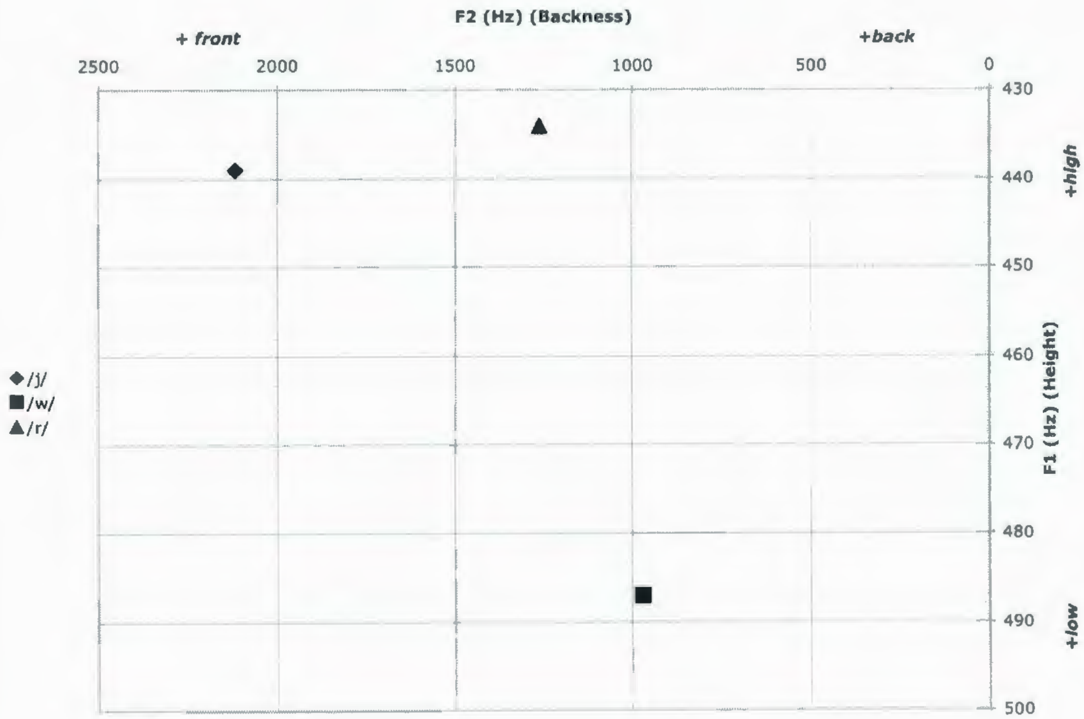

Figure 9: F1 vs. F2 'Vowel Space' graph for /j/, /w/, /r/

Given that Marshall's /r/ is palatal-like, it has much in common with the bunched /r/ which is produced with a more palatal tongue position than retroflex /r/ (refer to §4.2.2). However, further articulatory analysis using technology such as ultrasound equipment would be required to determine the exact articulatory pattern for /r/ used by Marshall.

The results for Marshall's /r/ versus /w/ *t*-test for the acoustic parameter F3-F0 (corresponding to *rhoticity*) are shown below in Table 15.

Table 15: R vs. W Independent *t*-test for F3-F0 (*rhoticity*)

		/r/	/w/
F3-F0 t-test	Mean	2388.4	2512.6
	Variance	67178.5	45175.8
	Standard deviation	259.19	212.55
	Sample size	46	65
	Degrees of Freedom	85	
	t Stat	-2.67	
	p value	0.009	

Because one of the acoustic correlates of *rhoticity* is a lowered F3, the acoustic parameter should negatively correlate with *rhoticity*, in other words, the lower the F3-F0 value, the more *rhotic* the segment. It is clear from the results of the *t*-test in Table 15 that Marshall's /r/ is significantly more *rhotic* than his /w/.

Table 16 below shows the results of the t-test for Marshall's /r/ versus /w/ on the acoustic parameter F3-F2, which also corresponds to the feature *rhotic*.

Table 16: R vs. W Independent *t*-test for F3-F2 (*rhoticity*)

		/r/	/w/
F3-F2 t-test	Mean	1246.9	1675.5
	Variance	239506.9	123585.1
	Standard deviation	489.39	351.55
	Sample size	46	65
	Degrees of Freedom	77	
	t Stat	-5.08	
	p value	0.000003	

Another acoustic correlate for the feature *rhoticity* is a close proximity of F2 and F3; therefore, the acoustic parameter F3-F2 has a negative correlation with *rhoticity*. The lower the value for F3-F2, the more *rhotic* the segment. The results in Table 16 show that

Marshall /r/ is significantly lower for the F3-F2 parameter. Hence, Marshall's /r/ is again more *rhotic* than his /w/.

The results shown in both Table 15 and Table 16 show that Marshall's /r/ is more *rhotic* than his /w/. These results would seem obvious for a normal adult English speaker; however, for Marshall's speech this is anything but obvious: the two segments sound identical to his listeners, yet, they are statistically proven to be acoustically different segments on all measures. Marshall's /r/ and /w/ differ significantly in the formant frequency values corresponding to *height*, *backness*, and *rhoticity*. While Marshall does not produce his /r/ like a normal adult speaker, he definitely produces a different segment for /r/ than for /w/. Therefore, a covert contrast is undeniably present in Marshall's speech. His /r/ is produced as a slightly rhotacized, *high*, *front*, rounded consonant, while his /w/ is produced as a *high*, *back*, rounded consonant, as in adult speech.

8 Discussion

The main hypothesis of this paper was (a) Marshall must have a contrast between /w/ and /r/ in singleton onsets and that (b) since we cannot hear a contrast, it must be covert. The results of this study confirm this hypothesis. Marshall was found to produce a significant difference between /w/ and /r/ in singleton onsets. However, because listeners cannot perceive the difference, the contrast is covert.

Four findings were of note in this paper. First, in §7.1 it was reported that all four of Marshall's approximant consonants were found to be significantly different from one another.

Second, with the exception of /j/, Marshall's approximants were acoustically different than the normal adult productions, as shown in §7.1. Marshall's /w/ had a higher F2 and F3 values, indicating that it may be somewhat less rounded than the adult production. Marshall's /l/ also had a higher F2 and F3 values, pointing toward a more acoustically *front* production than the adult production. Marshall's /r/ had F1 and F2 values very close to that of the adult productions, yet had a drastically higher F3 value.

Third, in §7.2 I discussed that by determining Marshall's neutral F3 value and using it to determine his overall F3 pattern for his approximants productions, Marshall's approximant consonants were, in fact, more similar to the adult normal productions than it first appeared. Marshall's F3 was consistently higher than the adults' production, especially for /r/, when the values were directly compared to the adults'. However, when his neutral F3 value was calculated, I found that Marshall was exhibiting an adult-like pattern of F3 lowering in his approximants. Marshall lowered his F3 value more for /r/ than for any other approximant F3; /r/ was followed by /w/, then /l/, then /j/ as would be expected in normal adult production. This indicates that although Marshall's productions are different from normal adult F3 measurements in their value they are displaying near-normal F3 proportions.

Finally, it was shown in §7.3 that Marshall's /r/ significantly differed from his /w/ on all measures obtained. Because both sounds were acoustically different on all formant and acoustic parameter measurements there is no doubt that Marshall is producing two acoustically distinct sounds. Also, because these two sounds are undistinguishable to listeners, the contrast between them must be covert. Marshall's /r/ production is certainly different from a normal adult /r/ production; yet, it is also undoubtedly different from his own /w/ production. All of the results discussed above point to the notion that Marshall has acquired the phonemic contrast between /w/ and /r/. However, as of yet, he cannot fully produce and/or coordinate all of the required acoustic cues of the contrast.

As mentioned in §5.2, children with AoS have been found to have more difficulty coordinating complex articulatory gestures than children with phonological disorders (Bahr, 2005). Also, during both normal and disordered acquisition children must learn to coordinate the production and timing of numerous phonetic cues in order for a contrast to be perceived, otherwise the contrast is covert (Scobbie, 1998). The evidence discussed above indicates that Marshall has acquired a consistent phonemic contrast between /r/ and /w/, it also seems that he is unable to properly coordinate the phonetic cues required for the contrast to be perceived by listeners.

Most of the data used in speech therapy diagnosis and treatment is based upon impressionistic transcriptions made by adult speakers. The existence of covert contrasts, such as the one found in this study, reveal that impressionistic transcriptions are not necessarily adequate in assessing disordered speech. They might be missing information relevant to clinical diagnosis and/or treatment (Scobbie et al., 1998). A speaker who simply does not have a particular contrast is quite different from a speaker for whom the contrast is present but covert. The speaker without the covert contrast needs to learn a phonological system (i.e. has a 'deeper' problem); the speaker with a covert contrast has acquired the system, but cannot implement it properly. The two speakers would certainly require different treatment regimens in a clinical situation. As Fangfang et al. (2009) point out, "…children who produce covert contrast have a better prognosis than children who produce no contrast at all" (Fangfang et al, 2009:112). Thus, when a contrast is being acquired, adult impressionistic transcriptions alone may not be sufficient. Transcriptions paired with the use of acoustic analysis to rule out covert contrast could be important in the clinical diagnosis and treatment of speech disorders. This seems especially true in the case of AoS speakers who appear to have a particular difficulty coordinating complex articulatory gestures (Bahr, 2005). In this case, acoustical analysis paired with transcriptions would help to objectively describe speech while limiting the perceptual bias.

The data in the present studies shows that Marshall is having articulatory difficulties with /r/ (and also other approximants). The acoustic measurements analysed here reveal what Marshall is doing articulatorily when he produces approximant consonants. This articulatory knowledge would serve well in a clinical setting in planning remediation strategies.

Covert contrast has largely been studied for stop consonants (Macken and Barton, 1979; Forrest et al., 1990; Gibbon, 1990; Gibbon at al., 1993; Edwards et al., 1997; Scobbie et al., 1998). As this paper shows, further research into covert contrast in different sound classes and different speaker types is needed in order to fully understand the role that covert contrast plays in both typical and disordered speech development.

Further research is also required in the area of AoS. The evidence obtained in the current study would seem to likely support the theory that AoS is an impairment of the motor programming function. Because Marshall is in fact producing a contrast between /r/ and /w/, he must have two distinct representative phonemes in his phonological system, yet he is unable to produce all of the required aspects of the contrast in an adult-like manner. More research of the type presented in this paper is still required to determine the precise underlying nature of the disorder. For example, the current study analysed only one type of syllabic environment (CVC), one vocalic environment (CɛC), one class of consonants (approximants), and one AoS speaker. This has certainly limited the scope of any conclusions that can be drawn from the evidence presented. However, the evidence does show that more research in this area is warranted. Future research of covert contrast in AoS should include longitudinal data, differing syllabic and vocalic environments, and different classes of consonants before solid conclusions about the deficit of AoS can be made. As well, more research and analysis of the data presented in the current study is also warranted. Particularly, to determine why /r/ is always perceived as [w] and not another approximant, such as [l] which has similar acoustic correlates to /r/.

Determining the nature of the deficit caused by AoS will certainly be important for the diagnosis and treatment of AoS in clinical settings. Additional studies of possible covert contrast in AoS would be beneficial to this pursuit because they would enhance the information obtained while removing much of the perceptual bias that is possible when dealing with covert contrast. As well, learning more about covert contrast and AoS will add to our understanding of both normal and disordered speech acquisition.

References

Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. Brain and Language, 88, 148-159.

Bahr, R. H. (2005). Differential Diagnosis of Severe Speech Disorders Using Speech Gestures. Topics in Language Disorders, 25(3), 254-265.

Bernhardt, B. H., & Stemberger, J. P. (1998). Handbook of Phonological Development: From the Perspective of Constraint-Based Nonlinear Phonology. San Diego: Academic Press.

Borden, G. J., Harris, K. S., & Raphael, L. J. (2003). Speech Science Primer: Physiology, Acoustics, and Perception of Speech (4th ed.). Philadelphia: Lippincott Williams & Wilkins.

Darley, F. L. (1969 unpublished). Input and Output Disturbances in Speech and Language Processing. Paper presented at the Paper Presented to the American Speech and Hearing Association, Chicago, IL.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). Motor speech disorders (1st ed.). Philadelphia: Saunders.

Delattre, P., & Freeman, D. C. (1968). A Dialect Study of American R's by X-Ray Motion Picture. Linguistics, 44, 29-68.

Edwards, H. T. (2003). Applied Phonetcis: The Sounds of American English (2nd ed.). San Diego, CA: Singular.

Edwards, J., Fourakis, M., Beckman, M. E., & Fox, R. A. (1999). Characterizing Knowledge Deficits in Phonological Disorders. Journal of Speech, Language, and Hearing Research, 42, 169-186.

Edwards, J., Gibbon, F., & Fourakis, M. (1997). On Discrete Changes in the Acquisition of the Alveolar/Velar Stop Consonant Contrast. Language and Speech, 40(2), 203-210.

Espy-Wilson, C. Y. (1992). Acoustic Measures for Linguistic Features Distinguishing the Semivowels /w j r l/ in American English. The Journal of the Acoustical Society of America, 92(2), 736-757.

Fabbro, F. (1999). Concise Encyclopedia of Language Pathology. Oxford: Elsevier Science Ltd.

Fangfang, L., Edwards, J., & Beckman, M. E. (2009). Contrast and Covert Contrast: The Phonetic Development of Voiceless Sibilant Fricatives in English and Japanese Toddlers. Journal of Phonetics, 37, 111-124.

Forrest, K., Weismer, G., Hodge, M., Dinnsen, D. A., & Elbert, M. (1990). Statistical Analysis of Word-Initial /k/ and /t/ Produced by Normal and Phonologically Disordered Children. Clinical Linguistics and Phonetics, 4, 327-340.

Gibbon, F. (1990). Lingual Activity in two Speech-Disordered Children's Attempts to Produce Velar and Alveolar Stop Consonants: Evidence from Electropalatographic (EPG) data. British Journal of Disorders of Communication, 25, 329-340.

Gibbon, F., Dent, H., & Hardcastle, W. (1993). Diagnosis and Therapy of Abnormal Alveolar Stops in a Speech-Disordered Child Using Electropalatography. Clinical Linguistics and Phonetics, 7(4), 247-267.

Glass, G. G., & Glass, E. W. (1976). Glass-Analysis for Decoding Only: Teachers Guide. Garden City, NY: Easier to Learn, Inc.

Hagiwara, R. (1995). Acoustic Realizations of American /R/ as Produced by Women and Men. UCLA Working Papers in Phonetics, 90, 1-187.

Jacks, A., Marquardt, T. P., & Davis, B. (2006). Consonant and Syllable Structure Patterns in Childhood Apraxia of Speech: Developmental Change in Three Children. Journal of Communication Disorders, 39, 424-441.

Kent, R. D., & Read, C. (1992). The Acoustic Analysis of Speech. San Diego: Singular Publishing Group, Inc.

Ladefoged, P. (2005). Vowels and Consonants: An Introduction to the Sounds of Languages (Second Edition ed.). Malden, Massachusetts: Blackwell Publishing.

Ladefoged, P., & Maddieson, I. (1996). The Sounds of the World's Languages. Cambridge, Massachusetts: Blackwell Publishing.

Maas, E., Robin, D. A., Wright, D. L., & Ballard, K. J. (2008). Motor Programming in Apraxia of Speech. Brain and Language, 106, 107-118.

Macken, M. A., & Barton, D. (1979). A Longitudinal Study of the Acquisition of the Voicing Contrast in American-English word-initial stops, as Measured by Voice Onset Time. Journal of Child Language, 7, 41-74.

Paddock, H. (Ed.). (1982). Newfoundland Dialects of English (Second Edition ed.). St. John's, Newfoundland: Department of Linguistics, Memorial University.

Peter, B., & Stoel-Gammon, C. (2008). Central Timing Deficits in Subtypes of Primary Speech Disorders. Clinical Linguistics and Phonetics, 22(3), 171-198.

Scobbie, J. M., Gibbon, F., Hardcastle, W. J., & Fletcher, P. (1998). Covert Contrast and the Acquisition of Phonetics and Phonology. In W. Ziegler & K. Deger (Eds.), Clinical Phonetics and Linguistics (pp. 147-156). London: Whurr Publishers.

Shriberg, L. D., Aram, D. M., & Kwiatkowski, J. (1997). Developmental Apraxia of Speech: I. Descriptive and Theoretical Perspectives. Journal of Speech, Language, and Hearing Research, 40, 273-285.

Sproat, R., & Fujimura, O. (1993). Allophonic Variation in English /l/ and its Implications For Phonetic Implementation. Journal of Phonetics, 21, 291-311.

Stevens, K. N. (1998). Acoustic Phonetics. Cambridge: Massachusetts Institute of Technology.

Willins, G. (In progress). The Effects of Orthography on Apraxia of Speech. Unpublished Master of Arts, Memorial University, St. John's.

Wray, A., Trott, K., & Aileen, B. (1998). Projects in Linguistics: A Practical Guide to Researching Linguistics. New York: Oxford University Press.

Appendix A

Below is the complete inventory of CɛC words extracted from Marshall's recorded tutorial sessions. These words all contain singleton approximant onsets. The manner in which this data was extracted and analysed is discussed in §6.5.

Legend:
Word= the word token extracted from recordings of Marshall's speech
Syllable= the syllable type
Rec. Date= the date of the recorded tutorial session from which the word was extracted
nw= nonsense word[7] (not a real word but one that Marshall made up)

/jɛC/ Word Inventory

Word	Syllable	Rec. Date
Yed (nw)	CɛC	5/4/04
Yes	CɛC	2/12/04
Yes	CɛC	2/12/04
Yes	CɛC	2/12/04
Yes	CɛC	2/12/04
Yes	CɛC	2/12/04
Yes	CɛC	2/12/04
Yes	CɛC	2/17/04
Yes	CɛC	3/9/04
Yes	CɛC	3/9/04
Yes	CɛC	3/9/04
Yes	CɛC	3/9/04
Yes	CɛC	3/11/04

[7] There are three nonsense words contained in the data that have either /r/ or /w/ in the onset position (two with /r/ and one with /w/). I determined which segment Marshall was intending to produce based on the context in which the word occurred. All nonsense words occurred during reading drills when Marshall came upon a word that was difficult for him. He would guess the word, often times making up a nonsense word in the process, but he would include the correct onset in his guesses. Willins always spoke the correct target word after Marshall's guesses, therefore making it possible to determine Marshall's target word while listening to the recordings.

Word	Syllable	Rec. Date
Yes	CεC	3/11/04
Yes	CεC	3/11/04
Yes	CεC	3/11/04
Yes	CεC	3/11/04
Yes	CεC	3/11/04
Yes	CεC	3/11/04
Yes	CεC	3/30/04
Yes	CεC	3/30/04
Yes	CεC	3/30/04
Yes	CεC	3/30/04
Yes	CεC	3/30/04
Yes	CεC	3/30/04
Yes	CεC	4/13/04
Yes	CεC	4/13/04
Yes	CεC	4/13/04
Yes	CεC	4/13/04
Yes	CεC	4/20/04
Yes	CεC	4/20/04
Yes	CεC	4/20/04
Yes	CεC	4/29/04
Yes	CεC	4/29/04
Yes	CεC	4/29/04
Yes	CεC	5/4/04
Yes	CεC	5/20/04
Yes	CεC	5/20/04
Yes	CεC	5/20/04
Yes	CεC	5/20/04
Yet	CεC	2/12/04
Yet	CεC	3/11/04
Yet	CεC	3/11/04
Yet	CεC	3/30/04
Yet	CεC	4/29/04
Yet	CεC	5/4/04
Yet	CεC	5/4/04
Yet	CεC	5/4/04

Word	Syllable	Rec. Date
Yet	CεC	5/4/04
Yet	CεC	5/4/04
Yet	CεC	5/4/04
Yet	CεC	5/4/04
Yet	CεC	5/4/04
Yet	CεC	5/4/04
Yet	CεC	5/4/04
Yet	CεC	5/4/04
Yet	CεC	5/4/04
Yet	CεC	5/4/04
Yet	CεC	5/4/04
Yet	CεC	5/20/04
Yet	CεC	5/20/04
Yet	CεC	5/20/04

/lεC/ Word Inventory

Word	Syllable	Rec. Date
Ledge	CεC	5/20/04
Leg	CεC	3/30/04
Leg	CεC	3/30/04
Leg	CεC	4/13/04
Leg	CεC	4/13/04
Leg	CεC	4/13/04
Leg	CεC	4/20/04
Leg	CεC	4/29/04
Leg	CεC	4/29/04
Leg	CεC	4/29/04
Leg	CεC	4/29/04
Lep (nw)	CεC	3/9/04
Lep (nw)	CεC	3/9/04
Lep (nw)	CεC	3/9/04
Lep (nw)	CεC	3/11/04
Less	CεC	3/11/04
Let	CεC	2/10/04

Word	Syllable	Rec. Date
Let	CεC	2/10/04
Let	CεC	2/10/04
Let	CεC	2/17/04
Let	CεC	2/17/04
Let	CεC	2/17/04
Let	CεC	2/17/04
Let	CεC	2/17/04
Let	CεC	2/17/04
Let	CεC	2/17/04
Let	CεC	2/17/04
Let	CεC	3/9/04
Let	CεC	3/9/04
Let	CεC	3/11/04
Let	CεC	3/11/04
Let	CεC	3/11/04
Let	CεC	3/11/04
Let	CεC	3/11/04
Let	CεC	3/11/04
Let	CεC	3/11/04
Let	CεC	3/11/04
Let	CεC	3/11/04
Let	CεC	3/30/04
Let	CεC	3/30/04
Let	CεC	3/30/04
Let	CεC	3/30/04
Let	CεC	3/30/04
Let	CεC	3/30/04
Let	CεC	3/30/04
Let	CεC	3/30/04
Let	CεC	3/30/04
Let	CεC	3/30/04
Let	CεC	4/13/04
Let	CεC	4/20/04
Let	CεC	4/20/04
Let	CεC	4/20/04
Let	CεC	4/29/04

Word	Syllable	Rec. Date
Let	CεC	5/4/04
Let	CεC	5/4/04
Let	CεC	5/4/04
Let	CεC	5/20/04

/rεC/ Word Inventory

Word	Syllable	Rec. Date
Red	CεC	3/30/04
Red	CεC	3/30/04
Red	CεC	3/30/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04

Word	Syllable	Rec. Date
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/13/04
Red	CεC	4/29/04
Red	CεC	4/29/04
Red	CεC	4/29/04
Red	CεC	4/29/04
Red	CεC	5/4/04
Red	CεC	5/4/04
Red	CεC	5/4/04
Red	CεC	5/4/04
Red	CεC	5/4/04
Red	CεC	5/4/04
Red	CεC	5/4/04
Red	CεC	5/4/04
Red	CεC	5/4/04
Red	CεC	5/6/04
Red	CεC	5/20/04
Reg (nw)	CεC	5/4/04
Ren (nw)	CεC	5/4/04

/wεC/ Word Inventory

Word	Syllable	Rec. Date
Weck (nw)	CεC	4/13/04
Well	CεC	3/9/04
Well	CεC	3/11/04
Well	CεC	3/11/04
Well	CεC	3/30/04
Well	CεC	3/30/04
Well	CεC	3/30/04
Well	CεC	3/30/04
Well	CεC	3/30/04
Well	CεC	3/30/04

Word	Syllable	Rec. Date
Well	CεC	3/30/04
Well	CεC	4/13/04
Well	CεC	4/13/04
Well	CεC	4/13/04
Well	CεC	4/13/04
Well	CεC	4/20/04
Well	CεC	4/20/04
Well	CεC	4/29/04
Well	CεC	4/29/04
Well	CεC	5/4/04
Well	CεC	5/4/04
Well	CεC	5/4/04
Well	CεC	5/6/04
Well	CεC	5/6/04
Well	CεC	5/20/04
Well	CεC	5/20/04
Well	CεC	5/20/04
Well	CεC	5/20/04
Well	CεC	5/20/04
Well	CεC	5/20/04
Well	CεC	5/20/04
Wet	CεC	4/13/04
Wet	CεC	4/13/04
Wet	CεC	4/13/04
Wet	CεC	4/20/04
Wet	CεC	5/20/04
When	CεC	3/30/04
When	CεC	3/30/04
When	CεC	3/30/04
When	CεC	3/30/04
When	CεC	4/20/04
When	CεC	4/20/04
When	CεC	4/20/04
When	CεC	4/20/04
When	CεC	4/20/04
When	CεC	4/20/04

Word	Syllable	Rec. Date
When	CɛC	4/20/04
Where	CɛC	3/9/04
Where	CɛC	3/11/04
Where	CɛC	3/11/04
Where	CɛC	3/11/04
Where	CɛC	3/11/04
Where	CɛC	3/11/04
Where	CɛC	3/11/04
Where	CɛC	3/30/04
Where	CɛC	4/20/04
Where	CɛC	4/20/04
Where	CɛC	5/4/04
Where	CɛC	5/4/04
Where	CɛC	5/4/04
Where	CɛC	5/4/04
Where	CɛC	5/4/04
Where	CɛC	5/20/04
Where	CɛC	5/20/04
Where	CɛC	5/20/04
Where	CɛC	5/20/04
Where	CɛC	5/20/04
Where	CɛC	5/20/04
Where	CɛC	5/20/04

Appendix B

Below is the inventory of CVC words used to measure Marshall's neutral F3 value as discussed in §7.2. As well, the individual F3 measurements for each word and the mean F3 value for each vowel and the entire inventory are also presented.

Legend:

Vowel= the vowel contained in the CVC word
Word= the token that was extracted from the recordings for measurement
F3 (Hz)= the F3 frequency in Hz measured for the given token
Mean (Hz)= the mean of the measured F3 frequencies for that particular vowel (i.e. /i/ or /e/ or /ɑ/, etc.)
Total F3 Mean= the combined mean of all the measured F3 frequencies for the inventory

Vowel	Word	Rec. Date	F3 (Hz)	Mean (Hz)
i	geek	03/09/04	3617.85	3263.98
	feet	03/09/04	3292.27	
	beet	03/09/04	3204.23	
	peas	02/17/04	3111.25	
	jeep	02/17/04	2888.98	
	deep	03/09/04	3556.47	
	deep	03/09/04	3459.68	
	teat	03/09/04	3263.80	
	seat	03/09/04	3367.57	
	meat	03/09/04	2877.70	

Vowel	Word	Rec. Date	F3 (Hz)	Mean (Hz)
e	take	03/09/04	3077.65	2922.80
	bake	03/09/04	2854.14	
	cake	03/09/04	2890.29	
	fake	03/09/04	2611.05	
	lake	03/09/04	2898.12	
	pain	03/09/04	3309.10	
	pain	03/09/04	2685.74	
	pain	03/09/04	2975.86	
	gate	02/17/04	3069.91	
	bate	03/09/04	2856.17	
ɑ	knock	03/09/04	2631.82	2744.10
	knock	03/09/04	2813.14	
	bawk	03/09/04	2659.01	
	pock	03/09/04	2583.27	
	dock	03/09/04	2865.50	
	sock	02/17/04	2965.25	
	what	03/09/04	2463.94	
	what	03/09/04	2785.26	
	dog	02/17/04	2826.18	
	ball	02/17/04	2847.61	

Vowel	Word	Rec. Date	F3 (Hz)	Mean (Hz)
o	goat	03/09/04	2865.72	2882.99
	goat	03/09/04	2685.92	
	note	03/09/04	2997.46	
	sote	03/09/04	2821.23	
	boat	03/09/04	3219.92	
	tote	03/09/04	2903.10	
	tote	03/09/04	2884.36	
	pole	03/09/04	2956.43	
	poke	03/09/04	2791.93	
	doke	03/09/04	2703.79	
u	soup	03/09/04	2642.75	2767.72
	soup	03/09/04	2959.87	
	soup	03/09/04	2718.09	
	poop	03/09/04	2684.82	
	dupe	03/09/04	2762.59	
	coop	03/09/04	2444.57	
	coop	03/09/04	2741.84	
	boom	03/09/04	2962.08	
	boom	03/09/04	3063.36	
	boom	03/09/04	2697.21	
Total F3 Mean	2916.32			

Appendix C

Results of the /r/ vs. /w/ formant *t*-Tests, discussed in §7.3. With the exception of F0 (as expected), all *t*-Test results found a statistically significant difference between /r/ and /w/.

/r/ vs. /w/ *t*-Test for F0

		/r/	**/w/**
F0 t-test	Mean	120.35	133.16
	Variance	979.91	1491.47
	Standard deviation	31.3	38.62
	Sample size	46	65
	Degrees of Freedom	107	
	t Stat	-1.93	
	p value	0.057	

/r/ vs. /w/ *t*-Test for F1

		/r/	**/w/**
F1 t-test	Mean	433.8	487.2
	Variance	11519.85	7151.06
	Standard deviation	107.33	84.56
	Sample size	46	65
	Degrees of Freedom	82	
	t Stat	-2.81	
	p value	0.006	

/r/ vs. /w/ *t*-Tests for F2

		/r/	/w/
F2 t-test	Mean	1261.8	970.2
	Variance	272647.7	78211.2
	Standard deviation	522.2	279.7
	Sample size	46	65
	Degrees of Freedom	63	
	t Stat	3.45	
	p value	0.001	

/r/ vs. /w/ *t*-Tests for F3

		/r/	/w/
F3 t-test	Mean	2508.8	2645.7
	Variance	71442.9	43118.2
	Standard deviation	267.3	207.6
	Sample size	46	65
	Degrees of Freedom	81	
	t Stat	-2.91	
	p value	0.005	